



**HAL**  
open science

**Étude des processus de biogenèse des petites particules ribonucléoprotéiques nucléolaires à boîtes C/D (snoRNP C/D) chez la levure *Saccharomyces cerevisiae* : caractérisation fonctionnelle et structurale d'une machinerie dédiée à l'assemblage de ces RNP**

Benjamin Rothé

► **To cite this version:**

Benjamin Rothé. Étude des processus de biogenèse des petites particules ribonucléoprotéiques nucléolaires à boîtes C/D (snoRNP C/D) chez la levure *Saccharomyces cerevisiae* : caractérisation fonctionnelle et structurale d'une machinerie dédiée à l'assemblage de ces RNP. Médecine humaine et pathologie. Université de Lorraine, 2012. Français. NNT : 2012LORR0013 . tel-01749161

**HAL Id: tel-01749161**

**<https://hal.univ-lorraine.fr/tel-01749161v1>**

Submitted on 29 Mar 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : [ddoc-theses-contact@univ-lorraine.fr](mailto:ddoc-theses-contact@univ-lorraine.fr)

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

[http://www.cfcopies.com/V2/leg/leg\\_droi.php](http://www.cfcopies.com/V2/leg/leg_droi.php)

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

**Ecole Doctorale BioSE (Biologie-Santé-Environnement)**

Thèse

**Présentée et soutenue publiquement pour l'obtention du titre de  
DOCTEUR DE L'UNIVERSITE DE LORRAINE  
Mention : « Sciences de la Vie et de la Santé »**

par **Benjamin ROTHÉ**

**Étude des processus de biogenèse des petites particules  
ribonucléoprotéiques nucléolaires à boîtes C/D (snoRNP C/D) chez la  
levure *Saccharomyces cerevisiae*. Caractérisation fonctionnelle et  
structurale d'une machinerie dédiée à l'assemblage de ces RNP.**

Le 30 mars 2012

**Membres du jury :**

**Rapporteurs :**

M. Denis LAFONTAINE	Professeur, Université Libre de Bruxelles, Belgique
M. Yves HENRY	Directeur de Recherche, CNRS, Toulouse

**Examineurs :**

M. Edouard BERTRAND	Directeur de Recherche, CNRS, Montpellier
M. Cosmin SAVEANU	Chargé de Recherche, Institut Pasteur, Paris
M <sup>me</sup> Christiane BRANLANT	Directeur de Recherche, CNRS, Nancy
M. Bruno CHARPENTIER	Professeur, Université Henri Poincaré, Nancy Directeur de thèse



## ***Remerciements***

Je tiens à exprimer ma profonde gratitude à Madame Christiane Branlant pour m'avoir accueilli au sein de son laboratoire et permis de réaliser ce travail dans les meilleures conditions. Je la remercie vivement de m'avoir fait part de son expérience et de ses conseils.

Je remercie tout particulièrement Bruno Charpentier, mon directeur de thèse, qui m'a formé, guidé et soutenu de mes débuts au laboratoire, jusqu'à l'aboutissement de cette thèse. C'est à travers ses cours que j'ai découvert la biologie moléculaire et que j'ai appris à aimer cette discipline. Je le remercie pour son enthousiasme et pour sa grande disponibilité. Nos nombreuses discussions et notre entente ont été des éléments clés de ma réussite. Travailler avec lui pendant toutes ces années aura été un honneur et un grand plaisir. Merci pour tout.

Je suis très sensible à l'honneur que me font Monsieur Denis Lafontaine et Monsieur Yves Henry d'avoir accepté d'être rapporteurs de cette thèse, je les en remercie sincèrement. Je remercie aussi vivement Monsieur Cosmin Saveanu et Monsieur Edouard Bertrand qui ont accepté d'examiner mon travail.

Je remercie également tous nos collaborateurs, avec lesquels j'ai eu si souvent l'occasion de discuter et d'échanger. Je pense notamment aux membres des équipes d'Edouard Bertrand, d'Alain Krol, de Solange Moréra, de Cosmin Saveanu, d'Alain Van Dorssealer et Sarah Sanglier-Cianferani. Un merci tout particulier à Jean-Michel Saliou pour sa grande contribution à ce projet. La collaboration que nous avons menée ensemble aura été une belle réussite.

Mon passage au laboratoire restera indissociable de l'estime et de l'amitié que j'ai pour Régis. Nous avons travaillé en parallèle sur le même sujet depuis notre stage de M2R. Notre entente et notre complicité aura été un fil conducteur essentiel de cette aventure. J'ai beaucoup apprécié travailler avec lui et je lui souhaite beaucoup de réussite pour la suite.

Je remercie vivement Xavier pour son amitié et pour le rôle important qu'il a joué pendant ma thèse. Nos longues discussions et nos débats passionnés ont été un immense enrichissement et ont fait naître en moi un grand intérêt pour la biologie structurale. Je lui suis très reconnaissant pour son écoute, pour ses conseils et pour sa bonne humeur communicative. Je n'oublierai pas non plus toutes les belles aventures que nous avons partagées, de Göttingen, à Porto-Vecchio, en passant par *Ciottulu di i Mori* ! Alors en attendant une nouvelle escapade, je lui adresse un grand merci. *Entre nous : "J'espère que ce topoguide te permettra, par la suite, de doubler quelques étapes..."*.

Je remercie également Nathalie Marmier-Gourrier avec qui j'ai travaillé lors de mon M2R. Je suis très reconnaissant pour tout le temps qu'elle m'a consacré et pour tout ce qu'elle m'a transmis.

Un grand merci à Christine avec qui j'ai eu le grand plaisir de travailler pendant une partie de ma thèse et qui m'a beaucoup apporté. Je la remercie pour sa grande motivation, pour sa gentillesse et pour sa bonne humeur.

Merci également à toutes les personnes du Box3 qui ont travaillé sur ce projet, je pense notamment à Claude, Magali, Matthieu et Clémence, qui avec Régis et Xavier, m'ont si souvent fait profiter de leur expérience dans la purification de protéines recombinantes et sans qui de nombreuses manipulations n'auraient pas pu voir le jour !

Un grand merci à Véronique Senty-Ségault pour tout le temps qu'elle m'a accordé et pour avoir si souvent pris le temps de répondre à mes innombrables questions. J'ai beaucoup appris à ses côtés et je lui en suis très reconnaissant. Merci aussi pour toutes nos discussions et pour les foudres chroniques que nous avons partagés.

Merci également à Nicolas avec qui j'ai partagé tant de débuts de week-end... Nos discussions et notre collaboration ont été très agréables.

Un immense merci à Aileen pour son investissement dans la correction de ce manuscrit. Sa rigueur, son expérience et ses conseils ont été précieux. Merci beaucoup.

Je remercie également Decebal qui a repris le flambeau et qui m'a beaucoup aidé à la fin de ma thèse. Merci et bonne chance pour la suite !

Je tiens également à remercier Valérie pour sa présence, pour sa bienveillance et pour sa bonne humeur. Travailler à ses côtés aura été un vrai plaisir.

Je veux également exprimer ma reconnaissance à toutes les personnes du Box1 que je n'ai pas encore remerciées. Je pense notamment à Thomas, Véronique, Guillaume, Jean-Baptiste, Marie-Christine. La bonne ambiance et l'esprit de partage qui régnaient au Box1 était un élément moteur de mon quotidien et je vous remercie pour tous ces bons moments partagés.

J'ai aussi une pensée amicale pour Sylvain avec qui j'ai eu maintes fois l'occasion de partager des discussions passionnantes.

Je remercie également tous les membres du laboratoire pour leur gentillesse et pour leur grande disponibilité. Merci à tous d'avoir toujours su m'accueillir, me faire part de votre expérience et répondre à mes questions.

Enfin, je ne saurais terminer sans avoir une pensée émue pour Jackie, que j'ai eu beaucoup de plaisir à côtoyer lors de mes débuts au laboratoire et qui nous manque beaucoup.

Ce long travail et tout cet investissement n'auraient bien-entendu pas été possibles sans le soutien sans faille de ma famille et de mes amis, et je tiens à leur faire part de ma reconnaissance.

Je remercie en particulier William et Greg qui ont été des confidents attentifs et qui m'ont souvent aidé à prendre les bonnes décisions. Leur amitié m'est précieuse.

Je remercie chaleureusement toute ma belle famille qui n'a cessé de me soutenir et de prier pour ma réussite. Merci à Driss, Fatiha, Meriem, Nabila, Ali, Khreïreddine, Noureddine et à tous les autres. Merci aussi à Jean-Luc et à Dalila qui nous manque tellement. *Choukrane jazilane lakoum wa baraka Allahou fikoum.*

Un immense merci à toute ma famille dont la présence à mes côtés est essentielle. Merci à mes parents Paule et Patrick, mes grands-parents Liliane, Marcel, Marie et Daniel, ma sœur Tiphaine, mon frère Antonin, Tonton Pierre, et Christopher, mon ami de toujours. Je ne pourrai jamais assez vous remercier pour votre soutien et pour vos encouragements. Vous êtes pour beaucoup dans ma réussite et je vous dois tant...

Enfin, je remercie Lamia, ma merveilleuse épouse, pour son soutien indéfectible. Elle a fait preuve de beaucoup de compréhension et m'a permis de surmonter les moments les plus difficiles. Je lui suis infiniment reconnaissant. Merci à toi, ainsi qu'à notre adorable petite Ilyana à qui je dédie cette thèse.

# ABBREVIATIONS





## Liste des abréviations :

3-AT	: 3-aminotriazol
A	: Adénine
A <sub>600nm</sub>	: Absorbance à la longueur d'onde de 600nm
AAA+	: <i>ATPase associated with diverse cellular processes</i>
ADN	: Acide désoxyribonucléique
ADN db	: Acide désoxyribonucléique double brin
ADNc	: Acide désoxyribonucléique complémentaire
APT	: <i>Associated with Pat1p</i>
ARN	: Acide ribonucléique
ARNm	: Acide ribonucléique messager
ARNnc	: ARN non codant
ARNr	: Acide ribonucléique ribosomique
ARNt	: Acide ribonucléique de transfert
ASF	: <i>A site finger</i>
ATP	: Adénosine 5' triphosphate
Bcd1	: <i>Box C/D snoRNA 1</i>
C	: Cytosine
CAB	: <i>Cajal bodies box</i>
CB	: <i>Cajal Bodies</i>
CBC	: <i>Cap Binding Complex</i>
Cbf5	: <i>Centromere binding factor 5</i>
CE	: <i>Capping Enzyme</i>
ChIP	: <i>Chromatin Immunoprecipitation</i>
Ci	: Curie
cpm	: coups par minute
CPSF	: <i>Cleavage and Polyadenylation Specificity Factor</i>
CRAC	: <i>Cross Linking and Analysis of cDNA</i>
CRM1	: <i>Chromosome Region Maintenance 1</i>
CstF	: <i>Cleavage and stimulation Factor</i>
CTD	: <i>Carboxy Terminal Domain</i>
Da, KDa	: Dalton, kilodalton
DAMP	: <i>Decreased abundance by mRNA perturbation</i>
DC	: <i>Dyskeratose Congénitale</i>
DES	: <i>GU rich Downstream Element Sequence</i>
DFC	: <i>Dense Fibrillar Center</i>
DMSO	: diméthylsulfoxyde
DNase	: Désoxyribonucléase
dNTP	: Désoxyribonucléotide triphosphate
DTT	: dithiothréitol
EDTA	: éthylène diamine tétraacétate
ETS	: <i>External Transcribed Spacer</i>
FC	: <i>Fibrillar Center</i>
G	: Guanine
GAR	: <i>Glycine and Arginine Rich</i>
GC	: <i>Granular Components</i>
GEM	: <i>Gemini des Cajal Bodies</i>
GST	: <i>Glutathione S-transferase</i>

## LISTE DES ABRÉVIATIONS

---

HEPES	: acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique
IBP160	: <i>Intron Binding Protein 160</i>
IGC	: <i>Interchromatin Granule Cluster</i>
IGS	: <i>Intergenic Spacer</i>
IPTG	: isopropyl $\beta$ -D-thiogalactoside
ITS	: <i>Intenal Transcribed Spacer</i>
Kan	: Kanamycine
Kb	: Kilobase
lncRNA	: <i>long non coding RNA</i>
Lsm	: <i>Sm Like</i>
mChIP	: <i>modified Chromatin Immunoprecipitation</i>
MOPS	: <i>3-(N-morpholino)propanesulfonic acid</i>
MRP	: <i>Mitochondrial RNA Processing</i>
Nat	: Nourseothricine
NB	: <i>Nucleolar bodies</i>
NES	: <i>Nuclear Export Signal</i>
Nhp2	: <i>Non-Histone Protein 2</i>
NLS	: <i>Nuclear Localization signal</i>
NMD	: <i>Non sens Mediated Decay</i>
NoLS	: <i>Nucleolar Localization Signal</i>
Nop	: <i>Nucleolar protein</i>
NOR	: <i>Nucleolar Organizing Region</i>
nt	: nucléotide
NTP	: Ribonucléotide triphosphate
NTS	: <i>Non Transcribed Spacer</i>
NUF1P	: <i>Nuclear FMRP Interacting Protein</i>
ORF	: <i>Open Reading Frame</i>
PAGE	: <i>polyacrylamide gel electrophoresis</i>
pb	: paire de base
PBS	: <i>Phosphate Buffer Saline</i>
PCR	: <i>Polymerase Chain Reaction</i>
PEG	: Polyéthylène glycol
pGal	: promoteur galactose dépendant
Pih1	: <i>Protein interacting with Hsp90-1</i>
PIKK	: <i>Phosphatidylinositol 3-kinase-related kinase</i>
PML	: <i>ProMyelocytic Leukemia Nuclear Body</i>
PPIase	: <i>Peptidyl-Propyl cis/trans Isomerase</i>
pré-ARNm	: précurseur de l'ARN messager
pré-ARNr	: précurseur de l'ARN ribosomique
PTC	: <i>Peptidyl Transferase center</i>
PUA	: <i>PseudoUridine synthase and Archaeosine transglycosylase</i>
R	: Purine (adénine ou guanine)
R2TP	: Rvb1p/Rvb2p/Tah1p/Pih1p
RBD	: <i>RNA Binding Domain</i>
Rex4	: <i>Rna exonuclease</i>
RMN	: Résonance Magnétique Nucléaire
RNase	: Ribonucléase
RNP	: <i>Ribonucléoparticle</i>
Rrp9p	: <i>Ribosomal RNA Processing 9</i>
Rsa1	: <i>Ribosome assembly 1</i>

## LISTE DES ABRÉVIATIONS

---

RT	: <i>Reverse Transcription</i>
Rtt106p	: <i>Regulator of Ty1 Transposition 106</i>
Rvb1/2	: <i>RuVB-like</i>
S	: <i>Svedberg</i>
SAM	: <i>S-Adenosyl-L-Methionine</i>
sb	: <i>simple brin</i>
SCAF	: <i>SR like CTD Associated Factors</i>
scaRNA	: <i>small Cajal Bodies RNA</i>
SDS	: <i>sodium dodecylsulfate</i>
SECIS	: <i>Selenocystein Insertion Sequence</i>
SELEX	: <i>Systematic Evolution of Ligands by Exponential enrichment</i>
SMA	: <i>Spinal Muscular Atrophy</i>
SMN	: <i>Survival of MotoNeuron</i>
snoRNA	: <i>small nucleolar RNA</i>
snoRNP	: <i>small nucleolar Ribonucleoparticle</i>
snRNA	: <i>small nuclear RNA</i>
snRNP	: <i>small nuclear RNP</i>
Snu13	: <i>small Nuclear ribonucleoprotein associated</i>
SR	: <i>Serine/Arginine Rich</i>
sRNA	: <i>small RNA</i>
sRNP	: <i>small RNP</i>
SSC	: <i>Saline-Sodium Citrate</i>
SSPE	: <i>Saline-Sodium Phosphate EDTA</i>
SSU	: <i>Small Subunit</i>
T	: <i>Thymine</i>
Tah1	: <i>TPR containing protein asociated with Hsp90</i>
TBE	: <i>Tris-Borate-EDTA</i>
TBP	: <i>TATA box Binding Protein</i>
TE	: <i>Tris-EDTA</i>
TGS1	: <i>TrimethylGuanosine Synthase 1</i>
TGS1-LF	: <i>TGS1-Long Form</i>
TGS1-SF	: <i>TGS1-Short Form</i>
TIP48/49	: <i>TBP Interacting Protein 48/49</i>
TPR	: <i>Tetratricopeptid Repeat</i>
TRAMP	: <i>Trf4p/Air2p/Mtr4p Polyadenylation</i>
t-UTP	: <i>U Three Protein Complex required for transcription</i>
UAS	: <i>Upstream activating sequences</i>
UE	: <i>Unité Enzymatique</i>
UsnRNA	: <i>Uridine rich small nuclear ribonucleic acid</i>
UsnRNP	: <i>Uridine rich small nuclear ribonucleoprotein particle</i>
Utp	: <i>U-three protein</i>
UTR	: <i>Untranslated region</i>
Y	: <i>Pyrimidine (cytosine, uracile ou thymine)</i>
Y2H/Y3H	: <i>Yeast two-hybrid/ Yeast three-hybrid</i>
YPD	: <i>Yeast extract Peptone Dextrose</i>
YPG	: <i>Yeast extract Peptone Galactose</i>
Ψ	: <i>Pseudouridine</i>

**Les abréviations internationales utilisées pour désigner les acides aminés (nomenclature à 3 lettres ou à une lettre) ne sont pas incluse dans cette liste.**

### ***N.B.* Nomenclature des protéines :**

Concernant les noms de protéines, nous avons respecté la nomenclature communément utilisée dans chaque organisme. Pour les protéines de levure seule la première lettre est en majuscule et le nom est suivi d'un « p » pour « *protein* » (*e.g.* Rsa1p). Pour les protéines humaines l'intégralité du nom apparaît généralement en lettre capitales (*e.g.* NUFIP). La séparation de deux noms de protéines par le symbole « / » fait simultanément référence à une protéine de levure et à son homologue humaine (*e.g.* Rsa1p/NUFIP). Les noms de protéines séparés par le symbole « - » fait référence à un complexe formé entre ces protéines (*e.g.* Snu13p-Rsa1p). Enfin, le nom des ORF de levure correspond au nom de la protéine en majuscule et en italique (*e.g.* *RSAl*).

**LISTE DES PUBLICATIONS  
SCIENTIFIQUES**



## Liste des publications scientifiques :

### **The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery**

Séverine Boulon , Nathalie Marmier-Gourrier, Bérengère Pradet-Balade, Laurence Wurth, Céline Verheggen, Beáta E. Jády, Benjamin Rothé, Christina Pescia, Marie-Cécile Robert, Tamás Kiss, Barbara Bardoni, Alain Krol, Christiane Branlant, Christine Allmang, Edouard Bertrand and Bruno Charpentier

*The journal of cell biology*, 2008 Feb 11;180(3):579-95  
PMID: 18268104

### **Formation of electrostatic and hydrophobic interactions between alpha helices of the Rsa1p assembly factor and the Snu13p snoRNP protein is required for efficient snoRNP assembly**

Benjamin Rothé\*, Régis Back\*, Xavier Manival, Marie-Cécile Robert, Magali Blaud, Christophe Romier, Michael Leichter, Bernard Maigret, Bruno Charpentier, Edouard Bertrand, and Christiane Branlant  
(\*contribution équivalente)

En préparation pour la revue *Nucleic Acid Research*

### **Functional link between proteins Rsa1 and Hit1 in yeast box C/D snoRNPs biogenesis**

Benjamin Rothé, Jean-Michel Saliou, Régis Back, Xavier Manival, Alain Van Dorssealer, Sarah Sanglier Cianferani, Christiane Branlant and Bruno Charpentier

En préparation pour la revue *Molecular and Cellular Biology*

### **The NMR high resolution structures of yeast Tah1 in a free form and in complex with the Hsp90 Cterminal tail explains the specific recognition of Hsp90 by the R2TP complex**

Régis Back, Cyril Dominguez, Benjamin Rothé, Claude Bobo, Chrystel Beauvils, Solange Moréra, Philippe Meyer, Bruno Charpentier, Christiane Branlant, Frédéric Allain and Xavier Manival

En préparation pour la revue *Structure*





**LISTE DES  
COMMUNICATIONS  
SCIENTIFIQUES**



**Liste des communications scientifiques :**

- 2010 :** **Etude du rôle des protéines Rsa1p, Hit1p et Pih1p dans la biogenèse des snoRNP à boîtes C/D chez *Saccharomyces cerevisiae***
- Benjamin Rothé\*, Régis Back, Xavier Manival, Christophe Romier, Edouard Bertrand, Christiane Branlant & Bruno Charpentier
- Exposé oral\*  
8<sup>ème</sup> Rencontre SifrARN, Dourdan 22-24 novembre 2010
- + Prix de la meilleure communication orale décerné par la Société Française de Biochimie et de Biologie Moléculaire (attestation jointe)
- 2009 :** **Rsa1p interacts with helix  $\alpha$ 3 of Snu13p: identification of the aminoacids involved in this interaction**
- Benjamin Rothé\*, Régis Back, Bernard Maigret, Xavier Manival, Edouard Bertrand, Christiane Branlant & Bruno Charpentier
- Exposé par affiche\*  
8<sup>th</sup> International Conference on Ribosome Synthesis, Regensburg, Allemagne, 26-30 août 2009
- 2008 :** **Décryptage des mécanismes d'assemblage des snoRNP à boîtes C/D chez *Saccharomyces cerevisiae***
- Benjamin Rothé\*, Régis Back, Xavier Manival, Nicolas Rolland, Véronique Senty-Ségault, Bérangère Pradet-Balade, Céline Verheggen, Edouard Bertrand, Christiane Branlant & Bruno Charpentier
- Exposé oral\*  
Colloque Levure, Modèle et Outils VII, La Colle sur Loup, 26 octobre-30 octobre 2008



SOCIÉTÉ FRANÇAISE DE BIOCHIMIE  
ET BIOLOGIE MOLECULAIRE

Association loi 1901 créée en 1914  
reconnue d'utilité publique (décret du 27/04/1933)

**PRIX DE LA MEILLEURE COMMUNICATION ORALE**

---

Attribué par La SOCIÉTÉ FRANÇAISE DE BIOCHIMIE ET BIOLOGIE MOLÉCULAIRE

à Benjamin ROTHÉ

Pour sa présentation à la 8<sup>ème</sup> Rencontre SifrARN de la SFBBM, du 22 au 24 novembre 2010 dans le  
Club Belambra de Dourdan.

Le Président de la SFBBM

A handwritten signature in black ink, appearing to read 'F. Dardele', is written over the printed name.

Frédéric DARDEL

# LISTE DES FIGURES & DES TABLEAUX



## *Liste des figures*

### AVANT-PROPOS

---

**Figure 1 :** Cascade d'événements impliquant des ARNnc dans le processus d'expression des gènes chez les eucaryotes.

### INTRODUCTION - CHAPITRE 1

---

**Figure 2 :** Représentation schématique des principales structures nucléaires rencontrées dans les cellules eucaryotes.

**Figure 3 :** Voies de transport nucléocytoplasmique des principales molécules d'ARN chez les eucaryotes.

**Figure 4 :** Visualisation des speckles et des paraspeckles par microscopie électronique dans des noyaux de cellules HeLa.

**Figure 5 :** Visualisation du nucléole par microscopie électronique dans des cellules HeLa et chez *Saccharomyces cerevisiae*.

**Figure 6 :** Visualisation des Cajal Bodies et des Nucleolar Bodies par microscopie électronique respectivement dans des cellules HeLa et chez *Saccharomyces cerevisiae*.

**Figure 7 :** Visualisation des GEMs, des PML et des Cleavage bodies par microscopie électronique dans des cellules HeLa.

**Figure 8 :** Bilan des phénomènes de couplages entre les différentes étapes de maturation des transcrits de l'ARN polymérase II.

**Figure 9 :** Structures des coiffes spécifiques de l'extrémité 5' des transcrits de l'ARN polymérase II.

**Figure 10 :** Mécanisme d'épissage par le spliceosome majeur.

**Figure 11 :** Voies de maturation des snoRNA à boîtes C/D hébergés dans des introns.

**Figure 12 :** Eléments impliqués dans la formation de l'extrémité 3' des snoRNA à boîtes C/D issus d'unités transcriptionnelles indépendantes chez la levure.

**Figure 13 :** Structure des unités transcriptionnelles codant les ARNr chez la levure *Saccharomyces cerevisiae*.

**Figure 14 :** Représentation des principales modifications retrouvées dans les ARNr eucaryotes.

**Figure 15 :** Localisation des nucléotides modifiés sur les ARNr 25S, 5.8S et 5S de *Saccharomyces cerevisiae*.

**Figure 16 :** Localisation des nucléotides modifiés sur l'ARNr 18S de *Saccharomyces cerevisiae*.

**Figure 17 :** Localisation structurale des nucléotides d'ARNr modifiés chez *Saccharomyces cerevisiae*.

**Figure 18 :** Maturation nucléolytique des ARNr 18S, 25S et 5.8S chez *Saccharomyces*.

**Figure 19 :** Formation co-transcriptionnelle des particules pré-ribosomiques chez *Saccharomyces cerevisiae*.

**Figure 20 :** Principales étapes de l'assemblage des sous-unités ribosomiques chez *Saccharomyces cerevisiae*.

### INTRODUCTION - CHAPITRE 2

---

**Figure 21 :** Homologies fonctionnelles et homologies de séquences au sein de la famille de protéines L7Ae/L30.

**Figure 22 :** Structure du motif en K-turn du snRNA U4 humain.

**Figure 23 :** Structures secondaires des motifs de type K-turn retrouvés dans différents ARN.

**Figure 24 :** Structures tridimensionnelles de la protéine L7Ae libre ou en complexe avec un ARN.

**Figure 25 :** Structures tridimensionnelles de la protéine Snu13p libre et de la protéine 15.5K complexée avec le motif en K-turn du snRNA U4.

**Figure 26 :** Structures tridimensionnelles de la Nhp2p de *S. cerevisiae* libre.

**Figure 27 :** Structure tridimensionnelle de la protéine Rpl30p (alias L30) de levure sous forme libre ou en complexe avec un motif en K-turn.

**Figure 28 :** Arbre phylogénétique des domaines L7Ae/L30-RBD.

**Figure 29 :** Représentation schématique de la di-snRNP U4/U6 humaine.

**Figure 30 :** Organisation structurale de la protéine PRP31 humaine.

**Figure 31 :** Structure tridimensionnelle du domaine 3' de la snRNP U4.

**Figure 32 :** Structure tridimensionnelle du complexe U4/15.5K/PRP31 humain.

**Figure 33 :** Représentation schématique de l'étape d'activation catalytique du spliceosome chez les mammifères.

**Figure 34 :** Représentation schématique d'un snoRNA à boîtes C/D.

**Figure 35 :** Structures tridimensionnelles des protéines Fibrillarine et NOP5.

**Figure 36 :** Représentation schématique des différents modèles proposés pour l'assemblage des snoRNP et des sRNP à boîtes C/D.

**Figure 37 :** Structure tridimensionnelle d'une sRNP à boîte C/D de *Sulfolobus solfataricus* obtenue par radiocristallographie.

**Figure 38 :** Complexe formé entre un motif C/D structuré en K-turn et les protéines L7Ae et NOP5 au sein de la sRNP à boîte C/D de *Sulfolobus solfataricus*.

**Figure 39 :** Vue détaillée du domaine catalytique d'une sRNP à boîte C/D de *Sulfolobus solfataricus*.

**Figure 40 :** Structure secondaire du snoRNA U3 (snR17A) de *Saccharomyces cerevisiae* obtenue par modifications chimiques et enzymatiques en solution.

**Figure 41 :** Modèle d'assemblage de la snoRNPU3 humaine.

**Figure 42 :** Représentation schématique de l'appariement formé entre le snoRNA U3 et le pré-ARNr 35S chez la levure *Saccharomyces cerevisiae*.

**Figure 43 :** Représentation schématique d'un snoRNA à boîtes H/ACA.

**Figure 44 :** Structure tridimensionnelle du complexe formé par les protéines GAR1, CBF5 et NOP10 chez *Pyrococcus furiosus*.

**Figure 45 :** Structure tridimensionnelle d'une sRNP à boîte H/ACA de *Sulfolobus solfataricus* obtenue par radiocristallographie.

### INTRODUCTION - CHAPITRE 3

---

**Figure 46 :** Représentation schématique de la voie de biogenèse des snRNP U1, U2, U4 et U5 chez les animaux.

**Figure 47 :** Représentation schématique du processus d'assemblage des snoRNP à boîtes H/ACA.

**Figure 48 :** Comparaison structurale du domaine (109-232) de la protéine Naf1p de *Saccharomyces cerevisiae* et de la protéine GAR1 de *Pyrococcus furiosus*.

**Figure 49 :** Voie de biogenèse proposée pour la snoRNP U3 chez les vertébrés.

**Figure 50 :** Voie de biogenèse proposée pour la snoRNP U3 chez *Saccharomyces cerevisiae*.

**Figure 51 :** Représentations structurales des AAA+ ATPases Rvb1p/TIP49 et Rvb2p/TIP48.

**Figure 52 :** Structure et fonction des chaperonnes HSP90.

**Figure 53 :** Conservation de séquence des protéines Rsa1p et NUFIP.

### RÉSULTATS & DISCUSSION - CHAPITRE 1

---

**Figure 54 :** Bilan des interactions établies entre Rsa1p et les constituants protéiques des snoRNP à boîtes C/D.

**Figure 55 :** Rsa1p est requise pour l'assemblage optimal des protéines Nop56p et Nop58p sur le complexe U14-Snu13p.

### RÉSULTATS & DISCUSSION - CHAPITRE 2

---

**Figure 56 :** *In cellulo*, le domaine PEP de Rsa1p/NUFIP constitue une interface clé entre la machinerie d'assemblage et les pré-RNP

**Figure 57 :** Chez la levure *Saccharomyces cerevisiae*, Rsa1p n'interagit pas avec la protéine Nhp2p.

**Figure 58 :** Modèle de structure hybride entre une sRNP à boîte C/D archaea et les protéines humaines 15.5K et Fibrillarine.

**Figure 59 :** Implication des résidus C-terminaux de l'hélice  $\alpha_3$  des protéines L7Ae et 15.5K dans la formation des RNP

**Figure 60 :** Conservation des résidus situés à l'interface entre Snu13p/15.5K et les protéines Fibrillarine, NOP56, NOP58 et PRP31.

**Figure 61 :** Comparaison de l'environnement du motif L--ED de Snu13p/15.5K au sein des interfaces d'interaction Rsa1p-Snu13p et 15.5K-FIB-NOP.

**Figure 62 :** Contacts intermoléculaires impliquant l'extrémité C-terminale de l'hélice  $\alpha_3$  de la protéine L7Ae au sein des sRNP à boîtes H/ACA.

**Figure 63 :** Accessibilité des résidus L71, E74 et D75 de la protéine 15.5K au sein du modèle de RNP hybride sRNA-15.5K-hFIB-aNOP5.



**Figure 64 :** Proposition de modèle pour l'assemblage des snoRNP à boîtes CD.

### RÉSULTATS & DISCUSSION - CHAPITRE 3

---

**Figure 65 :** L'accumulation de certains snoRNA à boîtes C/D est indépendante de Rsa1p et Hit1p.

**Figure 66 :** Mise au point de l'approche d'immunoprécipitation de la chromatine à partir du facteur d'assemblage Naf1p.

**Figure 67 :** Les facteurs d'assemblage ne sont pas détectés sur les gènes de snoRNA à boîtes C/D en cours de transcription.

**Figure 69 :** Les facteurs d'assemblage Rsa1p, Pih1p et Bcd1p ne sont pas détectés sur le snoRNA intronique U18 en cours de transcription.

**Figure 70 :** Analyse fonctionnelle de la protéine Pih1p.

**Figure 71 :** Analyse fonctionnelle de la protéine Hit1p.

**Figure 72 :** Les sous complexes Rsa1p<sub>230-375</sub>-Hit1p et Pih1p-Tah1p interagissent avec Nop58p de façon non compétitive.

**Figure 73 :** Implication potentielle de Rsa1p et Hit1p dans des événements de régulation.

### RÉSULTATS & DISCUSSION - CHAPITRE 4

---

**Figure 74 :** Le taux de snoRNA à boîtes C/D n'est pas dépendant de Hsp82p et de sa cochaperone Tah1p.

**Figure 75 :** Le contrôle du taux de protéine Pih1p par Tah1p n'est pas dépendant de Hsp82p.

### RÉSULTATS & DISCUSSION - CHAPITRE 5

---

**Figure 76 :** *In vitro* l'assemblage de la protéine Rrp9p sur le motif B/C du snoRNA U3 est indépendante de la machinerie d'assemblage.

**Figure 77 :** Rsa1p améliore la fixation de la protéine Snu13p sur le motif C/D du snoRNA U3.

**Figure 78 :** Le profil de digestion enzymatique de l'ARN U3 $\Delta$ 2,3,4 est modifié en présence de Rsa1p.

**Figure 79 :** Le domaine 5' de l'ARN U3 $\Delta$ 2,3,4 est requis pour le recrutement des protéines Snu13p et Rsa1p<sub>230-381</sub> au niveau du motif C/D.

**Figure 80 :** Mise au point d'une technique de détection des précurseurs du snoRNA U3 par RT-PCR.

**Figure 81 :** Mise en évidence du caractère précoce de l'épissage du snoRNA U3 chez la levure *S. cerevisiae*.

**Figure 82 :** *In vivo*, Rsa1p semble associée aux pré-snoRNA U3 +12 et +18.

**Figure 83 :** L'absence de Rsa1p n'inhibe pas l'incorporation de Nop58p au sein de la snoRNP U3.

**Figure 84 :** L'interaction entre Rsa1p et Snu13p est requise pour la biogenèse optimale de l'ARN U3 $\Delta$ 2,3,4.

**Figure 84 :** Nouveau modèle de biogenèse proposé pour la snoRNP U3 chez *Saccharomyces cerevisiae*.

**Figure 86 :** Cartographie des sites d'interaction potentiels de la protéine 15.5K avec différents partenaires protéiques.

### RÉSULTATS & DISCUSSION - CHAPITRE 6

---

**Figure 87 :** *In vitro* Bcd1p interagit avec les snoRNA à boîtes C/D, indépendamment de la présence de motifs structurés en K-trun.

**Figure 88 :** Bcd1p interagit avec les ARN de façon peu spécifique.

**Figure 89 :** Le domaine à doigt de zinc de Bcd1p est nécessaire mais non suffisant pour interagir avec l'ARN.

**Figure 90 :** Recherche des partenaires de Bcd1p *in cellulo*.

**Figure 91 :** Bcd1p interagit par double-hybride avec le facteur de remodelage des histones Rtt106p.

**Figure 92 :** La portion N-terminale (1-168) contenant le domaine à doigt de zinc constitue le fragment minimal fonctionnel *in vivo* de la protéine Bcd1p.

**Figure 93 :** Le domaine à doigt de zinc de Hit1p est fonctionnel dans la protéine Bcd1p.

## CONCLUSION & PERSPECTIVES

---

**Figure 94 :** Cartographie détaillée des interactions physiques entre les facteurs d'assemblage et les constituants des snoRNP à boîtes C/D.

**Figure 95 :** Modèle proposé pour l'assemblage des snoRNP à boîtes C/D chez la levure *Saccharomyces cerevisiae*.

**Figure 96 :** Bilan des fonctions d'assemblage de complexes macromoléculaires impliquant le complexe R2TP et HSP90 chez l'homme.

## *Liste des tableaux*

### AVANT-PROPOS

---

**Tableau 1 :** Principales classes d'ARN non codants retrouvés dans le monde vivant.

### INTRODUCTION

---

**Tableau 2 :** Principaux effets observés suite à la perte de modification de nucléotides au niveau des ARNr 18S et 25S chez *S. cerevisiae*.

### MATÉRIEL & MÉTHODES

---

**Tableau 3 :** Plasmides recombinants utilisés au cours de la thèse.

**Tableau 4 :** Oligonucléotides utilisés au cours de la thèse.

### RÉSULTATS & DISCUSSION

---

**Tableau 5 :** Tableau récapitulatif des liens établis entre les facteurs d'assemblage et les composants des RNP de la famille L7Ae chez la levure *S. cerevisiae* et chez l'homme.

**Tableau 6 :** Résultats du crible GIM réalisé à partir de la souche de levure BY4742- $\Delta$ RSAL.

**Tableau 7 :** Résultats du crible GIM réalisé à partir de la souche de levure BY4742- $\Delta$ PIH1.

**Tableau 8 :** Résultats du crible GIM réalisé à partir de la souche de levure BY4742- $\Delta$ TAH1.

### CONCLUSION & PERSPECTIVES

---

**Tableau 9 :** Bilan final des liens établis entre les facteurs d'assemblage et les composants des RNP de la famille L7Ae chez la levure *S. cerevisiae* et chez l'homme.

# TABLE DES MATIÈRES



# TABLE DES MATIÈRES

## AVANT-PROPOS

L'infrastructure ARN de la cellule eucaryote .....	1
Monde ARN et évolution .....	1
ARN et maturation.....	2
ARN interférence et régulation de l'expression des gènes.....	3
ARN, la matière noire de la cellule eucaryote ?.....	3
Du monde ARN vers un monde RNP .....	4
La biogenèse des RNP, un processus clé du métabolisme cellulaire .....	5

## INTRODUCTION

<b>CHAPITRE 1 : Le noyau et la biogenèse des ARN.....</b>	<b>8</b>
<b>1. Organisation structurale et fonctionnelle du noyau .....</b>	<b>8</b>
1.1. <i>L'enveloppe nucléaire</i> .....	9
1.2. <i>Les territoires chromosomiques et les espaces interchromatiniens</i> .....	10
1.3. <i>Les speckles et les paraspeckles</i> .....	12
1.4. <i>Le nucléole</i> .....	14
1.4.1. <i>Organisation du nucléole et biogenèse des ribosomes</i> .....	14
1.4.2. <i>Les autres fonctions du nucléole</i> .....	16
1.5. <i>Les corps de Cajal et les corpuscules nucléolaires</i> .....	17
1.5.1. <i>Les corps de Cajal</i> .....	17
1.5.2. <i>Les corpuscules nucléolaires</i> .....	20
1.6. <i>Les autres corpuscules nucléaires</i> .....	21
<b>2. Biogenèse des ARN transcrits par l'ARN polymérase II .....</b>	<b>23</b>
2.1. <i>Transcription par l'ARN polymérase II et rôle du domaine CTD</i> .....	23
2.2. <i>Maturation de l'extrémité 5' des transcrits de l'ARN polymérase II</i> .....	26
2.2.1. <i>La coiffe 7-méthylguanosine (<sup>m7</sup>G)</i> .....	26
2.2.2. <i>La coiffe triméthylée 2,2,7-méthylguanosine (<sup>m2,2,7</sup>G)</i> .....	27
2.2.3. <i>Maturation nucléolytique de l'extrémité 5' de certains snoRNA</i> .....	29
2.3. <i>Maturation des transcrits de l'ARN polymérase II par épissage</i> .....	29
2.3.1. <i>Le spliceosome, description du mécanisme d'épissage</i> .....	30
2.3.2. <i>Maturation des sous produits d'épissage : le cas des snoRNA</i> .....	32
2.3.2.1. <i>Voie majeure</i> .....	33

2.3.2.2.	Voie mineure .....	34
2.4.	<i>Maturation de l'extrémité 3' des transcrits de l'ARN polymérase II</i> .....	36
2.4.1.	Mécanisme de clivage/polyadénylation.....	36
2.4.2.	Terminaison indépendante de la polyadénylation.....	38
2.4.2.1.	Maturation de l'extrémité 3' des ARNm d'histones.....	38
2.4.2.2.	Maturation de l'extrémité 3' des UsnRNA.....	38
2.4.2.3.	Maturation de l'extrémité 3' des snoRNA.....	39
2.4.2.4.	Maturation de l'extrémité 3' de l'ARN télomérique.....	43
3.	<b>Production des ARNr et biogenèse des ribosomes</b> .....	<b>44</b>
3.1.	<i>Organisation des gènes codant les ARNr</i> .....	44
3.2.	<i>Modifications chimiques des nucléotides des ARNr</i> .....	46
3.3.	<i>Maturation nucléolytique des ARNr</i> .....	48
3.3.1.	Maturation conduisant à la production des ARNr 18S, 5.8S et 25S de <i>S. cerevisiae</i> .....	48
3.3.1.1.	Maturation du pré-ARNr 35S.....	48
3.3.1.2.	Maturation du pré-ARNr 20S.....	50
3.3.1.3.	Maturation du pré-ARNr 27SA2.....	50
3.3.2.	Maturation de l'ARNr 5S de <i>S. cerevisiae</i> .....	51
3.4.	<i>Assemblage des sous-unités ribosomiques</i> .....	51
3.4.1.	Du pré-ribosome 90S à la sous-unité 40S.....	52
3.4.2.	Assemblage de la sous-unité 60S.....	55
3.4.3.	Assemblage des ribosomes et consommation d'énergie.....	57
3.5.	<i>Ribosomopathie et défauts de la biogenèse des ribosomes</i> .....	58
 <b>CHAPITRE 2 : Les RNP de la famille L7Ae</b> .....		<b>60</b>
1.	<b>Les protéines de la famille L7Ae/L30</b> .....	<b>61</b>
1.1.	<i>Les motifs ARN en K-turn et en K-loop</i> .....	61
1.2.	<i>Description de différentes classes de motifs en K-turn</i> .....	63
1.2.1.	Les motifs en <i>K-turn</i> dans les ARNr et les ARNm.....	63
1.2.2.	Le motif en <i>K-turn</i> du snRNA U4.....	64
1.2.3.	Les motifs en <i>K-turn</i> des sRNA et des snoRNA à boîtes C/D.....	64
1.2.4.	Les motifs en <i>K-turn</i> particuliers du snoRNA U3.....	65
1.2.5.	Les motifs en <i>K-turn</i> des sRNA à boîtes H/ACA d'archée.....	66
1.2.6.	Les motifs en <i>K-turn</i> des éléments SECIS.....	66
1.3.	<i>Le domaine L7Ae/L30-RBD et ses propriétés d'interaction avec l'ARN</i> .....	66
1.3.1.	La protéine L7Ae d'archée.....	67
1.3.2.	La protéine Snu13p/15.5K.....	69
1.3.3.	La protéine Nhp2p/NHP2.....	72
1.3.4.	La protéine SBP2.....	73
1.3.5.	La protéine L30.....	75
1.4.	<i>Protéines de la famille L7Ae/L30 et évolution</i> .....	76
2.	<b>La snRNP U4</b> .....	<b>78</b>
2.1.	<i>Composition et organisation structurale de la snRNP U4</i> .....	78
2.1.1.	Le snRNA U4.....	78
2.1.2.	Les protéines de la snRNP U4.....	79
2.1.2.1.	Les protéines Sm.....	79
2.1.2.2.	La protéine Prp31p/PRP31.....	80
2.1.2.3.	Les protéines CYPH, Prp4p/60K et Prp3p/90K de la di-snRNP U4/U6.....	81
2.1.3.	Structure et assemblage de la snRNP U4.....	81
2.1.3.1.	Assemblage du domaine 3' de la snRNP U4.....	81

2.1.3.2.	Assemblage des protéines spécifiques de la snRNP U4.....	82
2.2.	Fonction de la snRNP U4 dans le mécanisme d'épissage des ARNm.....	84
2.3.	Des mutations dans le gène codant PRP31 à l'origine de rétinites pigmentaires.....	85
2.4.	Des mutations touchant le motif en K-turn du snRNA U4atac à l'origine du syndrome de Taybi-Linder 86	
<b>3.</b>	<b>Les snoRNP à boîtes C/D et H/ACA .....</b>	<b>87</b>
3.1.	Les snoRNP à boîtes C/D.....	87
3.1.1.	Organisation et fonction des snoRNP à boîtes C/D.....	88
3.1.1.1.	Les snoRNA à boîtes C/D.....	88
3.1.1.2.	Les protéines des snoRNP à boîtes C/D.....	90
3.1.1.2.1.	La 2'-O-méthylase Nop1p/Fibrillarine.....	90
3.1.1.2.2.	Les protéines Nop56p/NOP56 et Nop58p/NOP58.....	91
3.1.1.3.	Structure et fonction des snoRNP à boîtes C/D.....	93
3.1.1.3.1.	Pré-requis pour l'assemblage des particules.....	93
3.1.1.3.2.	Modèles symétriques et asymétriques.....	94
3.1.1.3.3.	Bases moléculaires de la 2'-O-méthylation guidée par l'ARN.....	97
3.1.2.	La snoRNP U3, une particule singulière impliquée dans la maturation nucléolytique du pré-ARNr 99	
3.1.2.1.	Le snoRNA U3.....	99
3.1.2.2.	Les protéines de la snoRNP U3.....	101
3.1.2.3.	Structure de la snoRNP U3.....	103
3.1.2.4.	Assemblage de la snoRNP U3 et différenciation cellulaire.....	105
3.1.2.5.	Fonction de la snoRNP U3.....	105
3.2.	Les snoRNP à boîtes H/ACA.....	107
3.2.1.	Les snoRNA à boîtes H/ACA.....	107
3.2.2.	Les protéines des snoRNP à boîtes H/ACA.....	108
3.2.2.1.	La protéine Cbf5p/Dyskérine.....	109
3.2.2.2.	La protéine Gar1p/GAR1.....	110
3.2.2.3.	La protéine NOP10.....	110
3.2.3.	Structure et fonction des snoRNP à boîtes H/ACA.....	111
3.2.4.	Le complexe télomérase chez les vertébrés.....	114
3.2.5.	Des mutations à l'origine de la Dyskératose Congénitale touchent des composants des particules à boîtes H/ACA.....	114
<b>CHAPITRE 3 : Biogenèse des RNP de la famille L7Ae.....</b>		<b>116</b>
1.	Principes généraux régissant l'assemblage des macro-complexes <i>in vivo</i> .....	117
2.	Biogenèse des UsnRNP, un modèle pour le contrôle de la spécificité d'assemblage.....	118
2.1.	Voies de biogenèse des UsnRNP.....	118
2.1.1.	Biogenèse nucléocytoplasmique des snRNP U1, U2, U4 et U5.....	118
2.1.2.	Biogenèse nucléaire de la snRNP U6.....	121
2.2.	Rôle du complexe SMN dans l'assemblage spécifique des UsnRNP.....	122
2.3.	Assemblage des UsnRNP et syndrome d'amyotrophie spinale.....	125
3.	Biogenèse des snoRNP à boîtes H/ACA, un modèle d'assemblage co-transcriptionnel.....	126
3.1.	Transcription et maturation des snoRNA à boîtes H/ACA.....	127
3.2.	Transport et localisation des snoRNP à boîtes H/ACA.....	127
3.3.	Assemblage des snoRNP à boîtes H/ACA.....	128
3.3.1.	Rôle de la protéine Naf1p/NAF1.....	129
3.3.2.	Rôle de la protéine Shq1p/SHQ1.....	131

<b>4.</b>	<b>Biogenèse des snoRNP à boîtes C/D .....</b>	<b>133</b>
4.1.	<i>Transcription et maturation des snoRNA à boîtes C/D .....</i>	134
4.2.	<i>Transport et localisation des snoRNP à boîtes C/D .....</i>	135
4.2.1.	Chez la levure .....	135
4.2.2.	Chez les vertébrés.....	136
4.2.3.	L'export cytoplasmique de la pré-snoRNP U8 .....	138
4.3.	<i>Importance des modifications post-traductionnelles pour la biogenèse des snoRNP à boîtes C/D ?</i> 139	
4.4.	<i>La snoRNP U3, un modèle d'étude pour la biogenèse des snoRNP .....</i>	140
4.4.1.	Biogenèse de la snoRNP U3 chez la levure.....	141
4.4.2.	Biogenèse de la snoRNP U3 chez les vertébrés .....	143
4.5.	<i>Vers un modèle d'assemblage pour les snoRNP à boîtes C/D ? .....</i>	145
4.6.	<i>Mise en évidence de nouveaux facteurs impliqués dans l'assemblage des snoRNP à boîtes C/D</i>	148
4.6.1.	Les AAA <sup>+</sup> -ATPase Rvb1p/TIP49 et Rvb2p/TIP48 .....	148
4.6.1.1.	Données structurales .....	149
4.6.1.2.	Données fonctionnelles .....	150
4.6.2.	La protéine Nop17p/NOP17 (alias Pih1p).....	153
4.6.3.	Le complexe R2TP, cofacteur des chaperonnes HSP90 .....	154
4.6.3.1.	Les chaperonnes HSP90 .....	154
4.6.3.2.	Identification du complexe R2TP .....	156
4.6.4.	La protéine Cwc24p .....	158
4.6.5.	La protéine Bcd1p .....	159
4.6.6.	Les protéines Rsa1p et NUFIP .....	159
4.6.6.1.	Rsa1p, un nouveau partenaire de la snoRNP U3 .....	159
4.6.6.2.	NUFIP, un nouveau partenaire de la protéine 15.5K .....	161
4.6.6.3.	Rsa1p et NUFIP, homologues fonctionnels ? .....	163

## OBJECTIFS

OBJECTIFS DE LA THÈSE .....	165
-----------------------------	-----

## MATÉRIEL & MÉTHODES

<b>1.</b>	<b>Matériel .....</b>	<b>170</b>
1.1.	<i>Les souches de cellules.....</i>	170
1.1.1.	Les souches de la bactérie <i>Escherichia coli</i> .....	170
1.1.2.	Les souches de la levure <i>Saccharomyces cerevisiae</i> .....	170
1.2.	<i>Les plasmides.....</i>	172
1.3.	<i>Les oligonucléotides .....</i>	174
1.4.	<i>Les milieux de culture.....</i>	174
1.4.1.	Cultures de bactéries.....	174
1.4.2.	Cultures de levures.....	174
1.5.	<i>Les tampons et solutions d'usage courant.....</i>	175
1.6.	<i>Les anticorps et les sérums .....</i>	175
<b>2.</b>	<b>Méthodes .....</b>	<b>176</b>
2.1.	<i>Techniques de génie génétique .....</i>	176



2.1.1.	Clonages .....	176
2.1.2.	Mutagenèse dirigée par PCR.....	176
2.1.3.	Production et purification de protéines recombinantes .....	177
2.1.3.1.	Surproduction de protéines chez <i>E. coli</i> .....	177
2.1.3.2.	Purification de protéines fusionnées à la GST.....	177
2.1.3.3.	Purification de protéines fusionnées à une étiquette 6-His.....	178
2.1.4.	Transcription/traduction <i>in vitro</i> de protéines radiomarquées .....	178
2.1.5.	Transcription <i>in vitro</i> et marquage radioactif des ARN .....	178
2.1.5.1.	Transcription <i>in vitro</i> par l'ARN polymérase T7.....	178
2.1.5.2.	Marquage au <sup>32</sup> P des ARN en 5' .....	179
2.2.	<i>Techniques de manipulation des levures</i> .....	179
2.2.1.	Transformation de cellules de levures compétentes .....	179
2.2.2.	Modification génétique des souches de levure .....	179
2.2.3.	Isolement de nouvelles souches de levure par sporulation et dissection de tétrades .....	180
2.2.4.	Utilisation de systèmes génétiques permettant de réprimer l'expression de gènes essentiels à la viabilité cellulaires .....	181
2.2.5.	Traitement de cellules en culture par les inhibiteurs de chaperonnes HSP90 et HSP70 ....	181
2.2.6.	Techniques de double-hybride et de triple-hybride .....	181
2.3.	<i>Techniques d'extraction de matériel biologique</i> .....	182
2.3.1.	Préparation des extraits cellulaires.....	182
2.3.2.	Extraction des ARN totaux.....	183
2.4.	<i>Techniques de détection à partir d'échantillons biologiques</i> .....	183
2.4.1.	Technique de Western-blot.....	183
2.4.2.	Technique de Northern-blot.....	183
2.4.2.1.	Pour la détection des snoRNA .....	183
2.4.2.2.	Pour la détection des ARNr .....	184
2.4.3.	Technique de RT-PCR.....	184
2.5.	<i>Techniques d'immunosélection</i> .....	185
2.5.1.	Approches standard d'immunoprécipitation.....	185
2.5.1.1.	Pour l'étude des snoRNP matures.....	185
2.5.1.2.	Pour l'étude des pré-snoRNP.....	185
2.5.2.	Purification de complexes pour analyse protéomique.....	185
2.5.3.	Immunoprécipitation de la chromatine.....	186
2.6.	<i>Techniques d'étude de complexes macromoléculaires in vitro</i> .....	187
2.6.1.	Technique de retard sur gel.....	187
2.6.2.	Technique d'empreintes enzymatiques .....	188
2.6.3.	Technique d'étude d'interactions protéiques par His pull-down.....	189
2.6.4.	Technique d'assemblage de RNP <i>in vitro</i> par His pull-down .....	189
2.6.5.	Technique d'assemblage de RNP en extrait cellulaire .....	189

## RÉSULTATS & DISCUSSION

### CHAPITRE 1 :

Mise en évidence d'une machinerie protéique dédiée à la biogenèse des RNP de la famille L7Ae.....	191
1. Présentation des travaux.....	191
2. Résultats .....	192
2.1. Publication n°1.....	192
2.2. Données complémentaires .....	194
2.2.1 Rsa1p interagit avec Nop58p et Prp31p .....	194

2.2.2	Rsa1p facilite l'assemblage de la snoRNP U14 .....	195
3.	Conclusion & Perspectives.....	198
3.1.	<i>Bilan</i> .....	198
3.2.	<i>Quel mode d'action pour la machinerie d'assemblage des RNP ?</i> .....	199
3.3.	<i>Questions liées à la mise en évidence de la machinerie d'assemblage</i> .....	200

## CHAPITRE 2 :

### Caractérisation structurale et fonctionnelle de l'interface d'interaction entre Rsa1p/NUFIP et Snu13p/15.5K..... 205

1.	Présentation des travaux.....	205
2.	Résultats .....	206
2.1.	<i>Publication n°2</i> .....	206
2.2.	<i>Données complémentaires</i> .....	208
2.2.1.	<i>In cellulo</i> , les résidus W245 et R249 de Rsa1p sont requis pour son recrutement sur les snoRNA.....	208
2.2.2.	Chez l'homme, le recrutement de PIH1 sur les snoRNA est également dépendant de l'interaction entre NUFIP et 15.5K.....	209
2.2.3.	Chez la levure, Rsa1p n'interagit pas avec la protéine Nhp2p.....	210
3.	Conclusion & Perspectives.....	211
3.1.	<i>Bilan</i> .....	211
3.2.	<i>Éléments de réflexion sur la structure des snoRNP à boîtes C/D</i> .....	213
3.3.	<i>Rsa1p/NUFIP et stratégie de mimétisme moléculaire ?</i> .....	216
3.4.	<i>Analogie avec le système H/ACA</i> .....	217
3.5.	<i>Proposition d'un modèle pour l'assemblage des snoRNP à boîtes C/D</i> .....	218

## CHAPITRE 3 :

### Caractérisation fonctionnelle de la machinerie d'assemblage et mise en évidence du rôle joué par la protéine Hit1p ..... 224

1.	Présentation des travaux.....	224
2.	Résultats .....	225
2.1.	<i>Publication n°3</i> .....	225
2.2.	<i>Données complémentaires</i> .....	227
2.2.1.	Rsa1p et Hit1p semblent spécifiques d'une sous-population de snoRNA à boîtes C/D .....	227
2.2.2.	Etude des liens entre transcription et assemblage des snoRNP.....	228
2.2.3.	Cartographie des domaines d'interaction protéiques impliqués dans le mécanisme d'assemblage des snoRNP.....	231
2.2.3.1.	Rsa1p interagit avec Pih1p par l'intermédiaire de son domaine C-terminal fonctionnel 232	
2.2.3.2.	Le domaine fonctionnel de Pih1p est responsable de l'interaction avec Nop58p et Rsa1p 233	

2.2.3.3.	Le domaine fonctionnel de Hit1p interagit avec le domaine fonctionnel de Rsa1p....	234
2.2.3.4.	Les interactions Pih1p-Nop58p et Rsa1p-Nop58p ne semblent pas mutuellement exclusives	235
2.2.4.	Cartographie des interactions génétiques à partir des souches de levures $\Delta RSA1$ , $\Delta PIH1$ et $\Delta TAH1$	236
2.2.4.1.	Description de l'approche GIM.....	236
2.2.4.2.	Description du réseau d'interactions génétiques impliquant l'ORF <i>RSAL</i> .....	237
2.2.4.3.	Description du réseau d'interactions génétiques impliquant l'ORF <i>PIH1</i> .....	239
2.2.4.4.	Description du réseau d'interaction génétique impliquant l'ORF <i>TAH1</i> .....	242
3.	Conclusion & Perspectives.....	243
3.1.	<i>Bilan</i> .....	243
3.2.	<i>Rsa1p, Hit1p et régulation métabolique ?</i> .....	244

## CHAPITRE 4 :

### Etude du rôle de Hsp82p et de sa co-chaperonne Tah1p dans le processus de biogenèse des snoRNP à boîtes C/D ..... 247

1.	Présentation des travaux.....	247
2.	Résultats.....	248
2.1.	<i>Publication n°4</i> .....	248
2.2.	<i>Données complémentaires</i> .....	250
2.2.1.	HSP90 et Tah1p ne sont pas requis pour l'accumulation des snoRNA à boîtes C/D.....	250
2.2.2.	Tah1p contrôle la stabilité de Pih1p indépendamment de HSP90.....	252
3.	Conclusion & Perspectives.....	253
3.1.	<i>Bilan</i> .....	253
3.2.	<i>Précision des relations fonctionnelles entre Pih1p, Tah1p et HSP90</i> .....	254
3.3.	<i>HSP90, Tah1p et maintenance du processus d'assemblage ?</i> .....	256

## CHAPITRE 5 :

### Etude de la biogenèse de la snoRNP U3 ..... 259

1.	Présentation des travaux.....	259
2.	Résultats.....	261
2.1.	<i>Etude in vitro de la RNP U3B/C</i> .....	261
2.1.1.	<i>In vitro</i> , l'assemblage de la protéine Rrp9p sur le snoRNA U3 ne requiert pas de facteur d'assemblage.....	261
2.1.2.	<i>In vitro</i> , Rsa1p <sub>230-381</sub> est recrutée sur la particule U3 B/C en présence de Rrp9p.....	263
2.2.	<i>Etude in vitro de la RNP U3C'/D</i> .....	263
2.2.1.	L'efficacité de fixation de Snu13p sur le motif U3C'/D est influencée par Rsa1p <i>in vitro</i> ..	264
2.2.2.	Le domaine 5' du snoRNA U3 est important pour le recrutement de Rsa1p sur le motif U3 C'/D	265
2.3.	<i>Etude de la voie de maturation du snoRNA U3</i> .....	267
2.3.1.	L'épissage du snoRNA U3 est un processus précoce.....	268
2.3.2.	Rsa1p est associée aux pré-snoRNA U3 +12 et +18.....	269

2.3.3. Des phénomènes de compensation impliquant Rrp9p pourraient assurer la production constante de snoRNP U3 .....	271
3. Conclusion & Perspectives.....	274
3.1. Proposition d'un nouveau modèle pour la biogenèse de la snoRNP U3.....	274
3.2. Spécificités liées à l'assemblage de la snoRNP U3.....	276

## CHAPITRE 6 :

Etude fonctionnelle de la protéine Bcd1p .....	280
1. Présentation des travaux.....	280
2. Résultats .....	281
2.1. Etude de l'interaction entre Bcd1p et les snoRNA à boîtes C/D.....	281
2.1.1. In vivo Bcd1p n'est pas stablement associée aux snoRNA matures .....	281
2.1.2. In vitro Bcd1p interagit de façon peu spécifique avec l'ARN .....	282
2.1.3. Le domaine à doigt de zinc de Bcd1p n'est pas un module d'interaction avec l'ARN .....	283
2.2. Recherche des partenaires protéiques associés à Bcd1p.....	284
2.2.1. Rsa1p interagit avec Rrp9p, Pih1p et Rvb2p dans des tests de double-hybride.....	284
2.2.2. In vivo, Bcd1p est associée aux composants des pré-snoRNP.....	285
2.2.3. Rtt106p, un nouveau partenaire de Bcd1p en lien avec la régulation de la transcription ..	287
2.3. Cartographie fonctionnelle des domaines de Bcd1p .....	289
2.3.1. La région N-terminale de Bcd1p, contenant un domaine à doigt de zinc, est requise pour la fonctionnalité de la protéine .....	289
2.3.2. Le domaine à doigt de zinc de Hit1p est fonctionnel dans Bcd1p .....	290
3. Conclusion & Perspectives.....	293

## CONCLUSION & PERSPECTIVES

1. Mise à jour d'une machinerie impliquée dans l'assemblage des RNP.....	297
2. Nouveaux éléments de réflexion sur l'assemblage des snoRNP C/D.....	298
3. Assemblage des snoRNP à boîtes C/D et régulations.....	301
4. Spectre d'action de la machinerie d'assemblage chez la levure .....	302
5. Diversification fonctionnelle de la machinerie d'assemblage chez l'homme .....	303

# AVANT-PROPOS



## AVANT-PROPOS

### L'infrastructure ARN de la cellule eucaryote

#### Monde ARN et évolution

« A un moment donné, une molécule remarquable s'est formée par accident. Elle n'était peut-être pas la plus grande ni la plus complexe, mais elle avait l'extraordinaire propriété de fabriquer des copies d'elle-même ». Cette première brique du monde vivant, telle que l'avait définie R. Dawkins (Dawkins, 1976), a longtemps été au centre de toutes les hypothèses sur l'origine de la vie. Et de façon tout à fait paradoxale, c'est au cœur du métabolisme contemporain que réside vraisemblablement une partie de la réponse à la plus fondamentale des questions. En effet, la plupart des évolutionnistes s'accordent aujourd'hui sur le fait que notre monde tripolaire ADN-ARN-protéine ait été précédé par un monde dans lequel l'ARN jouait à la fois le rôle de catalyseur et celui de matériel génétique. Ce scénario, baptisé *The RNA world* par W. Gilbert (Gilbert, 1986), repose sur un certain nombre d'observations biochimiques telles que la synthèse des dNTPs de l'ADN à partir des NTP de l'ARN, la réplication de l'ADN à partir d'amorces ARN ou encore la découverte, dans les années 1980, d'ARN à activités catalytiques, les ribozymes (Joyce, 2002; Maurel, 2003). A ce titre, l'un des principaux éléments appuyant cette théorie fut sans contexte la mise en évidence par l'équipe de T.A. Steitz, du rôle catalytique de l'ARNr dans la formation de la liaison peptidique au sein du ribosome (Moore & Steitz, 2002; Nissen *et al.*, 2000). Ainsi, la prévalence de l'ARN dans la synthèse de l'ADN et des protéines, ainsi que la sélection *in vitro* d'ARN doués d'autoréplication (Lincoln & Joyce, 2009) sont autant d'arguments en faveur de l'antériorité des acides ribonucléiques et de leurs grandes capacités fonctionnelles.

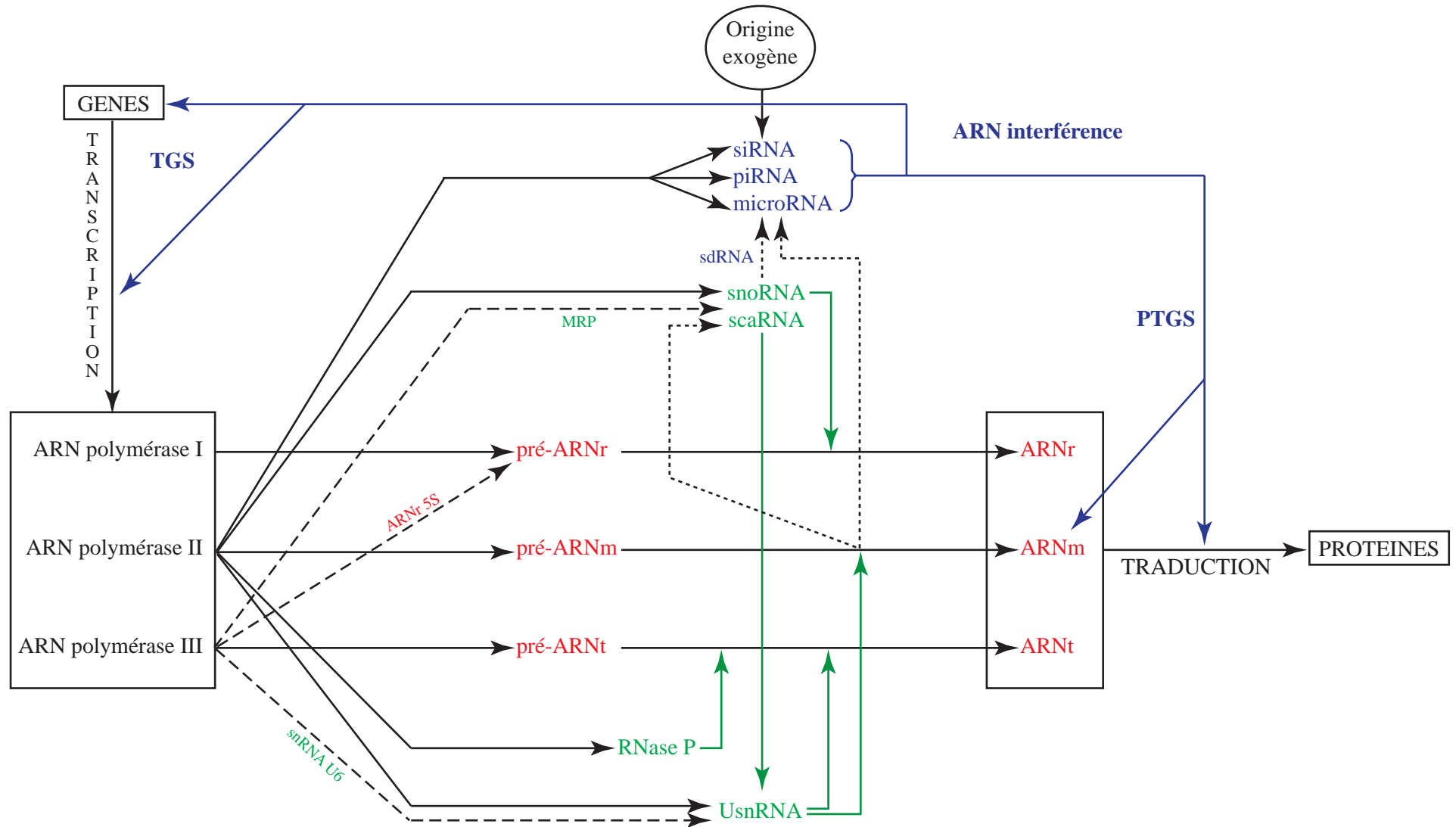
Dans le monde contemporain, les ARN occupent un rôle central dans les fonctions cellulaires les plus fondamentales telles que la transcription de l'information génétique en ARNm, son décodage par les ARNt et la synthèse des protéines par les ARNr. La conservation de ces processus dans l'ensemble du monde vivant témoigne d'une infrastructure ARN ancestrale, inhérente à la cellule moderne et à son métabolisme. Néanmoins les dernières décennies de

recherches ont progressivement dévoilé de nouvelles potentialités pour la molécule d'ARN, avec la découverte de nouvelles classes d'ARN non codants (ARNnc) impliqués dans de nombreuses fonctions telles que la maturation d'autres ARN, la régulation de l'expression génique ou le remodelage de la chromatine (**Fig. 1 ; Tableau 1**). Ainsi, la quantité et la variété de ces nouveaux rôles suggère que l'infrastructure ARN moderne n'est pas simplement le produit fixe d'un ancien monde ARN, mais bien un système dynamique au service de l'adaptabilité et de la diversification du vivant.

### **ARN et maturation**

Dans les années 1970-80 un certain nombre d'ARN non codants dédiés à la maturation des trois principales classes d'ARN (ARNm, ARNt, ARNr) ont été découverts chez les eucaryotes et pour certains, chez les archées et les bactéries (Collins & Penny, 2009) (**Tableau 1**). Leur mode de fonctionnement repose généralement sur un appariement de bases spécifique avec l'ARN cible et sur des mécanismes d'excisions ou de modifications de nucléotides. Cette catégorie d'ARN « au service de l'ARN » inclue notamment les snoRNA (*small nucleolar RNA*) qui, associés à des protéines sous forme de snoRNP (*small nucleolar RiboNucleoParticle*), sont impliqués dans la maturation des ARNr par l'introduction de modifications chimiques et de clivages endonucléolytiques. L'épissage des pré-ARNm implique pour sa part l'intervention du spliceosome, un large complexe composé d'environ 170 protéines et de 5 UsnRNA (*U rich small nuclear RNA*). Enfin, la maturation des précurseurs des ARNt nécessite l'intervention d'un ARN à activité ribozyme universel présent au sein du complexe de la RNase P, qui permet l'élimination d'une séquence excédentaire en 5' par clivage endonucléolytique. D'autre part, il existe des liens forts entre ces différentes machineries qui dépendent souvent les unes des autres pour leur propre biogenèse (**Fig. 1**). En effet, la formation des UsnRNA matures implique la modification chimique spécifique de certains nucléotides par des snoRNP ou par des scaRNP (*small cajal bodies RNP*), et de nombreux snoRNA, hébergés dans des introns de pré-ARNm, requièrent l'intervention du spliceosome pour leur propre production.





**Figure 1 : Cascade d'événements impliquant des ARNnc dans le processus d'expression des gènes chez les eucaryotes.** Les 3 grandes classes d'ARN directement impliqués dans la synthèse protéique (ARNr, ARNm, ARNt) apparaissent en rouge. Les ARN impliqués dans des processus de maturation sont indiqués en vert (snoRNA, scaRNA, snRNA, RNase P). Les ARN impliqués dans des étapes de régulation du niveau d'expression des gènes sont indiqués en bleu (siRNA, piRNA, microRNA et sdRNA). Les flèches noires indiquent l'origine et le devenir des différentes molécules. Les traits en pointillés larges représentent des voies spécifiques d'une molécule unique (l'ARNr 5S, le snoRNA de la RNase MRP et le snRNA U6 sont transcrits par l'ARN polymérase III). Les traits en pointillés fins représentent des voies de production d'ARNnc à partir d'autres catégories d'ARN. Les flèches vertes représentent les activités de maturation. Les flèches bleues indiquent les activités de régulation. TGS : Transcriptional Gene Silencing; PTGS : Post Transcriptional Gene Silencing.

<b>Classe</b>	<b>Nom complet</b>	<b>Fonction</b>	<b>Conservation</b>
<b>rRNA</b>	ribosomal RNA	Synthèse des protéines	bactéries, archées, eucaryotes
<b>tRNA</b>	transfert RNA	Décodage de l'information génétique	bactéries, archées, eucaryotes
<b>RNase P</b>	Ribonuclease P	Maturation des ARNt	bactéries, archées, eucaryotes
<b>SRP RNA</b>	Signal Recognition Particle RNA (7SL RNA chez les mammifères)	Traduction des protéines membranaires ou sécrétoires	eucaryotes
<b>snoRNA</b>	small nucleolar RNA	Maturation des ARNr et du snRNA U6 au niveau du nucléole	archées, eucaryotes
<b>scaRNA</b>	small cajal-bodies RNA	Maturation des snRNA au niveau des Cajal-Bodies	eucaryotes
<b>snRNA</b>	small nuclear RNA	Epissage des pré-ARNm	eucaryotes
<b>TERC</b>	Telomerase RNA Component	Réplication des télomères	eucaryotes
<b>7SK</b>	/	Régulation transcriptionnelle par inhibition de PTEF-b	eucaryotes
<b>U7</b>	/	Maturation des ARNm codant les histones	eucaryotes
<b>siRNA</b>	small interfering RNA	Régulation transcriptionnelle ou traductionnelle de l'expression génique	eucaryotes
<b>miRNA</b>	micro RNA	Régulation transcriptionnelle ou traductionnelle de l'expression génique	eucaryotes
<b>piRNA</b>	piwi-associated RNA	Régulation transcriptionnelle de l'expression génique	eucaryotes
<b>PASR</b>	Promoter-Associated Long RNA	Fonction inconnue	eucaryotes
<b>PALR</b>	Promoter-Associated Short RNA	Fonction inconnue	eucaryotes
<b>TASR</b>	Termini-Associated Short RNA	Fonction inconnue	eucaryotes
<b>TSSaRNA</b>	Transcription Start Sites antisense RNA	Fonction inconnue	eucaryotes
<b>PROMPTs</b>	Promoter upstream transcripts	Fonction inconnue	eucaryotes
<b>tiRNA</b>	tiny RNA	Régulations de l'aposition de marques épigénétiques ?	eucaryotes
<b>spliRNA</b>	splice-site RNA	Régulations de l'aposition de marques épigénétiques ?	eucaryotes
<b>lncRNA</b>	long non coding RNA	Longs ARNnc associés à diverses fonctions (ex: Xist, NEAT1, Evf-2, Vault, ...)	eucaryotes
<b>gRNA</b>	guide RNA	ARN guides d'édition	trypanosome
<b>CRISPR RNA</b>	Clustered Regularly Interspaced Short Palindromic Repeats RNA	Ciblage et dégradation d'ADN exogène d'origine phagique ou plasmidique	bactéries, archées
<b>sRNA</b>	small RNA	Régulation de l'expression génique par des mécanismes variés ( <i>cis</i> ou <i>trans</i> ) et spécifiques	bactéries, archées

**Tableau 1 : Principales classes d'ARN non codants retrouvés dans le monde vivant.**

## ARN interférence et régulation de l'expression des gènes

Au cours des années 1990, une série de découvertes, en particulier celle de petits ARN dits « interférents », a fait rentrer la biologie moléculaire dans une nouvelle ère, celle des riborégulateurs. Plus petites (20 à 30 nucléotides), moins abondantes individuellement, mais beaucoup plus variées, ces molécules d'ARN exercent leur activité sur la régulation de l'expression des gènes à différents niveaux (**Fig. 1 ; Tableau 1**) (Moazed, 2009; Siomi & Siomi, 2009). Découvert fortuitement chez le pétunia (Jorgensen, 1990), puis chez *C. elegans* (Fire *et al.*, 1998), le mécanisme d'interférence ARN (RNAi) ou *RNA silencing*, implique de petits ARN non codants (siRNA et miRNA) qui, associés à des protéines de la famille Argonaute, sont capables de réprimer spécifiquement l'expression d'un gène. Cette répression repose sur un appariement du petit ARNnc avec l'ARNm cible et s'effectue soit à un niveau post-transcriptionnel par une voie dite PTGS (*Post-Transcriptional Gene Silencing*), soit à un niveau transcriptionnel par une voie dite TGS (*Transcriptional Gene Silencing*). Dans les mécanismes de type PTGS, le petit ARNnc s'apparie avec un ARNm cible au niveau du cytoplasme pour réprimer sa traduction ou induire sa dégradation. Dans le cas des mécanismes de type TGS, l'ARNnc est adressé dans le noyau afin de cibler spécifiquement une région génomique et de réprimer sa transcription (Gonzalez *et al.*, 2008; Verdel *et al.*, 2004). Ces différents ARNnc peuvent avoir une origine exogène, généralement virale (i.e siRNA), ou endogène à la cellule (i.e. siRNA et miRNA). Dans ce dernier cas, les siRNA sont généralement issus de rétro-transposons, tandis que les miRNA, qui peuvent également posséder leur propre unité transcriptionnelle, sont principalement hébergés dans des introns de gènes codants des protéines (Kim *et al.*, 2009a; Rodriguez *et al.*, 2004). De façon intéressante, des travaux récents ont également rapporté l'existence de miRNA dérivant de snoRNA (Ender *et al.*, 2008; Ono *et al.*, 2010; Scott *et al.*, 2009). La mise à jour de ces molécules, rebaptisées sdrRNA (*snoRNA derived RNA*), révèlent donc un nouveau lien évolutif inattendu entre deux classes d'ARN non codants (Taft *et al.*, 2009) (**Fig. 1**).

## ARN, la matière noire de la cellule eucaryote ?

Au cours des années 2000, l'analyse du génome humain a fait l'objet d'une grande surprise, puisqu'elle a révélé l'existence de seulement 20000 à 25000 gènes, soit autant que dans d'autres organismes moins complexes comme *C. elegans* (Clamp *et al.*, 2007; Goodstadt &

Ponting, 2006). Ces données révélèrent alors que la complexité du développement et de la physiologie d'un organisme n'est pas strictement corrélée à son nombre de gène. Et si les phénomènes d'épissage alternatif et de modifications post-traductionnelles augmentent considérablement la diversité et la fonctionnalité du protéome, il n'en demeure pas moins que les territoires non codants représentent environ 98,2% du génome. De façon remarquable, les récents progrès en matière de séquençage à haut débit ont permis de démontrer que ces incroyables quantités de séquences non codantes, retrouvées au sein des génomes eucaryotes, sont transcrites de façon diffuse (Carninci *et al.*, 2005; Katayama *et al.*, 2005). Des études ont en effet révélé que plus de 90% du génome humain est transcrit et que les cellules de mammifères produisent en général de quelques dizaines de milliers, à plusieurs centaines de milliers d'ARN non codants (Birney *et al.*, 2007). Les connaissances actuelles sur les multiples rôles de l'ARN et ces récentes découvertes laissent donc entrevoir l'importance d'une infrastructure jusqu'ici sous-estimée et des phénomènes de régulation beaucoup plus complexes qu'ils n'avaient été imaginés jusqu'alors. Plus qu'un simple « bruit transcriptionnel », les milliers d'ARN non codants encore non caractérisés pourraient donc constituer la matière noire de la cellule eucaryote. Leur étude, la compréhension des mécanismes moléculaires sous-jacents, ainsi que leur importance pour la mise en place des fonctions supérieures constituent donc un champ de recherche majeur pour les années à venir.

### **Du monde ARN vers un monde RNP**

Hormis certains éléments ARN agissant en *cis* et certains transcrits régulateurs anti-sens, la grande majorité des ARNnc ne fonctionnent pas seuls mais associés à des protéines au sein de complexes ribonucléoprotéiques, les RNP. D'un point de vue évolutif, les RNP modernes sont souvent considérées comme les témoins d'un transfert d'activité de l'ARN vers les protéines (Cech, 2009; Hogg & Collins, 2008; Poole *et al.*, 1998). Le monde ARN, tel qu'il a été proposé, aurait en effet évolué vers un « monde RNP », avant d'aboutir à la situation actuelle dans laquelle RNP et protéines se partagent les fonctions de catalyse. H. Noller a notamment proposé qu'à un certain stade de l'évolution, les ribozymes aient pu bénéficier de la présence de petits peptides générés dans leur environnement, pour former les premiers complexes ARN-protéines (Noller, 2004). Une telle association aurait pu constituer un avantage sélectif important en élargissant leurs répertoires structuraux et fonctionnels. Dans ce contexte, et tenant compte de l'émergence de polypeptides plus complexes, certaines RNP auraient

évoluées en intégrant les nouvelles potentialités de leurs composants protéiques, tandis que d'autres auraient perdu leur sous-unité ARN pour donner naissance à des enzymes purement protéiques (Cech, 2009; Jeffares *et al.*, 1998). En dépit de cette tendance évolutive, de nombreuses enzymes RNP ont donc été conservées pour des rôles clé du métabolisme cellulaire. La prévalence de l'ARN dans ces fonctions repose entre autre sur des caractéristiques propres comme la capacité d'adopter des structures variées et modulables au cours de l'acte catalytique, ou encore la capacité d'interagir avec d'autres ARN par appariement de base spécifique. Au sein des RNP, l'ARN joue donc des rôles relativement variés : il peut être porteur d'activités catalytiques (ribosome, RNase P, MRP ...), il peut servir de guide pour cibler un substrat spécifique (snoRNP, UsnRNP, miRNA, siRNA, ...), il peut servir de matrice pour la synthèse de polymères (TERC), il peut réguler certaines fonctions protéiques (7SK, Evf-2, ...), il peut également jouer un rôle de plate-forme architecturale (Xist, NEAT1, ...). Dans un environnement cellulaire riche en ribonucléases, les protéines sont quant à elles nécessaires pour assurer la stabilité de l'ARN et permettre son accumulation dans la cellule. Elles jouent également des rôles majeurs dans les processus de transport et de localisation des RNP. Enfin, sur la base de leurs propriétés structurales et catalytiques, certaines d'entre elles prennent une part active dans la fonction des particules.

### **La biogenèse des RNP, un processus clé du métabolisme cellulaire**

Au sein des RNP, ARN et protéines ont donc co-évolué pour aboutir à une relation d'interdépendance fonctionnelle et structurale. Au delà des propriétés spécifiques de chaque constituant, le fonctionnement des particules repose sur un arrangement structural précis qui conditionne sa fonction. L'assemblage des différents composants pour former une entité fonctionnelle repose sur la mise en place d'interactions spécifiques qui s'établissent selon des mécanismes souvent complexes. Ces derniers s'intègrent dans le cadre plus large de la biogenèse des RNP, qui comprend également la traduction des protéines dans le cytoplasme et leur acheminement vers les sites d'assemblage, la transcription et la maturation des ARN, ainsi que le transport des particules vers leurs sites fonctionnels ou vers des zones de stockage. Ces événements se déroulent de façon séquentielle et ordonnée à travers différents domaines subcellulaires et intègrent généralement des étapes de contrôle qualité visant à assurer la production de particules pleinement fonctionnelles. Leur mise en place requiert l'intervention en *trans* de nombreux facteurs protéiques ou ribonucléoprotéiques, comme des

enzymes de modification, des ribonucléases, des hélicases, des chaperonnes, des facteurs de transport ou encore des protéines assurant le couplage entre les différentes étapes du processus. De ce fait, le nombre de facteurs requis pour la production d'une RNP excède souvent très largement le nombre d'éléments composant la particule elle-même. A titre d'exemple, le ribosome eucaryotes, composé de 4 ARNr et d'environ 80 protéines ribosomiques, fait intervenir pour sa synthèse plus de 180 facteurs d'assemblage et une centaine de snoRNP au cours d'un processus séquentiel se déroulant successivement dans le nucléole, le nucléoplasme, puis le cytoplasme (Henras *et al.*, 2008).

Le laboratoire du Dr Christiane Branlant étudie les mécanismes moléculaires liés à la biogenèse des RNP et à leur fonction. Au cours de mon travail de thèse, dirigé par le Pr Bruno Charpentier, nous nous sommes tout particulièrement intéressés aux RNP contenant des protéines de la famille L7Ae/L30. Ces protéines, qui comptent des représentants dans l'ensemble du monde vivant, sont des composants essentiels de nombreuses RNP eucaryotes, comme les snoRNP à boîtes C/D et H/ACA, la snRNP U4, le complexe télomérase des vertébrés, ou encore les mRNP chargées de l'incorporation des sélénocystéines dans les protéines. Leur implication dans de nombreux processus fondamentaux comme la synthèse des ribosomes, l'épissage, la traduction ou la réplication des télomères, en font donc un modèle de choix pour l'étude de la biogenèse des RNP. De plus, la construction de ces particules sur la base d'éléments structuraux communs constitue un aspect tout à fait intéressant pour aborder la notion de spécificité des mécanismes d'assemblage. En collaboration avec les équipes d'Edouard Bertrand à Montpellier et d'Alain Krol à Strasbourg, nous avons contribué à mettre à jour l'existence de facteurs conservés chez les eucaryotes et dédiés à l'assemblage des RNP appartenant à cette famille. Puis, en utilisant le modèle des snoRNP à boîtes C/D chez la levure *Saccharomyces cerevisiae*, nous avons entrepris le décryptage des mécanismes liés à leur fonctionnement, par diverses approches de biochimie, de biologie moléculaire et de génétique. Nous avons ainsi décrit les liens physiques et fonctionnels s'établissant entre les facteurs d'assemblage et nous avons apporté de nouveaux éléments de compréhension quant à leur mode d'action sur les composants des RNP.

L'introduction de mon travail de thèse se divisera en trois parties. Dans un premier temps, je présenterai l'architecture nucléaire de la cellule eucaryote et j'en décrirai deux processus majeurs : la biogenèse des ARN transcrits par l'ARN polymérase II et la biogenèse des ribosomes. Dans le cadre de notre étude, la compréhension détaillée de ces phénomènes

apparaît primordiale. Les snoRNA, transcrits par l'ARN polymérase II, sont en effet soumis à des événements de maturation spécifiques, directement liés à leur mode de production et intimement couplés au processus d'assemblage des snoRNP. De plus, une fois assemblées, les snoRNP sont acheminées dans le nucléole, au sein duquel elles jouent un rôle déterminant pour la biogenèse des sous-unités ribosomiques. Dans un second volet, j'aborderai sous la forme d'une étude comparative, la composition, la structure et la fonction des RNP contenant des protéines de la famille L7Ae. Cet examen détaillé permettra de déterminer de façon précise les points communs et les spécificités liés à l'assemblage de chaque type de RNP. Enfin, dans une troisième partie je passerai en revue l'état des connaissances sur les mécanismes de biogenèse de ces RNP au début de ma thèse. J'aborderai notamment le cas de l'assemblage des UsnRNP et des snoRNP H/ACA, qui ont déjà fait l'objet de nombreuses investigations et qui constituent des exemples de référence pour notre travail sur les snoRNP à boîtes C/D. L'ensemble de cette partie bibliographique sera centrée sur notre organisme de référence, *Saccharomyces cerevisiae*, néanmoins j'ai veillé tant que possible à établir des parallèles avec les modèles vertébrés, en particulier avec le système humain.





# INTRODUCTION



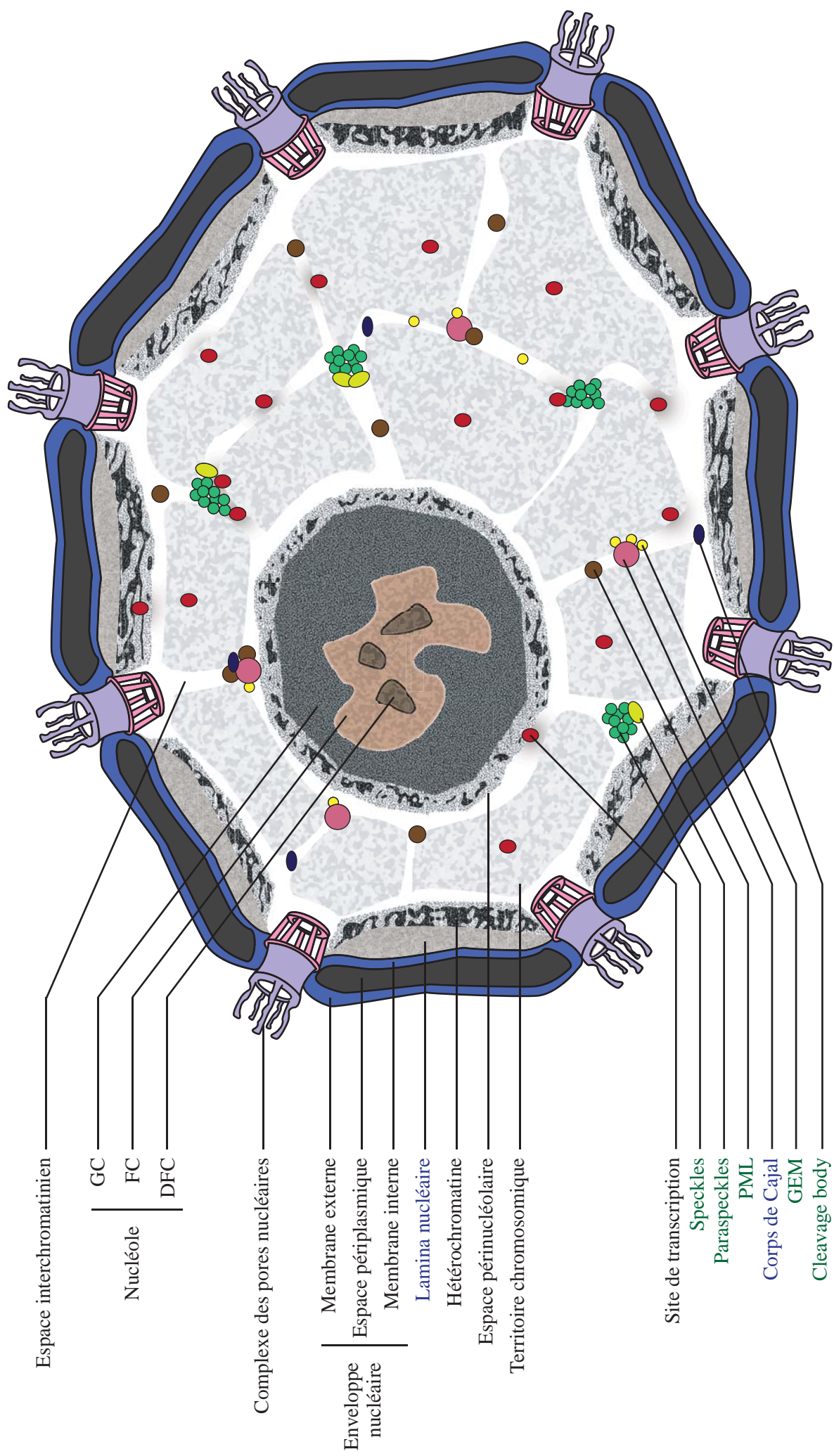
## CHAPITRE 1

### Le noyau et la biogenèse des ARN

Le noyau des cellules eucaryotes centralise les fonctions de stockage, de réplication, de maintenance et de transcription de l'information génétique. Il constitue ainsi le centre névralgique du métabolisme cellulaire, au sein duquel sont initiés la majorité des processus liés à l'expression des gènes. Il contient notamment trois machineries spécialisées, les ARN polymérases I, II et III, qui se partagent les fonctions de transcription (**Fig. 1**). Le noyau constitue également un site privilégié pour les phénomènes de maturation des ARN et pour leur assemblage au sein de RNP. Sur la base de leurs modes de transcription et de leurs classes, les molécules d'ARN subissent des traitements spécifiques leur permettant d'acquérir une fonctionnalité. A l'issue de ces événements, certains ARN sont ensuite exportés vers le cytoplasme, tandis que d'autres demeurent dans le noyau au sein duquel ils exercent leurs fonctions. Ces différents processus, qui font l'objet de nombreux couplages, sont hautement régulés et leur déroulement apparaît intimement lié à l'architecture nucléaire.

#### 1. Organisation structurale et fonctionnelle du noyau

Le noyau, observé pour la première fois par F. Bauer en 1802, constitue la première structure intra-cellulaire décrite. Délimité par une enveloppe nucléaire, son architecture interne repose sur l'existence de nombreux domaines sub-nucléaires, comme les territoires chromosomiques, le nucléole, les speckles ou encore les corps de Cajal (**Fig. 2**). Ces différentes structures sont dépourvues de membranes et sont formées par l'accumulation spécifique de molécules impliquées dans des fonctions communes. Bien que fonctionnellement très compartimenté, le noyau eucaryote constitue néanmoins une organelle extrêmement dynamique, dont l'organisation est remodelée au cours du cycle cellulaire et en fonction de la physiologie de la cellule. De plus, comme nous allons le voir, l'architecture nucléaire est intimement liée aux processus de biogenèse des ARN, dont le déroulement dépend de plusieurs compartiments hautement spécialisés.



**Figure 2 : Représentation schématique des principales structures observées dans le noyau des cellules eucaryotes.** GC : Composants Granulaires; FC : Centres Fibrillaires; DFC : Centres Fibrillaires Denses; PML : Promyelocytic Leukemia nuclear bodies; GEM : Gemini des Cajal Bodies. Certaines structures (annotées en noir) sont conservées chez tous les eucaryotes, d'autres (annotées en bleu) sont partiellement conservées, enfin certaines (représentées en vert) sont spécifiques de certains groupes d'espèces (animaux et plantes pour l'essentiel).

## 1.1. L'enveloppe nucléaire

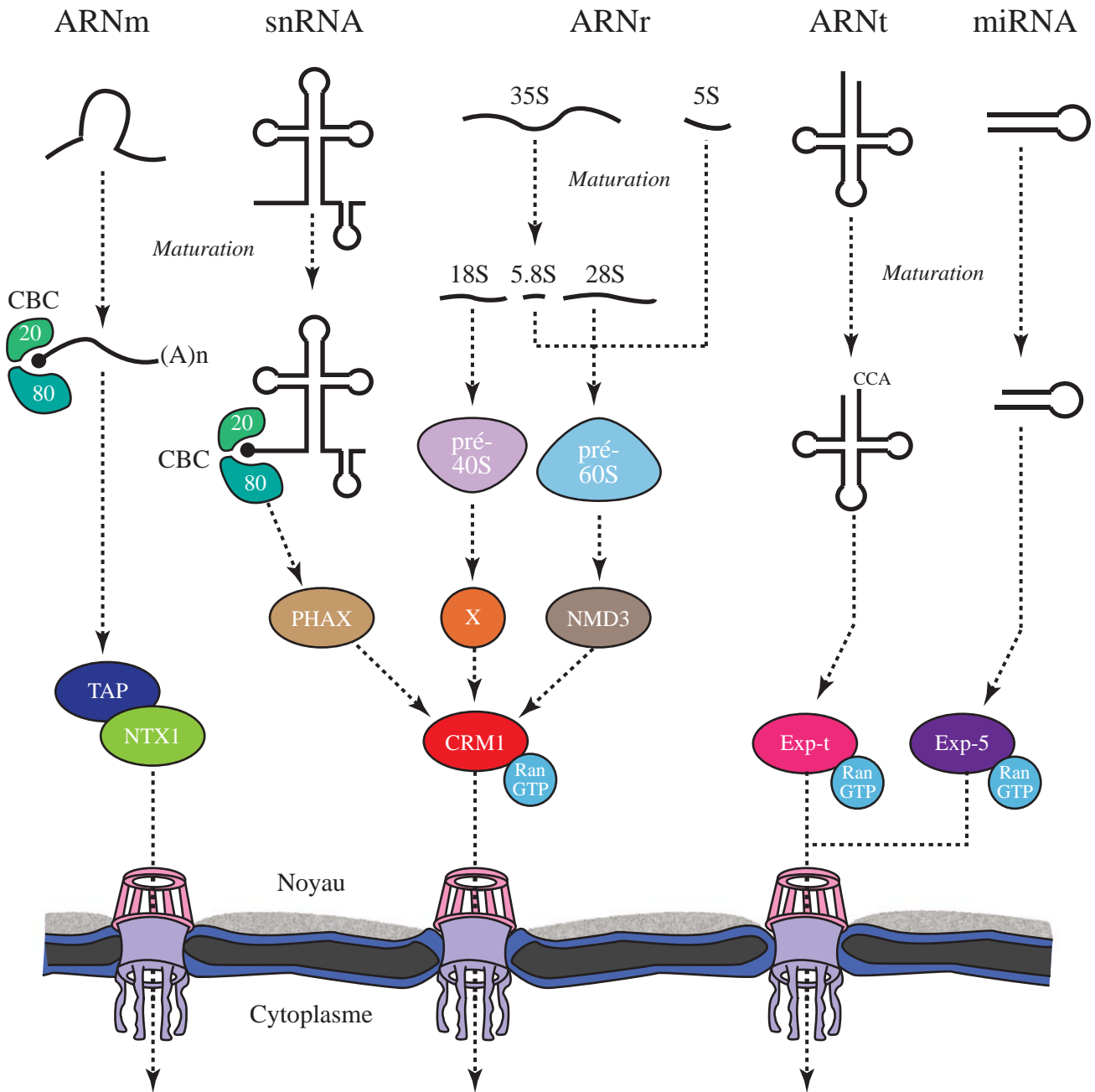
L'enveloppe nucléaire est une bicouche lipidique renfermant un espace périplasmique et contenant plus d'une centaine de protéines différentes (Hetzer & Wentz, 2009). La membrane externe est en continuité avec la membrane du réticulum endoplasmique. La membrane interne permet quant à elle l'ancrage des chromosomes et joue un rôle prédominant dans l'organisation du noyau. Chez les métazoaires, ce rôle repose sur l'existence d'un réseau rigide de protéines transmembranaires fibrillaires, les lamines, qui forment une structure périnucléaire appelée lamina nucléaire (Gruenbaum *et al.*, 2005) (**Fig. 2**). Les autres eucaryotes ne possèdent pas de lamines, néanmoins plusieurs études ont mis en évidence l'existence de protéines assurant des rôles similaires (Grund *et al.*, 2008; Taddei *et al.*, 2004).

L'enveloppe nucléaire est interrompue par des pores nucléaires qui constituent des zones d'échanges entre le noyau et le cytoplasme (**Fig. 2**). Ces pores sont formés par des complexes multiprotéiques de grande taille (40 à 125 MDa en fonction des espèces), composés de protéines appelées nucléoporines (D'Angelo & Hetzer, 2008). Relativement bien conservés au cours de l'évolution, les complexes des pores nucléaires assurent le contrôle du trafic cyto-nucléoplasmique en induisant une perméabilité sélective. Ils permettent en effet la diffusion passive de petites molécules et limitent le passage des molécules supérieures à 40 KDa qui nécessitent un transport actif (Stewart, 2007). Il existe un grand nombre de voies de transport en fonction du type de molécules, mais la plupart d'entre elles impliquent des protéines de la famille des karyophérines. Celles-ci sont capables d'interagir avec les complexes des pores nucléaires et facilitent le transit des protéines cargo qui leur sont associées (Stewart, 2007). Elles se divisent en deux grandes catégories : les importines, qui assurent le transport cytoplasme-noyau, et les exportines, qui assurent le transport noyau-cytoplasme. Les karyophérines ciblent les protéines cargo présentant des séquences conservées appelées signaux d'adressage. Les protéines candidates à l'export sont en effet porteuses de séquences hydrophobes, les NES (*Nuclear Export Signal*), reconnues par les exportines, tandis que les candidates à l'import sont porteuses de séquences basiques, les NLS (*Nuclear Localisation Signal*), reconnues par les importines. L'ensemble de ces processus est régulé par un gradient de Ran-GTP, faiblement concentré dans le cytoplasme et fortement concentré dans le noyau, qui assure la directionnalité du transport.

L'export des ARN vers le cytoplasme implique différents types de karyophérines qui établissent des interactions avec certaines protéines des RNP (Cullen, 2003; Erkmann & Kutay, 2004) (**Fig. 3**). Les ARNt suivent une voie spécifique impliquant l'exportine-t (Los1p chez la levure), les précurseurs de miRNA sont pris en charge par l'exportine-5, l'export des sous-unités ribosomiques, des UsnRNA et de la majorité des autres ARN impliquent le facteur général CRM1 (Crm1p/Xpo1p/Kap124p chez la levure). Les ARNm suivent en revanche une voie spécifique qui n'implique ni karyophérine, ni Ran-GTP, mais les protéines de la famille NFX, qui peuvent également interagir avec les nucléoporines. Le facteur de transport d'ARNm le mieux caractérisé est la protéine TAP (alias NFX1, Mex67p chez la levure), dont le recrutement au niveau des complexes des pores nucléaire requiert l'interaction avec la protéine NXT1 (alias p15, Mtr2p chez la levure). De plus, même si la protéine TAP peut interagir directement avec les ARNm, leur mécanisme d'export fait intervenir un certain nombre d'adaptateurs comme la protéine REF (alias Aly, Yra1p chez la levure) qui interagit à la fois avec TAP et avec les ARNm épissés. L'ensemble de ces facteurs sont ainsi responsables de phénomènes de couplages entre les grandes étapes de la biogenèse des ARNm et participent à des mécanismes de contrôle qualité assurant l'export des ARNm correctement maturés.

## 1.2. Les territoires chromosomiques et les espaces interchromatiniens

Dans les cellules eucaryotes, l'ADN est associé à des protéines histones pour former la chromatine. Celle-ci est facilement visualisée sous la forme de chromosomes individuels dans les cellules en mitose. En revanche, dans les cellules en interphase, les chromosomes ont une structure relâchée qui peut adopter deux états distincts en fonction du degré de condensation de la chromatine. Les zones correspondant à des structures décondensées forment une région transcriptionnellement active appelée euchromatine, tandis que les zones condensées sont associées à une faible activité transcriptionnelle et forment une structure appelée hétérochromatine. Au sein des régions hétérochromatiniennes on distingue l'hétérochromatine constitutive, pauvre en gène, qui contient des séquences répétées principalement issues des centromères et des télomères, et l'hétérochromatine facultative qui contient des régions codantes pouvant adopter les caractéristiques structurales et fonctionnelles de



**Figure 3 : Voies de transport nucléocytoplasmique des principales molécules d'ARN chez les eucaryotes (adapté de Cullen et al., 2003).** L'export des ARNm implique l'hétérodimère TAP-NTX1 ainsi que divers adaptateurs assurant le transport spécifique des ARNm correctement maturées. L'export des snRNA implique la protéine PHAX qui interagit à la fois avec l'exportine CRM1 et avec le "Cap Binding Complex" (CBC, formé par l'hétérodimère CBC20-CBC80) qui se fixe sur les coiffes m<sup>7</sup>G (●). Rq : il n'existe pas d'équivalent de PHAX chez *S. cerevisiae*. L'export de la sous-unité pré-60S implique CRM1 et l'adaptateur NMD3 qui interagit avec la protéine ribosomique RPL10. L'export de la sous-unité pré-40S est également dépendant de CRM1 mais la nature de l'adaptateur demeure inconnue (X). L'export des ARNt suit une voie spécifique impliquant l'Exportine t (Exp-t), qui fixe spécifiquement les ARNt correctement maturés.

l'hétérochromatine, en fonction des événements de régulation (Ex : chromosome X inactif chez les femelles mammifères).

Pendant longtemps, le noyau a été considéré comme une organelle chaotique contenant de grandes quantités d'ADN désorganisé. Depuis le début des années 1980, de nombreuses études ont montré que la disposition des chromosomes en interphase n'est pas un processus aléatoire, mais que chacun d'entre eux occupe un volume nucléaire spécifique baptisé territoire chromosomique (Cremer *et al.*, 2006; Parada *et al.*, 2004b). Néanmoins, cette topologie ne correspond pas à une organisation figée, puisqu'elle subit de profonds remaniements au cours des processus de différenciation et qu'elle varie de façon drastique en fonction des types cellulaires (Parada *et al.*, 2004a). Cette dynamique architecturale permet une réorganisation topologique des gènes en fonction des besoins transcriptionnels et constitue donc un niveau supplémentaire de régulation de l'expression génique (Heard & Bickmore, 2007). D'autre part, plusieurs études ont montré que l'organisation globale des chromosomes au sein du noyau est préservée à la suite de la mitose, contribuant ainsi au maintien des profils d'expression spécifiques des différents lignages cellulaires (Cremer *et al.*, 2006; Gerlich *et al.*, 2003).

L'étude de cette topologie a révélé une forte tendance à la séparation des régions génomiques en fonction de leur activité transcriptionnelle (Lanctot *et al.*, 2007) (**Fig. 2**). L'hétérochromatine se répartit essentiellement en périphérie du noyau où elle se trouve ancrée par la lamina, ainsi que dans la région péri-nucléolaire. L'euchromatine investit quant à elle l'espace nucléoplasmique en suivant un gradient qui semble répartir les régions les plus transcriptionnellement actives vers l'intérieur du noyau. Les territoires chromosomiques sont séparés par des lacunes formant un ensemble appelé espace interchromatinien, dont le diamètre peut varier de plusieurs micromètres à quelques nanomètres dans ces branches les plus fines (Cremer & Cremer, 2001). Cet espace interchromatinien s'invagine au sein des territoires chromosomiques de façon comparable à un réseau spongieux et permet l'accès de certaines régions génomiques à la machinerie de transcription (Francastel *et al.*, 2000). Certaines études ont même montré que par le biais de cette porosité, certaines régions d'hétérochromatine permettent l'accès à des facteurs de transcription pour assurer une expression localisée au sein de régions globalement silencieuses (Chen *et al.*, 2005b).

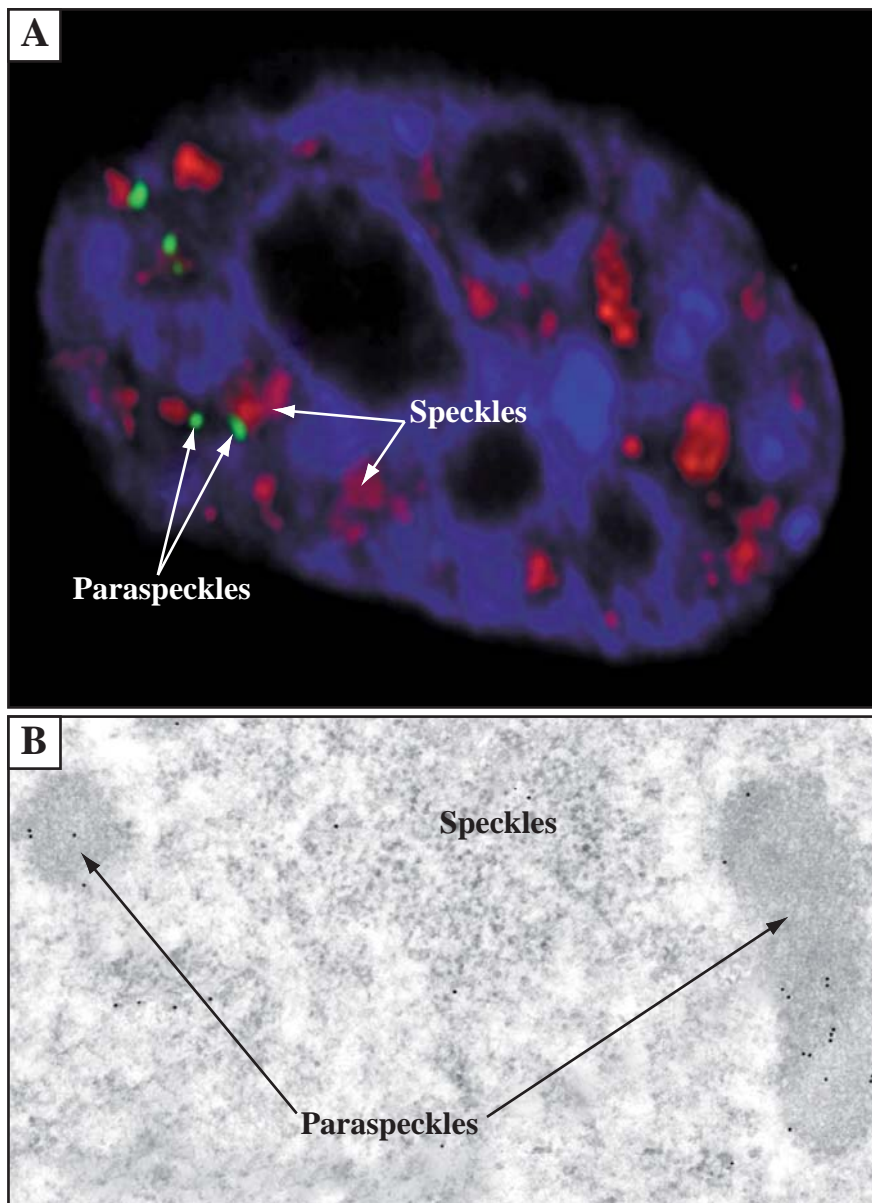


Il semble donc que l'exposition des gènes au niveau de l'espace interchromatinien constitue un élément déterminant pour leur expression. Il a notamment été montré que cet espace contient des zones focalisées de haute concentration en ARN polymérase II, baptisées *transcription factory* ou sites de transcription (Iborra *et al.*, 1996) (**Fig. 2**). Ces unités, numériquement inférieures au nombre de gènes exprimés, co-localisent avec un grand nombre d'unités transcriptionnelles et assurent l'expression de groupes de gènes physiquement proches, aussi bien que de gènes localisés sur des chromosomes distincts (Sutherland & Bickmore, 2009). Ces observations donnent lieu à de nouvelles perspectives puisque des résultats récents mettent en évidence que des interactions entre chromosomes au sein des espaces interchromatiniens pourraient contribuer à réguler l'expression des gènes. En effet, ces travaux suggèrent l'existence d'un véritable réseau de communication entre chromosomes qui, par des mécanismes d'action en *trans*, pourrait participer au contrôle des activités de transcription et d'épissage à l'échelle de groupes de gènes (Fraser & Bickmore, 2007; Gondor & Ohlsson, 2009; Rodley *et al.*, 2009).

### 1.3. Les speckles et les paraspeckles

Retrouvés au niveau de l'espace interchromatinien dans le noyau des cellules de mammifères (**Fig. 2**), les speckles sont des structures enrichies en facteurs d'épissage des pré-ARNm, comme les UsnRNP et les protéines SR (Lamond & Spector, 2003). Les speckles, parfois appelés compartiments de facteurs d'épissage, ont une taille variable de 1 à quelques  $\mu\text{m}$ . Ils s'assemblent sous la forme de grappes, désignées sous le terme d'IGC (*Interchromatin Granule Clusters*) et composées de granules de 20 à 25 nm, reliées entre elles par de fines fibrilles (Thiry, 1995) (**Fig. 4**). Les speckles ont également été identifiés chez le xénope et chez la drosophile, en revanche aucune structure similaire n'a été observée chez la levure.

Des études d'hybridation *in situ* ont montré que les speckles ne contiennent pas ou peu d'ADN et que leur assemblage ne s'effectue pas autour de *loci* génomiques. En revanche, ils sont localisés à proximité de gènes transcriptionnellement actifs (Thiry, 1995). Plusieurs équipes ont ainsi proposé que les speckles pourraient constituer des compartiments de stockage, d'assemblage, de modification et de recyclage des facteurs d'épissage dédiés à l'approvisionnement des sites de transcription (Misteli *et al.*, 1997; Saitoh *et al.*, 2004). Cette



**Figure 4 : Visualisation des speckles et des paraspeckles par microscopie électronique dans des noyaux de cellules HeLa.** (A) Localisation des speckles grâce à des anticorps anti-SC35 (en rouge) et des paraspeckles grâce à des anticorps anti-PSPC1 (en vert). La chromatine, marquée au DAPI, apparaît en bleu. (B) Les speckles (au centre) sont formés de granules regroupés sous forme de grappes. Les paraspeckles apparaissent de part et d'autre sous la forme de corpuscules plus homogènes. Images tirées de Fox *et al.*, 2002 (A) et Bond *et al.*, 2009 (B).

coopérativité entre sites de transcription et speckles pourrait notamment être régulée par des cycles de phosphorylation/déphosphorylation, assurant un contrôle des échanges de facteurs entre les deux compartiments (Misteli *et al.*, 1997; Saitoh *et al.*, 2004).

Les speckles sont des structures dynamiques qui varient en forme, en nombre et en taille en fonction des types cellulaires et en fonction de l'activité transcriptionnelle des cellules. Il a notamment été décrit que leur volume augmente de façon significative lorsque la transcription ou l'épissage sont inhibés (Melcak *et al.*, 2000; O'Keefe *et al.*, 1994). Les speckles sont dissociés au cours de la mitose, puis ils se ré-assemblent pendant la phase G1 sur la base d'interactions entre protéines, régulées par des cycles de phosphorylation/déphosphorylation (Lamond & Spector, 2003).

Des études plus récentes ont montré que les speckles peuvent également être associés à des structures distinctes baptisées paraspeckles (Bond & Fox, 2009; Fox *et al.*, 2002) (**Fig. 4**). Ces dernières sont des corpuscules dynamiques mesurant 0,5 à 1  $\mu\text{m}$ , variant en nombre en fonction des types cellulaires et dont la formation requiert un état actif de la transcription par l'ARN polymérase II. Ils ont pour particularité de s'assembler autour d'un lncRNA (*long non coding RNA*), l'ARN NEAT1, et constituent donc des corpuscules ribonucléoprotéiques (Chen & Carmichael, 2009; Clemson *et al.*, 2009).

D'autres travaux ont également montré que les paraspeckles constituent des sites de rétention d'ARN hyper-édités (conversion d'adénosines en inosines). Des protéines de la famille DBHS s'accumulent en effet dans les paraspeckles et pourraient piéger certains ARNm ayant subi une édition de leur séquence 3'UTR. Cette rétention apparaît néanmoins réversible, puisqu'en réponse à certains signaux cellulaires, les ARNm pourraient être libérés suite au clivage de la séquence éditée (Chen & Carmichael, 2009; Chen *et al.*, 2008; Osenberg *et al.*, 2009; Prasanth *et al.*, 2005). Encore peu connu, ce processus de régulation de l'expression génique par rétention d'ARNm semble impliqué dans les processus de différenciation cellulaire, la formation des paraspeckles étant notamment associée à une perte de pluripotence des cellules embryonnaires (Chen & Carmichael, 2009). Enfin, d'autres données révèlent une corrélation entre la formation des paraspeckles et certaines infections virales, suggérant ainsi que des virus pourraient utiliser les paraspeckles pour leur production ou bien que ces régions pourraient être impliquées dans une forme de défense anti-virale (Saha *et al.*, 2006).

## 1.4. Le nucléole

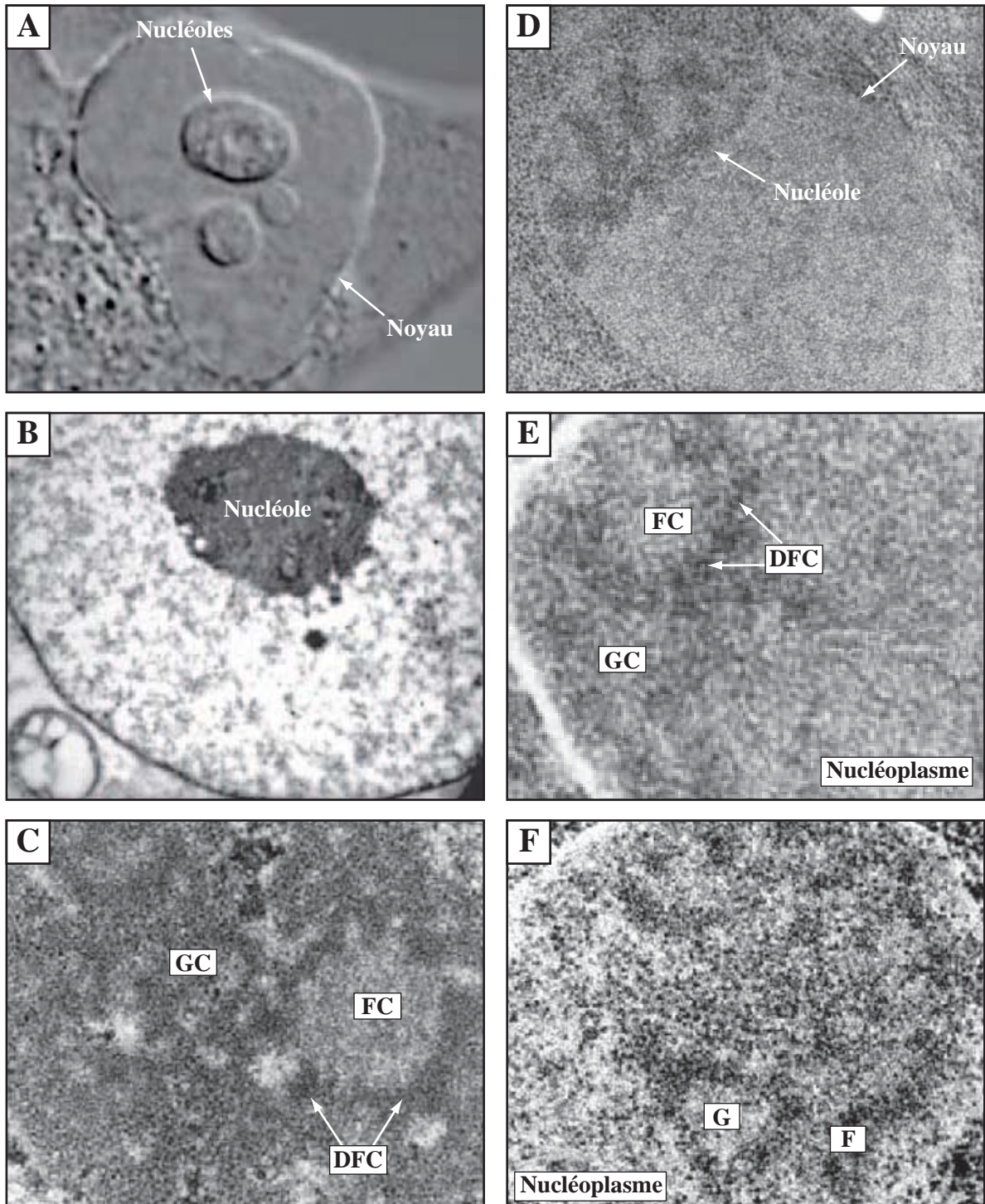
Le nucléole, initialement décrit sous le nom de *macula* germinative en 1835 par R. Wagner, a été la première structure sub-nucléaire observée. Chez les mammifères on compte de 1 à 5 nucléoles qui occupent généralement une position centrale au sein du noyau (**Fig. 2, 5A et 5B**). Ils forment une structure très dynamique, qui se dissocie lors de la mitose et qui se réassemble en fin de télophase (Boisvert *et al.*, 2007). Chez la levure *S. cerevisiae*, le nucléole adopte en revanche un positionnement plus périphérique, il adopte une structure en forme de croissant qui occupe environ un tiers du noyau et qui établit des contacts étendus avec l'enveloppe nucléaire (Oakes *et al.*, 1998; Taddei *et al.*, 2010) (**Fig. 5D**). De plus, comme c'est le cas pour les organismes ne présentant pas de rupture de l'enveloppe nucléaire lors de la mitose, le nucléole de *S. cerevisiae* persiste tout au long du cycle cellulaire.

### 1.4.1. Organisation du nucléole et biogenèse des ribosomes

Le nucléole est le site de production des ribosomes (Kressler *et al.*, 2010) (cf. Introduction, Chapitre 1, 3). Il s'organise autour de clusters génomiques, les NORs (*Nucleolar Organizing Region*), formés par la répétition en tandem des unités transcriptionnelles codant les ARNr. Chez les eucaryotes, le nombre de gènes ribosomiques varie d'une centaine à plusieurs centaines de copies. Les gènes d'ARNr sont transcrits dans le nucléole par l'ARN polymérase I sous la forme d'un long précurseur dont la maturation nucléolytique va permettre la libération des ARNr matures 5.8S, 18S et 28S (25S chez la levure). Au cours de ce processus, certains nucléotides sont modifiés de façon spécifique par les snoRNP à boîtes C/D et H/ACA, et les ARNr sont associés aux protéines ribosomiques pour former les sous-unités du ribosome. L'ARNr 5S, transcrit par l'ARN polymérase III est également incorporé dans la future sous-unité 60S, au sein du nucléole. Les particules pré-ribosomiques ainsi formées transitent ensuite vers le nucléoplasme avant d'être exportées vers le cytoplasme à travers les complexes des pores nucléaires (**Fig. 3**). Le nucléole, dont la morphologie semble être influencée par les besoins en traduction de la cellule, est souvent considéré comme une véritable usine à ribosomes, chaque cellule de levure génère ainsi une moyenne de 40 nouveaux ribosomes par seconde (Carmo-Fonseca *et al.*, 2000).

## Cellules HeLa

## *Saccharomyces cerevisiae*



**Figure 5 : Visualisation du nucléole par microscopie électronique dans des cellules HeLa et chez *Saccharomyces cerevisiae*.** (A) Noyau de cellule HeLa comportant plusieurs nucléoles en position centrale. (B) Noyau de cellule HeLa comportant un nucléole unique en position centrale. (C) Vue des 3 régions nucléolaires principales observées dans une cellule HeLa. (D) Noyau de cellule de levure *S. cerevisiae*. Le nucléole, en forme de croissant, occupe environ un tiers du volume nucléaire et adopte une position périphérique. (E, F) Vues des différentes régions nucléolaires observées chez *S. cerevisiae*. Images tirées de Boisvert *et al.*, 2007 (A, B) ; Tiry *et al.*, 2005 (C et F) et Trumtel *et al.*, 2000 (D et E). FC : Centres fibrillaires ; DFC : Centres Fibrillaires Denses ; GC : Composants Granulaires ; F : réseau de brins fibrillaires ; G : région granulaire.

Les processus complexes de formation des ribosomes impliquent plusieurs régions du nucléole qui peuvent être distinguées en microscopie électronique sur la base de différences morphologiques. Le nucléole de mammifères peut être subdivisé en trois sous-compartiments (Thiry & Lafontaine, 2005) : (i) une ou plusieurs structures centrales composées de fines fibrilles, appelées centres fibrillaires (FC, *Fibrillar Center*), (ii) une région plus périphérique formée de fibres compactes, appelée centres fibrillaires denses (DFC, *Dense Fibrillar Center*), (iii) une troisième région de structure granulaire entourant les deux précédentes, appelée composants granulaires (GC, *Granular Components*) (**Fig. 2 et 5C**). Cette architecture apparaît intimement liée à la synthèse des ribosomes qui se déroule de façon vectorielle depuis le centre vers la périphérie du nucléole. L'ADNr et les facteurs de transcription sont localisés dans les FC, la transcription se déroule vraisemblablement à l'interface des FC et des DFC (Huang, 2002), la maturation des pré-ARNr et l'assemblage des protéines ribosomiques débutent dans les DFC, puis se poursuivent dans les GC, avant la sortie des sous-unités ribosomiques dans le nucléoplasme et l'export vers le cytoplasme (Tschochner & Hurt, 2003).

Si l'organisation tripartite du nucléole est conservée chez les mammifères, elle ne constitue toutefois pas une caractéristique constante chez l'ensemble des eucaryotes. De précédentes études de microscopie, réalisées notamment chez *S. cerevisiae*, avaient suggéré la conservation de ce type d'organisation au sein du *phylum* (Leger-Silvestre *et al.*, 1999; Trumtel *et al.*, 2000) (**Fig. 5E**). Cependant, une étude comparative inter-organismes plus récente a permis de mettre en évidence que seuls les vertébrés amniotes possèdent un véritable nucléole tripartite, tandis que les autres organismes possèdent un nucléole bipartite composé d'un réseau de brins fibrillaires (F) entrelacé dans une région granulaire (G) (Thiry & Lafontaine, 2005) (**Fig. 5F**). Les auteurs de cette étude proposent que la division du réseau fibrillaire (F) en deux régions clairement distinctes (FC et DFC) chez les amniotes soit corrélée à une augmentation drastique de la taille des régions inter-géniques (IGS, *Intergenic Spacer*) séparant les clusters de gènes ribosomiques. Ils suggèrent ainsi, que l'émergence d'un troisième compartiment aurait pu permettre une régulation transcriptionnelle plus fine en permettant une meilleure séparation physique entre les régions transcrites et les IGS par un jeu de boucles chevauchant les FC et les DFC.

### 1.4.2. Les autres fonctions du nucléole

Des fonctions distinctes de la biogenèse des ribosomes ont également été identifiées au sein du nucléole (Boisvert *et al.*, 2007). Plusieurs autres classes d'ARN non codants y transitent notamment afin de subir des étapes de maturation (clivages, modifications, assemblage). C'est notamment le cas du snRNA U6 qui subit des réactions de 2'-O-méthylation et de pseudouridylation (Ganot *et al.*, 1999), de certains ARNt qui sont maturés par la RNase P nucléolaire (Bertrand *et al.*, 1998), ou encore de l'ARN télomérique de mammifères (Lin *et al.*, 2008a) et de l'ARN de la particule SRP de mammifères (Jacobson & Pederson, 1998).

D'autre part, plusieurs analyses protéomiques ont été menées afin de caractériser le protéome nucléolaire, elles ont permis de dénombrer plus de 700 protéines dans le nucléole humain (Andersen *et al.*, 2005; Andersen *et al.*, 2002; Scherl *et al.*, 2002). Ces études ont révélées une très grande conservation du protéome nucléolaire puisque environ 90% de ces protéines possèdent un homologue chez la levure *S. cerevisiae* (Andersen *et al.*, 2005). De façon intéressante, leur classification fonctionnelle a en revanche révélé que seulement 30% d'entre elles présentent une fonction relative à la biogenèse des ribosomes, suggérant ainsi l'importance des autres rôles joués par le nucléole (Hinsby *et al.*, 2006; Leung *et al.*, 2003). Parmi ces protéines on retrouve en effet des facteurs de contrôle du cycle cellulaire, des facteurs de réplication et de réparation de l'ADN, et également beaucoup de protéines impliquées dans l'épissage des pré-ARNm. De plus, ces études ont montré que la composition du protéome nucléolaire subit de larges changements en fonction des états physiologiques de la cellule, ce qui renforce l'idée d'une structure dynamique s'adaptant de façon morphologique et fonctionnelle aux besoins de la cellule. Il a ainsi été proposé que le nucléole puisse jouer un rôle clé dans la régulation du cycle cellulaire, par le biais de la séquestration transitoire de certains facteurs (Boisvert *et al.*, 2007). De nombreuses protéines apparaissent en effet associées au nucléole de façon spécifique à différents stades du cycle cellulaire. C'est par exemple le cas de la protéine Cdc14p de *S. cerevisiae*, une phosphatase essentielle pour promouvoir la sortie de mitose qui est piégée dans le nucléole en interphase avant d'être relarguée dans le nucléoplasme en début d'anaphase (Azzam *et al.*, 2004; Shou *et al.*, 1999).

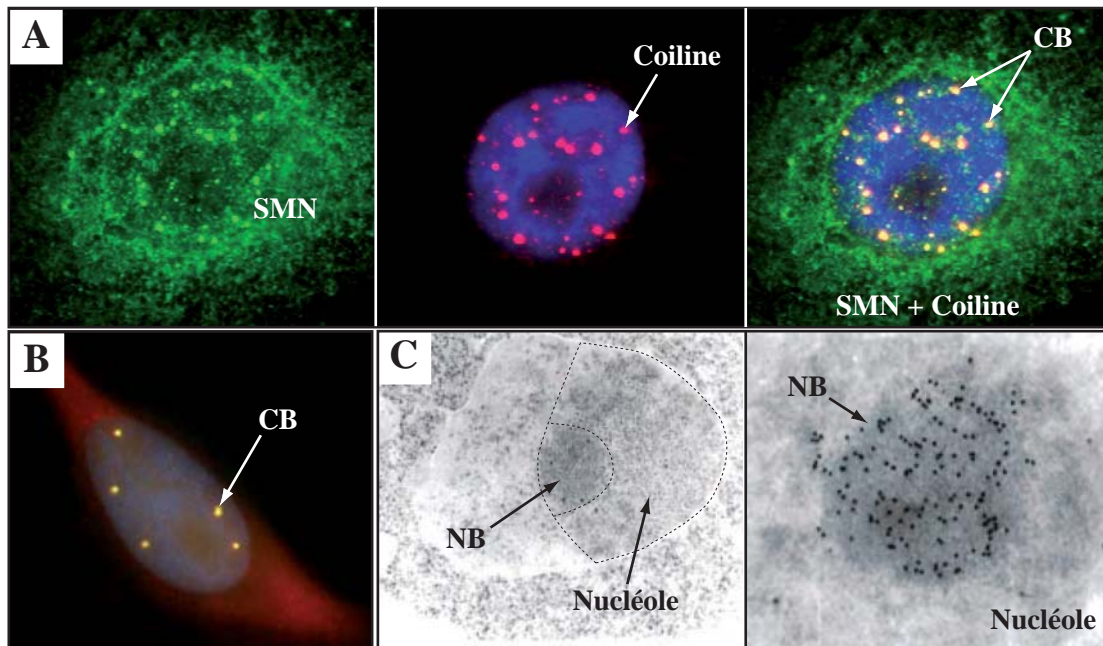
## 1.5. Les corps de Cajal et les corpuscules nucléolaires

### 1.5.1. Les corps de Cajal

Les corps de Cajal (CB, *Cajal Bodies*) ont été mis en évidence pour la première fois en 1903 par Ramon y Cajal dans des cellules neuronales de rat. Initialement décrits en tant que « corps accessoires du nucléole », ils ont été renommés *coiled bodies* dans les années 1960, du fait de leur structure rappelant un collier de perles, avant d'acquérir leur nom définitif en 1999 en hommage à leur découvreur (Gall *et al.*, 1999). Ces structures, retrouvées à la fois chez les animaux et chez les plantes, présentent une taille variant de 0,2 à 2  $\mu\text{m}$  et comptent 1 à 10 exemplaires par noyau, en fonction des espèces et des types cellulaires (Cioce & Lamond, 2005; Richard *et al.*, 2003) (**Fig. 2, 6A et 6B**). D'autre part, les CB qui sont désassemblés lors de la mitose, varient en taille et en nombre au cours du cycle cellulaire pour atteindre un maximum lors de la transition G1/S (Andrade *et al.*, 1993; Carmo-Fonseca *et al.*, 1993). L'abondance des CB apparaît ainsi corrélée aux besoins métaboliques et les cellules ayant un taux élevé d'expression génique ont généralement des CB plus développés (Andrade *et al.*, 1993; Carmo-Fonseca *et al.*, 1993).

Des études menées chez différents organismes ont également montré que les CB présentent un caractère très dynamique durant l'interphase (Cioce & Lamond, 2005). Ils peuvent fusionner et se diviser, ils peuvent aussi effectuer de larges mouvements au sein du nucléoplasme, par diffusion passive ou en suivant un mécanisme de transport actif en étroite association avec la chromatine (Boudonck *et al.*, 1999; Gorisch *et al.*, 2004; Platani *et al.*, 2002; Platani *et al.*, 2000). Ils peuvent ainsi se déplacer à la périphérie du nucléole ou même à l'intérieur de celui-ci (Platani *et al.*, 2000). Ils sont également détectables au niveau de zones précises du nucléoplasme, généralement associés à des *loci* génomiques codant les histones (Frey & Matera, 1995; Schul *et al.*, 1999b), ou à des *loci* codant certains UsnRNA (Frey & Matera, 2001; Smith & Lawrence, 2000) et snoRNA (Gao *et al.*, 1997; Schul *et al.*, 1999a). Enfin, ils apparaissent parfois en étroites relations avec d'autres corps nucléaires comme les corps PML (*Promyelocytic leukemia nuclear bodies*), suggérant ainsi l'existence de phénomènes d'échanges dynamiques entre ces différentes structures sub-nucléaires (Cioce & Lamond, 2005).





**Figure 6 : Visualisation des Corps de Cajal et des Corpuscules nucléolaires par microscopie électronique respectivement dans des cellules HeLa et chez *Saccharomyces cerevisiae*.** (A) Visualisation des CB (en jaune) dans une cellule HeLa par co-localisation des protéines SMN (en vert) et Coiline (en rose). Dans cette expérience de co-localisation, la protéine SMN est largement surexprimée et induit la formation d'un grand nombre de CB. La chromatine, marquée au DAPI, apparaît en bleu. La protéine SMN diffuse dans le cytoplasme. (B) Visualisation des CB dans une cellule HeLa dans un contexte d'expression normale. On dénombre seulement 5 CB. (C) Visualisation des NB chez *S. cerevisiae* grâce à la localisation d'un snoRNA (U14-MS2) par hybridation *in situ* avec une sonde fluorescente. La concentration du snoRNA matérialise une zone correspondant au NB, située à l'intérieur du nucléole (régions entourées de pointillés). Images tirées de Morris *et al.*, 2008 (A) ; Whittom *et al.*, 2008 (B) et Verheggen *et al.*, 2001 (C). CB : Cajal Bodies; NB : Nucleolar Bodies.

L'étude de la composition moléculaire des CB a révélé une grande diversité. L'un de ses principaux composants est la Coiline (également appelée p80-Coiline), une protéine découverte en 1991 à partir d'un sérum auto-immun (Raska *et al.*, 1991). Fréquemment utilisée en temps que marqueur des CB (**Fig. 6**), la Coiline joue un rôle essentiel pour leur formation et pour le maintien de leur intégrité. Des expériences réalisées chez la souris ont en effet montré que son absence conduit à une diminution de la viabilité des organismes, associée à la formation de CB résiduels incomplets et hétérogènes en composition (Bauer & Gall, 1997; Tucker *et al.*, 2001). Cette propriété repose sur sa capacité à s'oligomériser et à interagir avec d'autres facteurs protéiques afin de favoriser la concentration locale de nombreux composants. De plus, certaines études ont mis en évidence que cette fonction peut être modulée selon le métabolisme cellulaire par le biais de modifications post-traductionnelles. Il a notamment été montré que l'hyperphosphorylation de la Coiline, qui réduit sa capacité d'oligomérisation, est corrélée au désassemblage des CB durant la mitose (Carmo-Fonseca *et al.*, 1993). D'autres données indiquent également que la méthylation de certaines arginines est requise pour l'interaction avec un autre composant majeur des CB, la protéine SMN (*Survival of Moto Neurons*) (Hebert *et al.*, 2002).

La principale fonction des CB semble liée à la biogenèse de certaines classes de RNP contenant des ARN non codants. Ils constituent notamment des sites privilégiés pour la modification et l'assemblage des UsnRNP en provenance du cytoplasme ou recyclées après leur intervention dans une réaction d'épissage (Morris, 2008). Les CB sont en effet enrichis en UsnRNP (Sleeman & Lamond, 1999a) et en facteurs impliqués dans leur biogenèse. Parmi ces facteurs on retrouve notamment la protéine SMN, qui joue un rôle essentiel pour l'assemblage des UsnRNP (Meister *et al.*, 2002; Pellizzoni *et al.*, 2002b), ainsi que les scaRNP (*small cajal bodies snoRNP*), des RNP spécifiques des CB réunissant les caractéristiques des snoRNP à boîtes C/D et H/ACA, et responsables de la modification des UsnRNA par 2'-O-méthylation et pseudouridylation (Darzacq *et al.*, 2002; Jady & Kiss, 2001). A la différence des speckles, le rôle des CB dans l'épissage semble en revanche se limiter à la production des UsnRNP, puisqu'ils sont dépourvus d'autres facteurs d'épissage majeurs, comme les protéines SR, et qu'aucun ARN polyadénylé n'y a été détecté.

Plusieurs études ont également mis en évidence la présence transitoire de snoRNP à boîtes C/D et H/ACA au sein des CB (Boulon *et al.*, 2004; Jady *et al.*, 2004; Narayanan *et al.*, 1999b; Samarsky *et al.*, 1998; Verheggen *et al.*, 2002). Cette accumulation pourrait être

notamment médiée par la protéine NOPP140, qui interagit à la fois avec la Coiline et avec les snoRNP à boîtes C/D et H/ACA (Isaac *et al.*, 1998; Yang *et al.*, 2000). Cette protéine, également détectée en abondance dans le nucléole, semble transiter entre les deux compartiments et pourrait ainsi faciliter le transport et la localisation des snoRNP (Verheggen *et al.*, 2001; Yang *et al.*, 2000). Cette fonction de transport intranucléaire des snoRNP implique aussi les facteurs PHAX et CRM1, qui s'accumulent également dans les CB (Boulon *et al.*, 2004; Frey & Matera, 2001). Il a ainsi été établi que les snoRNP séjournent dans les CB pour y subir certaines étapes de maturation et d'assemblage avant de rejoindre le nucléole. Chez l'homme, l'assemblage des protéines Fibrillarine, NOP56 et NOP58 sur le snoRNA U3 semble se dérouler au sein des CB (Verheggen *et al.*, 2002). De plus, les étapes finales de maturation nucléolytique de l'extrémité 3' de cet ARN et l'hyperméthylation de sa coiffe par l'enzyme TGS1 sont également effectuées dans ces corpuscules (Verheggen *et al.*, 2002).

Le composant ARN de la télomérase humaine (hTERC) est également détecté dans les CB, en particulier dans les cellules cancéreuses en interphase (Jady *et al.*, 2004). Cet ARN contient en effet des caractéristiques communes avec les ARN à boîte H/ACA puisqu'il possède, dans sa région 3', un motif H/ACA qui s'avère essentiel à sa stabilité, à sa fonction, mais aussi à sa localisation au sein des CB (Fu & Collins, 2003; Jady *et al.*, 2004; Mitchell *et al.*, 1999a). Plusieurs études ont ainsi mis en évidence que le domaine H/ACA de hTERC renferme une séquence d'adressage spécifique des CB, la CAB box (*Cajal Bodies box*), également présente dans les scaRNA (Fu & Collins, 2003; Jady *et al.*, 2004; Richard *et al.*, 2003; Theimer *et al.*, 2007). De plus, il a été montré que la protéine SMN, impliquée dans l'assemblage des UsnRNP, interagit également avec l'enzyme hTERT du complexe télomérase (Bachand *et al.*, 2002). S'il semble vraisemblable que les CB constituent un site d'assemblage et de maturation de la télomérase, l'importance fonctionnelle de cette localisation n'a cependant pas encore été totalement résolue. A la différence des snoRNP et des UsnRNP qui séjournent transitoirement dans les CB, la télomérase semble y résider de façon permanente (Jady *et al.*, 2004). Il a notamment été montré qu'une partie des télomères associés à hTERC co-localisent avec les CB, suggérant que ces derniers pourraient être directement impliqués dans la fonction de maintenance des télomères, en délivrant la télomérase au niveau de ses sites d'activité (Jady *et al.*, 2006). Enfin, les CB qui sont fréquemment associés aux *loci* codant les histones, présentent également une forte accumulation de la snRNP U7 qui joue un rôle essentiel dans la maturation de l'extrémité 3' des ARNm codant les protéines histones (Frey & Matera, 1995; Pillai *et al.*, 2001).

Outre les complexes RNP, les CB contiennent d'autres composants comme des facteurs de transcription et des protéines impliquées dans la régulation du cycle cellulaire. Les CB ne sont pas directement impliqués dans les mécanismes de transcription, néanmoins ils sont parfois détectés à proximité des sites de transcription. Ils pourraient donc permettre de mobiliser certains facteurs de transcription et, par le biais de leur mobilité, les délivrer au niveau de *loci* précis comme ceux codant les UsnRNA, les snoRNA ou encore les histones. Ils contiennent notamment une sous-unité du facteur de transcription PTF $\gamma$ , spécifique des gènes de UsnRNA, ainsi que certains composants de l'ARN polymérase II, comme TBP, TFIIH et TFIIIF (Schul *et al.*, 1998). La forme hypophosphorylée de la grande sous-unité de l'ARN polymérase II est détectée dans les CB, tandis que sa forme hyperphosphorylée se localise en périphérie du compartiment (Schul *et al.*, 1998). Les facteurs de clivage et polyadénylation CstF77 et CPSF100 sont également retrouvés dans les CB (Gall *et al.*, 1999). Enfin, des sous-unités des ARN polymérase I et III ont également été mis en évidence dans les CB d'ovocytes de xénope (Gall *et al.*, 1999). Concernant la régulation du cycle cellulaire, les CB contiennent plusieurs complexes cycline/kinase, comme le complexe Cdk2/Cyline E qui intervient dans la régulation de la transcription des gènes d'histones lors de l'entrée en interphase (Liu *et al.*, 2000; Ma *et al.*, 2000; Ye *et al.*, 2003). En conditions de stress, ils peuvent également contenir la protéine p53 qui s'y accumule par l'intermédiaire d'une interaction avec la protéine SMN (Young *et al.*, 2002).

### 1.5.2. Les corpuscules nucléolaires

Chez la levure *S. cerevisiae*, il n'existe pas de CB tels qu'ils ont été décrits chez les animaux ou chez les plantes. En revanche, des travaux de l'équipe d'E. Bertrand ont rapporté l'existence de structures présentant des caractéristiques semblables, situées à l'intérieur même du nucléole (Verheggen *et al.*, 2001) (**Fig. 6C**). Ces structures d'aspect sphérique, baptisées corpuscules nucléolaires (NB, *Nucleolar Bodies*), sont généralement présentes en un exemplaire unique par noyau et localisent en bordure des DFC, qui constituent les sites d'action des snoRNP matures. Il a en effet été montré que des snoRNA à boîtes C/D surexprimés dans les cellules de levures s'accumulent de façon marquée au niveau des NB. En revanche ceux-ci semblent spécifiques des snoRNP à boîtes C/D, puisqu'aucune accumulation de snoRNA à boîtes H/ACA ou de UsnRNA n'a été constatée. D'autre part, les protéines

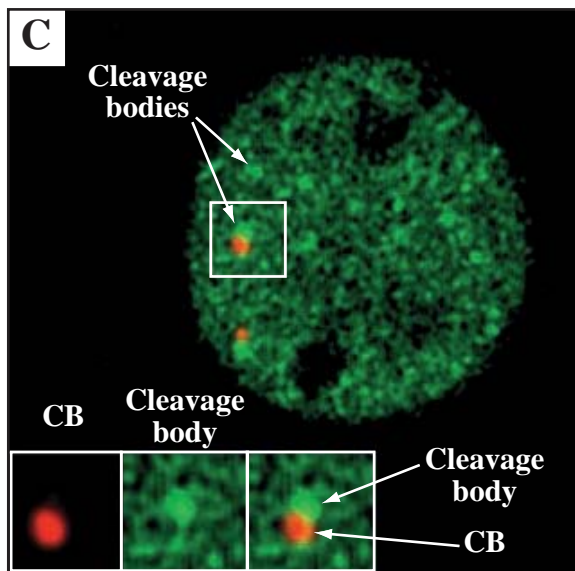
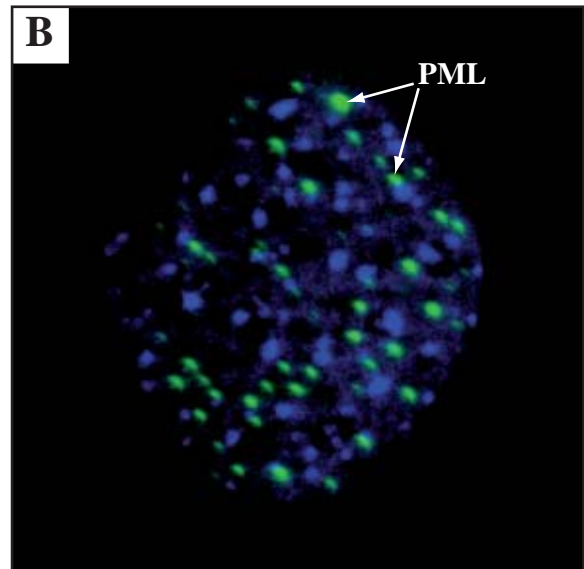
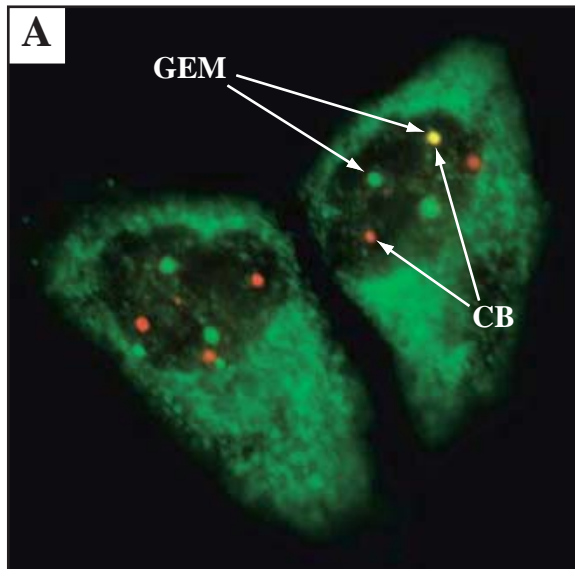
constitutives des snoRNP à boîtes C/D (Snu13p, Nop1p, Nop56p et Nop58p) qui s'accumulent également au niveau des NB, sont requises pour la localisation des snoRNA dans le nucléole et dans les NB.

En dépit de l'absence d'homologue des protéines Coiline et SMN chez *S. cerevisiae*, les auteurs ont comparé certaines caractéristiques des CB et des NB (Verheggen *et al.*, 2001). Ils ont notamment montré que la protéine humaine SMN exprimée de façon hétérologue dans la levure co-localise avec les snoRNA au niveau des NB. La protéine Srp40p, l'homologue de NOPP140, s'accumule également au niveau des NB et semble requise pour la rétention des snoRNA à boîtes C/D dans ces compartiments. A l'inverse, ils ont montré que la protéine Nsr1p, l'homologue de la Nucléoline (un composant majeur du nucléole), présente un effet antagoniste en favorisant la dispersion des snoRNA à travers le nucléole. De façon similaire aux CB, les NB renferment également l'hyperméthylase Tgs1p, ainsi que des précurseurs du snoRNA U3 non maturés en 3' (Verheggen *et al.*, 2002). Enfin, à l'image des CB qui constituent des structures dynamiques, les NB sont observables dans des conditions de croissance particulières (Verheggen *et al.*, 2002). La culture des cellules en milieu solide, associée à d'importants échanges gazeux, mais à un accès limité aux nutriments, est favorable à leur formation. A l'inverse, la culture des cellules en milieu liquide, associée à des échanges gazeux faibles, mais à un niveau d'accès élevé aux nutriments, ne semble pas propice à leur formation. Cette observation suggère donc que la mise en place des NB et la biogenèse des snoRNA sont adaptées selon des besoins spécifiques, liés à l'état physiologique des cellules.

## 1.6. Les autres corpuscules nucléaires

En plus du nucléole, des speckles ou des CB déjà bien caractérisés, les études de microscopie ont révélé l'existence de nombreux autres corpuscules nucléaires parmi lesquels les GEMs, les PML ou encore les *cleavage bodies* (Spector, 2006; Zhao *et al.*, 2009) (**Fig. 2 et 7**).

Les GEMs, encore appelés *Geminis des Cajal Bodies* sont des corpuscules nucléaires enrichis en protéine SMN mais dépourvus en UsnRNP (Liu & Dreyfuss, 1996; Matera & Frey, 1998). Ils colocalisent souvent avec les CB mais ils peuvent également apparaître de façon distincte (**Fig. 7A**). Cette association réversible est régulée par le taux de diméthylation de la Coiline



**Figure 7 : Visualisation des GEMs, des PML et des Cleavage bodies par microscopie électronique dans des cellules HeLa. (A)** Visualisation des GEMs grâce à des anticorps anti-SMN (en vert), les CB sont marqués grâce à des anticorps anti-coiline (en rouge). Le signal jaune indique la colocalisation des deux structures. SMN apparaît de façon diffuse dans le cytoplasme. **(B)** Visualisation des PML grâce à des anticorps anti-PML (en vert), la chromatine, marquée au DAPI, apparaît en bleu. **(C)** Visualisation des Cleavage bodies grâce à des anticorps anti-CstF64 (en vert). Certains d'entre-eux sont adjacents aux CB marqués à l'aide d'anticorps anti-coiline (en rouge). Images tirées de Gubitz *et al.*, 2004 (A) ; Bernardi *et al.*, 2007 (B) et Schull *et al.*, 1999 (C). GEM : Gemini des Cajal Bodies ; CB : Cajal Bodies ; PML : Promyelocytic Leukemia bodies.

qui influe sur son interaction avec SMN. La Coiline est ainsi hypométhylée dans les cellules où les GEMs sont séparés des CB (Boisvert *et al.*, 2007; Hebert *et al.*, 2002; Young *et al.*, 2001). La fonction précise des GEMs demeure inconnue et il n'existe pas d'équivalent chez la levure *S. cerevisiae*, qui ne possède pas d'homologue de la protéine SMN.

Les PML (*ProMyelocytic Leukemia bodies*), parfois appelés ND10 (*Nuclear Domain 10*) ou corps de Kremer, sont des corpuscules sphériques de 0,3 à 1 µm de diamètre, présents en 10 à 30 exemplaires par noyau dans les cellules de mammifères (Maul *et al.*, 2000) (**Fig. 7B**). Ce sont des corpuscules très mobiles qui peuvent s'associer à d'autres structures nucléaires comme les CB ou les *cleavage bodies* (Muratani *et al.*, 2002) et qui peuvent également adopter une localisation cytoplasmique (Bellodi *et al.*, 2006). Les PML contiennent plus d'une quarantaine de facteurs, parmi lesquels on retrouve notamment des suppresseurs de tumeurs, comme la protéine PML qui constitue le composant majeur de ces structures, le facteur p53 ou encore la protéine Rb (*RetinoBlastoma protein*) (Dellaire *et al.*, 2003). Ils sont impliqués dans de nombreux processus tels que la régulation de la transcription (Marcello *et al.*, 2003; Ruggero *et al.*, 2000), le contrôle de la croissance cellulaire (Marcello *et al.*, 2003; Ruggero *et al.*, 2000), l'apoptose (Bernardi & Pandolfi, 2003; Guo *et al.*, 2000; Quignon *et al.*, 1998), la réparation de l'ADN (Carbone *et al.*, 2002; Conlan *et al.*, 2004; Dellaire & Bazett-Jones, 2004) ou encore la réponse anti-virale médiée par l'interféron (Regad & Chelbi-Alix, 2001). Les PML sont également impliqués dans le mécanisme ALT (*Alternative Lengthening of Telomeres*), un processus d'élongation des télomères par recombinaison homologue, alternatif à l'action de la télomérase (Grobelny *et al.*, 2000; Yeager *et al.*, 1999).

Les *cleavage bodies* sont des foci de 0,2 à 1 µm, présents au nombre de 1 à 4 par cellules et concentrant des facteurs impliqués dans la maturation de l'extrémité 3' des ARNm (Schul *et al.*, 1996; Schul *et al.*, 1999b) (**Fig. 7C**). Identifiés dans les cellules humaines, ces structures contiennent en effet des facteurs comme CstF64 et CPSF 100, impliqués dans les mécanismes de clivage et polyadénylation, ainsi que pour 20% d'entre eux, des ARNm néo-synthétisés (Schul *et al.*, 1996; Schul *et al.*, 1999b). Les *cleavage bodies* sont souvent adjacents aux CB et sont fréquemment associés aux gènes codant les histones, suggérant un rôle commun des deux compartiments dans la biogenèse des ARNm issus de ces *loci* (Schul *et al.*, 1999b).

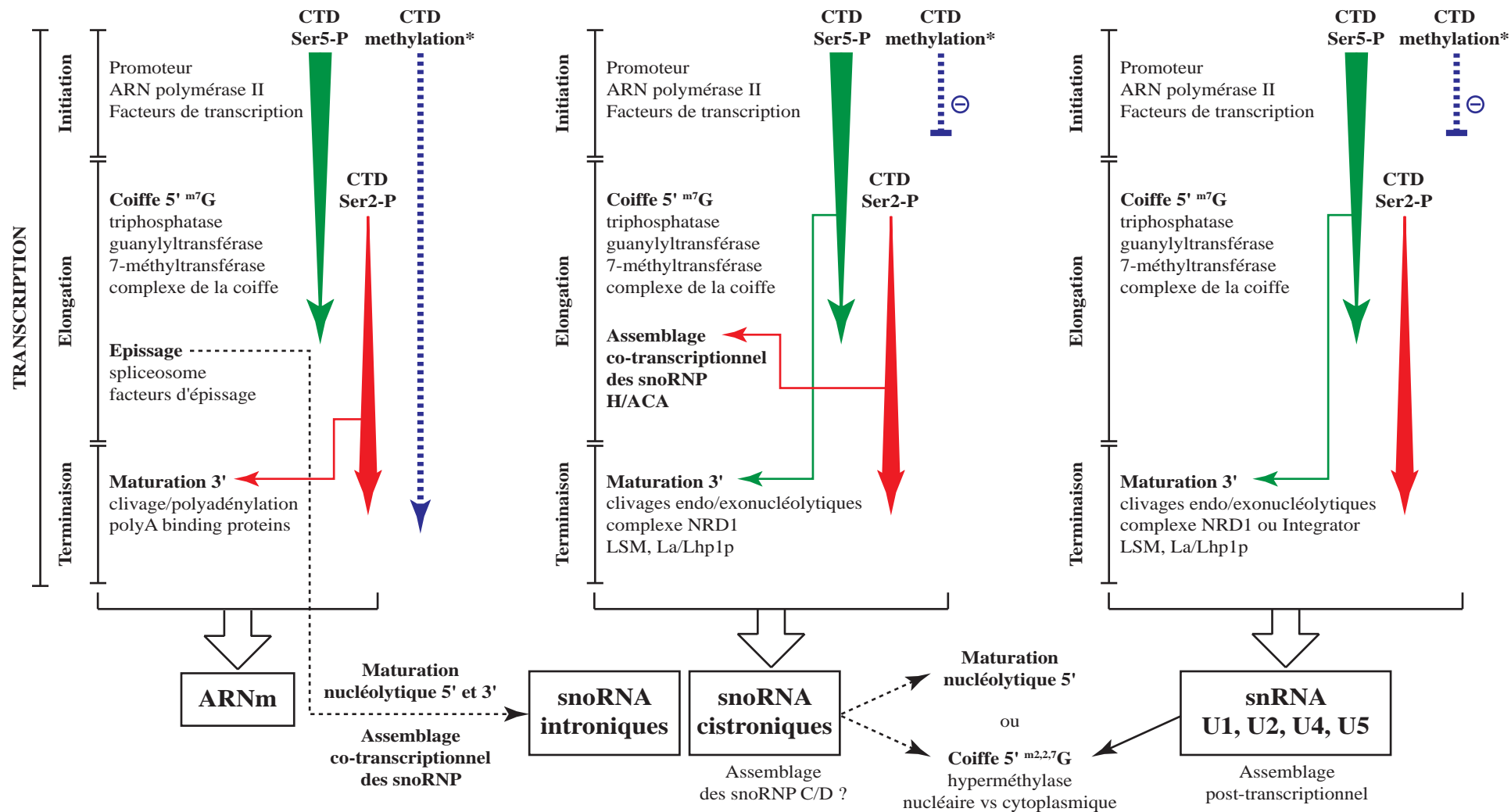
## 2. Biogenèse des ARN transcrits par l'ARN polymérase II

Chez les eucaryotes, l'ARN polymérase II est la machinerie nucléaire responsable de la transcription des ARNm. L'export des transcrits et leur traduction au niveau du cytoplasme requiert le déroulement préalable d'événements de maturation spécifiques, comme l'ajout d'une coiffe en 5', la polyadénylation de l'extrémité 3' et l'épissage des séquences introniques (**Fig. 8**). L'ensemble de ces étapes nécessite l'intervention d'un grand nombre de facteurs, dont certains s'associent de façon durable avec les transcrits et constituent des éléments d'information pouvant influencer sur leur export, leur localisation, leur stabilité et leur traduction. De nombreuses études ont ainsi révélé que les différentes étapes de la maturation des ARNm font l'objet de couplages visant à coordonner leur déroulement et à assurer l'efficacité et la fidélité du processus d'expression des gènes. L'ARN polymérase II est également responsable de la transcription de nombreux ARN non codants comme les snoRNA à boîtes C/D et H/ACA, les UsnRNA ou les microRNA. Parmi ces molécules, certaines possèdent leur propre unité transcriptionnelle et sont engagées, avec certaines spécificités que nous aborderons, dans les voies de maturation spécifiques des produits de l'ARN polymérase II (**Fig. 8**). D'autres, en revanche, sont générées au cours de la maturation de pré-ARNm et constituent donc des produits dérivés de l'activité de l'ARN polymérase II (**Fig. 8**).

### 2.1. Transcription par l'ARN polymérase II et rôle du domaine CTD

A la différence des bactéries et des archées qui possèdent une seule ARN polymérase, les eucaryotes possèdent trois ARN polymérases spécialisées (**Fig. 1**). L'étude de leurs caractéristiques a néanmoins permis de révéler le caractère fondamental et universel du processus de transcription, puisque les ARN polymérases présentes dans les 3 *phyla* partagent 5 sous-unités communes qui sont à la base de leur fonction enzymatique (Kwapisz *et al.*, 2008). En plus de ces éléments communs, les ARN polymérases eucaryotes sont constituées d'un grand nombre d'autres polypeptides apportant des spécificités différentes à chaque enzyme. L'ARN polymérase II est la machinerie la plus complexe puisqu'elle implique une soixantaine de protéines (Hahn, 2004).





**Figure 8 : Bilan des phénomènes de couplages entre la transcription et les différentes étapes de maturation des transcrits de l'ARN polymérase II.** CTD : Carboxy Terminal Domain (Domaine C-terminal de l'ARN polymérase II) ; Ser5-P : phosphorylation de la Sérine 5 (en vert) ; Ser2-P : phosphorylation de la Sérine 2 (en rouge). Une étude récente indique que la méthylation de l'arginine 1810\* du domaine CTD humain pourrait contribuer à réguler négativement l'expression des snoRNA et des snRNA (Sims *et al.*, 2011). De nombreuses études ont révélé le caractère co-transcriptionnel de l'assemblage des snoRNP. La question demeure en revanche posée dans le cas des snoRNA à boîtes C/D issus d'unités transcriptionnelles indépendantes. Enfin, l'assemblage des snRNP, qui implique une phase cytoplasmique, constitue pour sa part, un processus essentiellement post-transcriptionnel.

L'ARN polymérase II possède notamment un domaine spécifique, le domaine C-terminal de la sous-unité RPB1 (CTD, *Carboxy-Terminal Domain*), qui joue un rôle majeur dans la régulation de la transcription et dans l'intégration des phénomènes de maturation des transcrits (Hirose & Ohkuma, 2007; Hocine *et al.*, 2010). Ce domaine sert de plateforme pour le recrutement de nombreux facteurs, dont certains possèdent des domaines d'interaction qui lui sont dédiés : les *CTD Interacting Domain* (CIDs), les domaines WW et les domaines FF (Noble *et al.*, 2005; Smith *et al.*, 2004; Verdecia *et al.*, 2000). De nombreuses études ont mis en évidence que ces interactions et les fonctions qui leurs sont associées font l'objet de régulations fines. Le domaine CTD est en effet composé de répétitions en tandem  $Y_1S_2P_3T_4S_5P_6S_7$  (26 chez la levure, 52 chez l'humain), dont les Serines 2 et 5 constituent des sites de phosphorylation (Hirose & Ohkuma, 2007). La phosphorylation de la Serine 5 par la kinase CDK7 (Kin28p chez la levure), associée au facteur de transcription TFIIF, s'effectue lors de l'initiation de la transcription (Hengartner *et al.*, 1998; Komarnitsky *et al.*, 2000). Elle contribue notamment au recrutement des enzymes impliquées dans l'ajout de la coiffe (Cho *et al.*, 1997; Ho & Shuman, 1999) et de facteurs de terminaison de la transcription spécifiques de petits transcrits non codants, comme les snoRNA (Vasiljeva *et al.*, 2008) (**Fig. 8**). La phosphorylation de la Serine 2 s'effectue de façon plus tardive sous l'effet de la kinase CDK9 (Ctk1p chez la levure) du complexe PTEF-b (Komarnitsky *et al.*, 2000). Elle joue un rôle majeur dans la progression de l'élongation et permet le recrutement de la machinerie de clivage/polyadénylation de l'extrémité 3' des transcrits (Ahn *et al.*, 2004) (**Fig. 8**). Le domaine CTD est également impliqué dans le couplage entre la transcription et l'épissage. Le facteur d'élongation TAT-SF2, qui interagit avec l'ARN polymérase II par l'intermédiaire de son association avec PTEF-b, peut notamment former un complexe avec les UsnRNP du spliceosome et stimuler à la fois la transcription et l'épissage (Fong & Zhou, 2001). Il a également été montré que plusieurs facteurs d'épissage établissent des interactions directes avec le domaine CTD. C'est notamment le cas des protéines SCAF (*SR-like CTD Associated Factors*) (Yuryev *et al.*, 1996) ou encore de la protéine de levure Prp40p, qui appartient à la snRNP U1 (Morris & Greenleaf, 2000). Plusieurs études ont également mis en évidence l'implication du domaine CTD dans les événements de remodelage de la chromatine. Il a notamment été montré chez la levure que le recrutement de la protéine Set1p, responsable de la triméthylation de l'histone H3 dans la région 5' des gènes transcrits, implique la phosphorylation de la Serine 5 (Ng *et al.*, 2003; Santos-Rosa *et al.*, 2002). Enfin, le domaine CTD joue un rôle clé dans la biogenèse de certaines RNP. Comme nous le verrons plus tard, plusieurs études ont mis en évidence son implication dans l'assemblage des snoRNP à boîtes

H/ACA (Ballarino *et al.*, 2005; Darzacq *et al.*, 2006; Fatica *et al.*, 2002b; Richard *et al.*, 2006; Yang *et al.*, 2005). La phosphorylation de la Serine 2 pourrait notamment constituer un élément déterminant pour le déclenchement du processus d'assemblage co-transcriptionnel des particules H/ACA (Fatica *et al.*, 2002b).

L'ensemble de ces données illustre donc le rôle central joué par le domaine CTD dans les mécanismes de couplage entre transcription et maturation des ARN (**Fig. 8**). Ce domaine, conservé chez les eucaryotes, contribue de ce fait à la mise en place de processus de maturation spécifiques des transcrits de l'ARN polymérase II. Comme nous l'avons vu, ces événements sont soumis à une régulation par phosphorylation au niveau de sites conservés. De façon remarquable, le ratio de phosphorylation des Serines 5 et 2 évolue au cours de la transcription, puisque la phosphorylation de la Serine 5 diminue, tandis que la phosphorylation de la Serine 2 augmente au cours de l'élongation (Komarnitsky *et al.*, 2000). Il a notamment été proposé que cette relation puisse constituer un système métrique relatif au positionnement de la polymérase, permettant un ajustement précis des processus de maturation subséquent (Gudipati *et al.*, 2008). Des analyses structurales du domaine CTD ont notamment abouti à l'introduction de la notion de « Code CTD » (Buratowski, 2003). Ce code inclurait dans sa combinaison, l'état de phosphorylation des Serines 2, 5 et 7, ainsi que l'état d'isomérisation (*cis* ou *trans*) des Prolines 3 et 6, et permettrait d'adapter la conformation du domaine CTD en fonction de l'interaction avec ses partenaires successifs. Enfin, une étude récente a montré que l'activité de l'ARN polymérase II humaine peut être modulée par la méthylation de l'Arginine 1810 du domaine CTD (Sims *et al.*, 2011). Cette modification est catalysée par l'enzyme CARM1 et semble s'effectuer avant la phosphorylation de la Serine 5. De façon intéressante, les auteurs ont montré que la mutation du résidu R1810, comme la délétion du gène *CARM1*, provoque une augmentation drastique du taux de UsnRNA et de snoRNA (**Fig. 8**). Cette modification pourrait ainsi jouer un rôle régulateur en permettant de discriminer le recrutement de certains partenaires de l'ARN polymérase II en fonction du type de gène transcrit.

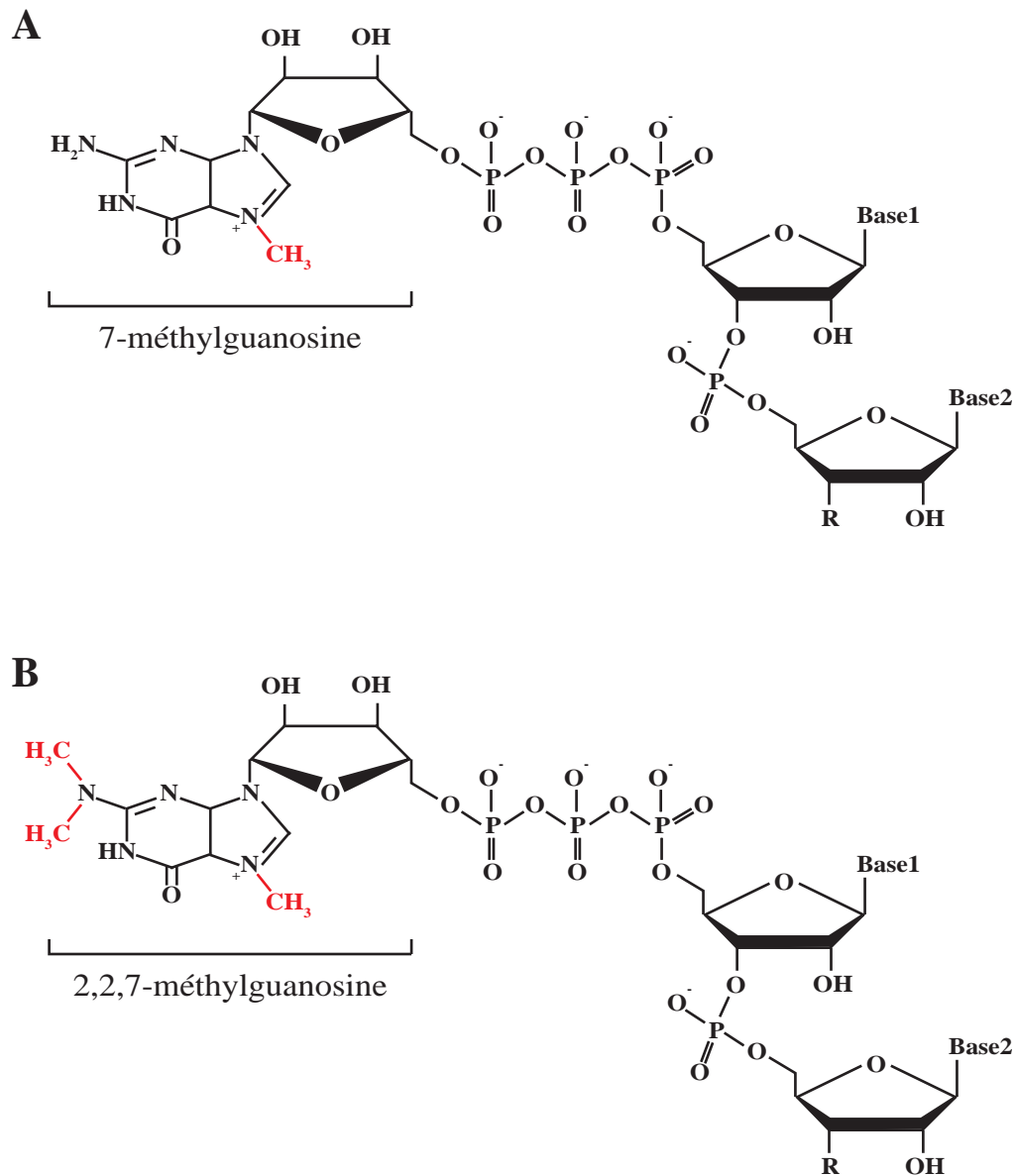
## 2.2. Maturation de l'extrémité 5' des transcrits de l'ARN polymérase II

L'ajout d'une coiffe  $m^7G$  au niveau de l'extrémité 5' des ARN est une caractéristique des transcrits de l'ARN polymérase II. Cette coiffe apporte de nouvelles propriétés à la molécule d'ARN. Elle est impliquée dans de nombreux processus et contribue aux couplages entre les événements de maturation. Parmi les transcrits coiffés, certains ARN non codants, comme les snRNA U1, U2, U4 et U5, les snoRNA ou l'ARN de la télomérase, subissent une étape de maturation supplémentaire qui consiste en une double méthylation en position 2 de la guanine. Cette coiffe triméthylée  $m^{2,2,7}G$  leur confère alors des propriétés distinctes de celles des ARNm. Chez la levure, la production de certains snoRNA implique en revanche, la maturation nucléolytique de l'extrémité 5' et la perte de la coiffe  $m^7G$ .

### 2.2.1. La coiffe 7-méthylguanosine ( $m^7G$ )

Le premier événement de maturation des transcrits de l'ARN polymérase II est la mise en place d'une coiffe 7-méthylguanosine au niveau de l'extrémité 5' (**Fig. 9A**). Cette étape requiert trois activités enzymatiques : une ARN triphosphatase qui hydrolyse le phosphate  $\gamma$  à l'extrémité 5' de l'ARN ( $5'pppN \rightarrow 5'ppN$ ), une ARN guanylyltransférase qui additionne un groupement GMP au niveau de la nouvelle extrémité 5' ( $5'ppN \rightarrow 5'GpppN$ ), et une ARN guanine-7-méthyltransférase qui catalyse la méthylation de la guanine au niveau de sa position N7 ( $5'GpppN \rightarrow 5'^{m^7}GpppN$ ) (Cowling, 2009). Chez *S. cerevisiae*, ces trois activités dépendent de trois protéines différentes, respectivement Cet1p, Ceg1p et Abd1p (Mao *et al.*, 1995; Shibagaki *et al.*, 1992; Tsukamoto *et al.*, 1997). En revanche, seulement deux facteurs sont impliqués chez les métazoaires, puisque les activités triphosphatase et guanylyltransférase sont portées par la même enzyme, baptisée RNGTT (*RNA guanylyltransferase-5'triphosphatase*) ou CE (*Capping Enzyme*), tandis que l'activité méthyltransférase est portée par l'enzyme RNMT (*RNA guanine-7-methyltransferase*) (Tsukamoto *et al.*, 1998; Yamada-Okabe *et al.*, 1998).

Comme nous l'avons vu précédemment, la mise en place de la coiffe, qui s'effectue de façon co-transcriptionnelle, est initiée par des contacts entre la machinerie de coiffage et l'ARN polymérase. Chez les mammifères, la RNGTT interagit directement avec le domaine CTD



**Figure 9 : Structures des coiffes spécifiques de l'extrémité 5' des transcrits de l'ARN polymérase II. (A) Coiffe 7-méthylguanosine ( $m^7G$ ) des ARNm. Sa mise en place nécessite l'intervention successive des activités triphosphatase, guanylyltransférase et méthyltransférase. (B) Coiffe 2,2,7-méthylguanosine ( $m^{2,2,7}G$ ) retrouvée sur les UsnRNA et sur certains snoRNA. L'hyperméthylation de la position 2 de la base requiert l'activité méthyltransférase de l'enzyme TGS1 (Tgs1p chez la levure). R : suite de la chaîne nucléotidique orientée dans le sens 5'-3'.**

phosphorylé par l'intermédiaire de son domaine guanylyltransférase (Cho *et al.*, 1997). De la même façon, chez la levure cette fonction implique le facteur Ceg1p (Rodriguez *et al.*, 2000).

Une fois ajoutée, la coiffe  $m^7G$  constitue une marque caractéristique de la transcription par l'ARN polymérase II et confère à l'ARN un certain nombre de propriétés. Elle est notamment reconnue par les protéines CBP20 et CBP80 (Cbc1p et Cbc2p chez la levure) qui forment le complexe de la coiffe (CBC, *Cap Binding Complex*) (Visa *et al.*, 1996). Concernant les ARNm, ce complexe d'origine nucléaire est remplacé à la suite du premier cycle de traduction par les facteurs de traduction eIF4E, eIF4G et eIF4A, qui forment le complexe cytoplasmique de la coiffe, appelé eIF4F (Gross *et al.*, 2003). En premier lieu, la coiffe augmente la stabilité du transcrit en le protégeant de la dégradation par les exonucléases 5'-3' (Hsu & Stevens, 1993; Walther *et al.*, 1998). Dans le cas d'un ARN coiffé, ce type de dégradation n'est en effet rendu possible qu'à la suite d'un processus spécifique, conservé chez les eucaryotes, impliquant l'hydrolyse de la coiffe par les facteurs DCP1-DCP2, suivie de la dégradation de l'ARN dans le sens 5'-3' par l'exonucléase XRN1 (Hsu & Stevens, 1993; Muhrad *et al.*, 1995). Le complexe nucléaire de la coiffe joue des rôles majeurs dans de nombreux processus. Il participe en effet au couplage entre les différentes étapes de production des transcrits de l'ARN polymérase II, en facilitant les événements d'épissage (Izaurrealde *et al.*, 1995; Ohno *et al.*, 1987) et les étapes de maturation de l'extrémité 3' des ARNm (Flaherty *et al.*, 1997). Il est essentiel au processus d'export des UsnRNA vers le cytoplasme (Izaurrealde *et al.*, 1995) (**Fig. 3**) et il joue un rôle important dans le contrôle qualité des ARNm par le système NMD (*Non-sens Mediated Decay*) (Hosoda *et al.*, 2005; Hwang *et al.*, 2010; Lejeune *et al.*, 2002). Concernant la traduction des ARNm, la coiffe joue également un rôle crucial puisque le facteur eIF4G, qui interagit avec la protéine PAB1 (*PolyA-Binding protein 1*), facilite la circularisation de l'ARNm et favorise la ré-initiation du processus de traduction (Kressler *et al.*, 2010; Tarun & Sachs, 1996; von der Haar *et al.*, 2004).

### 2.2.2. La coiffe triméthylée 2,2,7-méthylguanosine ( $m^{2,2,7}G$ )

Parmi les transcrits porteurs d'une coiffe  $m^7G$ , certains acquièrent, au cours de leur biogenèse, des caractéristiques distinctes des ARNm par le biais d'une modification supplémentaire. En effet, les UsnRNA, les snoRNA issus d'unités transcriptionnelles propres, ainsi que l'ARN du

complexe télomérase, sont porteurs d'une coiffe 2,2,7-triméthylguanosine ( $m^{2,2,7}G$ ) qui n'est pas reconnue par le complexe CBC (**Fig. 9B**). La conversion de la coiffe  $m^7G$  en coiffe  $m^{2,2,7}G$  est catalysée par l'hyperméthylase TGS1 (*TrimethylGuanosine Synthase 1*) (Mouaikel *et al.*, 2002). Chez l'homme, il existe deux isoformes principales de la protéine TGS1 : une isoforme longue appelée TGS1-LF (*TGS1 Long Form*) et une isoforme courte, produite par clivage protéolytique du domaine N-terminal et appelée TGS1-SF (*TGS1 Short Form*) (Girard *et al.*, 2008). TGS1-SF est localisée exclusivement dans le noyau où elle s'accumule dans les CB (Girard 2008). TGS1-LF est principalement cytoplasmique, bien qu'elle soit détectable au sein du noyau, principalement dans les CB (Girard *et al.*, 2008; Verheggen *et al.*, 2002). Chez la levure en revanche, il n'existe qu'une seule forme de la protéine, appelée Tgs1p. Celle-ci possède une localisation nucléolaire et apparaît particulièrement enrichie au niveau des NB (Mouaikel *et al.*, 2002; Verheggen *et al.*, 2002).

Chez l'homme, la triméthylation de la coiffe des UsnRNA s'effectue lors de la phase d'assemblage cytoplasmique des UsnRNP par l'isoforme TGS1-LF (Girard *et al.*, 2008; Mouaikel *et al.*, 2002). En revanche, chez *S. cerevisiae*, les UsnRNA, qui ne semblent pas transiter par le cytoplasme, sont triméthylés dans le noyau, vraisemblablement au niveau du nucléole (Mouaikel *et al.*, 2002). Concernant les snoRNA et l'ARN de la télomérase, la triméthylation est nucléaire et se déroule dans les CB (Jady *et al.*, 2004; Verheggen *et al.*, 2002; Zhu *et al.*, 2004). Chez la levure, en absence de CB, ces modifications s'effectuent dans le nucléole et plus particulièrement au niveau des NB dans le cas des snoRNA à boîtes C/D (Franke *et al.*, 2008; Verheggen *et al.*, 2002).

La fonction associée à la coiffe  $m^{2,2,7}G$  a été identifiée pour certaines catégories d'ARN non codants. Concernant les UsnRNP de mammifères, il a été montré que la coiffe triméthylée permet le ré-import des particules dans le noyau, par le biais d'une interaction avec l'importine- $\beta$ , médiée par la Snuportine 1 (Huber *et al.*, 1998; Palacios *et al.*, 1997). Une autre étude, réalisée chez l'homme, a également permis de montrer l'importance de la coiffe  $m^{2,2,7}G$  pour l'adressage nucléolaire des snoRNA à boîtes C/D (Boulon *et al.*, 2004). L'équipe d'E. Bertrand a en effet montré que les facteurs de transport PHAX et CRM1 sont requis pour l'acheminement séquentiel des particules vers les CB, puis vers le nucléole. En effet, PHAX fixe les ARN porteurs d'une coiffe  $m^7G$  par le biais d'une interaction avec le complexe CBC. Dans un second temps, la triméthylation de la coiffe au niveau des CB, induit la dissociation de ce complexe et favorise l'interaction avec CRM1, qui facilite alors l'adressage nucléolaire.

Chez la levure, l'importance de la triméthylation n'a pas encore été clairement établie, d'autant plus que la délétion du gène *TGS1* ne confère qu'un léger défaut de croissance à basse température (Mouaikel *et al.*, 2002). Une étude récente a néanmoins démontré que l'absence de Tgs1p peut affecter la réplication des télomères et contribuer à un vieillissement prématuré des cellules, suggérant ainsi l'importance fonctionnelle de la coiffe  $m^{2,2,7}G$  pour la fonction de la télomérase chez la levure (Franke *et al.*, 2008).

### 2.2.3. Maturation nucléolytique de l'extrémité 5' de certains snoRNA

Chez la levure, les snoRNA transcrits à partir d'unités indépendantes acquièrent une coiffe  $m^7G$  au cours de leur transcription. Certains, comme les snoRNA à boîtes C/D U3, snR4 ou snR13, conservent cette coiffe qui sera ensuite hyperméthylée (Samarsky & Fournier, 1999). En revanche, un grand nombre de snoRNA à boîtes C/D (ex : snR40) et quelques snoRNA à boîtes H/ACA subissent une maturation nucléolytique de leur extrémité 5' (Chanfreau *et al.*, 1998a; Lee *et al.*, 2003). Ce processus implique l'endonucléase Rnt1p qui reconnaît une structure en tige-boucle spécifique située dans l'extension 5' de ces ARN. La définition de l'extrémité 5' mature de ces ARN serait ensuite assurée par l'activité 5'-3' exonucléase des enzymes Xrn1p ou Rat1p.

### 2.3. Maturation des transcrits de l'ARN polymérase II par épissage

Chez les eucaryotes, la plupart des gènes codant les protéines sont transcrits sous la forme de pré-ARNm, dont les séquences traduites peuvent être interrompues par des introns (Wahl *et al.*, 2009). L'élimination de ces introns repose sur le mécanisme d'épissage, qui consiste en une double transestérification, associant à l'élimination des séquences introniques, la ligature des séquences codantes, les exons. Cette réaction, relativement simple sur le plan chimique, fait intervenir une machinerie ribonucléoprotéique très complexe, le spliceosome, composé d'environ 170 protéines et de 5 UsnRNA (Wahl *et al.*, 2009). En ciblant des éléments de séquences spécifiques, le spliceosome définit des introns et catalyse leur élimination. Chez les vertébrés, l'épissage des pré-ARNm est un phénomène courant puisque la majorité des gènes contiennent de nombreux introns. Chez la levure en revanche, seulement 3% des gènes, soit

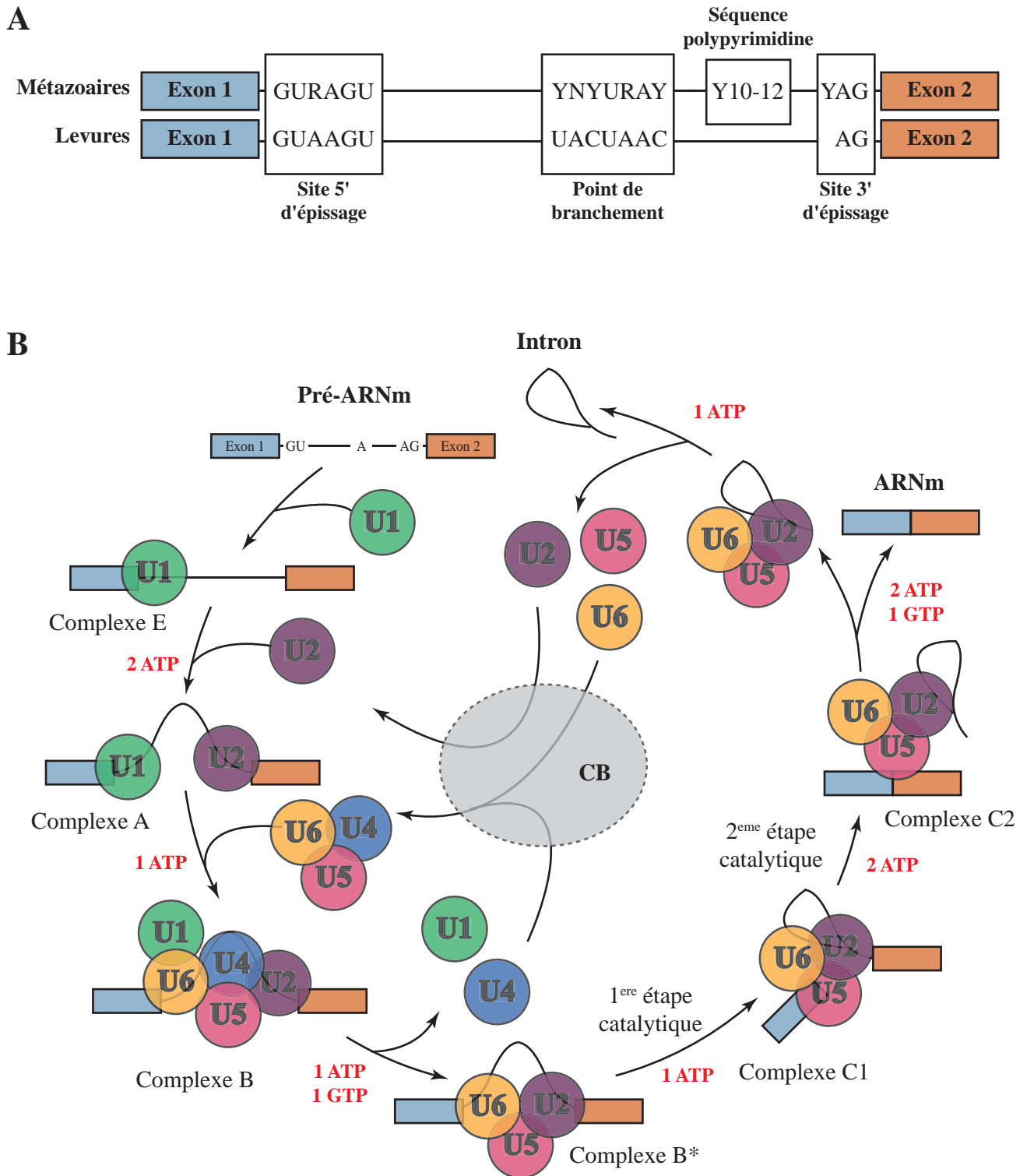


environ 25% des ARNm, possèdent un intron, généralement unique et de taille relativement courte (Hocine *et al.*, 2010). Outre sa fonction première, le mécanisme d'épissage permet une augmentation considérable de la taille et de la diversité du protéome, puisque la présence de sites alternatifs d'épissage permet la production d'une grande variété d'ARNm au contenu exonique variable. Ainsi, chez l'homme 92 à 94% des gènes sont soumis à l'épissage alternatif (Wang *et al.*, 2008). Il est également à l'origine de la production d'autres catégories d'ARN puisque de nombreux ARN non codants, dont certains snoRNA, sont hébergés dans des introns. Enfin, le processus d'épissage n'est pas réservé exclusivement aux ARNm, certains ARN non codants de grande taille, possédant leur propre unité transcriptionnelle, contiennent en effet des introns.

### 2.3.1. Le spliceosome, description du mécanisme d'épissage

Le spliceosome est une machinerie relativement bien conservée chez les eucaryotes. Il est composé des snRNP U1, U2, U4, U5, U6 et d'un grand nombre de protéines additionnelles. En comparaison avec les autres RNP cellulaires, comme le ribosome, dont la composition et la structure sont relativement stables au cours de la traduction, le spliceosome présente une incroyable flexibilité, tant dans sa composition que dans sa conformation (Wahl *et al.*, 2009). Le processus d'épissage se déroule en effet en plusieurs étapes, nécessitant de multiples échanges de facteurs à l'origine d'importants changements structuraux (**Fig. 10**).

L'intervention du spliceosome repose sur la présence de signaux d'épissage permettant de définir les séquences introniques (Wahl *et al.*, 2009) (**Fig. 10A**). Chaque intron contient ainsi à ses extrémités respectives des sites 5' et 3' d'épissage, ainsi qu'un point de branchement, localisé dans sa partie avale et séparé du site 3' par une séquence polypyrimidine chez les métazoaires. Dans un premier temps, la snRNP U1 se fixe sur le site 5' d'épissage tandis que les protéines SF1 et U2AF interagissent respectivement avec le point de branchement et avec le site 3'. Cet ensemble forme le complexe E. Dans un deuxième temps, la snRNP U2 est recrutée au niveau du point de branchement pour former le complexe A ou pré-spliceosome. Puis, la tri-snRNP U4/U6.U5 se fixe à son tour pour former le complexe B ou spliceosome pré-catalytique. Au cours de l'étape d'activation le snRNA U6 s'apparie avec le site 5' et déplace le snRNA U1. Cette étape s'accompagne de réarrangements majeurs qui conduisent



**Figure 10 : Mécanisme d'épissage par le spliceosome majeur (adapté de Wahl *et al.*, 2009).** (A) Eléments de séquence impliqués dans l'épissage chez les métazoaires et chez la levure. Les exons sont représentés par les rectangles de couleur. L'intron, qui porte les sites d'épissage, apparaît en trait fin. R : purine; Y : pyrimidine; N : purine ou pyrimidine. (B) Présentation des principales étapes de la réaction d'épissage. Seules les UsnRNP sont représentées. Après leur désassemblage, les snRNP peuvent transiter par les Cajal Bodies (CB) pour être recyclées. La consommation d'énergie sous forme d'ATP ou de GTP est indiquée en rouge pour chaque étape.

au relargage des snRNP U1 et U4, et à la formation du complexe B\* ou spliceosome activé. Dans cette configuration, le groupement 2'-hydroxyle d'un résidu A, situé dans la boîte de branchement, réalise alors une attaque nucléophile au niveau de la liaison phosphodiester en 5' du premier nucléotide de l'intron, pour former une structure intermédiaire en lasso. On parle alors de complexe C1. Puis, le groupement 3'-hydroxyle libéré à l'extrémité de l'exon 1 réalise à son tour une attaque nucléophile au niveau de la liaison phosphodiester en 3' du dernier nucléotide de l'intron. Le résultat de cette seconde étape est la libération de l'intron sous forme d'une structure en lasso et la formation d'une nouvelle liaison phosphodiester entre les deux exons. On parle alors de complexe C2. Ce complexe est ensuite dissocié et ses composants peuvent être recyclés pour participer à un nouveau cycle d'épissage. A la suite de l'épissage, l'intron suit une voie de dégradation spécifique. Celui-ci est tout d'abord linéarisé par l'enzyme de débranchement DBR1 (Dbr1p chez la levure), avant de subir une dégradation exonucléolytique à partir de ces deux extrémités (Kim *et al.*, 2000; Kramer & Keller, 1985).

L'ensemble de ces étapes requiert l'intervention de nombreux autres facteurs permettant de stabiliser ou de dissocier certaines interactions. Parmi eux, on retrouve notamment un grand nombre d'hélicases ATP-dépendantes qui agissent sur la formation des duplex ARN-ARN et qui catalysent de multiples réarrangements. Enfin, chez certains métazoaires, il existe une seconde machinerie, fonctionnellement analogue, appelée spliceosome mineur. Celle-ci possède ses propres signaux d'épissage et se compose des snRNP U11, U12, U4atac et U6atac, seule la snRNP U5 est commune aux deux machineries (Patel & Steitz, 2003).

En plus des sites d'épissage permettant la définition des introns, il existe de nombreux autres éléments de séquence permettant le recrutement de facteurs régulateurs activateurs ou répresseurs. Ceux-ci agissent directement sur l'épissage en favorisant ou en inhibant le recrutement du spliceosome. De façon générale, les éléments activateurs sont fixés par des protéines SR (*Serine/Arginine rich*), tandis que les éléments inhibiteurs recrutent plutôt des protéines hnRNP (*heterogenous Ribonucleoprotein*) (Singh & Valcarcel, 2005). Le recrutement de ces facteurs fait l'objet de phénomènes de compétition, dont le résultat contribue au choix des sites d'épissage et constitue un élément moteur de l'épissage alternatif. Chez la levure, l'épissage alternatif constitue toutefois un événement relativement peu fréquent. La plupart des protéines SR sont en effet absentes, les séquences régulatrices favorisant l'épissage sont rares et les introns, souvent uniques, sont définis de façon non ambiguë par des sites d'épissage canoniques (Izquierdo & Valcarcel, 2006).

Comme nous l'avons mentionné précédemment, de nombreuses études ont mis en évidence l'existence de mécanismes de couplage entre la transcription et l'épissage, impliquant en particulier le domaine CTD de l'ARN polymérase II. Ces données, ainsi que la co-localisation des facteurs d'épissage avec les sites de transcription, indiquent que l'épissage constitue un mécanisme essentiellement co-transcriptionnel (Moore & Proudfoot, 2009). Quelques travaux ont même montré que le recrutement de certains facteurs d'épissage sur les transcrits naissants peut influencer positivement sur le déroulement de l'étape d'élongation de la transcription (Fong & Zhou, 2001; Kameoka *et al.*, 2004; Lin *et al.*, 2008b). D'autre part, la coiffe joue également un rôle important dans la mise en place de la machinerie d'épissage. Chez la levure le complexe CBC interagit en effet avec certains facteurs d'épissage (Colot *et al.*, 1996), tandis que chez les mammifères, il est impliqué dans la reconnaissance par la snRNP U1 du premier site 5' d'épissage (Lewis *et al.*, 1996).

A l'image des autres processus liés à la synthèse et à la maturation des transcrits, l'épissage influe également de façon majeure sur le devenir des molécules d'ARN. Un certain nombre de facteurs protéiques demeurent en effet fixés sur le transcrit à la suite de l'intervention du spliceosome et constituent des « marques d'épissage ». Ces marques représentent ainsi une garantie de qualité du processus d'épissage et jouent un rôle déterminant pour l'export, la localisation, la traduction et la dégradation des ARNm (Moore & Proudfoot, 2009). Parmi les facteurs impliqués on retrouve notamment certaines protéines SR, le facteur d'export REF/Aly (Yra1p chez la levure), ou encore le complexe EJC (*Exon Junction Complex*), qui est déposé 20 à 24 nucléotides en amont des jonctions exon-exon, qui joue un rôle clé dans les mécanismes de contrôle qualité des ARNm et d'initiation de la traduction chez les mammifères (Le Hir *et al.*, 2003; Moore & Proudfoot, 2009).

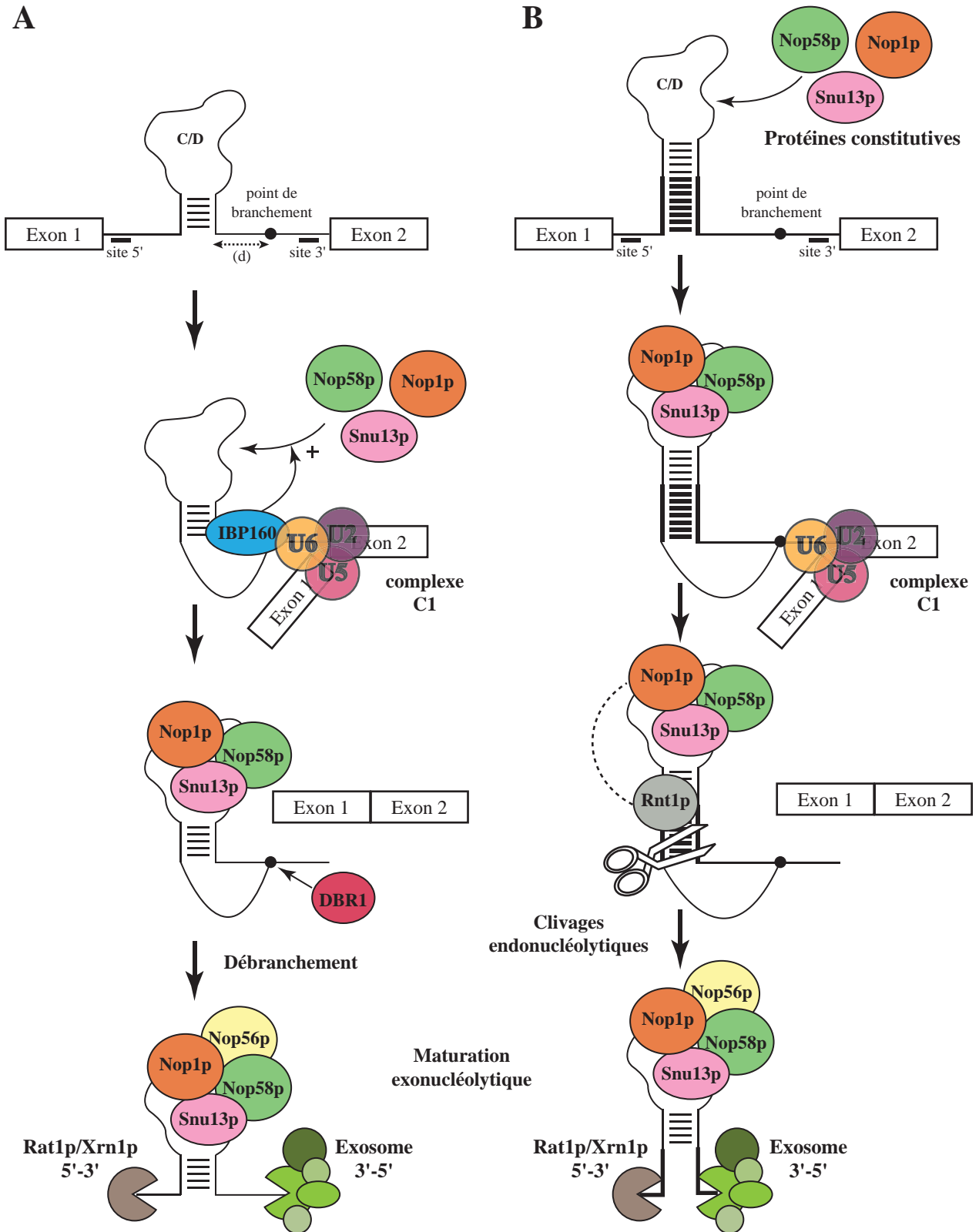
### **2.3.2. Maturation des sous produits d'épissage : le cas des snoRNA**

Parmi les nombreux ARN non codants retrouvés chez les eucaryotes, beaucoup ne sont pas produits à partir de leur propre unité transcriptionnelle, mais sont hébergés dans des introns de gènes. La libération de ces molécules implique des étapes de maturation spécifiques qui peuvent être dépendantes ou indépendantes de la machinerie d'épissage. La production des snoRNA à boîtes C/D introniques peut ainsi s'effectuer selon deux voies parallèles, une voie

majeure, dépendante de l'épissage, et une voie mineure, impliquant des clivages endonucléolytiques (**Fig. 11**).

### 2.3.2.1. Voie majeure

La voie de maturation principale, qui a été décrite chez différents organismes, implique le débranchement de la structure en lasso par l'enzyme DBR1, suivi de la maturation exonucléolytique des extrémités du snoRNA (Cavaille & Bachellerie, 1996; Kiss & Filipowicz, 1993; Ooi *et al.*, 1998; Petfalski *et al.*, 1998) (**Fig. 11A**). Chez la levure la maturation nucléolytique de l'extrémité 5' implique les facteurs Rat1p ou Xrn1p, tandis que la maturation en 3' nécessite l'intervention de l'exosome (Allmang *et al.*, 1999; Kenna *et al.*, 1993; Larimer *et al.*, 1992; Ooi *et al.*, 1998; Petfalski *et al.*, 1998; van Hoof *et al.*, 2000b; Villa *et al.*, 2000). Cette voie de maturation semble largement dépendante de la localisation du snoRNA à l'intérieur de l'intron. Dans les cellules de mammifères, les snoRNA à boîtes C/D sont en effet localisés de façon préférentielle 70 nucléotides en amont du site 3' d'épissage et au moins 41 nucléotides en amont du point de branchement (Hirose & Steitz, 2001). La diminution de cet espacement peut conduire à l'abolition de la production du snoRNA, sans néanmoins affecter l'épissage du pré-ARNm. Chez la levure, la situation semble légèrement différente puisque la diminution de l'espacement entre le snoRNA et le point de branchement, en deçà de la valeur seuil de 44 nucléotides, abolit la production du snoRNA mais affecte également la production de l'ARNm (Vincenti *et al.*, 2007). Des travaux réalisés dans des extraits de cellules HeLa par l'équipe de J.A. Steitz, ont permis d'améliorer la compréhension de ces phénomènes de couplage entre épissage et biogenèse des snoRNA à boîtes C/D. Il a en effet été montré que les protéines 15.5K et Fibrillarine, constitutives des snoRNP à boîtes C/D, sont recrutées après la première réaction de transesterification (complexe C1) sur les intermédiaires en lasso (Hirose *et al.*, 2003). Une seconde étude a ensuite permis de montrer que la protéine IBP160 (*Intron Binding Protein 160*), appartenant au complexe spliceosomal C1, interagit avec une zone de 33 à 40 nucléotides en amont du point de branchement (Hirose *et al.*, 2006). Cette protéine semble constituer le facteur de couplage entre les deux processus puisque sa déplétion abolit la formation de la snoRNP. De plus, lorsque le snoRNA est localisé trop près du point de branchement, le facteur IBP160 interagit directement avec la séquence du snoRNA et



**Figure 11 : Voies de maturation des snoRNA à boîtes C/D hébergés dans des introns (adapté de Hirose *et al.*, 2003).** (A) Voie majeure dépendante de la machinerie d'épissage et du débranchement de la structure en lasso par DBR1. Le positionnement du snoRNA dans l'intron est soumis à une contrainte de distance (d). Chez l'homme, (d) correspond au site de fixation de la protéine IBP160 qui favorise l'assemblage de la RNP lors de l'étape C1 de l'épissage. (B) Voie mineure indépendante de spliceosome. Le snoRNA n'a pas de localisation précise mais il est flanqué par une tige boucle externe (en gras). L'assemblage de la RNP est plus précoce et indépendant du spliceosome. La libération du snoRNA implique des clivages endonucléolytiques, catalysés par Rnt1p chez la levure. Dans ces deux voies, le mode de recrutement de la protéine constitutive Nop56p n'a pas été clairement établi. Les traits fins en pointillés représentent l'interaction entre Nop1p et Rnt1p.

empêche son repliement correct et l'assemblage de la particule. En revanche, aucune interaction n'a été identifiée entre IBP160 et les protéines 15.5K et Fibrillarine, suggérant l'existence possible de contacts avec d'autres protéines associées de façon transitoire ou permanente aux snoRNA à boîtes C/D. Enfin, les auteurs ont proposé que l'activité hélicase, portée par la protéine IBP160, puisse favoriser le repliement correct du snoRNA et/ou la dissociation de protéines fixées non spécifiquement.

### 2.3.2.2. Voie mineure

La seconde voie, indépendante de l'épissage, a également fait l'objet de nombreuses études. Cette dernière, semblable à la maturation des snoRNA polycistroniques (cf. paragraphe 2.4.2.3), repose sur des clivages endonucléolytiques réalisés en amont et en aval du snoRNA, suivis par la maturation exonucléolytique des deux extrémités (Caffarelli *et al.*, 1996; Ooi *et al.*, 1998; Petfalski *et al.*, 1998; Villa *et al.*, 1998) (**Fig. 11B**). Des études réalisées sur les snoRNA à boîtes C/D, ont montré que le déclenchement de cette voie repose notamment sur la formation d'une tige externe flanquant le snoRNA à ses deux extrémités (Darzacq & Kiss, 2000; Villa *et al.*, 2000). Cette structure pourrait en effet compenser la mauvaise localisation des snoRNA, en assurant leur production par la mise en place d'une voie alternative (Hirose *et al.*, 2003). La présence de cette tige externe, relativement stable, pourrait ainsi favoriser le repliement correct du snoRNA afin d'assurer l'assemblage précoce de la particule. L'équipe de J.A. Steitz a en effet montré que le recrutement des protéines 15.5K et Fibrillarine s'effectuait avant l'étape d'épissage (Hirose *et al.*, 2003). Enfin, une étude réalisée à partir du snoRNA U18 de levure a permis d'établir un lien entre ces événements et le processus d'excision de l'ARN (Giorgi *et al.*, 2001). Ce snoRNA présente un profil particulier puisque la voie majeure est responsable de 70% de sa production, tandis que la voie mineure assure la production des 30% restants (Villa *et al.*, 1998). En utilisant une approche de maturation dans un extrait cellulaire déficient pour l'épissage et dépourvu de l'activité de débranchement, l'équipe de I. Bozzoni a mis en évidence l'implication de l'endonucléase Rnt1p pour la production du snoRNA par la voie mineure. De plus, il a été montré que les protéines constitutives Nop1p et Nop58p sont associées à l'ARN avant sa libération du précurseur et que la protéine Nop1p établit une interaction physique avec l'enzyme Rnt1p (**Fig. 11B**). Cette interaction semble d'autant plus importante, que la libération du snoRNA par Rnt1p requiert

la présence de Nop1p. Les auteurs suggèrent ainsi que la protéine Nop1p permettrait le recrutement de l'enzyme Rnt1p, qui pourrait alors reconnaître la tige externe comme substrat non canonique et permettre l'excision du snoRNA dans un mécanisme indépendant de l'épissage (Giorgi *et al.*, 2001).

L'ensemble des données concernant ces deux voies de maturation semble indiquer que l'assemblage des snoRNP contenant un snoRNA à boîtes C/D introniques s'effectue de manière précoce, vraisemblablement de façon co-transcriptionnelle. Des données d'immunoprécipitation de la chromatine ont notamment révélé que les protéines cœurs Nop1p, Nop56p et Nop58p sont détectées au niveau des *loci* hébergeant les snoRNA introniques U18 et U24 (Morlando *et al.*, 2004; Vincenti *et al.*, 2007). Ce constat est par ailleurs corroboré par des données de microscopie, puisque dans les cellules humaines, les CB sont préférentiellement localisés au niveau des gènes contenant des snoRNA introniques (Schul *et al.*, 1999a). De plus, une autre étude a montré que la protéine 15.5K néosynthétisée, qui s'accumule initialement au niveau des speckles, progresse de façon unidirectionnelle et dépendante de la transcription, en direction du nucléole (Leung & Lamond, 2002). Nous reviendrons en détail sur ces aspects de la biogenèse des snoRNP dans le Chapitre 3.

Concernant les snoRNA à boîtes H/ACA introniques, beaucoup moins d'informations sont disponibles. Comme pour la majorité des snoRNA à boîtes C/D, leur libération semble dépendre du débranchement de la structure en lasso, suivi d'une maturation exonucléolytique (Kiss & Filipowicz, 1995; Ooi *et al.*, 1998). En revanche, ce processus présente certaines particularités puisqu'une analyse menée sur 80 snoRNA à boîtes H/ACA humains a révélé l'absence de localisation préférentielle au sein de la séquence intronique (Richard *et al.*, 2006). De plus, à la différence des snoRNA à boîtes C/D, leur biogenèse ne repose pas sur des contacts avec la machinerie d'épissage, mais implique très largement le domaine CTD de l'ARN polymérase II (Richard *et al.*, 2006). En effet, plusieurs études réalisées chez l'homme et chez la levure, ont mis en évidence la nature co-transcriptionnelle de l'assemblage des snoRNP à boîtes H/ACA (Ballarino *et al.*, 2005; Darzacq *et al.*, 2006; Richard *et al.*, 2006; Yang *et al.*, 2005). Nous reviendrons également en détail sur ce point dans le Chapitre 3.



## 2.4. Maturation de l'extrémité 3' des transcrits de l'ARN polymérase II

La terminaison de la transcription des ARNm par l'ARN polymérase II est caractérisée par un mécanisme impliquant le clivage et la polyadénylation de l'extrémité 3' du transcrit. Ce processus implique un grand nombre de facteurs et joue un rôle important dans le devenir de l'ARN. Comme c'est le cas pour la maturation de l'extrémité 5', certaines catégories d'ARN impliquent en revanche des mécanismes de terminaison alternatifs. C'est notamment le cas des ARNm d'histones, des UsnRNA, de certains snoRNA ou encore de l'ARN du complexe de la télomérase.

### 2.4.1. Mécanisme de clivage/polyadénylation

Chez les eucaryotes, la terminaison de la transcription des ARNm implique des séquences spécifiques localisées à l'extrémité 3' des transcrits. Ces séquences sont reconnues par une machinerie spécialisée qui effectue un clivage endonucléolytique, suivie d'une étape de polyadénylation à partir de l'extrémité 3' générée (Millevoi & Vagner, 2009). Chez les mammifères, le signal de clivage/polyadénylation est composé d'un élément AAUAAA, suivi en aval, d'une séquence riche en GU, appelée DES (*GU-rich Downstream Element Sequence*). Le processus de terminaison implique 5 facteurs : (i) le complexe CPSF (*Cleavage and Polyadenylation Specificity Factor*) qui interagit avec la boîte AAUAAA et dont la sous-unité CPSF73 porte l'activité endoribonucléase, (ii) le complexe CstF (*Cleavage stimulation factor*) qui se fixe sur l'élément DES, (iii et iv) les facteurs CFIm et CFII (*Cleavage Factor I/II*), et (v) le facteur PAP (*Poly(A) polymerase*) qui catalyse l'ajout de la queue poly(A) (Millevoi & Vagner, 2009). Cette machinerie forme un ensemble relativement complexe, puisque environ 85 facteurs ont déjà été identifiés chez l'homme (Shi *et al.*, 2009). Chez la levure, le signal de polyadénylation contient de nombreux éléments de séquences riches en pyrimidine. Trois facteurs impliqués dans le processus de terminaison ont été identifiés : les complexes CPF (*Cleavage and Polyadenylation Factor*) et CFIA (*Cleavage Factor IA*), homologues partiels des complexes CPSF et CstF de mammifère, ainsi que le complexe CFIB (*Cleavage Factor IB*), spécifique de la levure (Millevoi & Vagner, 2009). La protéine Ysh1p qui porte l'activité endoribonucléase et la poly(A) polymérase Pap1p appartiennent toutes deux au complexe CPF (Millevoi & Vagner, 2009). Comme nous l'avons vu précédemment,

le domaine CTD de l'ARN polymérase II est impliqué dans la mise en place de ce processus de terminaison. Une étude réalisée chez la levure a notamment montré que la protéine Pcf11p du complexe CFIA interagit avec le domaine CTD phosphorylé au niveau de la Serine 2 et participe au couplage entre transcription et maturation de l'extrémité 3' (Licatalosi *et al.*, 2002). Enfin, des travaux ont montré qu'à la suite du processus de clivage/polyadénylation, la partie d'ARN excédentaire, située en aval du site de coupure, est dégradée par l'activité 5'-3' des exoribonucléases XRN2 et Rat1p, respectivement chez l'homme et chez la levure (Kim *et al.*, 2004; West *et al.*, 2004).

Chez les mammifères comme chez la levure, la queue poly(A) naissante est fixée par un facteur spécifique, porteur de domaines RRM, la protéine PABP (*Poly A Binding Protein*, Pab1p chez la levure). Cette interaction, qui s'établit sur des séries de 11 à 14 résidus, contribue de façon majeure à la polyadénylation puisque la protéine PABP stimule la processivité de l'enzyme (de 80 fois environ) et assure un contrôle de la longueur de la queue poly(A) (Meyer *et al.*, 2002; Minvielle-Sebastia *et al.*, 1997). Celle-ci atteint ainsi une taille moyenne d'environ 200 à 300 nucléotides chez les mammifères. D'autre part, certains ARN possèdent plusieurs sites de terminaison de la transcription et sont sujets à des phénomènes de polyadénylation alternatifs (Tian *et al.*, 2005). Ce processus permet ainsi de générer différentes isoformes d'ARNm, souvent tissu-spécifique. Il peut être couplé à l'épissage alternatif et contribue aux phénomènes de régulation de l'expression génique et à l'accroissement de la diversité du protéome (Beaudoing & Gautheret, 2001; Danckwardt *et al.*, 2008; Tian *et al.*, 2005). Chez l'humain, on estime à plus de 50% le nombre de gènes soumis à ce processus (Tian *et al.*, 2005).

Comme c'est le cas pour la coiffe, la queue poly(A) des ARNm joue de multiples rôles. La présence persistante des protéines PABP augmente la stabilité du transcrit en le protégeant des exonucléases 3'-5'. Sa déadénylation constitue ainsi un mécanisme à part entière visant à réguler la demi-vie de l'ARNm. Il fait intervenir des facteurs spécialisés comme les complexes CCR4-NOT et PAN, ou la protéine PARN spécifique des vertébrés (Wiederhold & Passmore, 2010). La queue poly(A) est également requise pour permettre l'export des transcrits vers le cytoplasme (Hilleren & Parker, 2001) et contribue, grâce aux protéines PABP, au contrôle qualité des ARNm par le système NMD (Amrani *et al.*, 2004; Behm-Ansmant *et al.*, 2007). Enfin, comme nous l'avons déjà mentionné, les protéines PABP sont également impliquées dans la circularisation de l'ARNm lors de la traduction, par le biais de

leur interaction avec le facteur eIF4G (Kessler & Sachs, 1998; Tarun & Sachs, 1996; von der Haar *et al.*, 2004).

## 2.4.2. Terminaison indépendante de la polyadénylation

### 2.4.2.1. Maturation de l'extrémité 3' des ARNm d'histones

Bien que la majorité des ARNm possèdent une queue poly(A), ceci ne constitue pas une règle absolue. L'étude des mécanismes d'expression des gènes codant les histones chez les animaux a en effet révélé l'existence d'un mécanisme de terminaison de la transcription alternatif, indépendant de la polyadénylation et impliquant la snRNP U7 (Marzluff *et al.*, 2008). Au cours de ce processus, le snRNA U7 s'hybride au niveau d'un élément HDE (*Histone Downstream Element*), situé dans la région 3' des transcrits, afin de spécifier un site de coupure. Des études ont montré que l'enzyme responsable du clivage endonucléolytique est la sous unité CPSF73 du complexe CPSF, qui intervient donc dans les deux mécanismes de terminaison (Dominski *et al.*, 2005a; Mandel *et al.*, 2006). En dépit de l'absence d'une queue poly(A), les ARNm d'histones conservent au niveau de leur extrémité 3'-UTR, une structure en tige-boucle reconnue par la protéine SLBP (*Stem-Loop Binding Protein*), également appelée HBP (*Hairpin Binding Protein*) (Martin *et al.*, 1997; Wang *et al.*, 1996). La structure ainsi formée possède une fonction analogue à la queue poly(A), puisqu'elle contribue à la stabilité de l'ARNm (Pandey & Marzluff, 1987) et s'avère impliquée dans les processus de transport (Eckner *et al.*, 1991) et de traduction (Ling *et al.*, 2002).

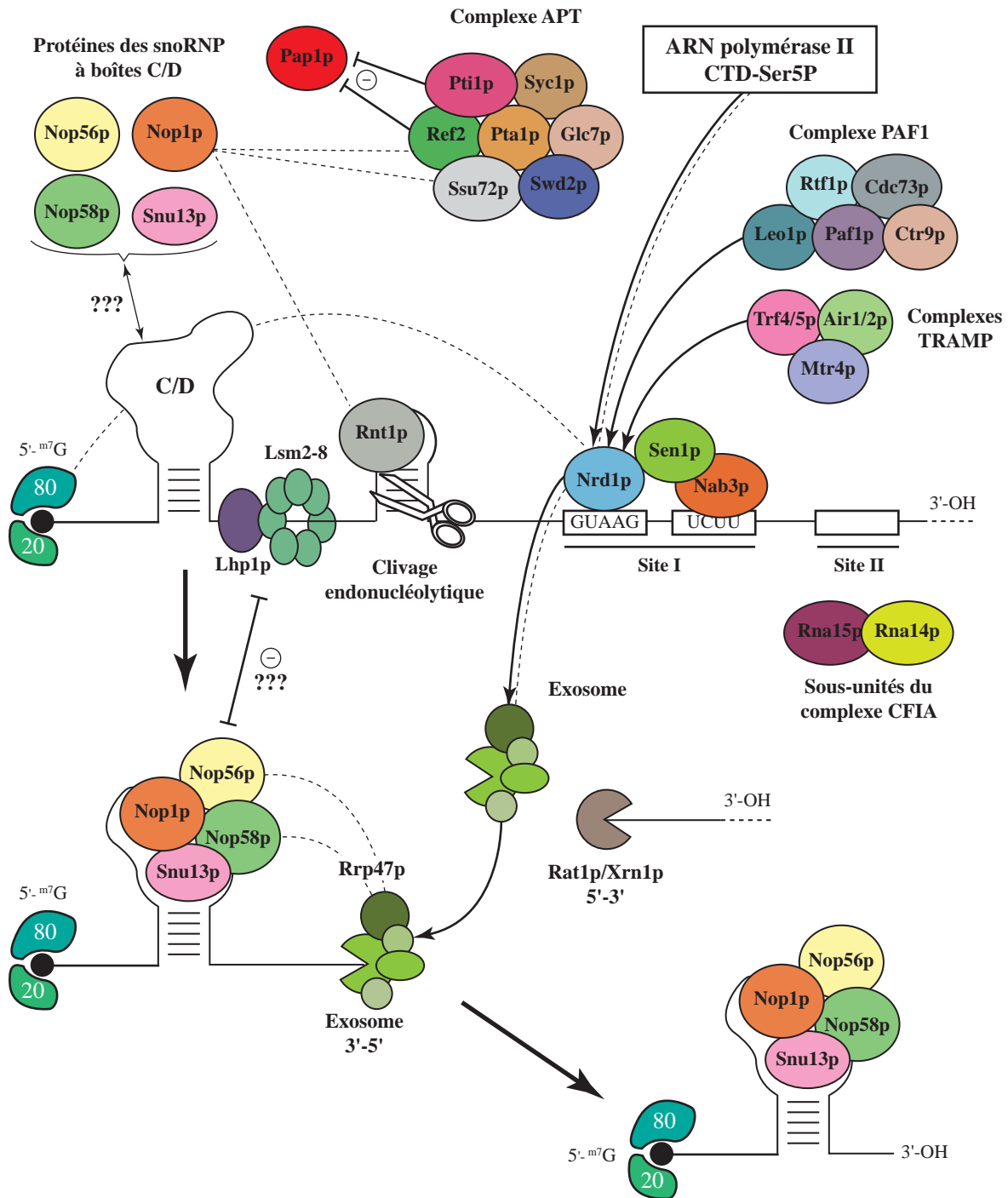
### 2.4.2.2. Maturation de l'extrémité 3' des UsnRNA

Parmi les transcrits de l'ARN polymérase II, les UsnRNA sont également dépourvus de queue poly(A). Il a été montré que la terminaison de leur transcription implique, chez les métazoaires, un mécanisme spécifique, dépendant du domaine CTD de l'ARN polymérase (Baillat *et al.*, 2005; Egloff *et al.*, 2007). En effet, au niveau des *loci* codant les UsnRNA, la phosphorylation de la Serine 7 du domaine CTD permet le recrutement d'un complexe de douze protéines, appelé *Integrator* (Baillat *et al.*, 2005). Ce complexe contient des sous-unités

similaires aux composants du complexe CPSF. La protéine Integrator 11 en particulier, présente environ 40% d'identité avec le domaine catalytique de l'endoribonucléase CPSF73 (Baillat *et al.*, 2005; Dominski *et al.*, 2005b). Le site de clivage permettant de définir l'extrémité 3' des UsnRNA est localisé entre une structure en tige-boucle située en amont et une boîte 3', située 9 à 19 nucléotides en aval (Ach & Weiner, 1987; Hernandez, 1985). Bien que ces éléments soient importants pour la formation de l'extrémité 3' des transcrits, aucun domaine de fixation à l'ARN, connu à ce jour, n'a été identifié parmi les protéines du complexe *Integrator*. Ceci suggère donc que d'autres facteurs pourraient intervenir dans ce processus ou qu'il existe au sein du complexe un domaine d'interaction avec l'ARN encore non décrit. Chez la levure *S. cerevisiae*, il n'existe vraisemblablement pas d'équivalent du complexe *Integrator*, suggérant que ce mécanisme de terminaison n'est pas conservé chez tous les eucaryotes. D'autres mécanismes ont néanmoins été décrits pour les UsnRNA de levure qui contiennent dans leur région 3', une structure en tige-boucle reconnue par l'endoribonucléase III Rnt1p (Abou Elela & Ares, 1998; Chanfreau *et al.*, 1997; Seipelt *et al.*, 1999). Après la coupure, l'extrémité 3' générée constitue un site d'entrée pour l'exosome qui poursuit alors la maturation jusqu'à l'extrémité 3' mature (Allmang *et al.*, 1999; van Hoof *et al.*, 2000b). D'autres facteurs semblent également impliqués dans ces processus. Concernant le snRNA U4, une étude a notamment mis en évidence l'implication des facteurs Nrd1p, Sen1p et Nab3p (Steinmetz *et al.*, 2001). Enfin, l'équipe de I. Bozzoni a également révélé que le facteur de clivage/polyadénylation CFIA est impliqué dans la maturation de l'extrémité 3' des snRNA U2 et U5 chez la levure (Morlando *et al.*, 2002). Les mécanismes de maturation de l'extrémité 3' des UsnRNA de levure présentent beaucoup de similarités avec celles des snoRNA. Ces mécanismes seront décrits plus en détail dans la section suivante.

#### 2.4.2.3. Maturation de l'extrémité 3' des snoRNA

Les snoRNA possédant leur propre unité transcriptionnelle sont relativement rares chez l'homme, tandis qu'ils constituent le cas de figure le plus fréquent chez *S. cerevisiae*. Pour cette raison, la majorité des investigations menées sur les mécanismes de terminaison de la transcription des snoRNA ont été réalisées chez la levure (**Fig. 12**). Leurs précurseurs sont la cible d'un clivage endonucléolytique au niveau de leur extension 3', suivie d'une dégradation exonucléolytique par l'exosome jusqu'à l'extrémité 3' mature (Allmang *et al.*, 1999;



**Figure 12 : Eléments impliqués dans la formation de l'extrémité 3' des snoRNA à boîtes C/D issus d'unités transcriptionnelles indépendantes chez la levure.** La terminaison de la transcription indépendante de la polyadénylation implique le trio de protéines Nrd1p-Nab3p-Sen1p, le complexe PAF1, le sous-complexe APT du facteur CPF et les sous-unités Rna14p et Rna15p du facteur CFIA. Le complexe APT joue un rôle clé dans ce processus en assurant l'exclusion de la polyA polymérase Pap1p. Le recrutement de Nrd1p est favorisé par la phosphorylation de la Serine 5 du domaine CTD (CTD Ser5-P), par le complexe PAF1 et par l'enzyme Trf4p du complexe TRAMP. La maturation de l'extrémité 3' implique un clivage endonucléolytique, qui pourrait être généré par l'enzyme Rnt1p, suivi d'une dégradation par l'exosome, dont le recrutement est favorisé par Nrd1p. Durant ces étapes, l'extrémité 3' pourrait être stabilisée par le complexe Lsm2-8 et par le facteur Lhp1p. L'équipe de D. Tollervey a par ailleurs montré que, dans le cas du snoRNA U3, leur présence sur l'ARN semble mutuellement exclusive avec celle des protéines constitutives Nop1p, Nop56p et Nop58p (Kufel *et al.*, 2000 & 2003). Les traits fins en pointillés représentent les interactions décrites entre les différentes protéines.

Chanfreau *et al.*, 1998a; Chanfreau *et al.*, 1998b; Perumal & Reddy, 2002; van Hoof *et al.*, 2000b). Comme c'est le cas pour les UsnRNA, la maturation de beaucoup de snoRNA nécessite l'intervention de l'endonucléase Rnt1p afin de générer les sites d'entrée pour l'exosome (Chanfreau *et al.*, 1998a; Chanfreau *et al.*, 1998b). Comme les enzymes DROSHA ou DICER, Rnt1p est une protéine orthologue à la ribonucléase III bactérienne, qui reconnaît les double-brins d'ARN et clive le duplex sur chacun des deux brins (Chanfreau *et al.*, 2000; Wu *et al.*, 2001). Rnt1p présente toutefois une spécificité pour les hélices surmontées d'une boucle à 4 résidus de type AGNN, située 13 à 16 paires de bases des sites de coupure (Chanfreau *et al.*, 2000; Wu *et al.*, 2001). D'autre part, une étude récente a montré que la sous-unité Rrp47p (alias Lrp1p) de l'exosome joue un rôle spécifique dans la maturation de l'extrémité 3' des snoRNA à boîtes C/D (Costello *et al.*, 2011). Rrp47p interagit en effet avec la sous-unité Rrp46p de l'exosome, ainsi qu'avec les protéines Nop56p et Nop58p, constitutives de la particule C/D, et pourrait ainsi contribuer à la définition de l'extrémité 3' mature des transcrits. De plus, des études réalisées sur le snoRNA U3 de levure ont montré que ses intermédiaires de maturation sont fixés de façon transitoire et coopérative par le complexe heptamérique Lsm 2-8 et la protéine Lhp1p (homologue de la protéine humaine LA) (Kufel *et al.*, 2000; Kufel *et al.*, 2003). Ces facteurs, également impliqués dans la maturation des transcrits de l'ARN polymérase III, semblent jouer un rôle de chaperon permettant de stabiliser l'extrémité 3' durant les étapes de clivage (Kucera *et al.*, 2011; Kufel *et al.*, 2000; Kufel *et al.*, 2003; Kufel *et al.*, 2002). Le complexe Lsm 2-8, qui s'assemble sous forme d'une structure en anneau, et la protéine Lhp1p interagissent directement avec l'ARN au niveau de régions riches en U (Achsel *et al.*, 1999; Ingelfinger *et al.*, 2002; Stefano, 1984). D'autres études ont permis d'étendre ces observations à d'autres ARN. L'utilisation de puces à ADN a notamment permis de montrer que la protéine Lhp1p est associée à un grand nombre de snoRNA chez la levure *S. cerevisiae* (Inada & Guthrie, 2004), tandis qu'un complexe hexamérique Lsm 2-7 a été observé en association avec le snoRNA à boîtes H/ACA snR5 dans le nucléole de levure (Fernandez *et al.*, 2004).

Chez quelques organismes comme *S. cerevisiae*, certains snoRNA sont générés à partir de transcrits polycistroniques. Dans de tels cas, la libération des snoRNA de taille mature implique des étapes de maturation nucléolytique supplémentaires, au niveau de régions internes séparant les différents ARN. Des études menées chez la levure ont permis de montrer que ces phénomènes de maturation impliquent également l'enzyme Rnt1p, qui permet de générer des sites d'entrée pour l'activité exonucléase 5'-3' des facteurs Rat1p et Xrn1p, et

pour l'activité exonucléase 3'-5' de l'exosome (Allmang *et al.*, 1999; Chanfreau *et al.*, 1998b; Petfalski *et al.*, 1998; van Hoof *et al.*, 2000b).

D'autres éléments intervenant en *cis* et en *trans* dans la terminaison de la transcription des snoRNA ont également été identifiés. C'est notamment le cas du trio de protéines Nrd1p, Nab3p et Sen1p, qui n'appartiennent pas à la machinerie de clivage/polyadénylation, mais qui interviennent néanmoins dans la terminaison de la transcription d'un certain nombre d'ARN, dont les snoRNA ou le snRNA U4 (Steinmetz *et al.*, 2001). Leur mutation affecte en effet la biogenèse de nombreux transcrits en provoquant l'accumulation de précurseurs non maturés en 3' (Steinmetz *et al.*, 2001). Ces trois facteurs sont essentiels à la viabilité cellulaire. Nrd1p et Nab3p appartiennent à la famille des protéines hnRNP (Conrad *et al.*, 2000; Steinmetz & Brow, 1996), tandis que Sen1p est une hélicase, initialement connue pour son association avec de nombreux snoRNA à boîtes C/D et H/ACA *in vivo* (Rasmussen & Culbertson, 1998; Ursic *et al.*, 1997). Ces protéines pourraient être recrutées directement sur l'extrémité 3' des précurseurs d'ARN, par le biais des éléments de type GUAAG et UCUU, reconnus respectivement par les protéines Nrd1p et Nab3p, dans un mécanisme coopératif (Carroll *et al.*, 2004; Hobor *et al.*, 2011; Lunde *et al.*, 2011; Steinmetz & Brow, 1998). Ces éléments constituent un terminateur de transcription bipartite, baptisé Site I, qui peut être suivi d'un second site de terminaison (*i.e.* Site II), similaire aux sites de clivage/polyadénylation des ARNm (Fatica *et al.*, 2000b; Steinmetz & Brow, 2003; Steinmetz *et al.*, 2006). Des travaux ont également montré que l'intervention du complexe Nrd1p-Nab3p-Sen1p est dépendante de la machinerie de transcription. La protéine Nrd1p peut en effet interagir avec le domaine CTD de l'ARN polymérase II, par le biais d'un domaine CID (Steinmetz & Brow, 1996; Steinmetz & Brow, 1998) (**Fig. 8**). Des données récentes indiquent en effet que la phosphorylation de la Serine 2 inhibe le recrutement de Nrd1p, tandis que la phosphorylation de la Serine 5 est déterminante pour son interaction avec la polymérase (Gudipati *et al.*, 2008; Kubicek *et al.*, 2011; Vasiljeva *et al.*, 2008). Enfin, une autre étude a montré que Nrd1p interagit à la fois avec la protéine Cbp80p du complexe de la coiffe et avec l'exosome, dont elle facilite le recrutement sur l'ARN (Vasiljeva & Buratowski, 2006). Ce dernier résultat souligne donc l'importance fonctionnelle de Nrd1p pour la maturation de l'extrémité 3' des transcrits et suggère l'existence d'un couplage entre le processus de terminaison de la transcription et la mise en place de la coiffe.

Chez la levure, un autre complexe, appelé PAF1, et composé des protéines Paf1p, Ctr9p, Cdc73p, Rtf1p et Leo1p, est également impliqué dans les mécanismes de terminaison de la transcription des snoRNA (Sheldon *et al.*, 2005). Ce complexe, conservé chez l'homme, intervient de façon générale au cours de l'élongation de la transcription par l'ARN polymérase II et induit la modification de certaines protéines histones (Krogan *et al.*, 2003a). Concernant la biogenèse des snoRNA, il a été montré que la mutation de ces composants conduit à l'accumulation de précurseurs non maturés en 3' et réduit l'association de Nrd1p avec les ARN (Sheldon *et al.*, 2005).

D'autres travaux ont également permis de montrer que certaines sous-unités de la machinerie de clivage/polyadénylation jouent un rôle spécifique dans la terminaison de la transcription des snoRNA. La mutation des sous-unités Rna14p et Rna15p du facteur CFIA conduit à l'inhibition du découplage clivage/polyadénylation pour les snoRNA à boîtes C/D et H/ACA, ainsi que pour les snRNA U2 et U5 (Fatica *et al.*, 2000b; Morlando *et al.*, 2002). De la même façon, le sous-complexe APT (*Associated with Pta1p*) du facteur CPF semble requis pour assurer ce découplage (Dheur *et al.*, 2003; Nedeia *et al.*, 2003). Ce complexe est composé des facteurs Pta1p, Pti1p, Swd2p, Glc7p, Ssu72p, Ref2p et Syc1p (Nedeia *et al.*, 2003). Parmi eux, les protéines Pti1p, Ref2p et Ssu72p semblent particulièrement importantes puisque leur mutation entraîne l'accumulation de snoRNA non maturés en 3' (Dheur *et al.*, 2003; Nedeia *et al.*, 2003) et que la présence de Pti1p et Ref2p dans la machinerie de terminaison de la transcription semble exclure l'enzyme Pap1p (Dheur *et al.*, 2003). Concernant la biogenèse des snoRNP à boîtes C/D, l'équipe d'I. Bozzoni a montré que la protéine Nop1p s'associe de façon co-transcriptionnelle au snoRNA monocistronique snR13 et pourrait faciliter le recrutement du complexe APT par le biais d'une interaction physique avec les facteurs Ref2p et Ssu72p (Morlando *et al.*, 2004). Ce résultat suggère donc l'existence probable d'un lien entre la terminaison de la transcription des snoRNA à boîtes C/D et l'assemblage des RNP.

L'ensemble de ces données met donc en lumière la complexité des phénomènes impliqués dans la formation de l'extrémité 3' des snoRNA. Ces processus intègrent à la fois des aspects de terminaison de la transcription et de maturation nucléolytique. Comme nous l'avons vu, la terminaison de la transcription des snoRNA implique la coopération entre des facteurs généraux du système de clivage/polyadénylation et des facteurs plus spécifiques, comme le trio Nrd1p-Nab3p-Sen1p. Sur la base d'interactions avec l'ARN polymérase II, avec le complexe de la coiffe et avec des éléments de séquences en *cis*, l'intervention de cette



machinerie contribue à la mise en place d'un mécanisme indépendant de la polyadénylation. D'autre part, différents facteurs comme la protéine Rnt1p, l'exosome ou les protéines Lsm et Lhp1p, participent à la définition de l'extrémité 3' mature des ARN.

Enfin, plusieurs études ont également rapporté que des défauts de cette machinerie de maturation (e.g. la mutation  $\Delta RRP6$ ) conduit à l'apposition transitoire d'une queue poly(A) (Egecioglu *et al.*, 2006; Grzechnik & Kufel, 2008; Houseley & Tollervey, 2006; LaCava *et al.*, 2005; van Hoof *et al.*, 2000b; Vanacova *et al.*, 2005; Wyers *et al.*, 2005). Cette polyadénylation, qui peut s'effectuer au niveau des sites I ou II de terminaison, fait intervenir les poly(A)-polymérase Trf4p et Trf5p, qui associées à l'hélicase Mtr4p et aux protéines Air1p et Air2p, forment les complexes TRAMP. Ces derniers sont impliqués dans la surveillance et la dégradation des ARN aberrants par ajout d'une queue poly(A), stimulant l'intervention de l'exosome (LaCava *et al.*, 2005; Vanacova *et al.*, 2005; Wyers *et al.*, 2005). Ces complexes pourraient même faire partie intégrante de la machinerie de terminaison/maturation 3' des snoRNA, puisque la protéine Trf4p s'est avérée requise pour le recrutement du facteur Nrd1p sur les pré-snoRNA (Grzechnik & Kufel, 2008). Par ailleurs, d'autres données indiquent que la poly(A)-polymérase classique, Pap1p, pourrait également être impliquée dans la polyadénylation transitoire des snoRNA (Grzechnik & Kufel, 2008). Ces nouveaux éléments suggèrent donc que la maturation de l'extrémité 3' des pré-snoRNA et leur dégradation font l'objet d'un équilibre fin, qui pourrait être à l'origine d'un système de contrôle visant à réguler le taux et/ou la qualité des molécules produites (Egecioglu *et al.*, 2006; Grzechnik & Kufel, 2008; Houseley *et al.*, 2006; Houseley & Tollervey, 2006; LaCava *et al.*, 2005; Vanacova *et al.*, 2005; Wyers *et al.*, 2005).

#### 2.4.2.4. Maturation de l'extrémité 3' de l'ARN télomérique

Concernant l'ARN du complexe de la télomérase (TERC), peu d'informations sont disponibles quant à la maturation de son extrémité 3'. Chez les levures *Saccharomyces cerevisiae* et *Schizosaccharomyces pombe*, des formes polyadénylées de l'ARN télomérique ont été détectées, suggérant l'ajout d'une queue poly(A) transitoire, qui serait ensuite rapidement éliminée (Box *et al.*, 2008; Chapon *et al.*, 1997). Chez *S. pombe*, un mécanisme original de maturation impliquant le spliceosome a par ailleurs été proposé (Box *et al.*, 2008). L'ARN télomérique de *S. pombe* contient en effet des signaux d'épissage dans sa région 3'.

Le spliceosome catalyse alors une réaction d'épissage incomplète permettant la libération de l'ARN mature (équivalent à l'exon 1) sans étape de religature des séquences exoniques.

### 3. Production des ARNr et biogenèse des ribosomes

Chez toutes les espèces, les ribosomes sont constitués de deux sous-unités : une petite sous-unité qui assure le décodage des ARNm, et une grande sous-unité, impliquée dans la formation de la liaison peptidique. Dans toutes les cellules, leur production constitue l'un des processus les plus complexes et les plus coûteux en énergie. Chez les eucaryotes, il implique en effet l'activité des 3 ARN polymérases et mobilise environ 70% de l'activité transcriptionnelle cellulaire (Warner, 1999). L'ARN polymérase II est en effet à l'origine de l'expression des protéines ribosomiques et de la quasi-totalité des facteurs impliqués en *trans* dans ce processus. L'ARN polymérase III assure la production du précurseur de l'ARNr 5S, tandis que l'ARN polymérase I assure la synthèse du précurseur des ARNr 5.8S, 18S et 28S (25S chez la levure). Au cours de leur maturation, les pré-ARNr s'assemblent avec des protéines ribosomiques et avec de nombreux facteurs pour former des particules pré-ribosomiques. La formation de ces complexes au sein du nucléole, est couplée à l'apposition de modifications chimiques spécifiques et au déroulement d'une série de clivages permettant la formation des ARNr de taille mature. L'ensemble de ces événements implique ainsi une quantité de facteurs excédant largement le nombre final de composants du ribosome. Parmi ces facteurs, on retrouve notamment plus d'une centaine de snoRNP qui jouent des rôles clés, à la fois dans les étapes de modifications de nucléotides, de clivages endonucléolytiques et de formation des particules pré-ribosomiques.

#### 3.1. Organisation des gènes codant les ARNr

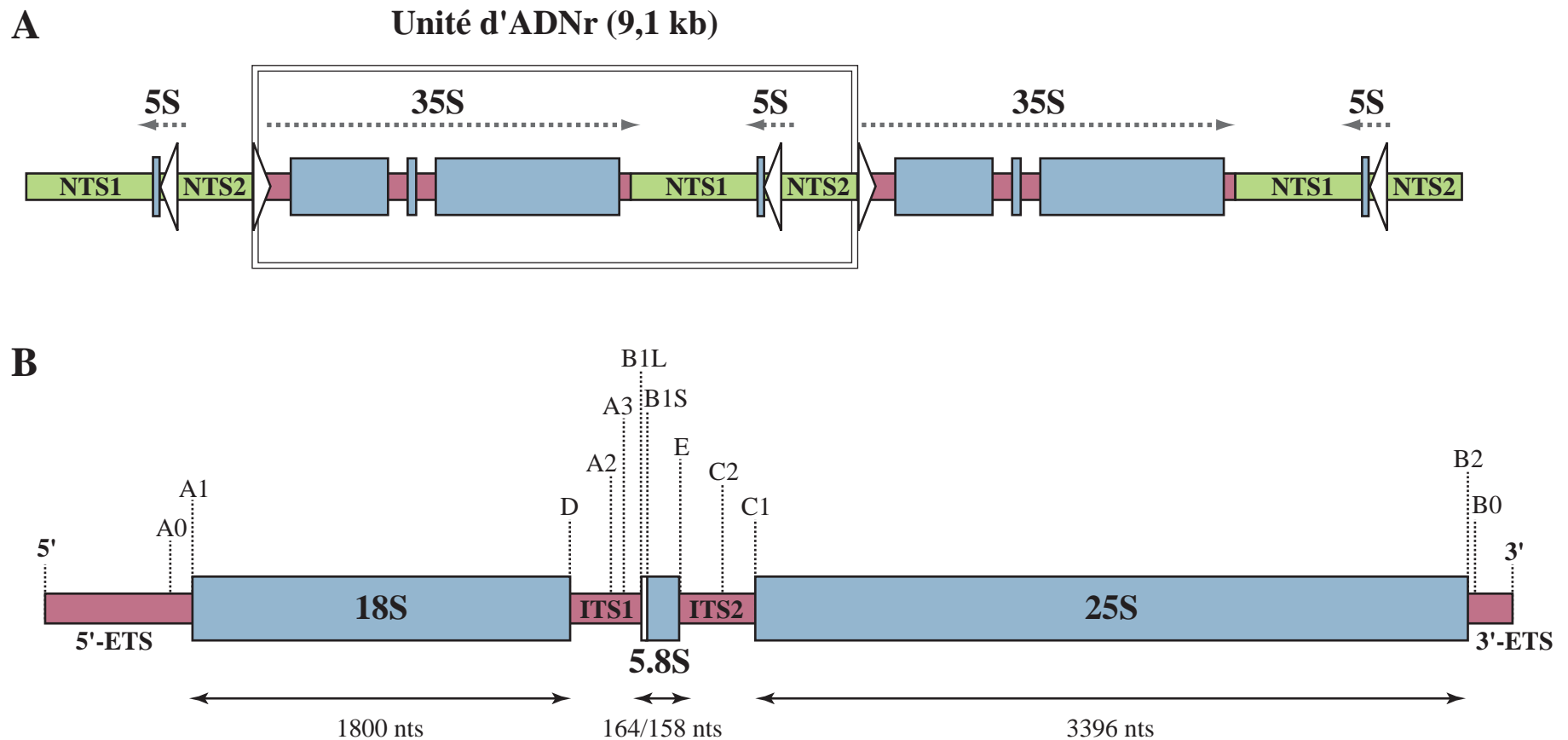
Les eucaryotes comportent 4 ARNr distincts : 3 sont présents au sein de la grande sous-unité, les ARNr 5S, 5.8S et 28S (25S chez la levure), tandis que la petite sous-unité contient l'ARNr 18S. Comme nous l'avons déjà mentionné, ces ARN sont codés au sein de deux unités transcriptionnelles distinctes. Les ARNr 18S, 5.8S et 25/28S sont en effet issus d'une même unité transcriptionnelle, transcrite par l'ARN polymérase I au niveau du nucléole. Le produit

de cette transcription est un grand transcrit unique, appelé pré-ARNr 35S chez la levure et pré-ARNr 45S chez l'homme. Au sein de ce précurseur, les séquences des ARNr sont bordées de régions « espaceurs » : les séquences 5'-ETS et 3'-ETS (*External Transcribed Spacer*), situées respectivement aux deux extrémités du transcrit, et les séquences ITS1 et ITS2 (*Internal Transcribed Spacer*), séparant les séquences 18S, 5.8S et 25/28S. L'ARNr 5S est quant à lui codé au sein d'une unité indépendante, transcrite par l'ARN polymérase III.

Chez *S. cerevisiae* ces deux unités transcriptionnelles sont physiquement rapprochées, puisqu'elles sont regroupées deux à deux au sein de segments de 9,1 kb (**Fig. 13**). Ces segments d'ADNr sont localisés sur le chromosome XII sous la forme de répétitions en tandem comptant 150 à 200 copies (Petes, 1979). Au sein de chaque segment, les deux unités sont transcrites en sens opposés et sont séparées par des régions non transcrites appelées NTS (*Non-Transcribed Spacer*). Dans cette configuration, la transcription de l'ARNr 5S par l'ARN polymérase III s'effectue directement au sein du nucléole (Haeusler & Engelke, 2006).

Cette organisation n'est cependant pas conservée chez tous les eucaryotes. Chez les mammifères, les deux unités transcriptionnelles ne sont pas juxtaposées et sont retrouvées, également sous forme de répétitions en tandem, dans des régions distinctes du génome. Chez l'homme, les unités transcriptionnelles codant les ARNr 18S, 5.8S et 28S sont ainsi réparties sur 5 paires de chromosomes distincts (les chromosomes 13, 14, 15, 21 et 22), tandis que les copies du gène codant l'ARNr 5S sont retrouvées sur le chromosome 1 (Sylvester *et al.*, 1986). Chez les mammifères l'ARNr 5S est donc transcrit par l'ARN polymérase III dans le nucléoplasme, puis importé dans le nucléole pour être incorporé au sein de la future sous-unité ribosomique 60S.

Enfin, des travaux récents ont mis à jour des phénomènes de régulation de la transcription du pré-ARNr 35S par le complexe TOR1 (Mayer & Grummt, 2006). La fixation de ce complexe au niveau du promoteur permet d'activer la transcription par l'ARN polymérase I et pourrait ainsi contribuer à adapter le taux de production des ribosomes en fonction des conditions environnementales.



**Figure 13 : Structure des unités transcriptionnelles codant les ARNr chez la levure *Saccharomyces cerevisiae*.** (A) Organisation génomique des unités d'ADNr retrouvées sous la forme de 150 à 200 répétitions en tandem au niveau du chromosome XII. Chaque unité de répétition de 9,1 kb contient deux unités transcriptionnelles en sens opposé. Les ARNr 18S, 5.8S et 25S sont codés au sein d'un même opéron transcrit sous la forme d'un pré-ARNr 35S par l'ARN polymérase I. L'ARNr 5S est codé au sein d'une petite unité transcrit par l'ARN polymérase III. Ces deux unités sont séparées par des séquences non transcrites NTS1 et NTS2. Les promoteurs transcriptionnels sont symbolisés par des triangles blancs. (B) Représentation schématique du pré-ARNr 35S. Les séquences correspondants aux ARNr sont séparées par les séquences espaceurs internes ITS1 et ITS2, le transcrit est borné par les séquences espaceurs externes 5'-ETS et 3'-ETS. Les sites de maturations nucléolytiques sont indiqués par des traits en pointillés. NTS1/2 : Non Transcribed Spacer 1/2 ; 5'/3'-ETS : 5'/3'-External Transcribed Spacer ; ITS1/2 : Internal Transcribed Spacer 1/2 ; nts : nucléotides.

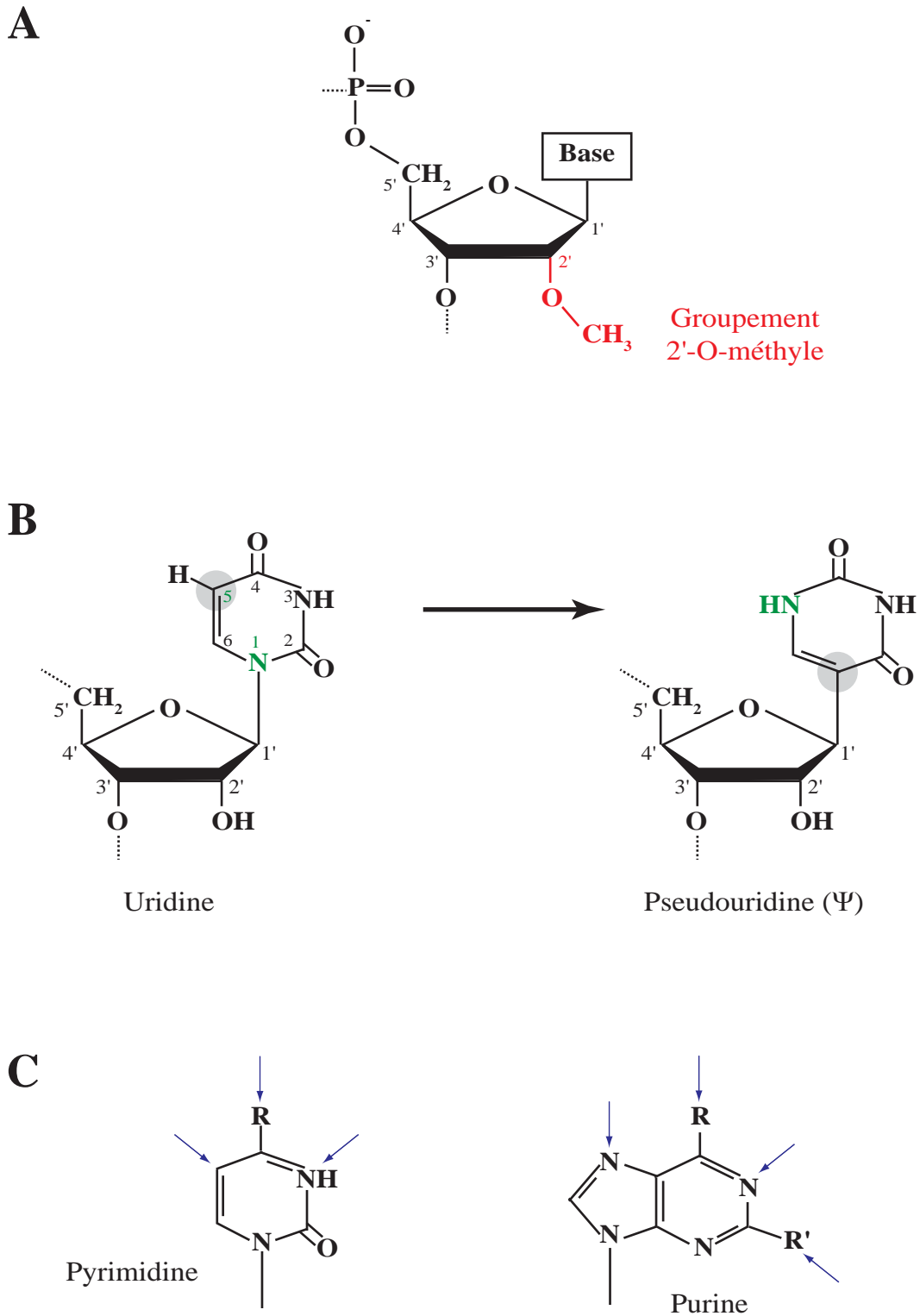
### 3.2. Modifications chimiques des nucléotides des ARNr

Chez tous les organismes, les ARNr portent de nombreux nucléotides modifiés (Piekna-Przybylska *et al.*, 2007). Le nombre et la nature de ces modifications varient cependant en fonction des espèces. Chez *E. coli*, on dénombre par exemple un total de 35 résidus modifiés, parmi lesquels 60% sont méthylés au niveau des bases, 9% seulement sont méthylés au niveau du ribose (2'-O-méthylation) et 31% concernent des cas de conversion d'uridines en pseudouridines (pseudouridylation). Chez les eucaryotes, leur fréquence augmente largement, puisque plus de 100 et 200 modifications ont respectivement été dénombrées chez la levure et chez les vertébrés. La nature des modifications rencontrées chez les eucaryotes est également différente, puisque 95% d'entre elles concernent des 2'-O-méthylations et des pseudouridylation. De plus, alors que les bactéries font exclusivement appel à des enzymes protéiques, de type méthyl-transférase et pseudouridyl-synthase, la plupart des 2'-O-méthylations et des pseudouridylation réalisées chez les eucaryotes, impliquent des complexes ribonucléoprotéiques, les snoRNP. Comme nous l'avons déjà mentionné, deux grandes familles de particules ont été déterminées sur des bases fonctionnelles et structurales : les snoRNP à boîtes C/D, qui catalysent les réactions de 2'-O-méthylation, et les snoRNP à boîtes H/ACA, qui catalysent les réactions de pseudouridylation. Nous reviendrons en détail sur la composition et le fonctionnement de ces RNP dans le chapitre suivant.

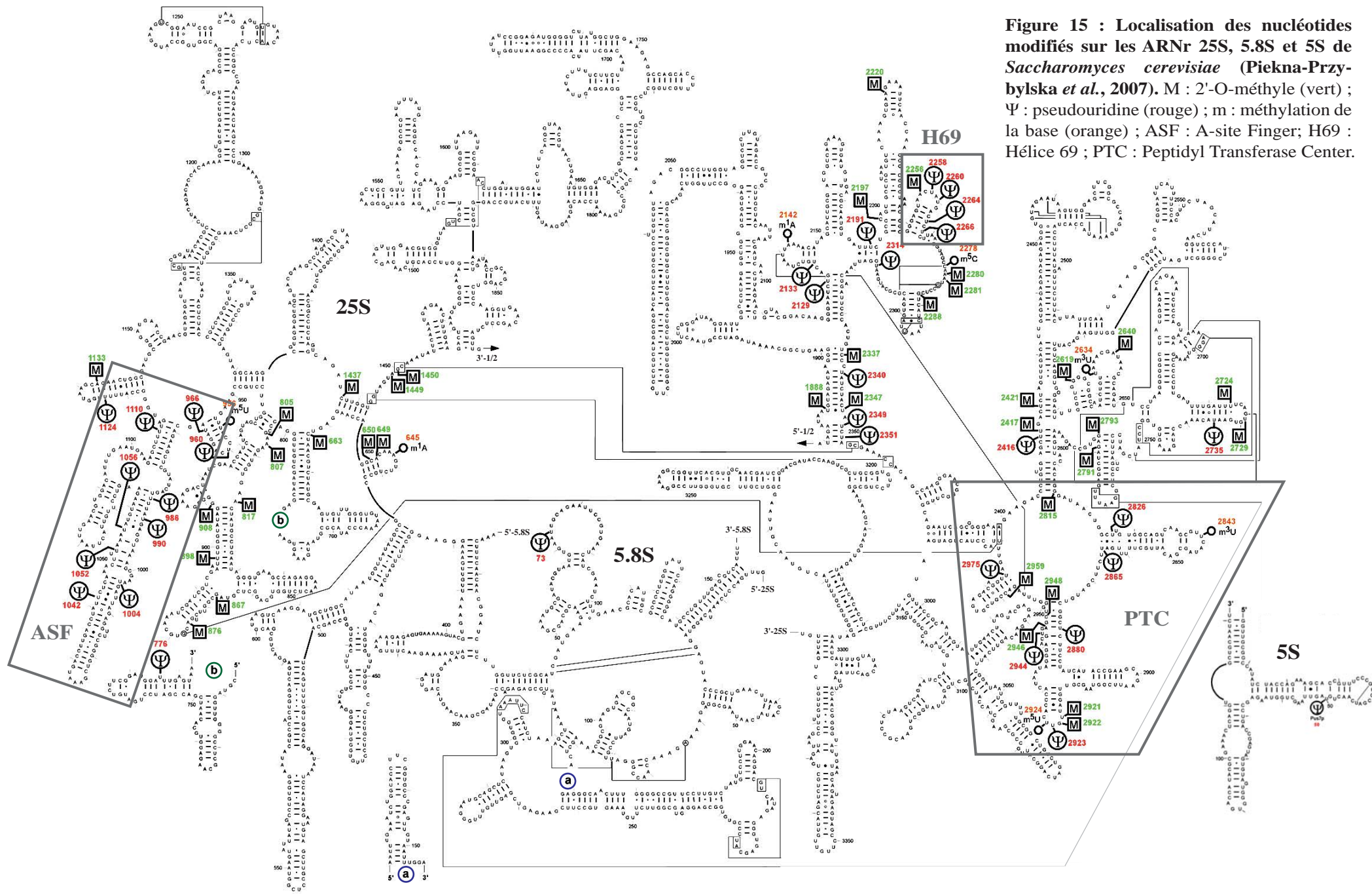
La majorité des nucléotides modifiés est recensée au niveau des ARNr 18S et 25S (Piekna-Przybylska *et al.*, 2007). A titre d'exemple, chez la levure *S. cerevisiae*, les ARNr 18S, 5.8S et 25S contiennent respectivement 35, 1 et 74 modifications, tandis que l'ARNr 5S contient une pseudouridine unique, ajoutée par l'enzyme Pus7p (Decatur & Schnare, 2008; Piekna-Przybylska *et al.*, 2007). Même si la nature co-transcriptionnelle ou post-transcriptionnelle du processus demeure incertaine, l'apposition des modifications par les snoRNP constitue un processus précoce, qui s'effectue sur le pré-ARNr 35S, au sein des DFC du nucléole (Decatur & Fournier, 2003; Hage & Tollervey, 2004). Seules quelques modifications sont ajoutées plus tardivement par des enzymes protéiques : après les premières étapes de clivage du pré-ARNr ou lors des étapes d'assemblage cytoplasmique (Hong *et al.*, 1997; Lafontaine *et al.*, 1994; Lapeyre & Purushothaman, 2004).

A l'échelle moléculaire, ces modifications confèrent aux résidus de nouvelles propriétés chimiques (Decatur & Fournier, 2003) (**Fig. 14**). La 2'-O-méthylation constitue en effet une protection majeure contre la dégradation endonucléolytique, en prévenant des attaques nucléophiles du groupement 3'-hydroxyle sur la liaison 3'-5' phosphodiester. Elle augmente l'hydrophobicité du résidu, favorise la conformation C3'-endo du ribose et peut également participer à de nouvelles interactions inter- ou intramoléculaires (Kawai *et al.*, 1992). La transformation d'un résidu uridine en pseudouridine améliore l'empilement de base et tend à rigidifier le squelette ribose-phosphate. De plus, la libération d'un groupement imine au niveau de l'hétérocycle fournit un site donneur de liaison hydrogène supplémentaire et modifie le spectre d'interactions potentielles de la base (Charette & Gray, 2000). Par le biais de ces nouvelles caractéristiques, le groupement 2'-O-méthyl et la pseudouridine peuvent ainsi exercer une réelle influence sur l'architecture de l'ARN. Ces modifications qui ont fait l'objet de nombreuses études en solution par RMN, présentent notamment un effet stabilisateur sur la structure tridimensionnelle des ARN (Cabello-Villegas & Nikonowicz, 2005; Davis, 1995; Rife & Moore, 1998).

La recherche des nucléotides modifiés au sein des ARNr a permis très tôt de constater leur accumulation au niveau de régions conservées et fonctionnellement importantes (Bakin *et al.*, 1994; Branlant *et al.*, 1981; Brimacombe *et al.*, 1993; Lane *et al.*, 1995; Smith *et al.*, 1992; Veldman *et al.*, 1981) (**Fig. 15 et 16**). Plus récemment, l'émergence de structures cristallographiques à haute résolution a ainsi confirmé la concentration de ces résidus modifiés au niveau des zones clés du ribosome, comme le centre peptidyl transférase de la grande sous-unité, le centre de décodage de la petite sous-unité et la surface d'interaction entre les deux sous-unités (Ban *et al.*, 2000; Ben-Shem *et al.*, 2010; Carter *et al.*, 2000; Nissen *et al.*, 2000; Piekna-Przybylska *et al.*, 2008a; Rabl *et al.*, 2011; Schmeing *et al.*, 2009; Selmer *et al.*, 2006; Wimberly *et al.*, 2000; Yusupov *et al.*, 2001) (**Fig. 17**). La fonction de ces modifications a depuis longtemps fait l'objet de beaucoup d'attentions. La répression de l'expression de snoRNA chez la levure a notamment permis d'étudier l'importance fonctionnelle de ces modifications à l'échelle individuelle (Ganot *et al.*, 1997a; Lowe & Eddy, 1999; Ni *et al.*, 1997; Qu *et al.*, 1999). Les résultats obtenus par ce type d'approche ont ainsi révélé, dans la majorité des cas, l'absence d'impact notoire. Seule la pseudouridine 2919 du centre peptidyl-transférase (King *et al.*, 2003) et la pseudouridine hypermodifiée 1191 du centre de décodage (Liang *et al.*, 2009b) ont une contribution quantifiable sur le métabolisme des cellules de levure (**Tableau 2**). En revanche, de nombreuses études ont permis d'établir



**Figure 14 : Représentation des principales modifications retrouvées dans les ARNr eucaryotes (adapté de Decatur *et al.*, 2002). (A) 2'-O-méthylation. (B) Conversion d'un résidu uridine en pseudouridine ( $\Psi$ ). La liaison N-glycosidique est remplacée par une liaison carbone-carbone formée entre le C1' du ribose et le C5 initial de l'uridine. (C) Positions des bases potentiellement ciblées par des méthylation.**



**Figure 15 : Localisation des nucléotides modifiés sur les ARNr 25S, 5.8S et 5S de *Saccharomyces cerevisiae* (Piekna-Przybylska *et al.*, 2007). M : 2'-O-méthyle (vert) ; Psi : pseudouridine (rouge) ; m : méthylation de la base (orange) ; ASF : A-site Finger; H69 : Hélice 69 ; PTC : Peptidyl Transferase Center.**



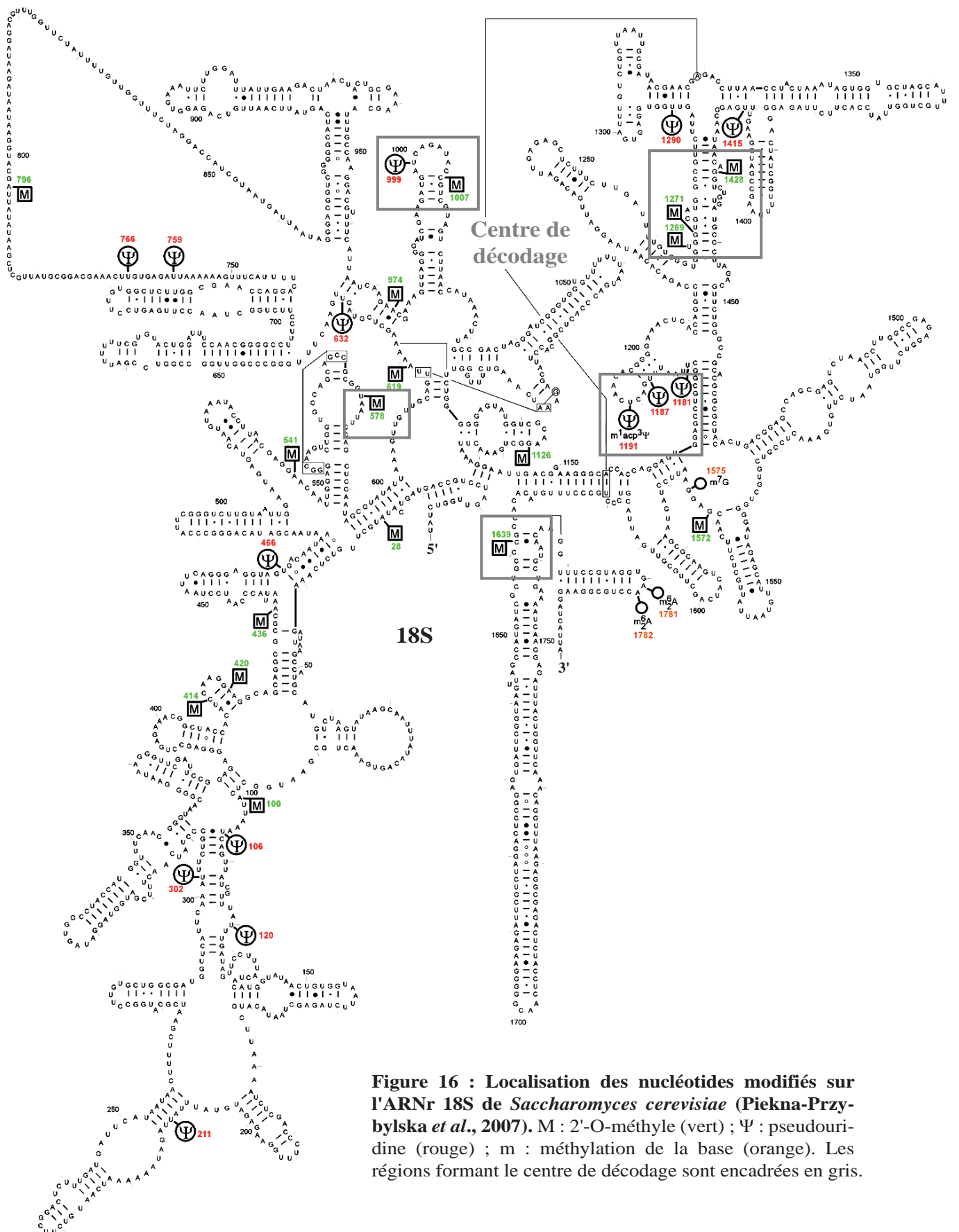
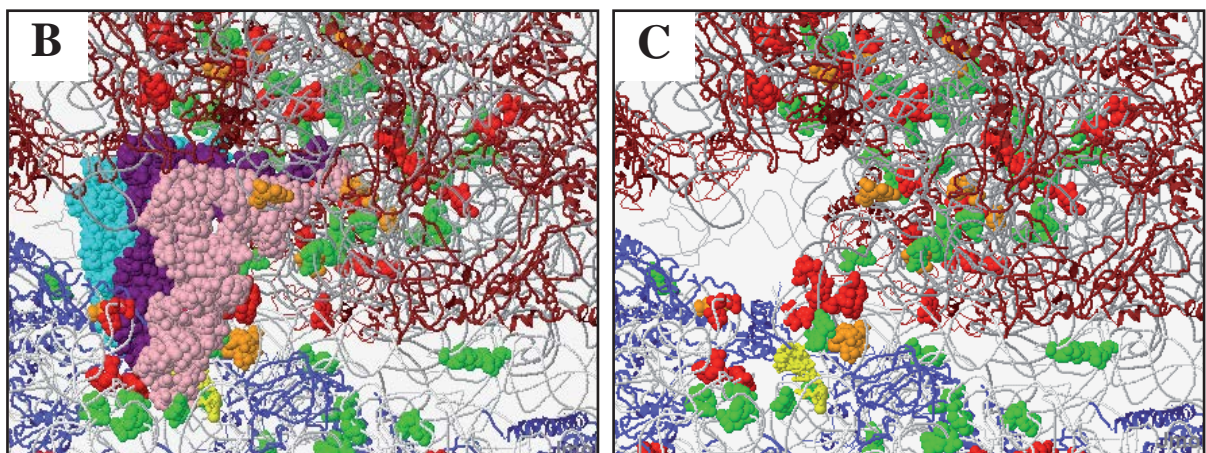
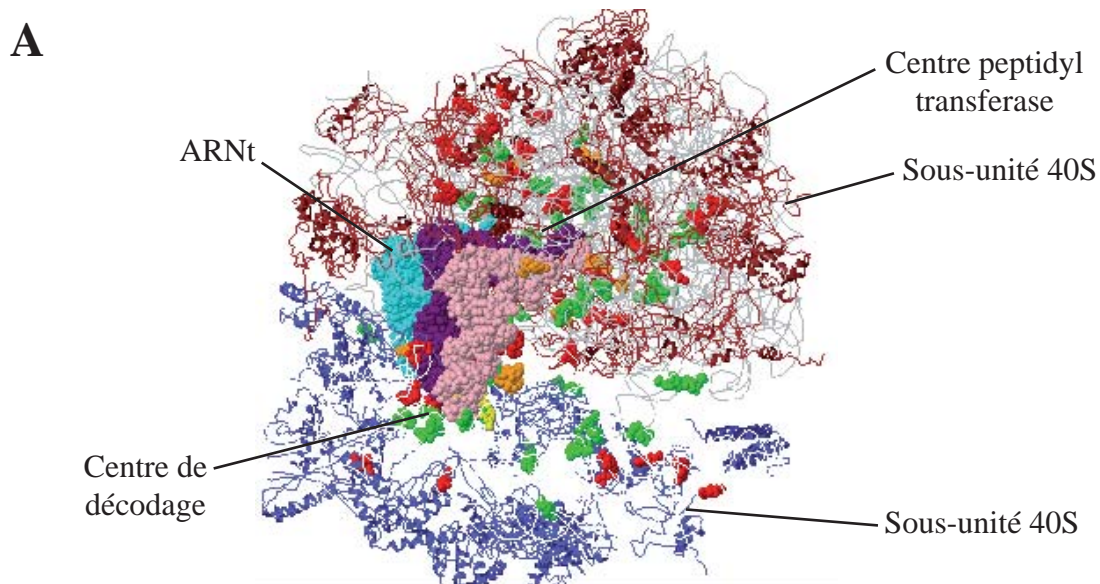


Figure 16 : Localisation des nucléotides modifiés sur l'ARNr 18S de *Saccharomyces cerevisiae* (Piekna-Przybylska *et al.*, 2007). M : 2'-O-méthyle (vert) ; Ψ : pseudouridine (rouge) ; m : méthylation de la base (orange). Les régions formant le centre de décodage sont encadrées en gris.



**Figure 17 : Localisation structurale des nucléotides d'ARNr modifiés chez *Saccharomyces cerevisiae* (d'après Piekna-Przybylska *et al.*, 2007).** (A) Structure complète du ribosome comprenant les deux sous-unités, l'ARNm et 3 ARNt occupant les sites A, P et E. (B) Site actif du ribosome contenant l'ARNm et les 3 ARNt. (C) Site actif du ribosome contenant uniquement l'ARNm. La structure ribosomique utilisée dans ces représentations a été obtenue chez la bactérie hyperthermophile *Thermus thermophilus* (Selmer *et al.*, 2006). La grande conservation du ribosome a permis de replacer sur cette structure, la plupart des modifications de nucléotides relevées chez *S. cerevisiae*. Les résidus 2'-O-méthylés sont représentés en vert, les pseudouridines en rouge et les méthylations de bases en orange. L'ARNm apparaît en jaune, le squelette principal des ARNr apparaît en gris, tandis que les protéines ribosomiques sont représentées en bleu (petite sous-unité) et en marron (grande sous-unité).

Modifications bloquées	Région	snoRNA guide	Phénotype	Références
$\psi$ 2919*	PTC	snR10	Défaut de traduction (20%)	King <i>et al.</i> , 2003
m1acp3 $\psi$ 1191	DC	snR35	Défaut de croissance (10%); Défaut de traduction (7%); Profil de polysomes anormal; Ralentissement du processus de maturation de l'ARNr 18S; Sensibilité accrue à la néomycine	Liang <i>et al.</i> , 2009
$\psi$ 1187	DC	snR36	Défaut de croissance (15%); Défaut de traduction (16%); Profil de polysomes anormal	Liang <i>et al.</i> , 2010
m1acp3 $\psi$ 1191	DC	snR35		
Cm1639	DC	snR70	Défaut de croissance (15%); Défaut de traduction (19%); Profil de polysomes anormal; Sensibilité accrue à la néomycine	Liang <i>et al.</i> , 2009
m1acp3 $\psi$ 1191	DC	snR35		
m1acp3 $\psi$ 1191	DC	snR35	Défaut de croissance (35%); Défaut de traduction (20%); Profil de polysomes anormal; Sensibilité accrue à la néomycine	Liang <i>et al.</i> , 2009 Baudin <i>et al.</i> , 2009
$\psi$ 1187	DC	snR36		
Gm1428	DC	snR56		
Cm1639	DC	snR70	Défaut de croissance (50%); Défaut de traduction (24%); Profil de polysomes anormal; Sensibilité accrue à la néomycine	Liang <i>et al.</i> , 2009 Baudin <i>et al.</i> , 2009
m1acp3 $\psi$ 1191	DC	snR35		
Gm1428	DC	snR56		
Cm1639	DC	snR70	Défaut de croissance (15%); Défaut de traduction (14%); Sensibilité accrue à la néomycine	Liang <i>et al.</i> , 2009
$\psi$ 1187	DC	snR36		
Gm1428	DC	snR56		
Am2256	H69	snR63	Défaut de croissance (35%); Défaut de traduction (35%); Diminution du taux d'ARNr (25%); Sensibilité accrue à la néomycine et à la sparsomycine	Baudin <i>et al.</i> , 2009 Liang <i>et al.</i> , 2007
$\psi$ 2264	H69	snR3		
$\psi$ 2266	H69	snR84		
$\psi$ 1004	ASF	snR56	Défaut de croissance; Défaut de traduction (20%); Profil de polysomes anormal; Sensibilité accrue à la paromomycine et à la sparsomycine	Piekna-Przybylska <i>et al.</i> , 2008
$\psi$ 1042	ASF	snR33		
$\psi$ 1052	ASF	snR81		
Am2256	H69	snR63	Défaut de croissance (17%); Défaut de traduction (25%); Diminution du taux d'ARNr (20%)	Baudin <i>et al.</i> , 2009 Liang <i>et al.</i> , 2007
$\psi$ 2258	H69	snR191		
$\psi$ 2260	H69	snR191		
$\psi$ 2266	H69	snR84		
Am2256	H69	snR63	Défaut de croissance (60%); Défaut de traduction (60%); Diminution du taux d'ARNr (50%); Ralentissement du processus de maturation des ARNr; Sensibilité accrue à la néomycine et à la sparsomycine	Liang <i>et al.</i> , 2007
$\psi$ 2258	H69	snR191		
$\psi$ 2260	H69	snR192		
$\psi$ 2264	H69	snR3		
$\psi$ 2266	H69	snR84		
$\psi$ 960	ASF	snR8	Défaut de croissance (23%); Défaut de traduction (10%); Profil de polysomes anormal; Sensibilité accrue à la paromomycine et à la sparsomycine	Baudin <i>et al.</i> , 2009 Piekna-Przybylska <i>et al.</i> , 2008
$\psi$ 986	ASF	snR8		
$\psi$ 1004	ASF	snR56		
$\psi$ 1042	ASF	snR33		
$\psi$ 1052	ASF	snR81		
$\psi$ 1124	ASF	snR56		

**Tableau 2 : Principaux effets observés suite à la perte de modification de nucléotides au niveau des ARNr 18S et 25S chez *S. cerevisiae*.**  $\psi$  : pseudo-uridine; Nm : nucléotide méthylé en position 2' du ribose; m1acp3 : modifications au niveau de la base; PTC : Peptidyl Transferase Center; DC : Decoding Center; H69 : Hélice 69; ASF : A-Site Finger. Les modifications conservées chez l'homme sont annotées en rouge. La région DC est située au niveau de la petite sous-unité (ARNr 18S), les régions PTC, H69 et ASF appartiennent à la grande sous-unité (ARNr 25S). H69 et ASF sont situées à l'interface avec la petite sous-unité. (\*)  $\psi$ 2919 =  $\psi$ 2923 dans la figure 15, d'après Piekna-Przybylska *et al.*, 2007).

l'existence de phénomènes de synergies entre les modifications, puisqu'un cumul de mutations dans une région donnée, peut conduire à de graves défauts de fonctionnement du ribosome (**Tableau 2**). En effet, des groupes de modifications, situés notamment au niveau du centre de décodage et à l'interface entre les deux sous-unités du ribosome, se sont révélés importants pour la production des ARNr 18S et 25S, pour la formation des polysomes, pour l'activité du ribosome, pour le contrôle de la fidélité de traduction et pour la croissance cellulaire (Baudin-Baillieu *et al.*, 2009; Baxter-Roshek *et al.*, 2007; King *et al.*, 2003; Liang *et al.*, 2007b; Liang *et al.*, 2009b; Piekna-Przybylska *et al.*, 2008b). Ces données récentes révèlent donc l'importance collective des modifications de nucléotides pour la biogenèse et la fonction des ribosomes. A défaut de forts impacts individuels, leur action synergique et leur grande conservation témoignent donc de l'avantage sélectif qu'elles ont pu apporter au cours de l'évolution, dans le cadre de l'optimisation du processus de traduction.

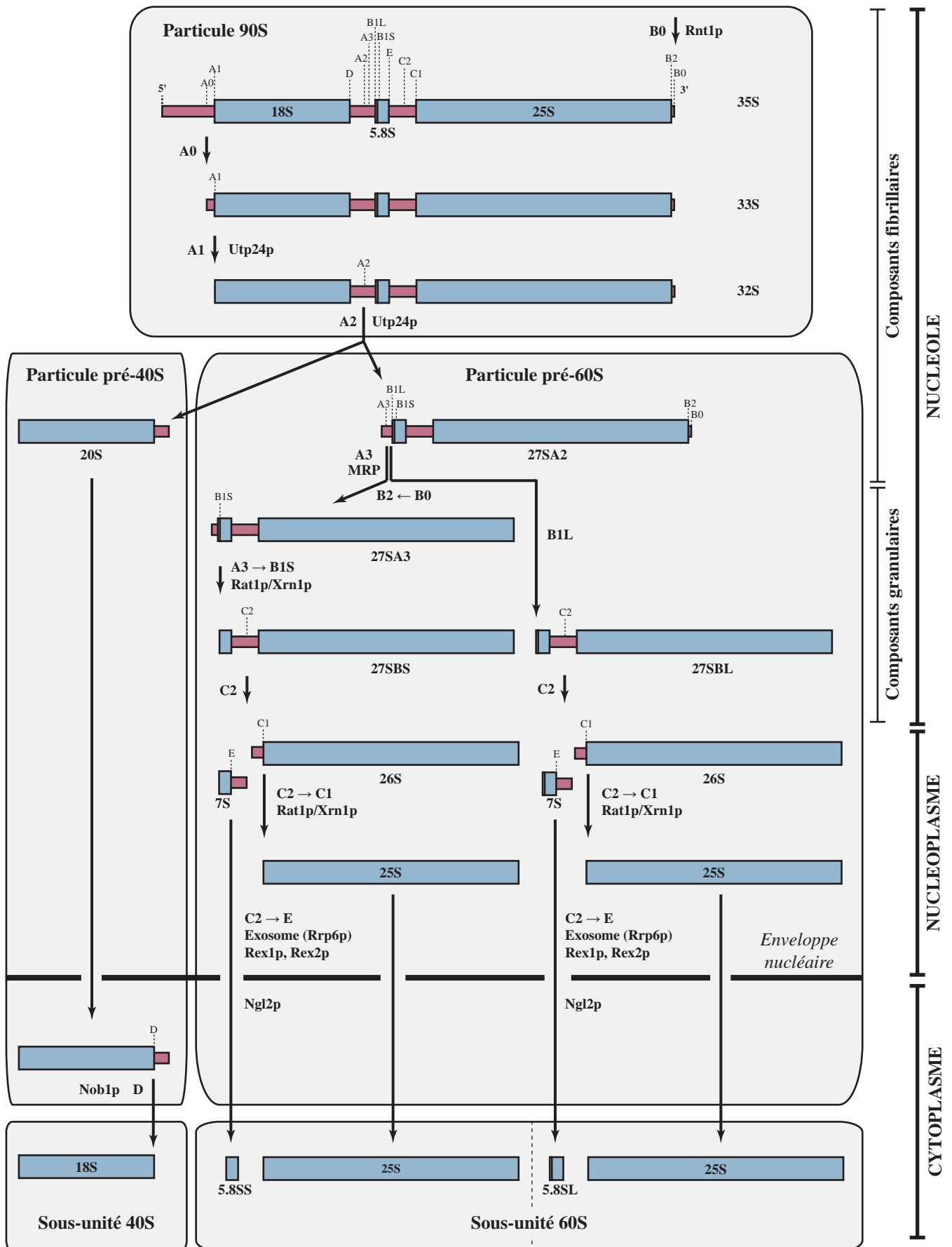
### **3.3. Maturation nucléolytique des ARNr**

L'élimination des séquences additionnelles contenues dans les précurseurs des ARNr est un processus complexe impliquant une série de clivages endo- et exonucléolytiques (**Fig. 18**). Ces événements de maturation reposent sur des éléments de séquence conservés au sein de l'ARN précurseur et sur l'intervention d'un grand nombre de facteurs agissant *in trans*. Ces mécanismes ont notamment fait l'objet d'études approfondies chez la levure *S. cerevisiae*.

#### **3.3.1. Maturation conduisant à la production des ARNr 18S, 5.8S et 25S de *S. cerevisiae***

##### **3.3.1.1. Maturation du pré-ARNr 35S**

La maturation nucléolytique du pré-ARNr 35S débute par l'élimination de la séquence 5'-ETS grâce à 2 clivages endonucléolytiques générés successivement aux sites de coupures A0 et A1 (**Fig. 18**). La coupure au site A0 aboutit à la formation de l'intermédiaire 33S (Hughes & Ares, 1991). Le clivage au site A1 génère ensuite un intermédiaire 32S et permet de délimiter l'extrémité 5' du futur ARNr 18S (Beltrame *et al.*, 1994). Puis, un troisième clivage,



**Figure 18 : Maturation par différentes activités nucléolytiques des précurseurs des ARNr 18S, 25S et 5.8S chez *Saccharomyces cerevisiae* (adapté de Henras *et al.*, 2008).** Les clivages exonucléolytiques sont matérialisés par les symboles  $\leftarrow$  (3'-5') et  $\rightarrow$  (5'-3'). Les facteurs responsables des clivages sont précisés lorsqu'ils sont connus.

réalisé dans la séquence ITS1, au site A2, permet la séparation du processus de maturation en deux voies distinctes : celle du pré-ARNr 20S, qui contient l'ARNr 18S, et celle du pré-ARNr 27SA2 qui renferme les ARNr 5.8S et 25S (De Jonge *et al.*, 1977; Henry *et al.*, 1994; Lindahl *et al.*, 1994). De plus, des études de microscopie électronique, réalisées chez la levure, ont permis de montrer que, dans des conditions de croissance normale, les coupures aux sites A0, A1, A2 et la libération du pré-ARNr 20S s'effectuent de façon co-transcriptionnelle (Osheim *et al.*, 2004). Ce scénario de clivages co-transcriptionnels ne semble toutefois pas conservé chez tous les eucaryotes, puisqu'il n'a été constaté que chez *S. cerevisiae* et chez *Dictyostelium* (Grainger & Maizels, 1980; Osheim *et al.*, 2004). Enfin, concernant la terminaison de la transcription, il a été montré qu'un clivage catalysé par l'endonucléase Rnt1p au niveau du site terminal B0, permet de définir l'extrémité 3' du précurseur (Henras *et al.*, 2004; Piper *et al.*, 1983).

Ces événements requièrent l'intervention de nombreux facteurs, parmi lesquels les snoRNP à boîtes C/D U3, U8, U14 et U22, ainsi que les snoRNP à boîtes H/ACA snR10 et snR30/U17 (Hughes & Ares, 1991; Kass *et al.*, 1990; Li *et al.*, 1990; Morrissey & Tollervey, 1993; Peculis & Steitz, 1993; Tycowski *et al.*, 1994; Zagorski *et al.*, 1988). Parmi ces ARN, seuls les snoRNA U3, U14 et snR30/U17 sont conservés chez tous les eucaryotes. Les snoRNA U8 et U22 sont en effet retrouvés uniquement chez les métazoaires, tandis que le snoRNA snR10 est spécifique de la levure. D'autre part, seuls les snoRNA U14 et snR10 présentent une double fonctionnalité, puisqu'ils sont également impliqués dans des réactions de modifications de nucléotides (Lestrade & Weber, 2006; Piekna-Przybylska *et al.*, 2007).

Le rôle joué par les snoRNA U3, U14, snR10 et snR30 dans la maturation nucléolytique du pré-ARNr a fait l'objet de nombreux travaux, en particulier chez la levure. On sait notamment que l'appariement du snoRNA U3 avec les régions 5'-ETS et 18S est essentiel pour les coupures précoces aux sites A0, A1 et A2 (Borovjagin & Gerbi, 1999; Hughes, 1996; Mereau *et al.*, 1997; Morrissey & Tollervey, 1993). L'appariement du snoRNA U14 avec une séquence de l'ARNr 18S est requis pour les clivages aux sites A1 et A2 (Beltrame *et al.*, 1994; Huang *et al.*, 1992; Jarmolowski *et al.*, 1990; Li & Fournier, 1992; Li *et al.*, 1990; Zagorski *et al.*, 1988). Enfin, l'appariement des snoRNA snR10 et snR30, respectivement au niveau des régions 5'-ETS et 18S, sont également requis pour le déroulement correct des coupures en A1 et A2 (Beltrame *et al.*, 1994; Fayet-Lebaron *et al.*, 2009; Liang *et al.*, 2010). Parmi ces ARN, on remarquera que les snoRNA U3, U14 et snR30 constituent, avec l'ARN

de la RNase MRP, les seuls snoRNA essentiels à la viabilité cellulaire chez la levure. Concernant les snoRNA U22 et U8, spécifiques des métazoaires, des études réalisées chez le xénope ont montré leur implication, respectivement dans la production de l'ARNr 18S (Tycowski *et al.*, 1994), et dans la maturation des ARNr 28S et 5.8S (Peculis, 1997; Peculis & Steitz, 1993). Nous reviendrons en détail dans le chapitre suivant sur le rôle joué par la snoRNP U3 dans ces processus.

### 3.3.1.2. Maturation du pré-ARNr 20S

La suite de la maturation de l'intermédiaire 20S se déroule après l'export de la particule pré-40S vers le cytoplasme. C'est une coupure endonucléolytique, générée au site D par la protéine Nob1p, qui permet de définir l'extrémité 3' mature de l'ARNr 18S (Fatica *et al.*, 2003; Lamanna & Karbstein, 2009; Trapman & Planta, 1976; Udem & Warner, 1972; Vanrobays *et al.*, 2001) (**Fig. 18**). Le fragment ITS1 libéré est alors dégradé par les exonucléases 5'-3' Rat1p et Xrn1p (Stevens *et al.*, 1991).

### 3.3.1.3. Maturation du pré-ARNr 27SA2

La maturation de l'intermédiaire 27SA2 se poursuit au sein du nucléole en suivant deux voies alternatives (**Fig. 18**). La voie majoritaire concerne environ 85% de la population d'ARN. Elle implique un clivage endonucléolytique au site A3, situé dans la région ITS1 résiduelle (Henry *et al.*, 1994; Lygerou *et al.*, 1996). Cette coupure est catalysée par la RNase MRP, une snoRNP porteuse d'une activité ribozyme, et aboutit à la formation de l'intermédiaire 27SA3 (Reilly & Schmitt, 1995). Cette étape permet ainsi de générer un site d'entrée pour les exonucléases 5'-3' Rat1p et/ou Xrn1p, qui dégradent l'ARN jusqu'au site B1S pour former l'intermédiaire 27SBS (Xue *et al.*, 2000b). Dans certains cas, le clivage au site A3 peut se produire indépendamment des clivages précoces en A0, A1 et A2. Dans ce cas, la formation de l'intermédiaire 27SA3 s'accompagne de la libération d'un intermédiaire 23S contenant l'ARNr 18S. Cet intermédiaire aberrant subit alors une dégradation exonucléolytique 3'-5' catalysée par l'exosome (Allmang *et al.*, 2000; Hughes & Ares, 1991). Cependant, des études plus récentes suggèrent qu'il pourrait également contribuer à la production de l'ARNr 18S

(Gallagher *et al.*, 2004). Dans la voie minoritaire, l'intermédiaire 27SA2 est clivé directement au site B1L, situé 6 nucléotides en amont du site B1S, sans subir de clivage au site A3. Cette voie, relativement peu décrite, abouti à la formation de l'intermédiaire 27SBL. La formation des intermédiaires 27SBS et 27SBL s'accompagne également de l'élimination de la séquence 3'-ETS, qui se déroule en parallèle des coupures aux sites B1 et qui permet de définir l'extrémité 3' mature du futur ARNr 25S au site B2 (Veldman *et al.*, 1980).

Les deux populations d'intermédiaires 27SB suivent ensuite une voie de maturation identique. Celle-ci débute par un clivage endonucléolytique au site C2, qui conduit à la libération des précurseurs 7S de l'ARNr 5.8S et du précurseur 26S de l'ARN 25S. Le fragment ITS2 situé en 5' de l'intermédiaire 26S est alors dégradé par les exonucléases 5'-3' Rat1p ou Xrn1p jusqu'au site C1 qui constitue l'extrémité 5' de l'ARNr 25S (Geerlings *et al.*, 2000). L'élimination du fragment ITS2 résiduel, situé en 3' des précurseurs 7S, requiert plusieurs coupures. Ce processus implique en effet l'activité exonucléase 3'-5' de l'exosome et des facteurs Rex1p et Rex2p, ainsi que l'endonucléase Ngl2p (Allmang *et al.*, 1999; Faber *et al.*, 2002; Mitchell *et al.*, 1996; van Hoof *et al.*, 2000a). Cette digestion s'effectue jusqu'au site E qui constitue l'extrémité 3' mature de l'ARNr 5.8S. Ce dernier est donc présent sous deux formes dans la cellule : une espèce courte 5.8SS majoritaire et une espèce minoritaire 5.8SL, plus longue de 6 nucléotides en 5'.

### 3.3.2. Maturation de l'ARNr 5S de *S. cerevisiae*

L'ARNr 5S est transcrit sous la forme d'un précurseur contenant une extension de 10 nucléotides à son extrémité 3'. Son élimination est assurée par l'exonucléase 3'-5' Rex1p, dont la progression semble être stoppée au niveau de l'extrémité 3' mature par la formation d'une structure secondaire impliquant les extrémités 5' et 3' de l'ARN (Lee & Nazar, 1997; van Hoof *et al.*, 2000a).

## 3.4. Assemblage des sous-unités ribosomiques

Les premières études sur l'assemblage des sous-unités ribosomiques ont été menées dans les années 1970 et ont conduit à l'identification de structures pré-ribosomiques 90S, 43S et 66S

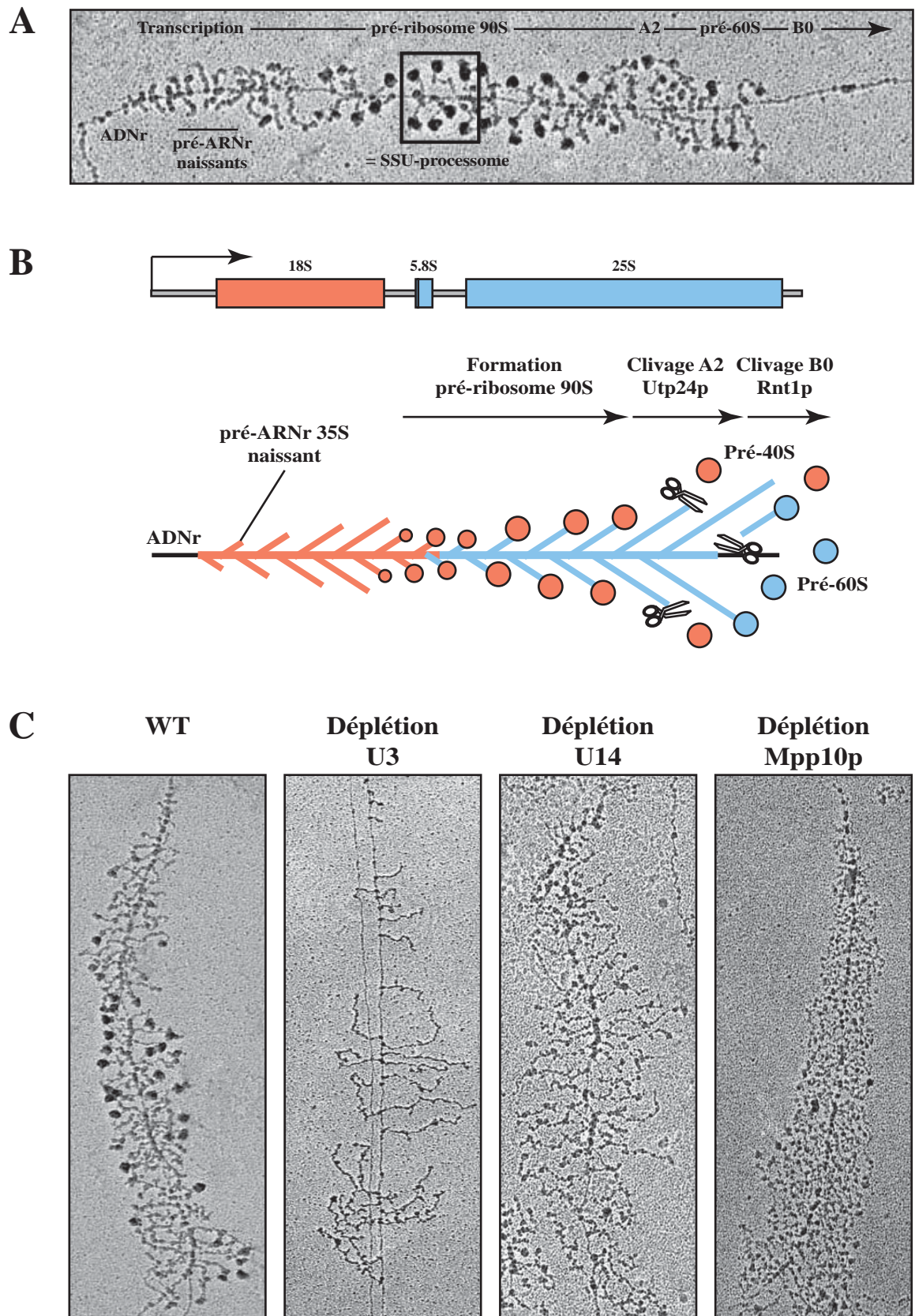


(Trapman *et al.*, 1975; Udem & Warner, 1972). Trente années supplémentaires ont néanmoins été nécessaires pour parvenir à en identifier les composants. C'est en effet le développement de méthodes de purification de haute affinité, comme la méthode TAP (*Tandem Affinity Purification*; (Rigaut *et al.*, 1999)), couplées à des approches d'analyses par spectrométrie de masse, qui ont permis une caractérisation biochimique approfondie de ces complexes. Ces études ont ainsi révélé l'intervention transitoire de nombreux facteurs non ribosomiques au cours d'un processus dichotomique, qui débute co-transcriptionnellement par la formation du pré-ribosome 90S (Tschochner & Hurt, 2003). Cette particule de haut poids moléculaire, qui est principalement composée des facteurs impliqués dans la biogenèse de l'ARNr 18S, est à l'origine de la particule pré-40S. La formation du complexe pré-60S se produit quant à elle, de façon secondaire et relativement indépendante. Son assemblage débute également de façon co-transcriptionnelle (Osheim *et al.*, 2004).

### 3.4.1. Du pré-ribosome 90S à la sous-unité 40S

Le processus d'assemblage des ribosomes débute de façon co-transcriptionnelle, par le recrutement sur le pré-ARNr naissant de certaines protéines ribosomiques, ainsi que de nombreux facteurs protéiques et ribonucléoprotéiques (Kressler *et al.*, 2010). Ces événements précoces se matérialisent notamment par la formation d'un bouton terminal à l'extrémité 5' du transcrit naissant (Miller & Beatty, 1969; Mougey *et al.*, 1993; Osheim *et al.*, 2004) (**Fig. 19**). La répétition de ce type d'éléments au niveau des régions d'ADNr activement transcrites forme une structure dite en « arbre de Noël », dont l'étude par microscopie électronique permet d'apprécier la cinétique des étapes initiales de biogenèse (Osheim *et al.*, 2004) (**Fig. 19A et 19B**). En 2002, deux groupes ont réalisé la purification et l'analyse de ces particules précoces (Dragon *et al.*, 2002; Grandi *et al.*, 2002).

En utilisant une approche de double purification (TAP-Nop58p, puis flag-Mpp10p), l'équipe de S. Baserga a caractérisé un complexe d'environ 2,2 MDa, composé de 28 protéines non ribosomiques et du snoRNA U3 (Dragon *et al.*, 2002). Parmi ces protéines, on retrouve les 4 protéines constitutives des snoRNP à boîtes C/D (Snu13p, Nop1p, Nop56p et Nop58p), 6 protéines spécifiques de la snoRNP U3 (Rrp9p, Imp3p, Imp4p, Mpp10p, Sof1p et Dhr1p), le facteur Rrp5p, ainsi que 17 nouvelles protéines, baptisées Utp (*U-three protein*) et dont la



**Figure 19 : Formation co-transcriptionnelle des particules pré-ribosomiques chez *Saccharomyces cerevisiae* (Extrait de Osheim *et al.*, 2004).** (A) Observation par microscopie électronique d'un gène d'ADNr en cours de transcription. (B) Représentation schématique des événements d'assemblage des structures pré-ribosomiques sur le pré-ARNr naissant. Chez la levure *S. cerevisiae* les clivages aux sites A0, A1, A2 s'effectuent de façon co-transcriptionnelle. (C) Impact de la déplétion des snoRNA U3, U14 et de la protéine Mpp10p sur la formation des structures pré-ribosomiques. La déplétion des snoRNA et de la protéine Mpp10p perturbent la formation des boutons terminaux et inhibent les clivages. L'effet est particulièrement marqué pour le snoRNA U3.

déplétion inhibe la production de l'ARNr 18S. En plus de ces facteurs, 5 protéines de la petite sous-unité ribosomique ont été identifiées (Rps4p, Rps6p, Rps7p, Rps14p et Rps28p). D'autre part, les auteurs ont également démontré que la déplétion du snoRNA U3 et de certaines protéines comme Utp7p, Imp3p ou Imp4p, conduit à la perte des boutons terminaux à l'extrémité 5' des transcrits naissants, suggérant la nature commune de ces structures et du complexe identifié (**Fig. 19C**). Les auteurs ont ainsi proposé que ce complexe ribonucléoprotéique de grande taille (environ 80S), formé sur la base de l'appariement du snoRNA U3 avec le pré-ARNr, constitue le cœur de la machinerie impliquée dans la maturation précoce de l'ARNr 18S. Cette nouvelle structure a ainsi été baptisée SSU processome (*Small Subunit processome*). On remarquera par ailleurs que les snoRNA U14, snR10 et snR30, qui sont présents dans les mêmes fractions que le pré-ribosome 90S dans les expériences de centrifugation en gradient de saccharose, n'ont pas été détectés dans les conditions de purification utilisées au cours de cette étude. Des travaux réalisés au laboratoire par V. Senty-Ségault ont néanmoins démontré que le snoRNA U14 est largement co-immunoprécipité par la protéine Mpp10p, suggérant ainsi la présence des snoRNA U3 et U14 au sein de structures communes (Communication personnelle).

En parallèle de ces travaux, l'équipe d'E. Hurt a caractérisé un complexe similaire (environ 80S) contenant certains facteurs non ribosomiques supplémentaires (dont Pwp2p, Enp1p, Nop14p, Krr1p, Kre31p, Bud21p, Noc4p et la GTPase Bms1p), quelques protéines de la petite sous-unité ribosomique et une minorité de protéines de la grande sous-unité (Grandi *et al.*, 2002). La majorité des nouveaux facteurs identifiés dans cette étude avaient précédemment été impliqués dans la maturation de l'ARNr 18S et dans la formation de la sous-unité 40S (Gelperin *et al.*, 2001; Liu & Thiele, 2001; Sasaki *et al.*, 2000; Shafaatian *et al.*, 1996; Wegierski *et al.*, 2001). Les auteurs ont également montré que deux d'entre eux, les protéines Krr1p et Kre31p, sont requis pour l'export cytoplasmique de la particule pré-40S, suggérant un phénomène de couplage entre les deux mécanismes. Enfin, ils ont mis en évidence que la plupart de ces facteurs non ribosomiques, à l'exception notoire de la protéine Enp1p, ne sont pas retrouvés sur le pré-ARNr 20S, suggérant leur dissociation de l'ARN après les étapes de coupures aux sites A0, A1 et A2.

Par la suite, d'autres composants du SSU-processome ont été caractérisés. C'est notamment le cas de la protéine YDR339c, baptisée Utp24p, qui a fait l'objet d'une attention particulière (Bleichert *et al.*, 2006). Ce facteur possède en effet un domaine de type PINc (*PilT N-*

*terminus domain*), porteur d'une activité endonucléase potentielle (Arcus *et al.*, 2004). De façon intéressante, la mutation des résidus catalytiques putatifs de Utp24p conduit à l'abolition des clivages aux sites A1 et A2 et suggère fortement que ce nouveau facteur pourrait être l'endonucléase impliquée dans la libération du pré-ARNr 20S (Bleichert *et al.*, 2006).

D'autres études ont également apporté des informations sur les événements d'assemblage du complexe pré-ribosomique 90S. Des données de microscopie électronique ont notamment permis de montrer que les boutons terminaux, situés à l'extrémité des transcrits naissants, évoluent en taille et en forme au cours du temps, suggérant un assemblage séquentiel de cette structure (Osheim *et al.*, 2004). Cette observation a été corroborée par la proposition d'un modèle d'assemblage hiérarchisé, faisant intervenir différents sous-complexes (Gallagher *et al.*, 2004; Krogan *et al.*, 2004; Perez-Fernandez *et al.*, 2007) (**Fig. 20**). Le premier module qui interagit avec le pré-ARNr est le complexe t-UTP (*U Three Protein complex required for transcription*). Ce complexe, composé de 7 protéines (Utp4p, Utp5p, Utp8p, Utp9p, Utp10p, Utp15p, et Nan1p), est recruté de façon précoce sur la chromatine, il contribue à l'activation de la transcription par l'ARN polymérase I et s'avère requis pour le recrutement des autres sous-complexes. La suite de l'assemblage implique deux voies indépendantes : d'une part, le recrutement de la snoRNP U3, du complexe UTP-B (Pwp2p, Dip2p, Utp21p, Utp13p, Utp18p, Utp6p), du complexe MPP10 (Mpp10p, Imp3p, Imp4p) et de 7 facteurs additionnels (Utp20, Bms1p, Kre33p, Nop14p, Enp2p, Noc4p, Krr1p), et d'autre part, le recrutement du facteur Rrp5p, suivi du complexe UTP-C (Utp22p, Rrp7p, Cka1p, Cka2p, Ckb1p, Ckb2p) et de l'hélicase Rok1p.

De nombreuses données concernant ces événements précoces d'assemblage ont donc émergé ces dernières années et ont révélé la grande complexité structurale et fonctionnelle du pré-ribosome 90S. Sa mise en place influe directement sur le processus de transcription, elle initie le processus d'assemblage de la petite sous-unité et s'avère essentielle à la maturation nucléolytique du pré-ARNr. Sur la base des multiples contacts qui s'établissent entre les facteurs non ribosomiques et le pré-ARNr 35S, plusieurs études ont également proposé que le SSU-processome constitue une chaperonne du pré-ARNr (Phipps *et al.*, 2011). La structuration correcte de l'ARN constitue en effet un élément crucial pour sa maturation nucléolytique et pour l'assemblage correct des protéines ribosomiques. Le rôle joué par le snoRNA U3 dans ces processus a été tout particulièrement étudié. Son appariement dans les

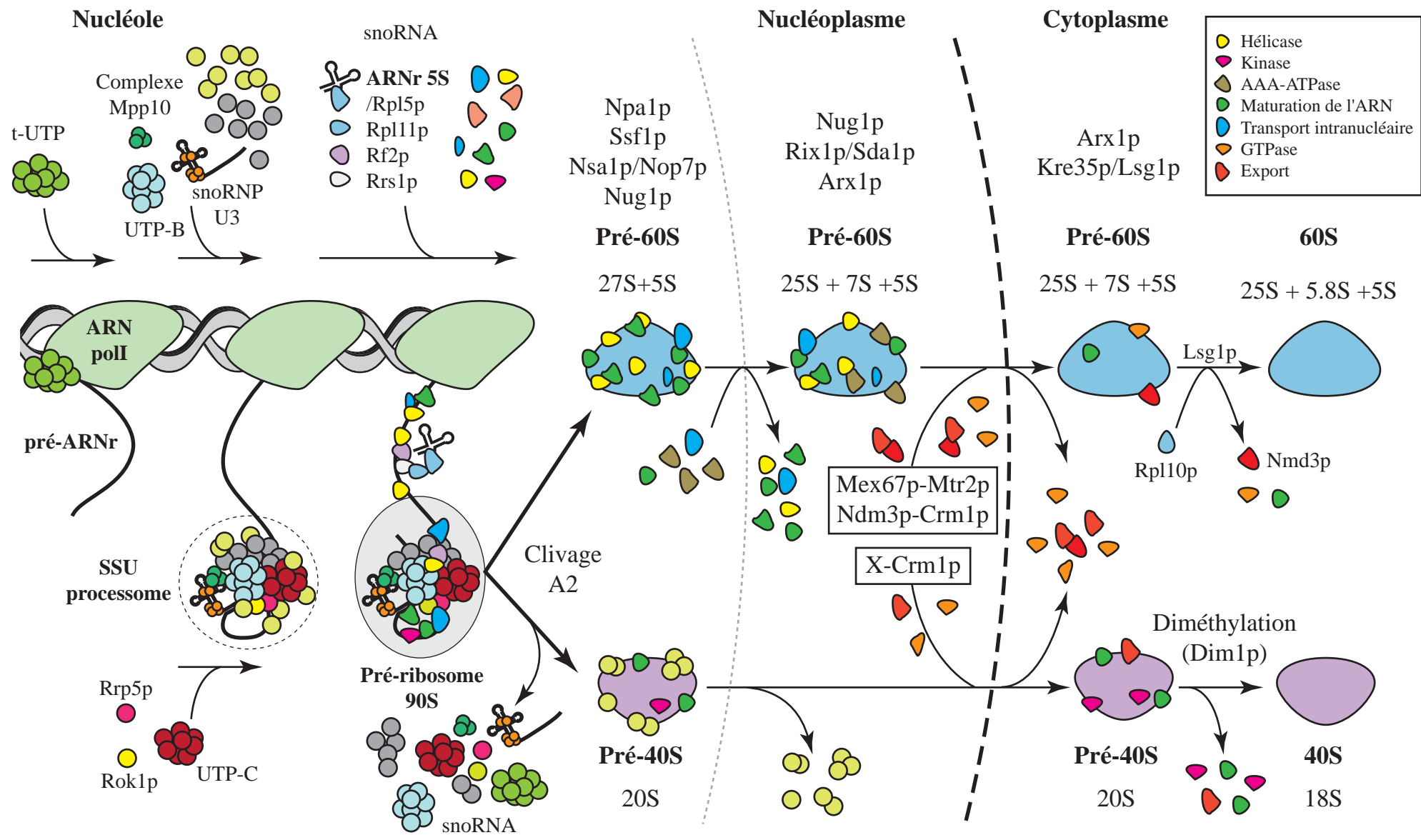


Figure 20 : Principales étapes de l'assemblage des sous-unités ribosomiques chez *Saccharomyces cerevisiae* (adapté de Perez-Fernandez *et al.*, 2007 et Tschochner *et al.*, 2003).

régions 5'-ETS et 18S permet en effet une exposition optimale des sites de coupures précoces et régule le repliement de la partie 5' du futur ARNr 18S (Beltrame & Tollervey, 1995; Gerczei *et al.*, 2009; Hughes, 1996; Mereau *et al.*, 1997; Sharma & Tollervey, 1999). Nous reviendrons sur ces aspects dans le chapitre suivant.

A la suite des clivages en A0, A1 et A2, la particule pré-40S est libérée. Certains facteurs sont conservés (Dim1p, Dim2p, Enp1p, Hrr25p, Nob1p, Prp43p, Rrp12p et Tsr1p), tandis que la grande majorité des autres composants non ribosomiques sont dissociés (Lebaron *et al.*, 2005; Schafer *et al.*, 2003). De nouvelles protéines ribosomiques sont recrutées, ainsi qu'un petit nombre de facteurs impliqués dans la biogenèse de la particule (dont Nob1p, Rio1p, Rio2p, Tsr1p, Fap7p) (Schafer *et al.*, 2003). La particule est alors rapidement transportée au niveau du cytoplasme en suivant la voie CRM1. Les étapes tardives de la biogenèse de la petite sous-unité impliquent la diméthylation conservée du pré-ARNr 20S au niveau de l'extrémité 3' de la région 18S (Brand *et al.*, 1977). Cette modification est catalysée par la méthylase Dim1p, l'un des rares facteurs présent depuis la formation du pré-ribosome 90S (Lafontaine *et al.*, 1994). Enfin, le pré-ARNr 20S est clivé au niveau du site D pour aboutir à la formation de l'ARNr 18S mature. Ce clivage, catalysé par la protéine Nob1p (Lamanna & Karbstein, 2009; Pertschy *et al.*, 2009), requiert l'intervention d'autres facteurs comme Fas7p, Rio1p et Rio2p (Geerlings *et al.*, 2003; Granneman *et al.*, 2005; Vanrobays *et al.*, 2001).

### 3.4.2. Assemblage de la sous-unité 60S

Chez la levure, l'assemblage de la sous-unité 60S est un processus qui débute également de façon co-transcriptionnelle. A la suite du clivage au site A2, de nouveaux bourgeons terminaux se forment en effet à l'extrémité 5' du transcrit pour donner naissance à la particule pré-60S (Osheim *et al.*, 2004) (**Fig. 19A et 19B**). Le complexe pré-60S nucléolaire a été la première particule pré-ribosomique purifiée grâce à la méthode TAP (Bassler *et al.*, 2001; Fatica *et al.*, 2002a; Harnpicharnchai *et al.*, 2001; Saveanu *et al.*, 2001). La comparaison des différentes données obtenues a révélé une certaine hétérogénéité de composition, suggérant l'existence de nombreux complexes intermédiaires. Par la suite, des études complémentaires ont permis de reconstituer cette voie d'assemblage en définissant des marqueurs non ribosomiques spécifiques des différents complexes (protéines annotées entre parenthèse dans le texte ci-dessous) (Dez *et al.*, 2004; Kressler *et al.*, 2008; Nissan *et al.*, 2002). Ces travaux

ont ainsi permis de distinguer des particules nucléolaires (Npa1p), (Ssf1p) et (Nsa1p, Nop7p), des particules nucléolaire/nucléoplasmiques (Nug1p), des particules nucléoplasmiques (Rix1p, Sda1p), des particules nucléoplasmique/cytoplasmiques (Arx1p) et des particules purement cytoplasmiques (Kre35p, Lsg1p) (**Fig. 20**). Un grand nombre de facteurs interviennent dans ces étapes, environ 50 protéines non ribosomiques sont associées aux particules nucléolaires, contre seulement 5 après l'export vers le cytoplasme (Nissan *et al.*, 2002).

Les particules pré-60S les plus précoces (Npa1p), qui contiennent le pré-ARNr 27SA2, se caractérisent notamment par la présence de nombreuses snoRNP et de quelques facteurs issus du pré-ribosome 90S (Dez *et al.*, 2004). Le complexe suivant (Ssf1p) contient un mélange de pré-ARNr 27SA et 27SB, mais apparaît dépourvu de snoRNP, suggérant que les modifications guidées ont déjà été effectuées lors des étapes précédentes (Fatica *et al.*, 2002a). La particule suivante (Nsa1p) se caractérise par la présence exclusive de la forme pré-ARNr 27SB et par la détection de l'ARNr 5S (Kressler *et al.*, 2008). Il a ainsi été proposé que le recrutement de l'ARNr 5S, associé sous forme d'une pré-RNP à la protéine Rpl5p, s'effectue au sein de cette particule. Néanmoins, une autre étude suggère que son incorporation constitue un événement antérieur, puisque l'ARNr 5S et la protéine Rpl5p ont également été détectés au sein du pré-ribosome 90S (Zhang *et al.*, 2007). Ce processus, ainsi que le recrutement de la protéine Rpl11p, dépend des facteurs Rpf2p et Rrs1p, dont la présence au sein des particules pré-60S est requise pour la maturation subséquente du pré-ARNr 27SB (Zhang *et al.*, 2007). La transition vers le nucléoplasme (Rix1p) s'accompagne ensuite de changements majeurs de composition et de la conversion presque totale du pré-ARNr 27SB en ARNr 25S et en pré-ARNr 7S (Nissan *et al.*, 2002). La particule suivante (Arx1p) est alors rendue compétente pour l'export vers le cytoplasme par la présence du facteur Nmd3p qui interagit avec Crm1p (Gadal *et al.*, 2001; Ho *et al.*, 2000) (**Fig. 3**), et de la protéine Arx1p qui interagit directement avec certaines nucléoporines (Bradatsch *et al.*, 2007; Hung *et al.*, 2008). Ce complexe pré-60S contient également, l'hétérodimère Mex67p-Mtr2p, impliqué dans l'export des ARNm (Yao *et al.*, 2007). En effet, l'équipe d'E. Hurt a récemment montré qu'une région du complexe Mex67p-Mtr2p, spécifique de la levure, peut interagir avec la particule pré-60S et promouvoir son export vers le cytoplasme, indépendamment du transport des ARNm (Yao *et al.*, 2007). Une fois dans le cytoplasme, la particule subit de nouveaux changements de composition et s'enrichit en ARNr 5.8S mature, dont la maturation finale requiert l'exonucléase cytoplasmique Ngl2p (Nissan *et al.*, 2002;

Pertschy *et al.*, 2007; Thomson & Tollervey, 2010). Enfin, la dernière particule pré-60S détectée (Kre35p, Lsg1p) contient toujours le facteur Nmd3p, mais présente une quantité réduite de Arx1p (Nissan *et al.*, 2002; Pertschy *et al.*, 2007; Ulbrich *et al.*, 2009). Le recyclage du facteur d'export Nmd3p constitue l'une des dernières étapes de l'assemblage de la particule 60S, il implique la GTPase Lsg1p dans un mécanisme couplé à l'incorporation de la protéine ribosomique Rpl10p (alias Qsr1p), requise pour l'association avec la sous-unité 40S (Dick *et al.*, 1997a; Dick *et al.*, 1997b; Eisinger *et al.*, 1997; Hedges *et al.*, 2005).

### 3.4.3. Assemblage des ribosomes et consommation d'énergie

En dépit de la grande conservation structurale et fonctionnelle des ribosomes, leur processus de biogenèse présente des différences notoires entre bactéries et eucaryotes. Chez les bactéries, ce processus qui peut être reproduit *in vitro* à partir d'ARNr et de protéines ribosomiques purifiées, implique relativement peu de facteurs non ribosomiques (Hage & Tollervey, 2004). Comme nous venons de le détailler, l'assemblage des ribosomes eucaryotes nécessite quant à lui l'intervention d'un grand nombre de ces facteurs et s'effectue au sein de plusieurs compartiments cellulaires. Les protéines ribosomiques, synthétisées dans le cytoplasme, doivent être, pour la plupart, importées dans le noyau pour être assemblées de façon séquentielle sur les ARNr en cours de maturation. La hiérarchisation de ce processus nécessite des événements de régulation permettant d'empêcher des associations trop précoces, de favoriser d'autres interactions ou à l'inverse, de dissocier certains complexes. En plus des coûts énergétiques liés au transport, l'ensemble de ces mécanismes de remodelage requière une grande consommation d'énergie et implique de nombreuses enzymes NTP-dépendantes (Kressler *et al.*, 2010; Strunk & Karbstein, 2009). Parmi ces facteurs on retrouve notamment des protéines kinases, des AAA<sup>+</sup>-ATPase, des GTPases, ainsi que de nombreuses ARN-hélicases de la famille des DExD/H-ATPases. De façon générale, les DExD/H-ATPases interviennent dans les étapes précoces du processus d'assemblage, tandis que les AAA<sup>+</sup>-ATPase et les GTPases sont essentiellement impliquées dans les étapes tardives de la biogenèse de la particule 60S. A titre d'exemple, le processus de relargage des snoRNP associées aux pré-ARNr fait intervenir certaines DExD/H-ATPases. Leur mutation conduit en effet à la rétention des snoRNA au sein des structures pré-ribosomiques et provoque généralement l'inhibition des coupures aux sites A0, A1 et A2 du pré-ARNr (Strunk & Karbstein, 2009). Parmi ces enzymes, on retrouve notamment le facteur Dbp4p, impliqué



dans la dissociation des snoRNA U14 et snR41 du pré-ARNr (Kos & Tollervey, 2005), la protéine Rok1p, qui joue un rôle essentiel dans le relargage du snoRNA snR30 (Bohnsack *et al.*, 2008), ainsi que les facteurs Has1p et Prp43p, également impliqués dans la dissociation de nombreux snoRNA (Bohnsack *et al.*, 2009; Liang & Fournier, 2006).

### **3.5. Ribosomopathie et défauts de la biogenèse des ribosomes**

Etant donné le rôle fondamental de la machinerie de traduction, la préservation de l'intégrité du processus de biogenèse des ribosomes apparaît essentielle pour la croissance et la survie cellulaire. Chez l'homme, des anomalies génétiques, à l'origine de défauts de production et de fonction des ribosomes, ont été identifiées et regroupées sous le terme de ribosomopathies (Freed *et al.*, 2010; Narla & Ebert, 2010). La plupart de ces maladies sont associées à des anomalies médullaires caractérisées par une diminution de la quantité de cellules sanguines et une prédisposition au cancer. L'impact d'un dysfonctionnement ribosomique sur les hématies est généralement important car avant la perte de leur noyau, ces cellules se divisent de façon active et synthétisent une grande quantité de ribosomes. Plusieurs mutations, recensées dans des gènes codant des protéines ribosomiques et des facteurs de biogenèse des ribosomes, ont ainsi été corrélées à différentes pathologies, dont l'anémie de Diamond-Blackfan, le syndrome de Shwachman-Diamond, le syndrome de Bowen-Conradi, l'hypoplasie du cartilage et des cheveux, ou encore la dyskératose congénitale.

Chez 30 à 40% des patients atteints d'anémie de Diamond-Blackfan, des mutations ont en effet été retrouvées dans les gènes RPL5, RPL11, RPL36, RPS15, RPS17, RPS19, RPS24 et/ou RPS27A (Choesmel *et al.*, 2008; Ganapathi *et al.*, 2007; Gazda *et al.*, 2008; Robledo *et al.*, 2008). Le syndrome de Shwachman-Diamond a été associé à une mutation de la protéine SBDS, dont l'homologue chez la levure (Sdo1p), est impliqué dans la dissociation cytoplasmique du facteur Tif6p de la sous-unité pré-60S (Menne *et al.*, 2007). Le syndrome de Bowen-Conradi résulte de mutations de la protéine EMG1 (Emg1p chez la levure), impliquée dans la biogenèse de la sous-unité 40S (Armistead *et al.*, 2009) L'hypoplasie du cartilage et des cheveux a été associée à des mutations survenant dans le snoRNA de la particule MRP (Martin & Li, 2007), tandis que la dyskératose congénitale peut impliquer des mutations dans les gènes codant les protéines Dyskérine (alias CBF5), NOP10 et NHP2,

constitutives des snoRNP à boîtes H/ACA (Marrone *et al.*, 2005; Vulliamy *et al.*, 2008; Walne *et al.*, 2007).

Néanmoins, dans ces deux derniers exemples, la multiplicité des rôles joués par ces facteurs ne permet pas de corréler de façon stricte le développement de la pathologie à un défaut de biogenèse des ribosomes. La particule MRP (*Mitochondrial RNA Processing*) est en effet impliquée dans la régulation du cycle cellulaire par le clivage de l'ARNm de la cycline B2 (Gill *et al.*, 2004), ainsi que dans le processus de réplication de l'ADN mitochondrial (Chang & Clayton, 1989; Lu *et al.*, 2011). Les protéines CBF5, NOP10 et NHP2 font, quant à elles, également partie du complexe de la télomérase et la découverte de mutations au niveau des composants TERT et TERC chez des patients atteints de dyskératose congénitale, suggère qu'un défaut de réplication des télomères pourrait être à l'origine du développement de la maladie (Armanios *et al.*, 2005; Marrone *et al.*, 2007; Trahan & Dragon, 2009; Vulliamy *et al.*, 2001). De plus, étant donné l'implication des particules H/ACA (snoRNP et scaRNP) dans la maturation des UsnRNP spliceosomales, des défauts touchant la machinerie d'épissage pourraient également constituer des déterminants liés à la survenue de cette pathologie.

Enfin, le syndrome de Prader-Willi, un autre type de maladie complexe, semble également impliquer un défaut de la machinerie de maturation des ribosomes. Cette pathologie, caractérisée par de lourds problèmes de développement, résulte dans une majorité de cas d'une délétion au niveau du *locus* q11-q13 du chromosome 15, soumis au phénomène d'empreintes génomiques. Relativement peu abondante en gènes codant des protéines, cette région du génome héberge de nombreux snoRNA à boîtes C/D (Cavaille *et al.*, 2000). Une étude approfondie a notamment montré qu'une perte d'expression des snoRNA du *cluster* de répétitions HBII-85 est responsable de la majorité des symptômes observés chez les patients (Sahoo *et al.*, 2008). En revanche les séquences ARN ciblées par ces snoRNA n'ont à ce jour pas encore été identifiées.

## CHAPITRE 2

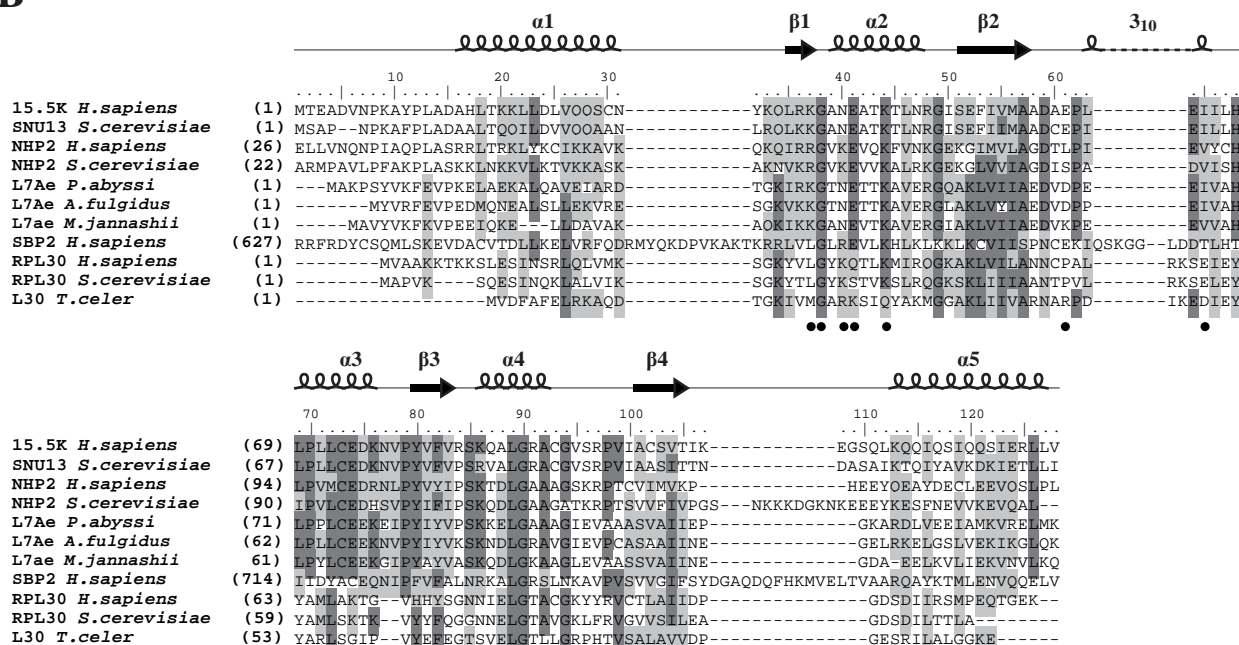
### Les RNP de la famille L7Ae

Les protéines de la famille L7Ae/L30, retrouvées dans l'ensemble du monde vivant, sont des composants essentiels de nombreuses RNP (**Fig. 21A**). Cette famille se définit par le partage d'un domaine commun de fixation à l'ARN, appelé L7Ae/L30-RBD (*L7Ae/L30 RNA Binding Domain*) (**Fig. 21B**), responsable de l'interaction avec des motifs de type *K-turn* ou *K-loop* présents dans de nombreux ARN (Allmang *et al.*, 2002; Koonin *et al.*, 1994; Kuhn *et al.*, 2002; Nolivovs *et al.*, 2005; Nottrott *et al.*, 1999). Cette famille, qui inclut notamment la protéine ribosomique RPS12 et le facteur d'élongation de la traduction eRF1, a été baptisée en référence aux protéines L7Ae et L30 (alias RPL7A et RPL30), qui sont retrouvées au sein de la grande sous-unité ribosomique (Ban *et al.*, 2000; Koonin *et al.*, 1994; Mager *et al.*, 1997). En plus de sa présence au sein du ribosome, la protéine L7Ae d'archée est également un constituant majeur des sRNP à boîtes H/ACA et C/D (Kuhn *et al.*, 2002; Rozhdestvensky *et al.*, 2003). Chez les eucaryotes, cette fonction est en revanche assurée par deux protéines distinctes de cette famille : la protéine NHP2, constituante des snoRNP à boîtes H/ACA (Henras *et al.*, 1998) et du complexe télomérase des vertébrés (Mitchell *et al.*, 1999a), et la protéine 15.5K (Snu13p chez la levure) constituante des snoRNP à boîtes C/D (Watkins *et al.*, 2000). De plus, la protéine Snu13p/15.5K eucaryote présente également une double fonctionnalité puisqu'elle avait initialement été identifiée en tant que facteur d'épissage appartenant à la snRNP U4 (Gottschalk *et al.*, 1999; Nottrott *et al.*, 1999; Stevens & Abelson, 1999). La protéine SBP2, impliquée dans la synthèse des sélénoprotéines chez certains eucaryotes, appartient elle aussi à cette famille (Allmang *et al.*, 2002; Copeland *et al.*, 2000). Enfin, plus récemment, la protéine L7Ae a également été identifiée comme un composant de la RNase P chez l'archée *Methanococcus maripaludis* (Cho *et al.*, 2010). Les RNP contenant les protéines de la famille L7Ae/L30 exercent donc de nombreux rôles et présentent des liens phylogénétiques étroits. Dans ce chapitre, nous nous intéresserons tout particulièrement à la sous-famille L7Ae, qui se distingue des facteurs affiliés à la protéine L30, et qui, en plus de la protéine L7Ae d'archée, comprend notamment les protéines eucaryotes Snu13p/15.5K et

A

Homologies fonctionnelles au sein de la famille L7Ae/L30 chez différents organismes				
Archées	<i>S. cerevisiae</i>	<i>H. sapiens</i>	RNP	Fonction
	Snu13p	15.5K	snRNP U4	Epissage
L7Ae	Snu13p	15.5K	sRNP - snoRNP C/D	Biogenèse des ribosomes
L7Ae	Nhp2p	NHP2	sRNP - snoRNP H/ACA	Biogenèse des ribosomes
L7Ae	Rpl7p	RPL7	grande sous-unité ribosomique	Traduction
L30	Rpl30p	RPL30	grande sous-unité ribosomique	Traduction
Rps12	Rps12p	Rps12	petite sous-unité ribosomique	Traduction
	eRF1	eRF1	ribosome en cours d'élongation	Elongation de la traduction
		SBP2	mRNP sélénocysteine	Traduction des sélénoprotéines

B



**Figure 21 : Homologies fonctionnelles et homologies de séquences au sein de la famille de protéines L7Ae/L30.** (A) Tableau résumant les principales fonctions assurées par des protéines de la famille L7Ae/L30 chez les archées et chez les eucaryotes. (B) Alignement de séquences des domaines RBD (RNA Binding Domain) de la famille L7Ae/L30 réalisé par ClustalW Multiple Alignment à l'aide du logiciel BioEdit (Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98). Les acides aminés identiques dans plus de 50% des séquences alignées sont représentés sur fond gris foncé. Les acides aminés similaires dans plus de 50% des séquences alignées sont représentés sur fond gris clair. Les éléments de structures secondaires reportés au dessus de l'alignement correspondent à ceux de la structure de la protéine humaine 15.5K en complexe avec le K-turn du snRNA U4 (Vidovic et al., 2000; PDB : 1E7K). Les pastilles noires indiquent les principaux acides aminés impliqués dans l'interaction avec l'ARN et qui font l'objet d'une discussion dans ce chapitre (cf. Figures 24 à 27).

Nhp2p/NHP2. Nous aborderons en détail la composition et la fonction des RNP contenant ces protéines : la snRNP U4, ainsi que les snoRNP à boîtes C/D et H/ACA.

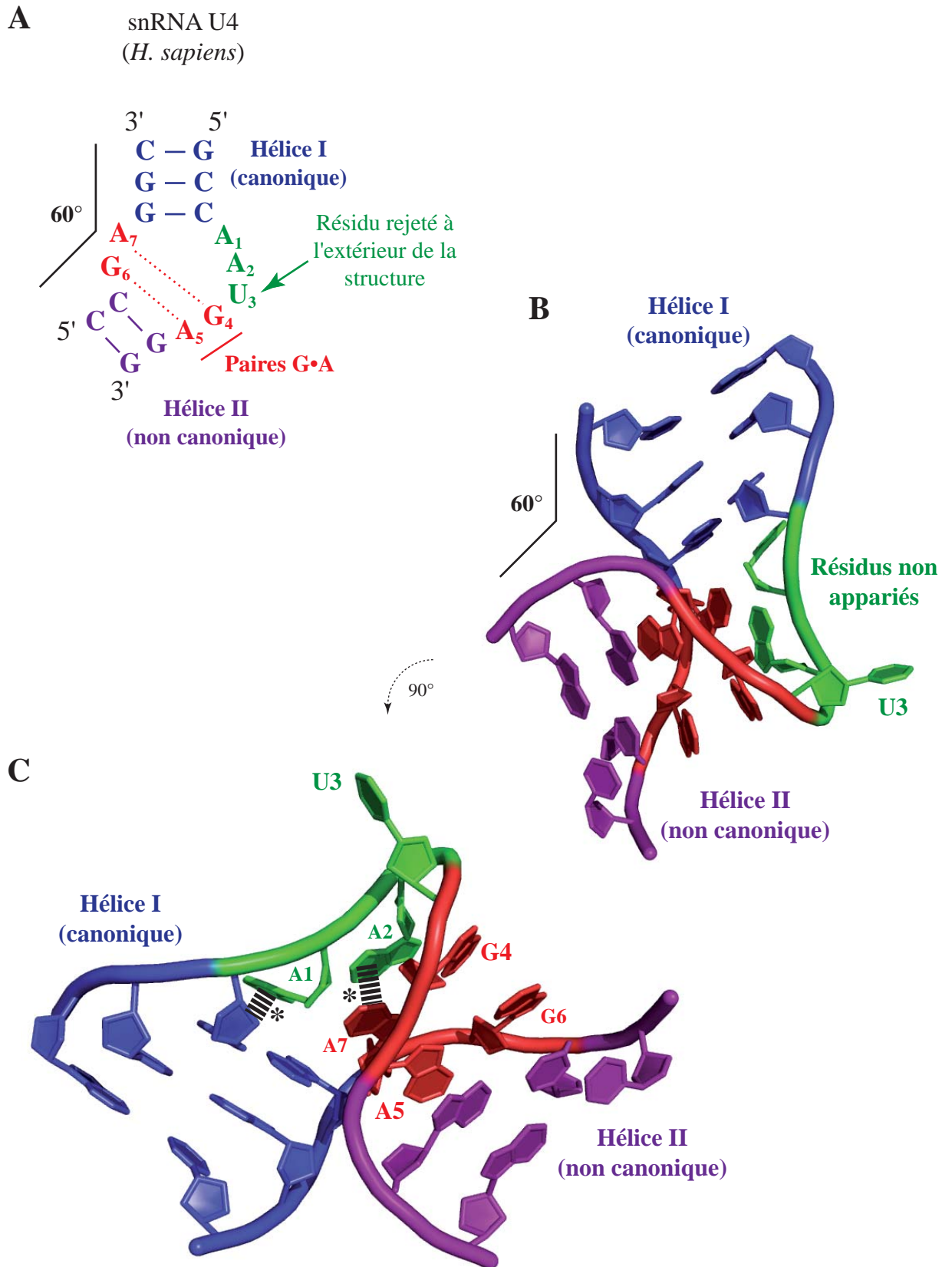
## 1. Les protéines de la famille L7Ae/L30

La formation d'un complexe entre un motif ARN structuré en *K-turn* et une protéine de la famille L7Ae/L30 constitue souvent un événement déterminant pour la formation de la RNP. Depuis le début des années 2000, de nombreuses études structurales ont permis de caractériser les bases moléculaires de ces interactions ARN-protéine.

### 1.1. Les motifs ARN en *K-turn* et en *K-loop*

Les ARN adoptent des structures contenant des hélices de type A et de nombreuses régions non hélicoïdales. Ces dernières, d'apparence moins structurées, forment néanmoins des structures 3D bien définies, contenant une information spatiale souvent essentielle à la fonction de l'ARN. De nombreux motifs retrouvés de façon récurrente dans les ARN ont ainsi été décrits, c'est notamment le cas des *tetraloops*, des *A-platforms*, des *A-turns*, des *S-turns* ou encore des *K-turns*. Ces derniers, également appelés motifs G•A, ont été initialement décrits par l'équipe de T. A. Steitz (Klein *et al.*, 2001), sur la base de l'analyse structurale des sous-unités ribosomiques de *Haloarcula marismortui* et de *Thermus thermophilus* (Ban *et al.*, 2000; Wimberly *et al.*, 2000).

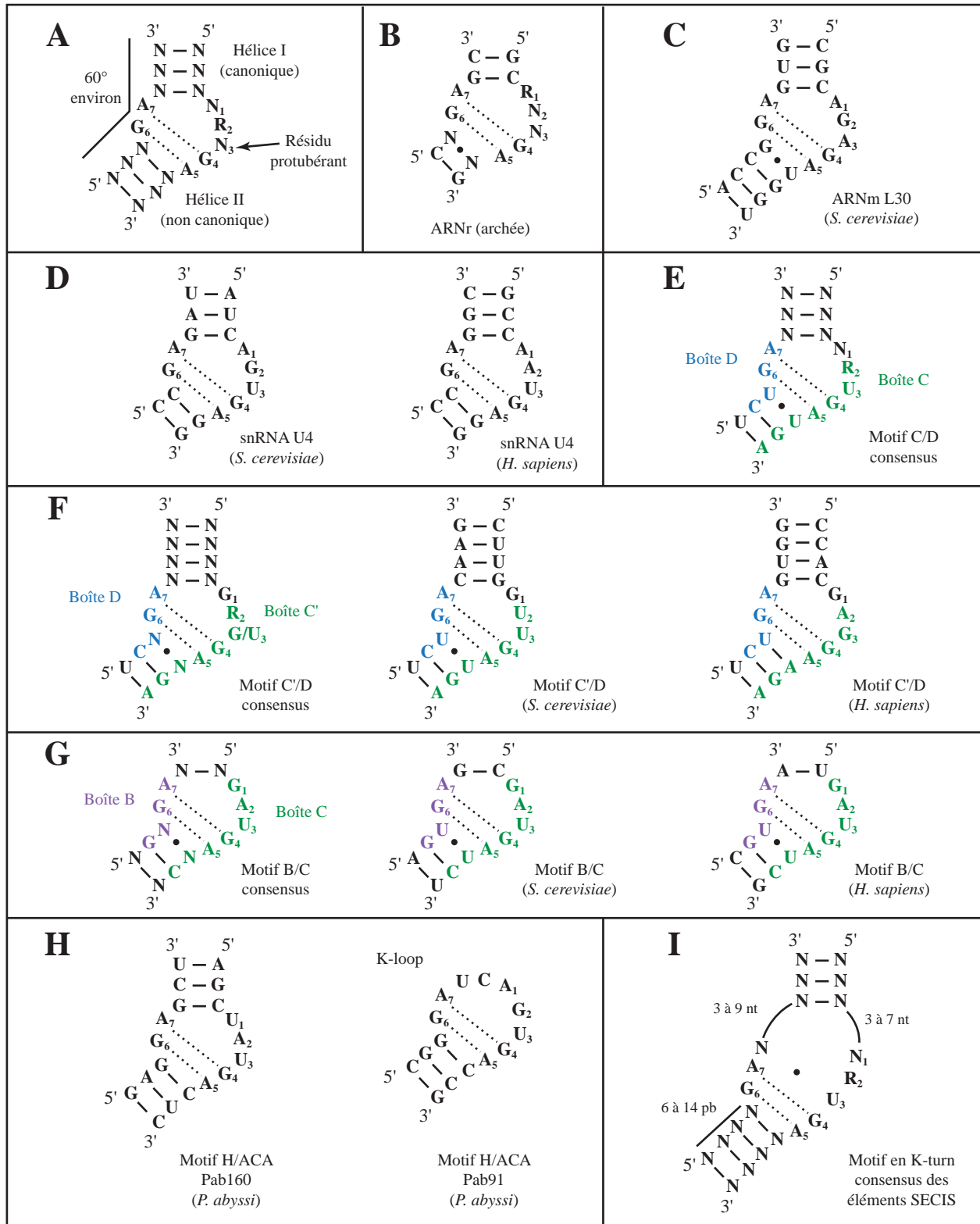
Un *K-turn* est un motif organisé en hélice-boucle-hélice comprenant environ 15 nucléotides (**Fig. 22A**). La boucle interne est formée de 3 résidus non appariés (aux positions 1, 2 et 3), situés sur le même brin, et de deux paires de bases non canoniques consécutives G•A et A•G. Cette boucle asymétrique est flanquée, en 5' des résidus non appariés, par une tige canonique (= hélice I) composée d'au moins 2 paires de bases G-C, tandis que les paires de bases non Watson-Crick G•A et A•G donnent naissance, côté 3', à une tige dite non canonique (= hélice II). Cet agencement particulier induit une forte courbure du squelette phosphodiester, avec un petit angle d'environ 60° mesuré entre les tiges, à l'origine du nom de motif en *kink-turn* ou *K-turn* (**Fig. 22B**). Cette conformation est stabilisée par des forces de *stacking* : l'empilement



**Figure 22 : Structure du motif en K-turn du snRNA U4 humain.** (A) Structure secondaire du motif en K-turn du snRNA U4 humain. (B, C) Structure tridimensionnelle du motif en K-turn du snRNA U4 humain au sein du complexe avec la protéine 15.5K (d'après Vidovic *et al.*, 2000 ; PDB : 1E7K). (\*) Empilement des résidus non appariés, avec plateaux de base des hélices I et II.

des paires G•A et A•G avec le restant de l'hélice II, ainsi que l'empilement des deux premiers résidus non appariés (en position 1 et 2), respectivement avec l'hélice I et avec l'hélice II (**Fig. 22C**). Seul le troisième résidu non apparié n'est pas impliqué dans une interaction de *stacking*. Ce dernier, situé à l'apex du grand angle, est projeté à l'extérieur de la structure dans une position totalement exposée au solvant et constitue un déterminant d'interaction privilégié avec des facteurs protéiques. Il a également été montré que 5 des 7 résidus de la boucle interne adoptent une conformation atypique 2'-endo au niveau du ribose (Moore *et al.*, 2004). Cette transition de la forme classique 3'-endo à la forme 2'-endo conduit à une exposition maximale des résidus localisés dans le sillon majeur. Plusieurs études ont par ailleurs démontré que la formation de la structure en *K-turn* est largement dépendante de l'environnement chimique et de l'interaction avec les partenaires protéiques (Cojocaru *et al.*, 2005; Falb *et al.*, 2010; Goody *et al.*, 2004; Matsumura *et al.*, 2003; Razga *et al.*, 2004; Turner *et al.*, 2005). Il apparaît en effet que sous sa forme libre, ce type de motif ARN est sujet à un échange dynamique entre la forme en *K-turn* et une forme plus relâchée. Le passage à la forme en *K-turn* est largement favorisé par la présence de cations métalliques et le motif adopte un état stable lors de la fixation d'un partenaire protéique. On parle de phénomène d'*induced-fit*.

Différents motifs en *K-turn* ont ainsi été identifiés (**Fig. 23**) : au sein des ARNr, dans les sRNA à boîtes H/ACA d'archée, dans les sRNA à boîtes C/D d'archée et dans les snoRNA à boîtes C/D eucaryotes, où ils sont formés par l'association des boîtes conservées C et D, ou encore au sein des éléments SECIS (*Selenocystein Insertion Sequence*), situés dans la région 3'-UTR des ARNm codant les sélénoprotéines chez les vertébrés. De plus, des motifs en *K-turn* ont également été trouvés au sein de certains éléments ribocommutateurs (*riboswitchs*) (Montange & Batey, 2006). Les motifs de type *K-turn* sont donc retrouvés dans l'ensemble du monde vivant. Comme nous l'aborderons ensuite, leur interaction avec des membres de la famille L7Ae/L30 a fait l'objet de nombreuses études et constitue l'exemple le plus détaillé. On remarquera néanmoins que leur spectre d'interaction ne se restreint pas aux protéines de cette famille et que d'autres partenaires potentiels ont été décrits. C'est par exemple le cas de la protéine hnRNPA2 de mammifères, qui semble capable de lier des motifs en *K-turn* présents dans certains ARNm, afin d'assurer leur transport dans les cellules neuronales (Munro *et al.*, 1999; Shan *et al.*, 2003; Tiedge, 2006).



**Figure 23 : Structures secondaires des motifs de type K-turn retrouvés dans différents ARN.** (A) Motif en K-turn consensus. (B) K-turn consensus défini pour les ARNr d'archées (d'après Klein *et al.*, 2001). (C) K-turn de l'ARNm codant la protéine L30 chez *S. cerevisiae* (d'après Li *et al.*, 1997). L30 interagit avec ce motif afin de réguler sa propre expression. (D) Motifs en K-turn du snRNA U4 chez *S. cerevisiae* et chez l'homme (d'après Mougin *et al.*, 2002). (E) K-turn consensus défini pour les snoRNA à boîtes C/D (d'après Marmier-Gourrier *et al.*, 2003). (F) K-turn consensus défini pour les motifs U3-C'/D, motif C'/D du snoRNA U3 chez *S. cerevisiae* et chez l'homme (d'après Marmier-Gourrier *et al.*, 2003). (G) K-turn consensus défini pour les motifs U3-B/C, motif B/C du snoRNA U3 chez *S. cerevisiae* et chez l'homme (d'après Marmier-Gourrier *et al.*, 2003). (H) Exemple de K-turn et K-loop retrouvés respectivement dans les sRNA à boîtes H/ACA Pab160 et Pab91 chez l'archée *P. abyssi* (d'après Muller *et al.*, 2008). (I) Motif en K-turn consensus retrouvé dans les éléments SECIS (d'après Allmang *et al.*, 2002).



Plus récemment, une forme variante de motif en *K-turn*, baptisé *K-loop*, a été décrite (Li & Ye, 2006; Nolivos *et al.*, 2005). Ce type de structure est essentiellement retrouvé dans les sRNA à boîtes C/D et H/ACA d'archée. Un *K-loop* est un motif structuralement similaire au *K-turn*, dans lequel la tige canonique est remplacée par une courte boucle terminale, formée généralement par les 3 nucléotides non appariés (**Fig. 23H**). Des expériences de retard sur gel, réalisées avec la protéine L7Ae, ont par ailleurs montré que la taille et la séquence de cette boucle terminale constituent des paramètres déterminants pour la reconnaissance par les partenaires protéiques (Nolivos *et al.*, 2005). Enfin, il a été proposé que l'abondance de ce type de motifs chez les archées hyperthermophiles de type *Pyrococcus*, soit le résultat d'un processus d'adaptation aux températures extrêmes (+100°C). En effet, bien que l'augmentation de la quantité d'appariement G-C soit largement privilégiée, la circularisation d'un ARN peut également contribuer à stabiliser sa structure (Galtier & Lobry, 1997). Ainsi, il a été proposé que l'abondance des structures en *K-loop* dans les sRNA de ces espèces constitue une adaptation supplémentaire dans le cadre de la résistance à la dénaturation des ARN par la chaleur (Nolivos *et al.*, 2005).

## 1.2. Description de différentes classes de motifs en *K-turn*

### 1.2.1. Les motifs en *K-turn* dans les ARNr et les ARNm

Au sein des ARNr, il existe de multiples *K-turn* qui constituent des sites de fixation pour de nombreuses protéines (**Fig. 23B**). A titre d'exemple, dans la grande sous-unité ribosomique de l'archée *H. marismortui*, 9 des 31 protéines ribosomiques (L4, L7Ae, L10, L15e, L19e, L24, L29, L32 et L37Ae) sont en interaction avec un motif en *K-turn* (Ban *et al.*, 2000; Klein *et al.*, 2001). L'analyse détaillée de ces motifs a révélé une grande diversité dans leur séquence et leur structure (Klein *et al.*, 2001). Bien que l'on retrouve généralement 3 résidus non appariés dans la boucle interne, ce nombre peut varier de 2 à 4, et il est même possible de compter un résidu non apparié supplémentaire de l'autre côté de la boucle. De la même façon, la nature du résidu protubérant à l'apex du *K-turn* varie également en fonction du partenaire protéique.

De tels motifs ont également été retrouvés dans des ARNm. C'est notamment le cas du transcrit codant la protéine L30, qui contient, dans son exon 1, un motif en *K-turn* situé à

proximité d'un site 5' d'épissage (**Fig. 23C**). Le recrutement spécifique de la protéine L30 sur ce motif conduit à la séquestration du site d'épissage au sein d'une structure stable et inhibe l'épissage du pré-ARNm (Chao & Williamson, 2004; Eng & Warner, 1991; Vilardell & Warner, 1994). Ce processus constitue ainsi un système de rétro-contrôle négatif permettant à la protéine L30 de réguler sa propre expression lorsqu'elle est présente en trop grande quantité dans la cellule.

### 1.2.2. Le motif en *K-turn* du snRNA U4

Le snRNA U4, retrouvé uniquement chez les eucaryotes, a pour particularité de présenter un motif en *K-turn* dans sa région 5' (Mougin *et al.*, 2002; Vidovic *et al.*, 2000) (**Fig. 23D**). Ce motif est caractérisé par la présence d'une courte hélice II, composée uniquement de deux appariements G-C. Une purine est également retrouvée en seconde position du *K-turn*. Enfin, le nucléotide protubérant en position 3 est une uridine. Comme nous le verrons par la suite, ce résidu est impliqué dans l'interaction avec la protéine Snu13p/15.5K (Nottrott *et al.*, 1999; Vidovic *et al.*, 2000).

### 1.2.3. Les motifs en *K-turn* des sRNA et des snoRNA à boîtes C/D

Chez les archées (sRNA), comme chez les eucaryotes (snoRNA), les ARN à boîtes C/D sont caractérisés par la présence de séquences conservées, appelées boîtes C et D, dont l'appariement forme un motif de type C/D qui adopte une structuration en *K-turn* (Baserga *et al.*, 1991; Caffarelli *et al.*, 1996; Huang *et al.*, 1992; Kuhn *et al.*, 2002; Marmier-Gourrier *et al.*, 2003; Watkins *et al.*, 1996; Watkins *et al.*, 2000) (**Fig. 23E**). La boîte C est définie par la séquence consensus 5'-RUGAUGA-3', tandis que la boîte D est définie par la séquence 5'-CUGA-3'. Chez les archées, ce motif permet le recrutement de la protéine L7Ae (Kuhn *et al.*, 2002), tandis que chez les eucaryotes il s'associe avec la protéine orthologue Snu13p/15.5K (Watkins *et al.*, 2002). Comme c'est le cas pour le snRNA U4, le résidu retrouvé de façon préférentielle en position 3 du *K-turn* (= second nucléotide de la boîte C) est donc une uridine. En revanche, contrairement au snRNA U4, les motifs C/D sont bordés par une hélice II plus longue, comportant 3 appariements, dont une paire de bases U•U non canonique empilée avec

les paires G•A, suivie d'une paire de bases C-G. En plus des boîtes C et D, certains snoRNA contiennent des séquences similaires moins bien conservées, appelées boîtes C' et D'. La distance séparant les séquences C' et D' est généralement comprise entre 3 et 12 nucléotides, ce qui peut, en fonction des cas, conduire soit à la formation d'une hélice I, soit à la formation d'un motif en *K-loop* (Kiss-Laszlo *et al.*, 1998; Kuhn *et al.*, 2002; Nolivos *et al.*, 2005).

#### 1.2.4. Les motifs en *K-turn* particuliers du snoRNA U3

Chez les eucaryotes, le snoRNA U3 présente une configuration particulière puisqu'il comporte deux motifs en *K-turn* atypiques : un motif C'/D qui correspond à un motif C/D non canonique, ainsi qu'un motif B/C qui est spécifique de ce snoRNA (Granneman *et al.*, 2002; Marmier-Gourrier *et al.*, 2003; Watkins *et al.*, 2000).

Le motif C'/D, qui remplit la même fonction que les motifs C/D classiques, fait l'objet, en fonction des organismes, de certaines déviations par rapport au consensus établi pour la boîte C (Marmier-Gourrier *et al.*, 2003) (**Fig. 23F**). Chez les levures à fission et chez les plantes, le premier résidu de la boîte C, qui correspond au second résidu non apparié du *K-turn*, est remplacé par un U. Chez les mammifères et chez les amphibiens, c'est en revanche le résidu protubérant du *K-turn* qui diverge du consensus, puisque le résidu U est remplacé par un G. D'autre part, la paire de bases U•U de l'hélice II peut être remplacée par une paire canonique AU ou UA. Enfin, alors que la position 1 du *K-turn* de type C/D est généralement occupée par un résidu variable, celle des motifs C'/D est constamment occupée par un G. Comme nous le verrons par la suite, certaines de ces spécificités peuvent diminuer l'affinité apparente des protéines Snu13p et 15.5K pour le motif C'/D.

Le motif B/C est quant à lui composé d'une boîte B de consensus 5'-GNGA-3' et d'une boîte C de consensus 5'-GAUGANC-3' (**Fig. 23G**). Il est généralement caractérisé par une hélice I très courte, comportant seulement un ou deux appariements. Dans l'hélice II, on retrouve en général une bonne conservation de l'appariement non canonique U•U, ainsi qu'une forte conservation d'une paire de base G-C, inversée par rapport à la configuration retrouvée dans les autres motifs en *K-turn*.

### 1.2.5. Les motifs en *K-turn* des sRNA à boîtes H/ACA d'archée

A la différence des snoRNA à boîtes H/ACA eucaryotes, qui recrutent la protéine NHP2, les sRNA à boîtes H/ACA d'archées, qui s'associent avec la protéine L7Ae, contiennent des motifs de type *K-turn* ou *K-loop* (Muller *et al.*, 2008; Rozhdestvensky *et al.*, 2003) (**Fig. 23H**). Etant donné leur grande diversité, aucun consensus n'a réellement été défini à ce jour pour ces motifs. Conformément aux spécificités d'interaction de la protéine L7Ae, on remarquera néanmoins la présence d'une uridine conservée au niveau du résidu protubérant.

### 1.2.6. Les motifs en *K-turn* des éléments SECIS

Comme nous l'avons déjà mentionné, le mécanisme d'incorporation des sélénocystéines au cours de la traduction implique un élément de séquence conservé appelé élément SECIS. Chez la majorité des eucaryotes et à l'exception notoire de la levure *S. cerevisiae*, cet élément forme, dans la région 3'-UTR des ARNm, une structure en tige-boucle qui contient, dans sa partie centrale, un motif de type *K-turn* permettant le recrutement de la protéine SBP2 (Allmang *et al.*, 2002; Copeland *et al.*, 2000; Fletcher *et al.*, 2001) (**Fig. 23I**). Bien qu'apparenté aux motifs précédemment décrits, l'élément responsable de la fixation de SBP2 présente des spécificités notoires (Allmang *et al.*, 2002). Les deux paires G•A sont en effet encadrées par deux autres appariements non canoniques et les résidus non appariés, présents de part et d'autre de la boucle interne, peuvent largement varier en nombre. Malgré ces différences, une purine et un résidu uridine sont néanmoins retrouvés en positions 2 et 3 du *K-turn*.

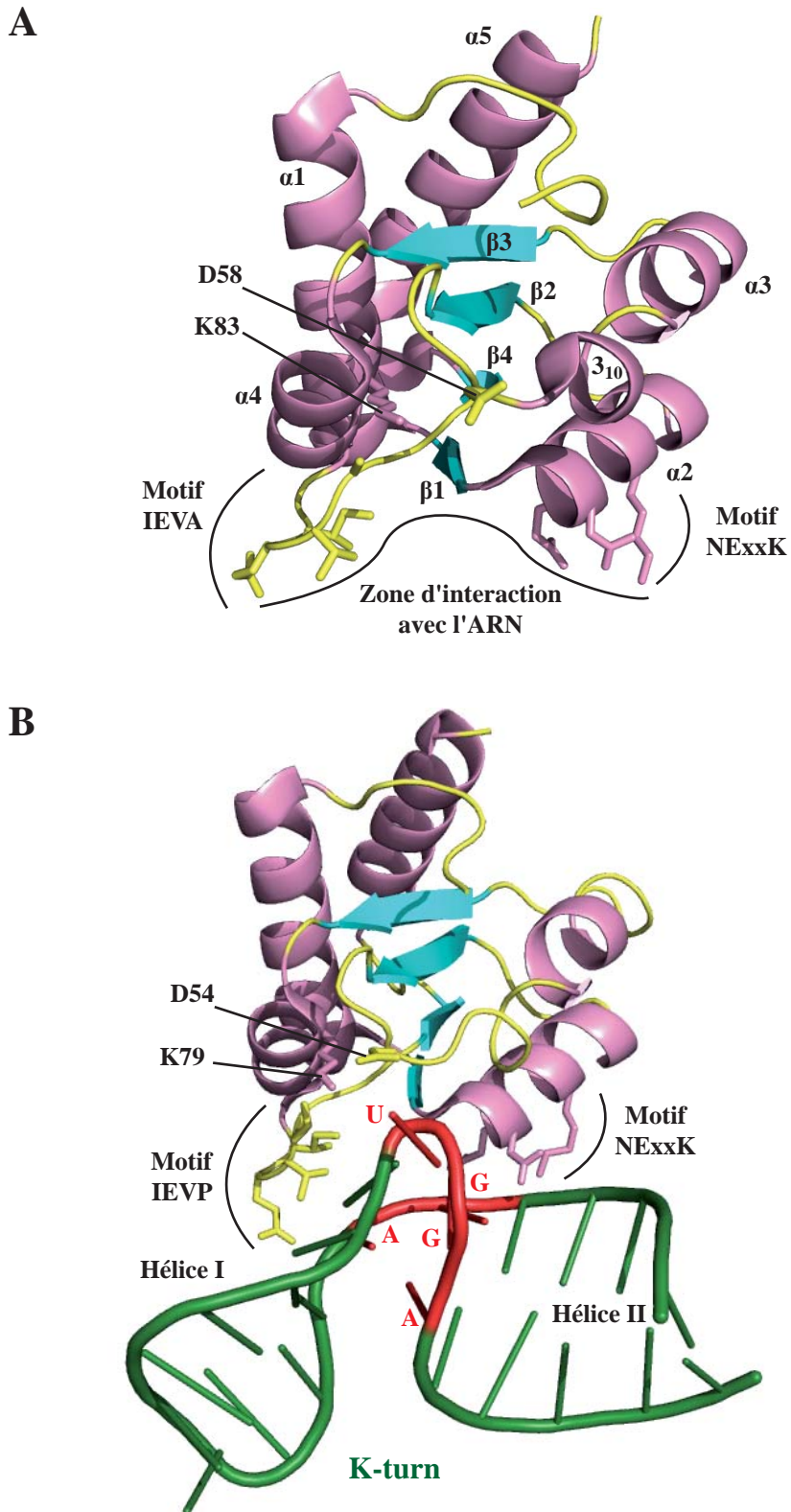
## 1.3. Le domaine L7Ae/L30-RBD et ses propriétés d'interaction avec l'ARN

Initialement, la famille L7Ae/L30 a été définie d'après un alignement de séquences protéiques de bactéries, d'archées et d'eucaryotes (Koonin *et al.*, 1994). L'homologie de séquence relevée sur environ 50 acides aminés incluait notamment les protéines ribosomiques L30, L7Ae et la protéine Nhp2p de levure. Par la suite, les facteurs Snu13p et 15.5K (Kuhn *et al.*,

2002), ainsi que la protéine SBP2 de mammifère (Allmang *et al.*, 2002; Copeland *et al.*, 2000) ont été intégrés à cette famille et le domaine d'homologie a été étendu à une centaine de résidus (**Fig. 21**). Pour certains de ces facteurs, le domaine L7Ae/L30-RBD représente la quasi-totalité de la protéine (ex : L7Ae, Snu13p, 15.5K), en revanche, pour d'autres, il ne constitue qu'un domaine intégré dans un polypeptide de plus grande taille (ex : SBP2). Un grand nombre de données structurales, issues d'études par cristallographie ou par RMN, sont disponibles pour les protéines de cette famille. Il existe notamment plusieurs structures des protéines L30 d'archée (Chen *et al.*, 2003; Fedorov *et al.*, 1999) et de levure (Chao & Williamson, 2004; Mao *et al.*, 1999), et de nombreuses structures de la protéine L7Ae d'archée obtenues sous forme libre (Charron *et al.*, 2004; Suryadi *et al.*, 2005), en complexe avec un motif ARN (Hamma & Ferre-D'Amare, 2004; Moore *et al.*, 2004) ou intégrée au sein d'une RNP (Ban *et al.*, 2000; Li & Ye, 2006; Liang *et al.*, 2007a; Liang *et al.*, 2009a; Ye *et al.*, 2009). Enfin, la structure des protéines Snu13p et Nhp2p de levure à l'état libre (Koo *et al.*, 2011; Oruganti *et al.*, 2005), ainsi que la structure de plusieurs complexes ARN/15.5K ont également été résolues (Biswas *et al.*, 2011; Vidovic *et al.*, 2000). D'après ces différentes études structurales le domaine L7Ae/L30-RBD est composé d'une alternance d'hélices  $\alpha$  et de feuillets  $\beta$ , formant une structure en sandwich  $\alpha$ - $\beta$ - $\alpha$ . L'exploitation de l'ensemble de ces données a également permis d'étudier les propriétés d'interaction des différents domaines L7Ae/L30-RBD avec l'ARN, afin de mieux comprendre les caractéristiques générales de cette famille, mais aussi les spécificités propres à chacun de ses membres.

### 1.3.1. La protéine L7Ae d'archée

L7Ae est une protéine d'environ 13 kD, capable d'interagir aussi bien avec les motifs en *K-turn* qu'avec les motifs en *K-loop* retrouvés dans les sRNA à boîtes C/D et H/ACA, ainsi que dans les ARNr (Charron *et al.*, 2004; Gagnon *et al.*, 2010; Kuhn *et al.*, 2002; Tran *et al.*, 2003). Sa structure tridimensionnelle est composée d'un feuillet  $\beta$  central, comprenant de 3 brins antiparallèles et un brin parallèle, bordé par 3 hélices  $\alpha$  d'un côté et par deux hélices  $\alpha$  de l'autre (Charron *et al.*, 2004) (**Fig. 24A**). La poche d'interaction avec l'ARN est formée par une région électropositive composée de l'hélice  $\alpha_2$ , du brin  $\beta_1$  et des boucles  $\beta_1$ - $\alpha_2$  et  $\alpha_4$ - $\beta_5$  (Hamma & Ferre-D'Amare, 2004; Moore *et al.*, 2004). La résolution d'un complexe entre L7Ae de *Archaeoglobus fulgidus* et un motif C/D structuré en *K-turn* a permis de montrer que



**Figure 24 : Structures tridimensionnelles de la protéine L7Ae libre ou en complexe avec un ARN.** (A) Structure de la protéine L7Ae de *P. abyssi* établie par radiocristallographie (Charron *et al.*, 2004; PDB : 1PXW). (B) Structure de la protéine L7Ae de *Archaeoglobus fulgidus* en complexe avec un motif C/D en K-turn établie par radiocristallographie (Moore *et al.*, 2004; PDB : 1RLG).  $\alpha$  : hélice  $\alpha$ ;  $\beta$  : feuillet  $\beta$ ;  $\eta$  : hélice  $3_{10}$ . Les paires de bases G•A et le résidu U protubérant du K-turn sont représentés en rouge.

la protéine établit des contacts primaires avec les squelettes phosphate de 6 résidus du motif, dont 5 sont situés au niveau de la boucle interne (Moore *et al.*, 2004) (**Fig. 24B**). Le résidu protubérant du motif en *K-turn*, un U dans les motifs C/D, est stabilisé par des contacts hydrophobes et établit deux liaisons hydrogènes, respectivement avec la lysine 79 et avec l'acide aspartique 54 (Moore *et al.*, 2004). Ces contacts jouent un rôle déterminant pour l'affinité de la protéine L7Ae pour l'ARN, puisque la substitution du résidu U en position 3 du *K-turn* diminue ou abolit totalement la formation d'un complexe stable entre les 2 partenaires *in vitro* (Charron *et al.*, 2004; Kuhn *et al.*, 2002). Dans les protéines Snu13p, 15.5K, NHP2 et SBP2, la lysine est conservée ou remplacée par une arginine, tandis que (à l'exception de NHP2) l'acide glutamique est conservé ou remplacé par un acide aspartique (**Fig. 21B**). Les deux autres nucléotides non appariés du *K-turn* sont également impliqués dans des contacts électrostatiques et hydrophobes participant à renforcer l'interaction (Moore 2004). La substitution du résidu en position 2 par une cytosine diminue l'affinité de L7Ae pour l'ARN, sans pour autant abolir la formation du complexe (Charron *et al.*, 2004). Enfin, l'une des séquences les mieux conservée au sein de la sous-famille L7Ae, le motif NExxK, situé dans l'hélice  $\alpha_2$ , est quant à lui majoritairement impliqué dans l'interaction avec les guanines des paires de bases G•A (Moore *et al.*, 2004). L'analyse structurale a en revanche révélé que les adénines des paires G•A ne forment pas d'interactions spécifiques avec les résidus de la protéine L7Ae, bien qu'elles soient nécessaires pour maintenir la structuration du *K-turn* (Moore *et al.*, 2004).

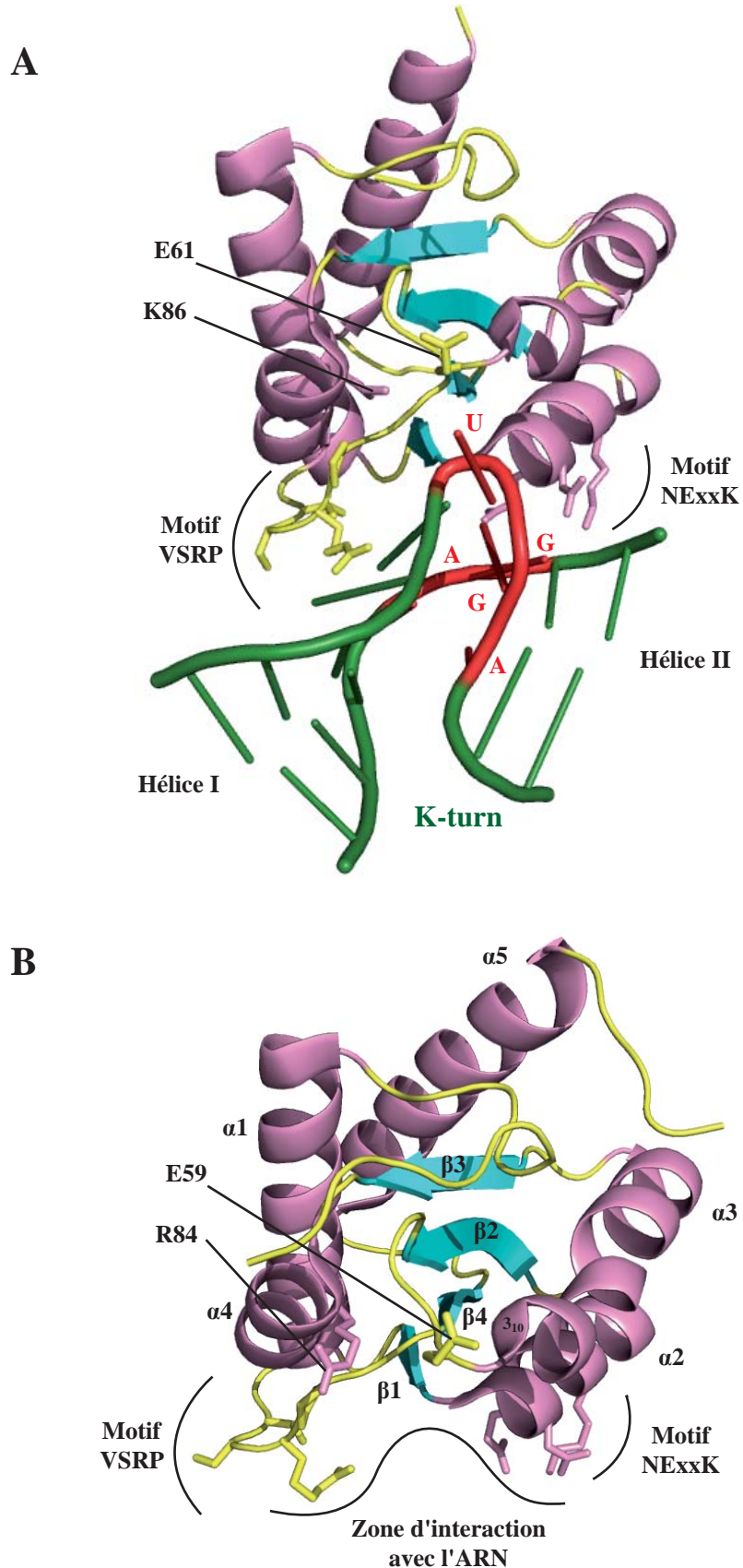
A la différence de ses homologues eucaryotes, la protéine L7Ae peut interagir avec des motifs structurés en *K-loop*. Récemment, une analyse des données structurales accumulées ces dernières années a permis de mettre en évidence une « signature » de 5 acides aminés, responsables de cette propriété (Gagnon *et al.*, 2010; Hamma & Ferre-D'Amare, 2004). En effet, la séquence L/IEVx, située directement en amont du brin  $\beta_4$ , est relativement bien conservée chez les archées, tandis que dans les protéines Snu13p/15.5K eucaryotes, elle est remplacée par un autre motif conservé, la séquence VSRP (**Fig. 21B**). Des études de retard sur gel ont ainsi démontré que la substitution de la séquence L/IEVx par les résidus VSRP abolit l'interaction de la protéine L7Ae avec les motifs *K-loop*, sans altérer sa propriété d'interaction avec les *K-turn* (Gagnon *et al.*, 2010). Au niveau structural, les résidus VSR de la protéine 15.5K sont impliqués dans des contacts avec les deux premiers résidus non appariés du motif en *K-turn* et le résidu R établit également une liaison électrostatique avec un phosphate de l'hélice I (Vidovic *et al.*, 2000). Dans la protéine L7Ae, les résidus L/I, E et V

jouent un rôle similaire mais n'établissent pas de contacts avec l'hélice I (Hamma & Ferre-D'Amare, 2004; Moore *et al.*, 2004). Ces légères variations structurales indiquent donc que la présence de l'hélice I n'est pas un paramètre déterminant pour l'interaction de la protéine L7Ae avec les motifs ARN. Enfin, un autre exemple de cette plasticité est la capacité de la protéine L7Ae à interagir avec des motifs en *K-turn* atypiques ne présentant qu'une seule paire de bases G•A, comme c'est le cas du motif KT-15 retrouvé dans l'ARNr 23S de l'archée *Pyrococcus abyssi* (Charron *et al.*, 2004).

### 1.3.2. La protéine Snu13p/15.5K

Bien que phylogénétiquement proches de la protéine L7Ae d'archée, les protéines Snu13p et 15.5K eucaryotes présentent un spectre d'interaction plus restreint avec l'ARN. Ces facteurs interagissent en effet uniquement avec les motifs en *K-turn* conventionnels et n'ont pas d'affinité pour les motifs en *K-loop* (Charron *et al.*, 2004; Gagnon *et al.*, 2010; Tran *et al.*, 2003). De plus, ils interagissent uniquement avec les snoRNA à boîtes C/D et avec un ARN absent chez les archées, le snRNA U4 (Nottrott *et al.*, 1999; Watkins *et al.*, 2000). Il est ainsi possible de distinguer 3 grandes classes de motifs en *K-turn* ciblés par ces protéines : le motif en *K-turn* localisé en 5' du snRNA U4, les motifs de type C/D et le motif B/C spécifique du snoRNA U3. La structure cristallographique d'un complexe entre la protéine 15.5K humaine et le motif en *K-turn* du snRNA U4 a été l'une des premières obtenues pour cette famille (Vidovic *et al.*, 2000) (**Fig. 25A**). Sa comparaison avec les structures de la protéine L7Ae d'archée a permis de mettre en évidence une structure hautement similaire, avec une conservation de l'organisation en sandwich  $\alpha$ - $\beta$ - $\alpha$ . Son domaine d'interaction avec l'ARN est composé des hélices  $\alpha$ 2 et  $\alpha$ 4, du brin  $\beta$ 1 et des boucles  $\beta$ 1- $\alpha$ 2,  $\beta$ 2- $\alpha$ 3 et  $\alpha$ 4- $\beta$ 4. Les acides aminés de ces régions interagissent préférentiellement avec les nucléotides de la boucle interne et avec les squelettes ribose-phosphate des résidus de l'hélice II. Le résidu U protubérant du *K-turn* est stabilisé dans une poche impliquant un résidu acide (E61 chez l'homme, E59 chez la levure) et un résidu basique (K86 chez l'homme, R84 chez la levure), comme nous l'avons précédemment décrit pour la protéine L7Ae. Cette stabilisation implique également des contacts hydrophobes avec les chaînes latérales des isoleucines 65 et 100. De plus, la substitution de cette uridine par un autre nucléotide abolit totalement l'interaction avec la protéine (Nottrott *et al.*, 1999). Comme chez l'archée, l'interaction avec les guanines





**Figure 25 : Structures tridimensionnelles de la protéine Snu13p libre et de la protéine 15.5K complexée avec le motif en K-turn du snRNA U4.** (A) Structure de la protéine 15.5K humaine en complexe avec le motif en K-turn du snRNA U4 établie par radiocristallographie (Vidovic *et al.*, 2000; PDB : 1E7K). (B) Structure de la protéine Snu13p de *S. cerevisiae* établie par radiocristallographie (Oruganti *et al.*, 2005; PDB : 2ALE).  $\alpha$  : hélice  $\alpha$ ;  $\beta$  : feuillet  $\beta$ ; 3.10 : hélice 3.10. Les paires de bases G•A et le résidu U protubérant du K-turn sont représentés en rouge.

des paires de bases G•A implique l'asparagine 40, l'acide glutamique 41 et la lysine 44 du motif conservé NExxK. Cette interaction implique également la glycine 38 qui constitue le seul résidu strictement conservé au sein de la famille L7Ae/L30 et dont la mutation abolit l'interaction de la protéine 15.5K avec le snRNA U4 (Nottrott *et al.*, 1999). De plus, les deux premiers nucléotides non appariés (deux A dans le *K-turn* du snRNA U4), font respectivement l'objet d'un *stacking* avec l'arginine 97, et avec la lysine 37 et la valine 95. Le remplacement de ces deux adénosines par des guanosines a peu d'effet sur l'interaction, en revanche, leur remplacement par des résidus pyrimidines inhibe totalement la formation du complexe (Nottrott *et al.*, 1999). Enfin, en utilisant comme modèle le snoRNA à boîtes C/D U14, Watkins *et al.* ont montré que si la formation des deux hélices du *K-turn* est un critère essentiel pour la formation d'un complexe avec la protéine 15.5K, la séquence de ces mêmes hélices ne semble pas constituer un élément déterminant pour le recrutement de la protéine (Watkins *et al.*, 2002).

Par la suite la structure de la protéine Snu13p de *S. cerevisiae* sous forme libre a également été résolue (Oruganti *et al.*, 2005) (**Fig. 25B**). Cette étude a confirmé que sa grande conservation de séquence avec la protéine humaine 15.5K (71% d'identité) se traduit également par une très forte conservation structurale. Par ailleurs, un autre travail avait déjà permis de montrer que la mutation dans la protéine Snu13p de certains résidus clés, comme l'acide glutamique (E59) responsable de la stabilisation du résidu U protubérant, abolit totalement la viabilité des cellules de levure (Dobbyn & O'Keefe, 2004). D'autre part, la comparaison des structures des deux protéines eucaryotes avec celles de L7Ae, ont conduit à formuler une hypothèse sur les différences de spécificité de ces protéines vis-à-vis des motifs ARN en *K-loop* et en *K-turn* (Oruganti *et al.*, 2005). La superposition des différentes structures révèle en effet que la surface d'interaction avec l'ARN est hautement similaire chez l'ensemble de ces facteurs, en revanche, de légères différences structurales relevées dans d'autres régions pourraient expliquer la plus grande plasticité des protéines d'archée. La première de ces différences est une insertion de deux résidus en C-terminal du brin  $\beta_4$  des protéines Snu13p et 15.5K, conduisant à la formation de liaisons hydrogènes supplémentaires qui allongent et stabilisent la partie du feuillet formée entre les brins  $\beta_2$  et  $\beta_4$ . La seconde différence est la formation d'un court brin  $\beta$  supplémentaire ( $\beta_0$ , non représenté dans la figure 25) en amont de l'hélice  $\alpha_1$  des protéines eucaryotes. Ce brin additionnel, absent dans les protéines L7Ae, contribue en effet à l'extension et à la stabilisation du feuillet  $\beta$  des protéines. Les auteurs proposent que ces deux différences tendent à rigidifier la structure des protéines

eucaryotes, limitant de ce fait les mouvements et donc l'adaptabilité de la zone de contact avec l'ARN. Cette observation pourrait ainsi expliquer les différences de spécificité observées entre les protéines L7Ae et Snu13p/15.5K.

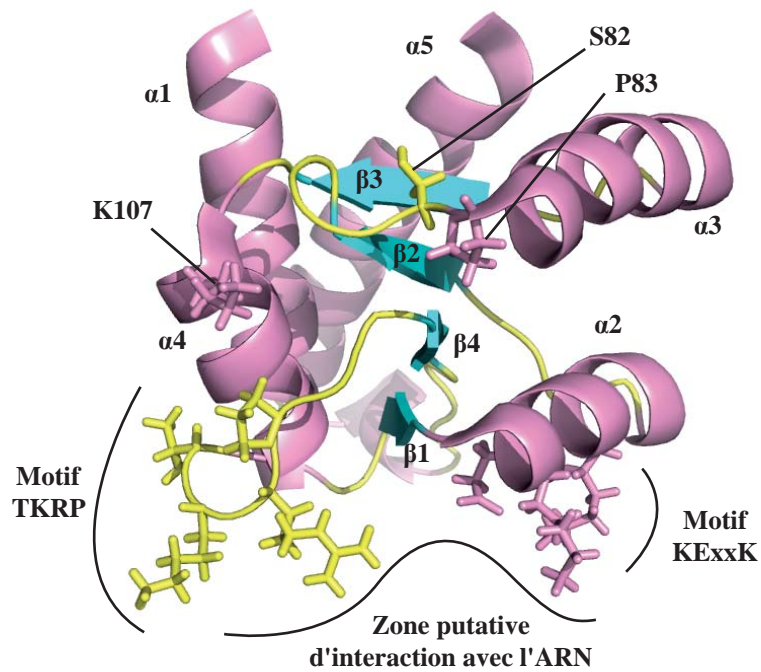
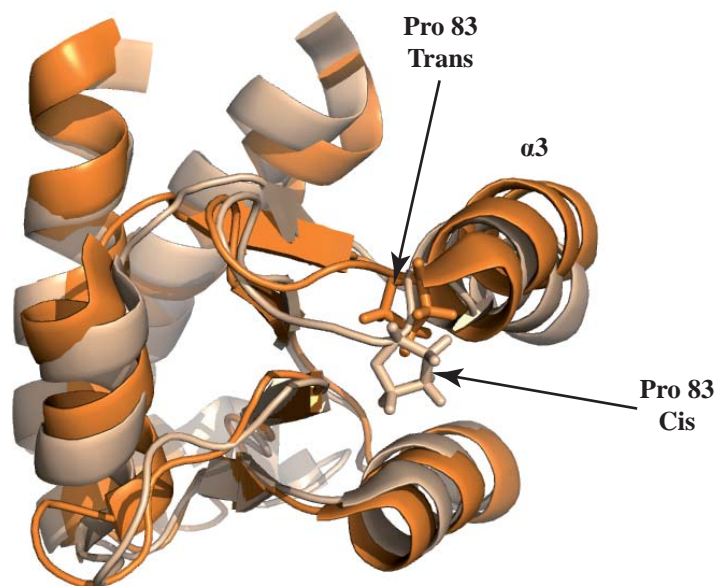
D'autre part, les protéines eucaryotes Snu13p et 15.5K sont des constituants essentiels de la snoRNP U3, dont l'ARN présente la caractéristique de posséder deux motifs en *K-turn* atypiques : les motifs C'/D et B/C (Granneman *et al.*, 2002; Marmier-Gourrier *et al.*, 2003; Watkins *et al.*, 2000). Dans le laboratoire de C. Branlant, le snoRNA U3 de levure a fait l'objet d'une attention toute particulière et l'interaction de la protéine Snu13p avec les motifs C'/D et B/C a été étudiée de façon approfondie (Clery *et al.*, 2007b; Marmier-Gourrier *et al.*, 2003) (**Fig. 23F,G**). Des expériences de retard sur gel ont notamment permis de montrer que les protéines Snu13p et 15.5K peuvent interagir indépendamment avec les motifs C'/D et B/C du snoRNA U3 (Marmier-Gourrier *et al.*, 2003; Watkins *et al.*, 2000). Dans ces deux cas, et comme cela avait également été montré pour l'interaction entre Snu13p et le snRNA U4, la pré-structuration du motif ARN en *K-turn* n'est pas un élément déterminant pour le recrutement de la protéine, puisque la formation de cette structure semble induite par la formation du complexe (Marmier-Gourrier *et al.*, 2003; Mougin *et al.*, 2002). De plus, il a été montré *in vitro* que la protéine Snu13p possède une affinité très faible pour le motif C'/D (Kd apparent  $> 1,5 \mu\text{M}$ ), tandis qu'elle présente une très forte affinité pour le motif B/C (Kd apparent  $< 250 \text{ nM}$ ) (Marmier-Gourrier *et al.*, 2003). Cette faible affinité naturelle de Snu13p vis-à-vis du motif C'/D s'explique par la présence d'un résidu U non canonique en position 2 du *K-turn*. En effet, sa substitution par une purine (A ou G) restaure une affinité substantielle (Kd apparent = 250 nM), alors que son remplacement par une cytosine abolit totalement l'interaction (Marmier-Gourrier *et al.*, 2003).

Concernant l'interaction de Snu13p avec le motif B/C, les résultats obtenus indiquent une très faible tolérance vis-à-vis des mutations réalisées dans la boucle interne. Cette observation, appuyée par une analyse phylogénique, a ainsi conduit à la proposition que la présence d'une hélice I courte (seulement un appariement G-C) au niveau de ce motif, induit une forte pression de sélection pour la conservation d'une séquence optimale au niveau des résidus non appariés du *K-turn* (Marmier-Gourrier *et al.*, 2003). Plus récemment, A. Cléry a réalisé au laboratoire une expérience de SELEX (*Systematic Evolution of Ligands by Exponential Enrichment*) avec la protéine Snu13p et un motif B/C dégénéré (Clery *et al.*, 2007b). Cette approche a notamment permis de confirmer les déterminants ARN majeurs pour l'interaction

avec Snu13p, puisque la quasi-totalité des 19 motifs sélectionnés présentaient les caractéristiques consensuelles d'un motif en *K-turn* : 2 hélices stables, une boucle asymétrique de 5+2 résidus, comportant deux paires de bases non canoniques G•A, ainsi qu'un résidu U en position 3. D'autre part, 18 des 19 ARN présentaient, au niveau de l'hélice II, une paire de bases non canonique U•U empilée avec les paires G•A. Comme nous l'avons abordé précédemment, cette paire de bases non canonique est une caractéristique des motifs C/D et se trouve remplacée, dans le snRNA U4, par un appariement canonique G-C. La sélection préférentielle d'une paire de base U•U à cette position révèle donc une différence notoire dans le mode d'interaction de Snu13p avec ces différents motifs et pourrait expliquer pourquoi la mutation de certains résidus de la protéine Snu13p (ex : E59) ont un effet drastique sur l'interaction avec le snRNA U4, sans pour autant affecter l'interaction avec le snoRNA U3 (Dobbyn & O'Keefe, 2004).

### 1.3.3. La protéine Nhp2p/NHP2

La protéine eucaryote Nhp2p/NHP2 est le seul membre de la famille L7Ae/L30 qui ne reconnaît pas spécifiquement les motifs de type *K-turn* ou *K-loop*. Bien que surprenante, cette propriété est toutefois cohérente avec l'absence de ce type de motifs au sein des snoRNA à boîtes H/ACA eucaryotes. Une étude réalisée par l'équipe de M. Caizergues-Ferrer et Y. Henry, a néanmoins permis de mettre en évidence la capacité de la protéine Nhp2p de levure à interagir directement avec l'ARN *in vitro* (Henras *et al.*, 2001). Malgré une légère préférence pour les structures en tige-boucle imparfaite, les résultats obtenus indiquent cependant une très faible spécificité d'interaction et aucune séquence cible spécifique des snoRNA à boîtes H/ACA n'a pu être identifiée. De plus, même si les principaux acides aminés normalement impliqués dans l'interaction avec les motifs en *K-turn* sont relativement bien conservés dans la protéine Nhp2p/NHP2, le résidu acide (E61 dans 15.5K humaine) responsable de la stabilisation du U protubérant n'est pas retrouvé dans cette protéine. Récemment, la structure tridimensionnelle de la protéine Nhp2p de *S. cerevisiae* a été résolue par RMN (Koo *et al.*, 2011) (**Fig. 26**). Cette étude, qui a permis de mettre en évidence un repliement typique de la famille L7Ae/L30 (**Fig. 26A**), a néanmoins révélé une légère différence structurale concernant la Proline conservée en position 83. En effet, alors que l'ensemble des données structurales obtenues pour L7Ae, Snu13p et 15.5K ont révélé la conformation *cis* majoritaire de ce résidu,

**A****B**

**Figure 26 : Structures tridimensionnelles de la protéine Nhp2p de *S. cerevisiae* (Koo et al., 2011).** (A) Structure de la protéine Nhp2p en solution obtenue par RMN (PDB : 2LBX). (B) Superposition des structures de Nhp2p wt (en orange) et du mutant S82/W (en saumon; PDB : 2LBW) réalisées à l'aide du logiciel PyMol. La mutation S82/W stabilise la conformation Cis de la proline et induit des changements de structure locale plus ou moins importants.  $\alpha$  : hélice  $\alpha$ ;  $\beta$  : feuillet  $\beta$ .

les travaux de Koo *et al.*, ont révélé l'existence d'une forte isomérisation *cis/trans* de la Proline 83 de la protéine Nhp2p (Koo *et al.*, 2011) (**Fig. 26B**). Or, ce résidu localisé à l'extrémité N-terminale de l'hélice  $\alpha_3$  est situé dans la zone putative d'interaction avec l'ARN et avec la protéine Nop10p de la particule H/ACA. Les auteurs ont ainsi montré que la stabilisation de cette Proline dans une conformation *cis* stabilise l'association de la protéine Nhp2p au sein des snoRNP *in vivo*. De façon remarquable, la Sérine 82 qui précède cette proline, est l'équivalent du résidu acide glutamique E61 de la protéine 15.5K. L'analyse structurale par RMN a ainsi montré que la conformation *cis* de la proline influe sur le positionnement de la Sérine 82 en orientant les groupements carbonyle et amide de sa chaîne principale en direction de la poche d'interaction avec l'ARN. Indépendamment de la nature de la chaîne latérale (ex : une lysine dans la protéine humaine) ce résidu pourrait donc contribuer à l'interaction avec l'ARN. Les auteurs proposent également que, malgré l'absence de véritable *K-turn*, ce résidu soit néanmoins impliqué dans la reconnaissance d'un résidu U, retrouvé de façon préférentielle à une distance de 10 ou 11 paires de bases de la poche de pseudouridylation des snoRNA à boîtes H/ACA. Cette configuration, gouvernée par l'état d'isomérisation de la Proline 83, pourrait ainsi être favorisée lors de la formation de la snoRNP. L'interaction de Nhp2p avec la protéine Nop10p au sein du complexe pourrait en effet stabiliser la conformation *cis* de la proline et permettre de ce fait, l'établissement de contacts optimaux entre Nhp2p et le snoRNA (Koo *et al.*, 2011). Ces résultats mettent donc en avant une différence majeure entre Nhp2p et les autres protéines de la famille L7Ae/L30, puisque sa capacité à interagir avec l'ARN pourrait être induite au sein de la RNP par le biais de contacts inter- et intramoléculaires.

#### 1.3.4. La protéine SBP2

La protéine SBP2 est une très grande protéine (854 acides aminés dans la forme humaine), composée de plusieurs domaines : un long domaine N-terminal non caractérisé (1-408), suivi d'un domaine putatif d'interaction avec le ribosome (408-515), puis d'un large domaine d'interaction avec les éléments SECIS (515-777), qui contient lui-même un module de type L7Ae/L30-RBD (670-754). Sur la base de la structure du complexe 15.5K/U4, l'équipe d'A. Krol a étudié le mode d'interaction de la protéine SBP2 humaine avec l'ARN, en testant l'effet de mutations générées au sein du domaine L7Ae/L30-RBD (Allmang *et al.*, 2002). Parmi les acides aminés mutés, douze sont impliqués dans l'interaction avec l'ARN, et 4

d'entre eux, les résidus G676, E679, E699 et R731, se sont révélés essentiels (**Fig. 21B**). De façon intéressante, ces résidus jouent également un rôle majeur dans les interactions décrites précédemment (Moore *et al.*, 2004; Vidovic *et al.*, 2000). En effet, la glycine 676 de SBP2, strictement conservée dans l'ensemble de la famille L7Ae/L30, est requise dans la protéine 15.5K pour l'interaction avec la guanine de la première paire de bases G•A (Vidovic *et al.*, 2000). L'acide glutamique 679 appartient au motif NExxK également impliqué dans l'interaction avec les guanines des paires de bases G•A (Moore *et al.*, 2004; Vidovic *et al.*, 2000). Enfin, l'acide glutamique 699 et l'arginine 731 (équivalent aux acides aminés E61 et K86 de la protéine 15.5K humaine) correspondent aux deux acides aminés majoritairement impliqués dans la stabilisation électrostatique du résidu U protubérant du *K-turn* (Moore *et al.*, 2004; Vidovic *et al.*, 2000). En revanche, le motif VSRP identifié dans les protéines Snu13p et 15.5K n'est pas conservé dans SBP2. En dépit de l'aspect non conventionnel des sites de fixation de la protéine SBP2, ces résultats suggèrent donc que le mode d'interaction de SBP2 avec l'ARN est similaire à celui de la protéine 15.5K.

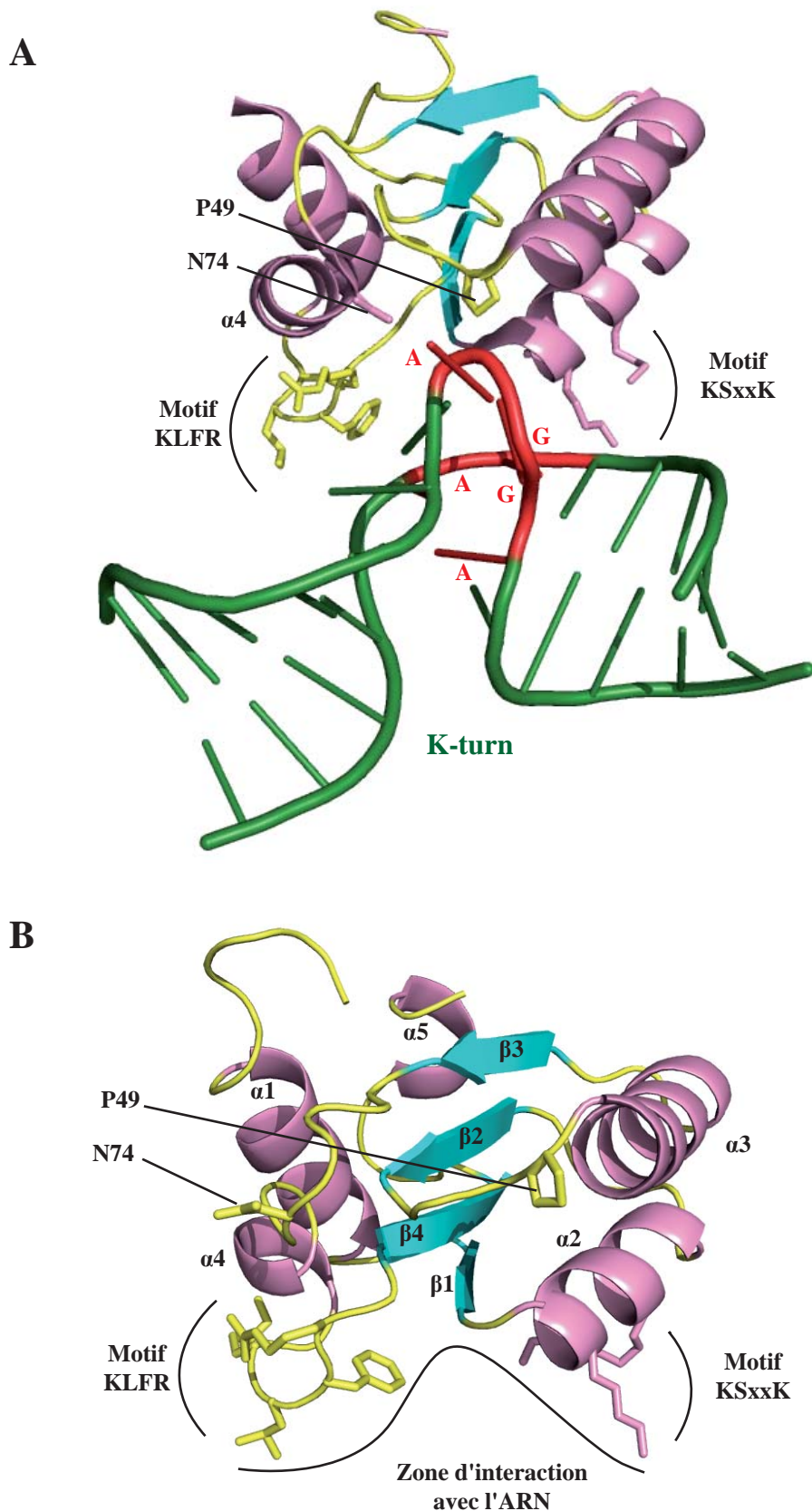
En utilisant à nouveau l'approche SELEX, et par comparaison avec la protéine Snu13p de levure, l'équipe de C. Branlant a entrepris la caractérisation des déterminants ARN requis pour la fixation de SBP2 (Clery *et al.*, 2007a). Cette étude a tout d'abord révélé que la protéine SBP2 n'interagit pas avec les motifs en *K-turn* du snRNA U4 ou de l'ARN U3B/C, tandis que la protéine Snu13p peut être recrutée sur l'élément SECIS de l'ARNm SelN avec la même spécificité que SBP2. De plus, la mise en place de l'approche SELEX a permis la sélection de seulement 7 ARN distincts, suggérant ainsi l'existence de fortes contraintes liées au mode d'interaction de SBP2 avec l'ARN. De façon remarquable, chacun de ces 7 motifs adopte une structure en *K-turn* canonique comprenant un tandem de paires G•A, 3 nucléotides non appariés au niveau de la boucle interne, dont un U en position 3. Dans les 7 structures les positions 1 et 2 du *K-turn* sont occupées par un di-nucléotide GA et les hélices II, relativement longues, comptent au moins deux appariements Watson-Crick consécutifs. On retrouve également de façon systématique une paire de bases non canonique (de préférence U•U) empilée avec les paires de bases G•A au niveau de l'hélice II. Cette observation pourrait par ailleurs expliquer l'absence d'affinité de SBP2 pour le snRNA U4 qui possède un appariement G-C à cette position. D'autre part, conformément à l'absence d'affinité de SBP2 pour le motif B/C, aucun des 7 ARN sélectionnés ne possède une hélice I courte ou un résidu en *bulged* sur le brin 3'. En revanche, les résultats obtenus indiquent que le motif optimal sélectionné par cette approche possède une affinité pour SBP2 trois fois plus faible que celle

observée dans le cas de l'ARNm SelN. Cette différence s'explique par la présence d'une boucle interne plus petite que celle observée dans les motifs SECIS naturels. En effet, l'insertion de résidus sur les brins 5' et 3' afin de reformer un large boucle interne, restaure la forte affinité de SBP2, tout en altérant la fixation de la protéine Snu13p. L'ensemble de ces résultats indiquent donc que si l'interaction de la protéine SBP2 avec l'ARN repose sur des bases communes aux protéines de la sous-famille L7Ae, elle implique néanmoins un certain nombre de caractéristiques spécifiques inédites. Comme nous allons l'aborder par la suite, cette divergence vis-à-vis de la sous-famille L7Ae pourrait tirer son explication dans la découverte récente de sa possible affiliation à la sous-famille L30 (Donovan & Copeland, 2009).

### 1.3.5. La protéine L30

La résolution des structures des protéines L30 d'archée et de levure a permis de mettre en évidence une topologie similaire à celles observées pour les protéines de la sous-famille L7Ae (Chao & Williamson, 2004; Chen *et al.*, 2003; Fedorov *et al.*, 1999; Mao *et al.*, 1999). En revanche l'étude de son mode d'interaction avec l'ARN révèle des différences de spécificité notoires (Chao & Williamson, 2004; Mao *et al.*, 1999) (**Fig. 27A**). Parmi les principaux déterminants protéiques identifiés dans les protéines L7Ae ou 15.5K, seule la glycine conservée joue un rôle similaire, puisqu'elle est retrouvée au sein de la poche de fixation à l'ARN. Le motif NExxK est remplacé, dans la protéine L30 de levure (alias Rpl30p), par un motif KSxxK qui interagit également avec les guanines des paires G•A. Les résidus acide glutamique/aspartique et lysine impliqués dans la stabilisation du U protubérant dans les protéines L7Ae ne sont pas conservés. Dans les protéines L30, ces résidus sont respectivement remplacés par une proline (P49 dans Rpl30p) et par une asparagine (N74 dans Rpl30p) qui stabilisent également le nucléotide protubérant. Enfin, le motif LEVA/VSRP identifié dans les protéines de la sous-famille L7Ae est remplacé par un motif KxF/YR dont la phénylalanine et l'arginine interagissent chez la levure avec le résidu en position 2 du *K-turn*. De plus, chez la levure, la comparaison des structures de la protéine Rpl30p sous sa forme libre et complexée avec l'ARN, révèle quelques différences topologiques locales (Chao *et al.*, 2003; Mao & Williamson, 1999) (**Fig. 27A et B**). De façon intéressante, la formation du complexe induit la structuration complète de l'hélice  $\alpha_4$ , qui fait partie intégrante de la zone





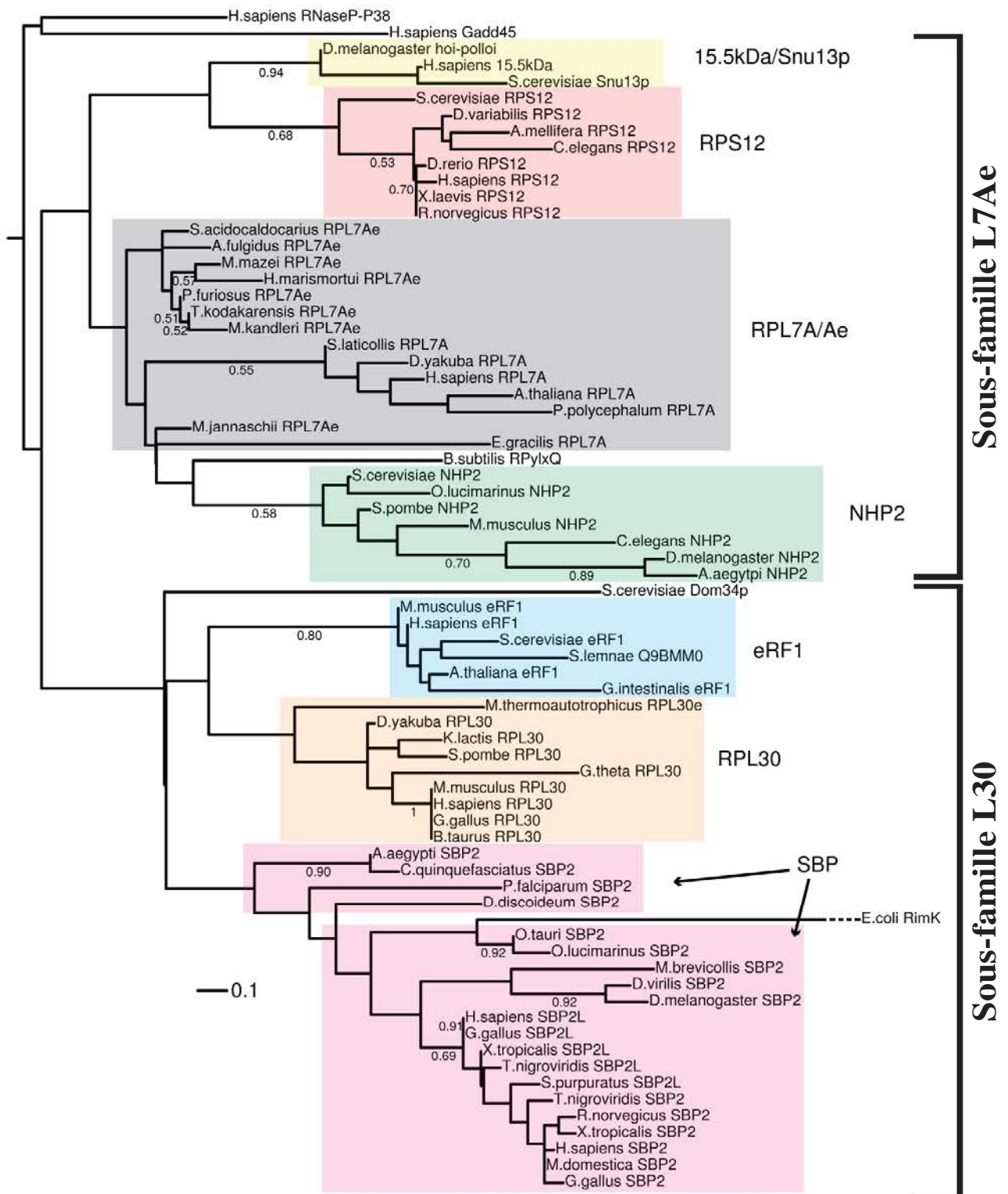
**Figure 27 : Structure tridimensionnelle de la protéine Rpl30p (alias L30) de levure sous forme libre ou en complexe avec un motif en K-turn.** (A) Structure de la protéine L30 de *S. cerevisiae* en complexe avec le motif en K-turn contenu dans le pré-ARNm codant la protéine L30, établie par radiocristallographie (Chao *et al.*, 2004; PDB : 1TOK). (B) Structure de la protéine L30 de *S. cerevisiae* en solution obtenue par RMN (Mao *et al.*, 1999b; PDB : 1CK2).  $\alpha$  : hélice  $\alpha$ ;  $\beta$  : feuillet  $\beta$ . Les paires de bases G•A et le résidu A protubérant du K-turn sont représentés en rouge.

d'interaction avec l'ARN et qui contient notamment l'asparagine (N74) impliquée dans la stabilisation du nucléotide protubérant. Ce repliement induit par la formation du complexe n'a pas été observé pour la protéine L7Ae et révèle donc une plus grande flexibilité de la protéine L30. Ces variations structurales se caractérisent par une différence dans la spécificité d'interaction avec l'ARN, dont la plus notable est la préférence pour un C ou un A en position 3 du *K-turn* (Li & White, 1997). Les déterminants ARN requis pour la fixation de la protéine L30 de levure ont également fait l'objet d'études approfondies basées sur l'utilisation des approches de SELEX et de mutagenèse dirigée (Li & White, 1997; White *et al.*, 2004). Par comparaison avec les protéines de la sous-famille L7Ae, ces travaux confirment une plus grande tolérance de la protéine L30 vis-à-vis du résidu protubérant en position 3 du *K-turn*, puisque cette position peut être occupée non seulement par un C ou un A, mais également par un U. Enfin, la fixation de la protéine L30 requiert également une purine en position 2 du *K-turn*, une hélice II contenant au moins deux paires G-C et un appariement non canonique (de préférence G•U) empilé avec les deux paires de bases G•A.

#### 1.4. Protéines de la famille L7Ae/L30 et évolution

Les données que nous venons de détailler indiquent clairement que les protéines de la famille L7Ae/L30 partagent un même mode d'interaction avec l'ARN, basé sur une organisation structurale relativement bien conservée. En revanche, nous avons pu observer que les déterminants impliqués dans ces interactions varient selon les protéines et selon leurs fonctions. Au-delà d'une origine commune, l'ensemble de ces résultats révèle donc la complexité des liens phylogéniques unissant les membres de cette famille. En se basant sur l'alignement des séquences des protéines archées et eucaryotes, et en utilisant l'approche du maximum de vraisemblance, une équipe a récemment proposé un arbre phylogénétique pour les domaines L7Ae/L30-RBD (Donovan & Copeland, 2009) (**Fig. 28**). Dans ce modèle, deux principaux *phyla* divergent à partir d'un même ancêtre commun : d'une part un ensemble assimilable au groupe communément appelé sous-famille L7Ae et qui contient les protéines L7Ae, RPS12, Nhp2p/NHP2 et Snu13p/15.5K, d'autre part, un groupe comprenant les protéines L30, eRF1 et SBP2, appelé sous-famille L30. La présence de SBP2 dans le second *phylum* peut paraître surprenante étant donné les similarités relevées précédemment avec les protéines de la sous-famille L7Ae, néanmoins une étude réalisée chez le rat a permis de

## Famille L7Ae/L30



**Figure 28 : Arbre phylogénétique des domaines L7Ae/L30-RBD (Extrait de Donovan *et al.*, 2009).** Cet arbre a été réalisé en utilisant l'approche du maximum de vraisemblance à l'aide du logiciel PhyML. 500 réplicats de bootstrap ont été réalisés et l'arbre a été enraciné en utilisant le groupe externe Gadd45/RNasePP38. Les valeurs de bootstrap inférieures à 0,5 ne sont pas représentées. L'échelle est proportionnelle au nombre de substitutions d'acides aminés par site. SBP2L est un paralogue de la protéine SBP2, elle constitue la seule forme retrouvée chez les invertébrés deutérostomiens.

démontrer une certaine proximité fonctionnelle entre les protéines SBP2 et L30 (Chavatte *et al.*, 2005). En effet, ces travaux ont mis en évidence que la protéine ribosomique L30 participe au processus d'incorporation des sélénocystéines au cours de la traduction, par le biais d'une interaction avec l'élément SECIS. Dans le modèle proposé, la protéine L30 serait en effet échangée avec la protéine SBP2 lors de l'association du ribosome sur l'ARNm. L'interaction de la protéine L30 avec le motif SECIS induirait alors un changement conformationnel qui pourrait être à l'origine de l'incorporation du Sec-ARNt<sup>Sec</sup> (ARNt chargé par la sélénocystéine) au sein du site A du ribosome. D'autre part, au sein du *phylum* L7Ae, l'appartenance de Nhp2p/NHP2 et L7Ae à un groupe monophylétique excluant Snu13p/15.5K, suggère que l'apparition d'une protéine spécifique des particules C/D eucaryotes pourrait constituer un événement relativement ancien. En revanche, les protéines Nhp2p/NHP2 auraient pu diverger plus tardivement du groupe L7Ae, en perdant leur capacité à interagir avec les motifs en *K-turn* et en relâchant, de ce fait, la pression de sélection imposée aux ARN à boîtes H/ACA pour la conservation d'un tel motif.

L'étude détaillée de la structure de la machinerie d'épissage eucaryote, a montré que le spliceosome moderne dérive probablement des ARN autocatalytiques de groupe II (Sharp 1991). Sur la base d'un schéma évolutif similaire, il a notamment été proposé que les protéines de la famille L7Ae/L30 aient évolué à partir de la machinerie primitive de traduction (Tran *et al.*, 2004). En effet, les protéines ribosomiques L7, L30 et RPS12 sont conservées dans l'ensemble du monde vivant, tandis que les rôles joués par les membres de cette famille dans d'autres processus, comme la biogenèse des ribosomes, l'épissage, la traduction des sélénoprotéines ou la réplication des télomères sont des fonctions retrouvées spécifiquement dans certains *phyla*. Dans un scénario possible, un motif en *K-turn* présent au sein d'un ARNr primitif aurait pu initier la formation d'une RNP fonctionnant en *cis* dans la biogenèse du ribosome. Tout en conservant sa fonction initiale, cette dernière aurait ensuite pu se désolidariser de l'ARNr pour former une RNP fonctionnant en *trans* qui constituerait le précurseur de certaines RNP modernes. La question de l'acquisition des différentes fonctions associées à ces RNP reste posée. Il est relativement probable que les fonctions liées à la maturation des ARNr tirent leur existence dans la fonction de cette particule primitive. En revanche, d'autres fonctions, comme l'acquisition de nouvelles cibles de modifications (UsnRNA, ARNm, ARNt) ou la participation au processus d'épissage des pré-ARNm (snRNP U4), pourraient résulter d'un processus de diversification s'étant déroulé de façon secondaire au cours de l'évolution.

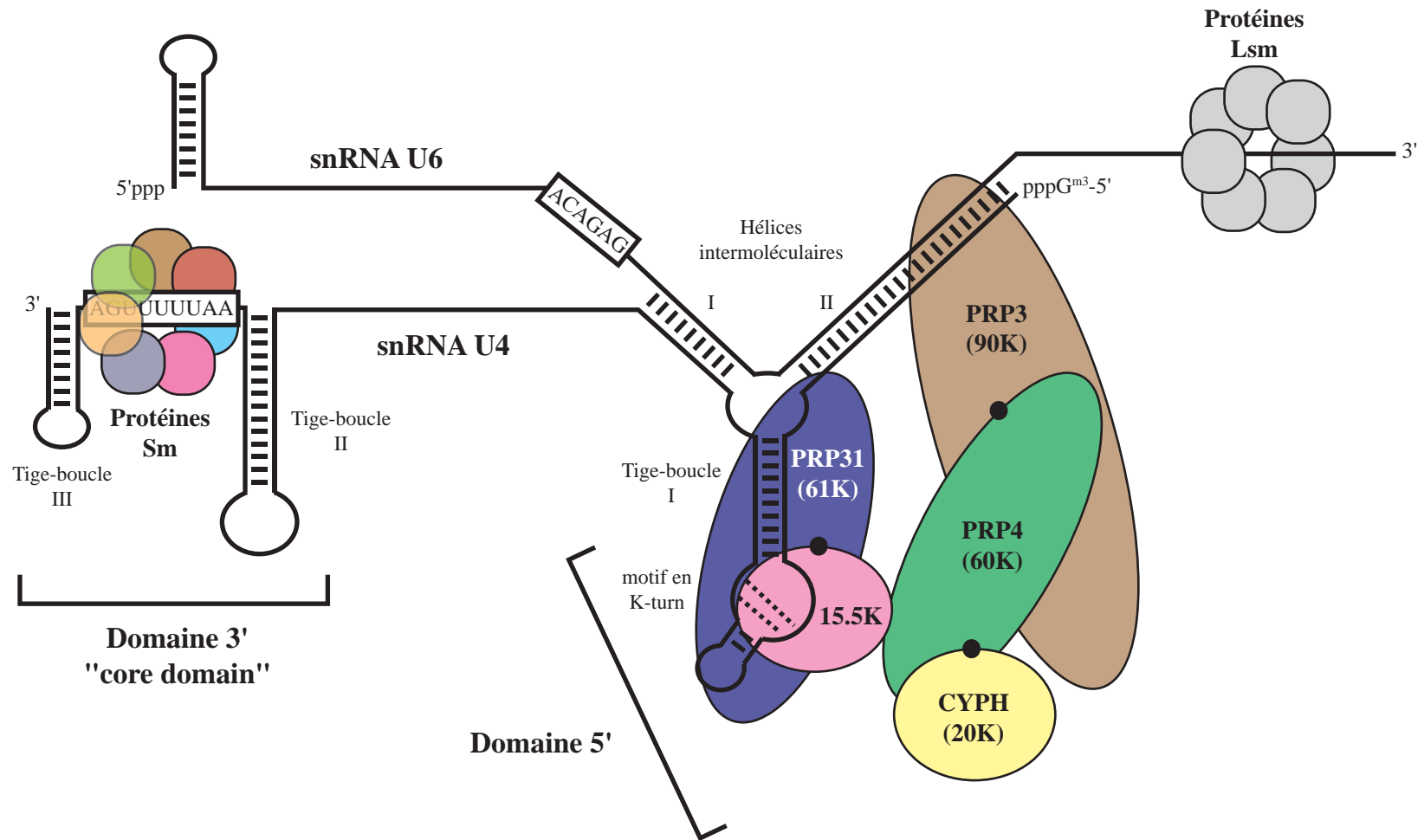
## 2. La snRNP U4

Comme nous l'avons vu dans le chapitre précédent, le spliceosome est une machinerie macromoléculaire hautement dynamique qui catalyse l'épissage des pré-ARNm chez les eucaryotes (**Fig. 10**). La tri-snRNP U4/U6.U5 est l'un des acteurs majeurs de ce macro-complexe et son incorporation au sein du spliceosome constitue une étape critique pour le mécanisme d'épissage. Au cours de ce processus, la snRNP U4 joue un rôle de chaperonne en permettant de délivrer le snRNA U6 dans une conformation optimale au sein du spliceosome. Cette fonction, qui implique la formation d'un duplex stable entre les 2 UsnRNA et qui requiert la présence de plusieurs protéines, constitue un élément déterminant pour la mise en place du site actif catalysant la réaction d'épissage.

### 2.1. Composition et organisation structurale de la snRNP U4

#### 2.1.1. Le snRNA U4

Le snRNA U4 comporte deux domaines distincts (**Fig. 29**). Sa partie 5' est formée par une structure en tige-boucle (tige-boucle I) qui contient à son extrémité, le motif en *K-turn* responsable du recrutement de la protéine Snu13p/15.5K (Nottrott *et al.*, 1999). Ce domaine 5' contient également les séquences complémentaires à la région 3' du snRNA U6. Lors de la formation du duplex, les deux UsnRNA subissent donc un remaniement conformationnel qui aboutit à la formation de deux hélices intermoléculaires (hélices I et II), de part et d'autre de la tige-boucle contenant le motif en *K-turn* (Mougin *et al.*, 2002). Le duplex adopte ainsi une structuration en « Y » qui est conservée entre l'homme et la levure (Mougin *et al.*, 2002). La région 3' du snRNA U4 est également structurée en tige-boucle (tige-boucle II et III) et contient un site Sm riche en U, responsable du recrutement des protéines Sm (Branlant *et al.*, 1982; Liautard *et al.*, 1982). Au cours de son cycle fonctionnel, le snRNA U4 est donc retrouvé alternativement sous forme libre ou associé au snRNA U6. Avant son incorporation au sein du spliceosome, la di-snRNP U4/U6 s'associe à la snRNP U5 pour former la tri-snRNP U4/U6.U5. Cette association ne repose en revanche pas sur la formation de nouveaux appariements intermoléculaires, mais implique des interactions entre les protéines des différentes particules (Liu *et al.*, 2006; Makarova *et al.*, 2002). Au cours de sa maturation, le



**Figure 29 : Représentation schématique de la di-snRNP U4/U6 humaine (adapté de Nottrott *et al.*, 2002).** Les pastilles noires symbolisent les interactions caractérisées entre les différents partenaires protéiques. Seule la protéine CYPH (20K) ne possède pas d'homologue chez la levure *S. cerevisiae*.

snRNA U4 fait également l'objet de réactions de pseudouridylation et de 2'-O-méthylation. Chez l'homme, il a notamment été montré que certaines d'entre elles, comme la 2'-O-méthylation du résidu A65, sont catalysées par des scaRNP au sein des CB (Darzacq *et al.*, 2002). Chez la levure en revanche, aucune modification n'a été décrite à ce jour (Massenet *et al.*, 1998).

### 2.1.2. Les protéines de la snRNP U4

La protéine Snu13p/15.5K est un constituant essentiel de la snRNP U4 puisque sa déplétion conduit à une disparition du snRNA U4 dans les cellules de levure (Watkins *et al.*, 2000). En plus de ce composant majeur, la snRNP U4 contient également un lot de 7 protéines Sm, ainsi que la protéine Prp31p/PRP31. Trois autres facteurs, les protéines 20K, Prp3p/90K et Prp4p/60K, sont recrutés spécifiquement lors de la formation de la di-snRNP U4/U6 (**Fig. 29**). Ces différentes protéines sont essentielles pour l'épissage et pour la viabilité cellulaire (Banroques & Abelson, 1989; Bjorn *et al.*, 1989; Lustig *et al.*, 1986; Nottrott *et al.*, 1999; Weidenhammer *et al.*, 1997).

#### 2.1.2.1. Les protéines Sm

Quatre des cinq composants majeurs du spliceosome, les snRNP U1, U2, U4 et U5, contiennent en commun un lot de 7 protéines Sm, appelées SmB/B', SmD1, SmD2, SmD3, SmE, SmF et SmG (Scofield & Lynch, 2008; Seraphin, 1995). Ces protéines appartiennent à la famille de protéines Sm, qui compte des représentants dans l'ensemble du monde vivant. Cette famille comprend également les protéines Lsm (1 à 8), qui sont des composants de la snRNP U6 et qui interviennent dans la maturation des snoRNA (Gottschalk *et al.*, 1999; Seraphin, 1995). Les protéines Sm sont caractérisées par la conservation de deux éléments de séquence, appelés motifs Sm, reliés par un *linker* de taille variable. Dans l'espace, ces deux motifs forment un « domaine Sm » composé d'une hélice  $\alpha$  en position N-terminale, suivie de 5 brins  $\beta$  anti-parallèles (Kambach *et al.*, 1999). Au sein d'une RNP, chaque protéine Sm interagit avec deux autres protéines Sm pour former un anneau composé de 6 ou 7 partenaires, au centre duquel se positionnent les séquences ARN riches en U avec lesquelles elles

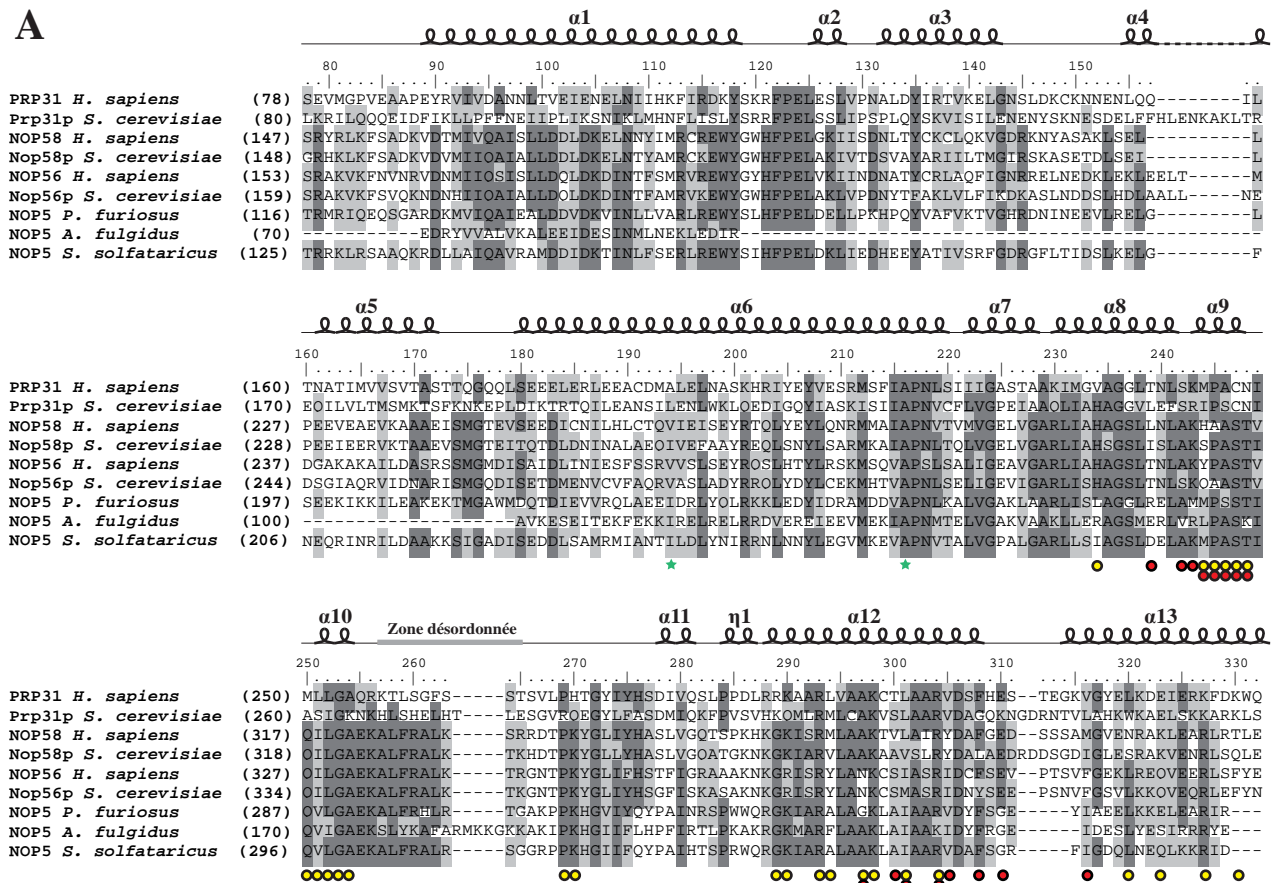
interagissent (Kambach *et al.*, 1999; Urlaub *et al.*, 2001). Les contacts entre protéines Sm sont établis par l'intermédiaire des brins  $\beta 4$  et  $\beta 5$ , qui forment le motif Sm2 (Kambach *et al.*, 1999).

### 2.1.2.2. La protéine Prp31p/PRP31

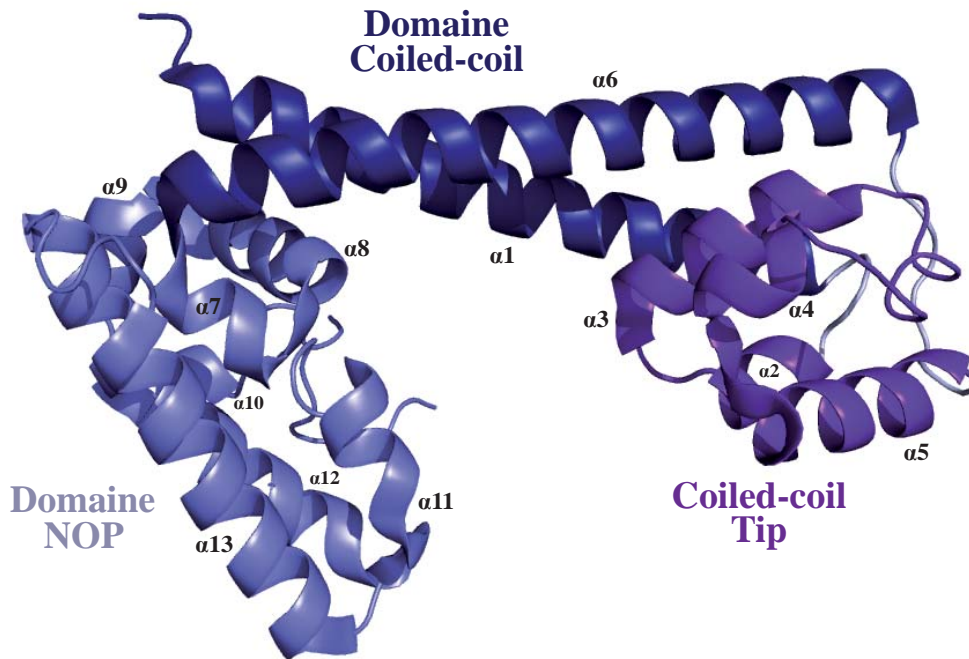
La fixation de la protéine Snu13p/15.5K au niveau du motif en *K-turn* du snRNA U4 permet le recrutement d'un autre composant de la snRNP U4, la protéine Prp31p/PRP31 (Nottrott *et al.*, 2002). Cette protéine, initialement baptisée 61K chez l'homme, est conservée chez les eucaryotes et s'avère requise pour l'incorporation de la tri-snRNP U4/U6.U5 au sein du spliceosome (complexe A) (Makarova *et al.*, 2002; Weidenhammer *et al.*, 1997; Weidenhammer *et al.*, 1996). Majoritairement composée d'hélices  $\alpha$ , cette protéine comporte trois principaux domaines : un long domaine *Coiled-coil* formé par les hélices  $\alpha 1$  et  $\alpha 6$ , un petit domaine globulaire formé par les hélices  $\alpha 2$  à  $\alpha 5$  (appelé *Coiled-coil tip*), et un domaine NOP, composé des hélices  $\alpha 7$  à  $\alpha 13$ , capable d'interagir avec le snRNA U4 et avec la protéine 15.5K (Liu *et al.*, 2007b) (**Fig. 30**). Récemment, une étude a mis en évidence des sites de phosphorylation dans la région C-terminale de la protéine PRP31 humaine, en aval du domaine NOP (Schneider *et al.*, 2010). Ces sites sont ciblés par la kinase à sérine PRP4K et sont conservés chez les eucaryotes, à l'exception de *S. cerevisiae* qui ne possède pas d'équivalent de cette enzyme. Cette même étude a également montré que l'association de la tri-snRNP U4/U5.U6 avec le pré-spliceosome est dépendante de PRP4K, suggérant ainsi que l'état de phosphorylation de PRP31 pourrait permettre de réguler son rôle dans l'assemblage du spliceosome. Enfin, comme nous l'aborderons par la suite, le domaine NOP est également retrouvé dans les protéines des snoRNP à boîtes C/D Nop56p/NOP56 et Nop58p/NOP58 (Gautier *et al.*, 1997; Makarova *et al.*, 2002; Vithana *et al.*, 2001) (**Fig. 30A**). Cette observation est tout à fait intéressante, puisque elle indique que les similarités structurales entre les snoRNP à boîtes C/D et la snRNP U4 ne se limitent pas au complexe formé par l'ARN et la protéine Snu13p/15.5K. Ceci suggère donc fortement que la snRNP U4 et les snoRNP auraient pu dériver d'une même particule ancestrale.



**A**



**B**



**Figure 30 : Organisation structurale de la protéine PRP31 humaine. (A)** Alignement de séquences des protéines partageant un domaine NOP réalisé par ClustalW Multiple Alignment à l'aide du logiciel BioEdit (Hall, T.A. 1999). Les acides aminés identiques dans plus de 50% des séquences alignées sont représentés sur fond gris foncé. Les acides aminés similaires dans plus de 50% des séquences alignées sont représentés sur fond gris clair. Les pastilles jaunes et rouges matérialisent les acides aminés impliqués respectivement dans des interactions avec l'ARN et la protéine 15.5K. Les étoiles vertes représentent les résidus mutés dans certains cas de rétinites pigmentaires. Les éléments de structures secondaires reportés au dessus de l'alignement correspondent à ceux relevés dans la structure de la protéine humaine PRP31 en complexe avec le K-turn du snRNA U4 et la protéine 15.5K (Liu et al., 2007; PDB : 2OZB) **(B)**.

### 2.1.2.3. Les protéines CYPH, Prp4p/60K et Prp3p/90K de la di-snRNP U4/U6

Chez l'homme, trois facteurs supplémentaires sont associés au snRNA U4 lorsqu'il est complexé au snRNA U6 : les protéines 20K, 60K et 90K, initialement baptisées en fonction de leur masse moléculaire (Horowitz *et al.*, 1997; Teigelkamp *et al.*, 1998). Ces 3 facteurs peuvent être purifiés indépendamment de la di-snRNP, sous la forme d'un complexe trimérique. Chez la levure *S. cerevisiae*, les facteurs 60K et 90K correspondent respectivement aux protéines Prp4p et Prp3p, en revanche, aucun homologue de la protéine 20K n'a pu être identifié (Anthony *et al.*, 1997; Banroques & Abelson, 1989). La protéine Prp4p/60K appartient à la famille des protéines WD, elle contient en effet 7 motifs WD-40 qui se structurent en un domaine en hélice de type  $\beta$ -propeller, responsable de l'interaction avec la protéine Prp3p/90K (Ayadi *et al.*, 1998; Gonzalez-Santos *et al.*, 2002). Chez l'homme, une interaction directe a également été détectée entre la protéine 60K (hPRP4) et le facteur 20K (Reidt *et al.*, 2003). Ce dernier, renommé Cyclophiline H, possède une structure similaire à celle de la Cyclosporine A et possède une activité Peptidyl-Prolyl *cis/trans* Isomérase (PPIase) (Reidt *et al.*, 2000). Chez l'homme, le spliceosome contient au moins 8 PPIases dont le rôle précis n'a pas encore été identifié (Wahl *et al.*, 2009). Néanmoins, il a été montré que la cyclosporine A, un inhibiteur des cyclophylines, affecte les deux étapes catalytiques de la réaction d'épissage (Horowitz *et al.*, 2002). Comme la plupart de ces protéines n'ont pas d'homologues chez la levure, il a notamment été proposé que ces dernières jouent un rôle dans le phénomène d'épissage alternatif en régulant la flexibilité du spliceosome chez certains eucaryotes (Wahl *et al.*, 2009). Enfin, des expériences de pontages aux UV réalisées *in vitro* ont montré que la protéine 90K (hPRP3) est capable d'interagir physiquement avec l'hélice II intermoléculaire formée par l'appariement entre les snRNA U4 et U6 (Nottrott *et al.*, 2002).

## 2.1.3. Structure et assemblage de la snRNP U4

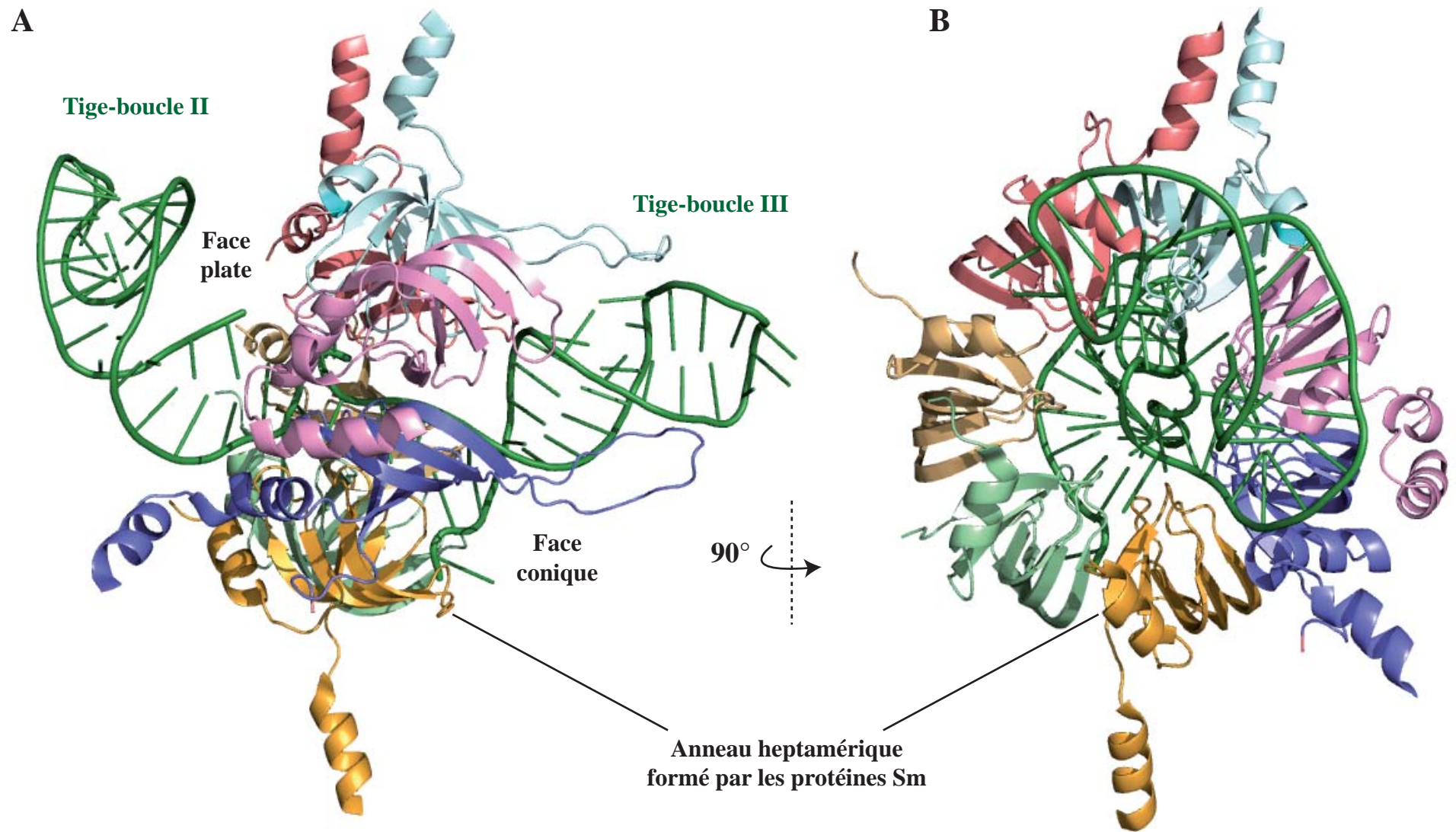
### 2.1.3.1. Assemblage du domaine 3' de la snRNP U4

Chez la majorité des eucaryotes, l'assemblage des protéines Sm sur le snRNA U4 s'effectue au niveau du cytoplasme, au cours d'un processus dirigé par le complexe SMN (Pellizzoni *et*

*al.*, 2002b). Récemment, l'équipe de J. Li a résolu la structure cristallographique du domaine 3' de la snRNP U4 humaine (Leung *et al.*, 2011) (**Fig. 31**). Les protéines Sm forment un anneau heptamérique possédant une face plate et une face conique. Cet anneau s'assemble autour du site en simple brin 5'-AAUUUUUGA-3' et interagit avec les phosphates et avec les plateaux de bases de l'ARN par l'intermédiaire des boucles des domaines Sm. Cette séquence est bordée en 5' et en 3' par les tige-boucle II et III, qui établissent respectivement des contacts avec les faces plate et conique de l'heptamère.

### 2.1.3.2. Assemblage des protéines spécifiques de la snRNP U4

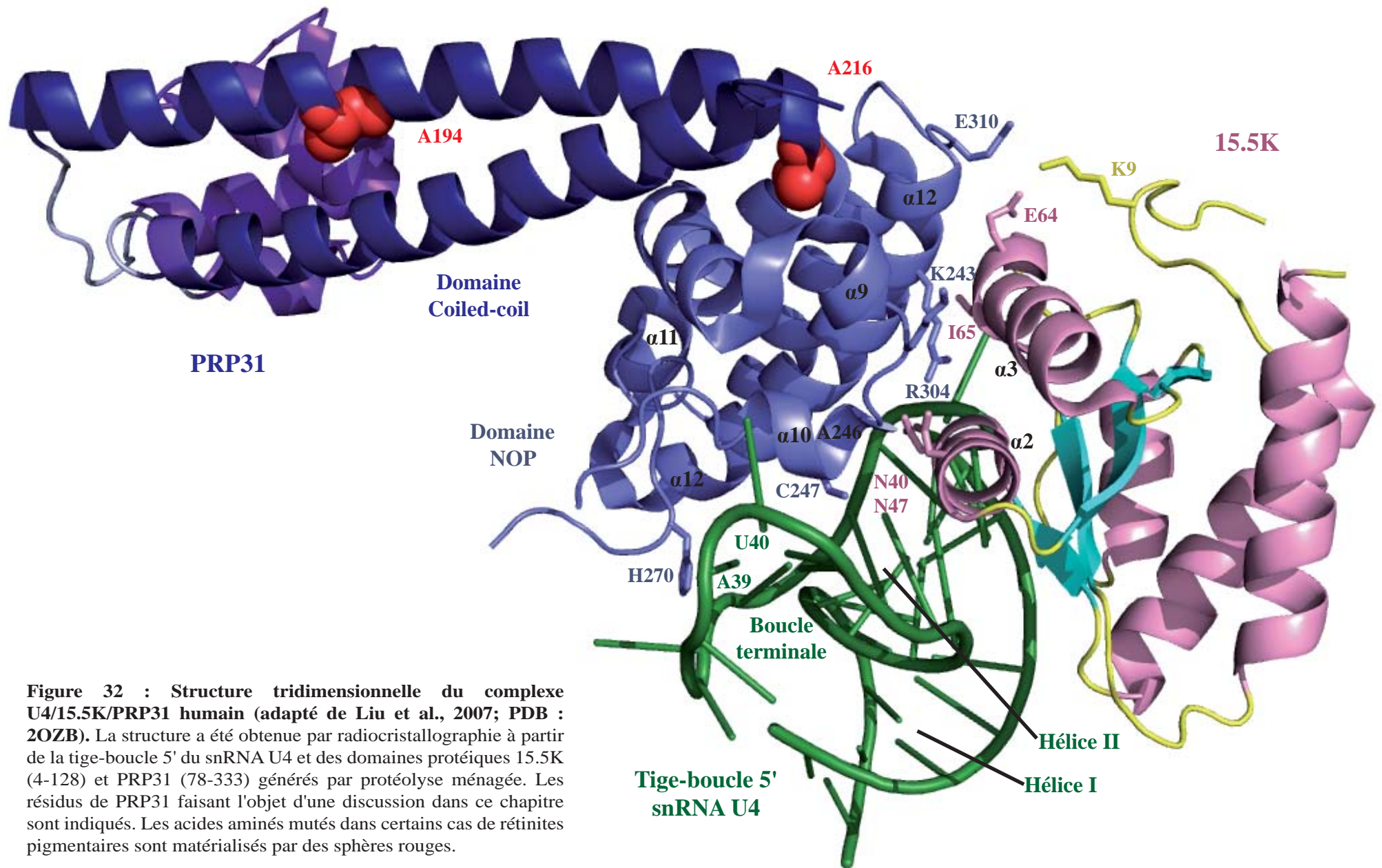
L'assemblage du domaine 5' s'effectue après le ré-import de la snRNP U4 dans le noyau, selon un mécanisme ordonné qui semble se dérouler essentiellement au niveau des CB. Il a ainsi été montré que la fixation de la protéine 15.5K sur le motif en *K-turn* du snRNA U4, est nécessaire pour le recrutement des autres protéines de la particule (Nottrott *et al.*, 2002). *In vitro*, la protéine PPR31 peut ainsi être recrutée sur un pré-complexe minimal formé par la protéine 15.5K et la partie 5' du snRNA U4, tandis que la fixation du complexe CYPH/PRP3/PRP4 nécessite la formation de l'hélice II intermoléculaire entre les snRNA U4 et U6 (Nottrott *et al.*, 2002). Ces résultats suggèrent donc que l'assemblage des protéines PRP31 et CYPH-PRP3-PRP4 repose à la fois sur la reconnaissance de déterminants ARN et protéines. Il a notamment été montré que la fixation de la protéine PRP31 sur la snRNP U4 implique des contacts avec la tige-boucle I du snRNA U4, ainsi qu'avec des résidus situés à sa base (Nottrott *et al.*, 2002; Schultz *et al.*, 2006a). Comme nous l'avons vu précédemment, l'une des caractéristiques du motif en *K-turn* contenu dans cette tige-boucle est de posséder une hélice II courte, puisque celle-ci ne comporte que deux appariements, tandis que les motifs C/D en comptent 3 (**Fig. 21**). Par une approche de mutagenèse dirigée, l'équipe de R. Lührmann a mis en évidence que cette caractéristique constitue un paramètre essentiel pour le recrutement de la protéine PRP31, puisque l'ajout d'une seule paire de bases dans l'hélice II abolit la formation du complexe trimérique ARN-15.5K-PRP31 (Schultz *et al.*, 2006a). De même, le raccourcissement de l'hélice I du *K-turn*, qui compte de façon invariable 8 paires de bases dans les snRNA U4, s'avère également délétère pour la fixation de PRP31 (Schultz *et al.*, 2006a).



**Figure 31 : Structure tridimensionnelle du domaine 3' de la snRNP U4 humaine (adapté de Leung *et al.*, 2011; PDB : 2Y9D).** La structure a été résolue par radio-cristallographie à partir de la partie 3' du snRNA U4 et des protéines SmB/B', SmD1, SmD2, SmD3, SmE, SmF et SmG. **(A)** Vue latérale du complexe. L'heptamère de protéines Sm reconnaît la séquence simple brin 5'-AAUUUUUGA-3', située entre les tiges-boucles II et III. **(B)** Vue du canal central de la structure en anneau formée par les protéines Sm autour de l'ARN.

En 2007, la résolution de la structure tridimensionnelle d'un complexe formé entre la tige-boucle 5' du snRNA U4, la protéine 15.5K et une partie de la protéine PRP31 a permis de mieux comprendre les bases moléculaires de cette interaction (Liu *et al.*, 2007b) (**Fig. 32**). En effet, les hélices  $\alpha 10$  et  $\alpha 12$  du domaine NOP établissent plusieurs contacts avec l'hélice II du motif en *K-turn*, tandis que la boucle située entre les hélices  $\alpha 10$  et  $\alpha 11$  interagit avec la boucle terminale surmontant le motif en *K-turn* (résidus 36 à 40). De façon intéressante, cette interaction, qui repose sur de nombreux contacts de van der Waals et qui implique plusieurs molécules d'eau, n'engage qu'une seule interaction spécifique, formée entre la cystéine 247 et les bases C41 et G43 de l'hélice II du *K-turn*. De plus, la boucle terminale est stabilisée par l'interaction du nucléotide U40 avec une zone hydrophobe de la protéine, et par l'empilement de l'histidine 270 avec la base du résidu A39 et la paire de bases terminale de l'hélice II. D'autre part, en utilisant une approche de modélisation moléculaire, les auteurs ont montré que l'hélice  $\alpha 10$  agit comme une butée, permettant de « mesurer » la longueur de l'hélice II du *K-turn*. En effet, l'ajout d'une paire de bases dans cette hélice provoque un encombrement stérique délétère pour la mise en place des contacts décrits précédemment. L'hélice  $\alpha 10$  de PRP31 pourrait ainsi constituer un élément important pour la discrimination du motif en *K-turn* du snRNA U4, afin d'assurer la spécificité d'assemblage de la snRNP U4. Malgré l'homologie de séquence relevée entre les domaines NOP de Prp31p/PRP31 et des protéines des snoRNP à boîtes C/D, ce résultat suggère donc que le mode de reconnaissance de l'ARN par les protéines Nop56p/NOP56 et Nop58p/NOP58 doit différer légèrement, afin de permettre l'accommodation à une hélice II plus longue. De plus, la structure du complexe révèle qu'une autre surface de la protéine PRP31, formée par les hélices  $\alpha 9$  et  $\alpha 12$ , établit des interactions avec les hélices  $\alpha 2$  et  $\alpha 3$  de la protéine 15.5K (**Fig. 32**). Ces contacts impliquent un réseau d'interactions entre les résidus K9, N40, N47, E64, I65 de la protéine 15.5K et les résidus K243, A246, R304, E310 de la protéine PRP31. De façon intéressante, les acides aminés de PRP31 impliqués dans cette interaction, sont extrêmement bien conservés au sein des protéines NOP56 et NOP58, suggérant que le mode de reconnaissance de la protéine Snu13p/15.5K par Prp31p/PRP31 et par les protéines des snoRNP à boîtes C/D pourrait impliquer les mêmes déterminants. Le domaine NOP partagé par ces différentes protéines présente donc des caractéristiques uniques et constitue un module à part entière, dédié à la reconnaissance d'interfaces ARN-protéines.

Comme nous l'avons déjà mentionné, le complexe CYPH-PRP3-PRP4 est recruté suite à l'appariement des snRNA U4 et U6, et constitue donc une sous-unité spécifique de la di-



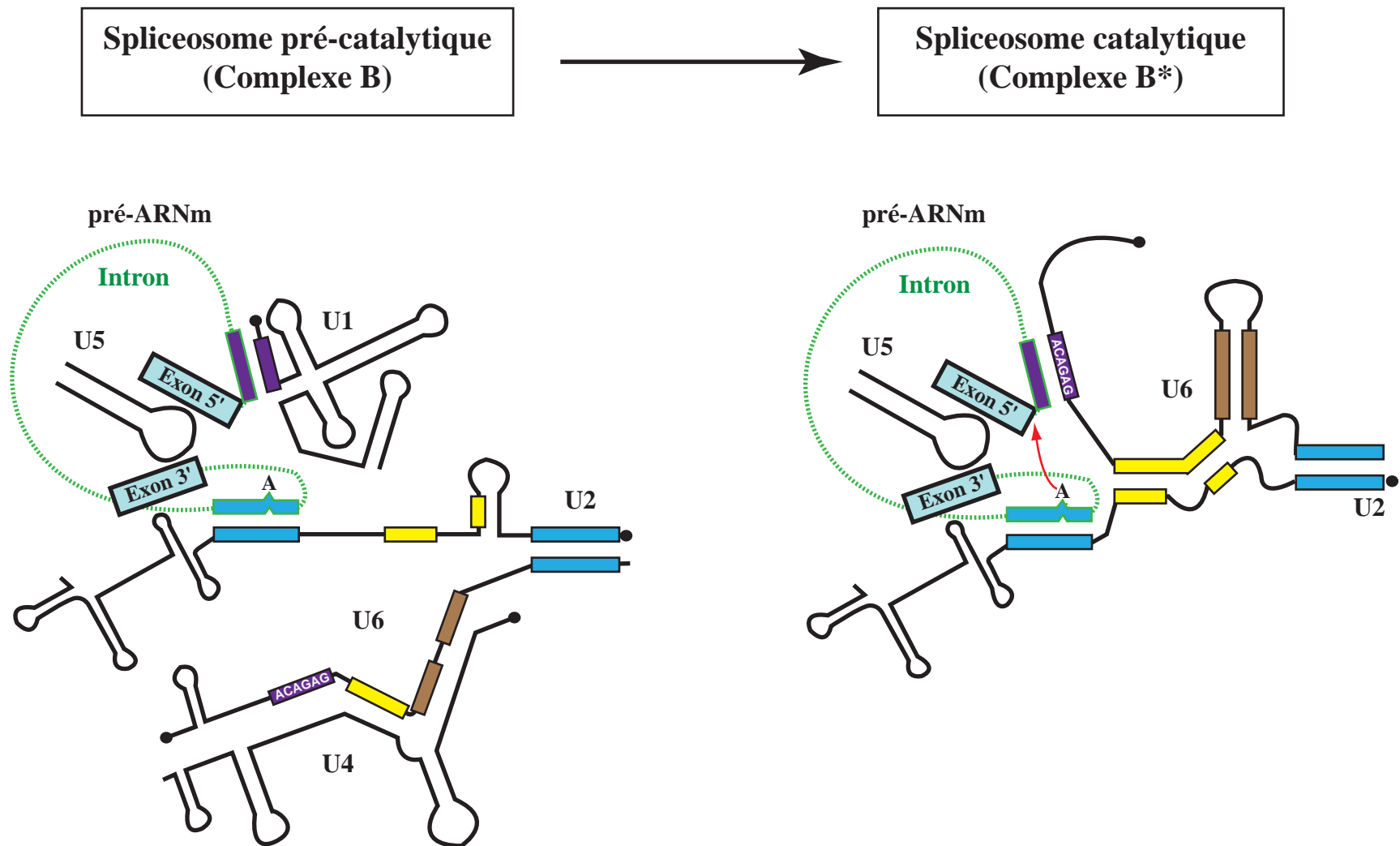
**Figure 32 : Structure tridimensionnelle du complexe U4/15.5K/PRP31 humain (adapté de Liu et al., 2007; PDB : 2OZB).** La structure a été obtenue par radiocristallographie à partir de la tige-boucle 5' du snRNA U4 et des domaines protéiques 15.5K (4-128) et PRP31 (78-333) générés par protéolyse ménagée. Les résidus de PRP31 faisant l'objet d'une discussion dans ce chapitre sont indiqués. Les acides aminés mutés dans certains cas de rétinites pigmentaires sont matérialisés par des sphères rouges.

UsnRNP (**Fig. 29**). Sa présence au sein de la particule implique notamment une interaction entre le facteur PRP3 et l'hélice II intermoléculaire du duplex snRNA U4/U6 (Nottrott *et al.*, 2002). Bien qu'aucune autre interaction n'ait été clairement décrite entre ces facteurs et les composants protéiques de la di-snRNP, le recrutement du complexe CYPH-PRP3-PRP4 pourrait impliquer des contacts directs avec la protéine 15.5K. En effet, l'équipe de R. Lührmann a montré que la mutation de certains résidus de surface, situés notamment sur les hélices  $\alpha 4$  et  $\alpha 5$  de la protéine 15.5K, perturbe le recrutement du complexe trimérique sur la RNP (Schultz *et al.*, 2006b). De façon intéressante, ces acides aminés sont situés à l'opposé de la surface d'interaction avec PRP31 et pourraient donc constituer des déterminants importants pour la stabilisation du complexe CYPH-PRP3-PRP4 au sein de la particule. D'autre part, les différentes expériences d'assemblage réalisées *in vitro*, indiquent que le recrutement de la protéine PRP31 et du complexe CYPH-PRP3-PRP4 constituent deux événements totalement indépendants (Nottrott *et al.*, 2002; Schultz *et al.*, 2006b).

Enfin, plusieurs études ont mis en évidence que l'assemblage correct des protéines spécifiques de la snRNP U4 et de la di-snRNP U4/U6 constitue un élément déterminant pour l'acquisition de la fonctionnalité. Chez l'homme, il a été montré que la protéine p110/SART3, qui joue un rôle clé dans la formation de l'appariement entre les snRNA U4 et U6, interagit à la fois avec l'ARN et avec la protéine PRP3 (Bell *et al.*, 2002; Medenbach *et al.*, 2004; Stanek *et al.*, 2003). La formation de la tri-snRNP U4/U6.U5 implique également le domaine 5' de la snRNP U4, puisque les protéines humaines PRP31 et PRP3 interagissent avec la protéine PRP6 de la snRNP U5 (Liu *et al.*, 2006), et que la protéine PRP3 interagit avec la protéine SNU66, qui est recrutée spécifiquement lors de la formation de la tri-snRNP (van Nues & Beggs, 2001).

## 2.2. Fonction de la snRNP U4 dans le mécanisme d'épissage des ARNm

Une fois formée, la tri-snRNP U4/U5.U6 est incorporée au sein du spliceosome, conduisant à la transition vers le complexe catalytique B\*, actif pour la première étape de la réaction d'épissage (**Fig. 33**). Ce processus implique des réarrangements conformationnels majeurs, en particulier la formation de nouveaux appariements intermoléculaires entre les UsnRNA et le pré-ARNm (Wahl *et al.*, 2009). Au sein du complexe pré-catalytique B, le snRNA U5 établit



**Figure 33 : Représentation schématique de l'étape d'activation catalytique du spliceosome chez les mammifères (adapté de Wahl *et al.*, 2009).** Lors de la transition entre les complexes B et B\*, le duplex U4/U6 est dissocié au profit d'un nouvel appariement entre les snRNA U6 et U2. De plus, la boîte 5'-ACAGAG-3' du snRNA U6 s'apparie avec le site 5' d'épissage à la place du snRNA U1. Les rectangles de même couleur représentent les régions complémentaires impliquées dans la formation d'appariements inter- ou intramoléculaires au sein du site actif du spliceosome. La flèche rouge symbolise l'attaque nucléophile à l'origine de la première étape de la réaction d'épissage. Le symbole ● représente les extrémités 5' modifiées des UsnRNA.



des interactions avec les exons 5' et 3', afin de permettre leur rapprochement dans l'espace, tandis que la région 3' du snRNA U6 s'apparie avec l'extrémité 5' du snRNA U2. Dans un second temps, la mise en place du complexe catalytique B\* implique la dissociation du duplex U4/U6, afin de libérer la séquence conservée ACAGAG du snRNA U6, qui interagit avec le site 5' d'épissage et déplace le snRNA U1 initialement apparié. Chez la levure *S. cerevisiae*, ces événements de remodelage impliquent notamment la protéine Prp28p, qui facilite l'échange entre les snRNA U1 et U6, ainsi que le facteur Brr2p, qui assure la dissociation du duplex U4/U6 et la formation d'un appariement étendu du snRNA U6 avec le snRNA U2. Le snRNA U4 joue donc un rôle majeur au cours de ce processus, puisqu'il permet de délivrer le snRNA U6 sous une forme inactive, assurant ainsi le déroulement ordonné des événements de mise en place du site actif du spliceosome. A la fin de la réaction d'épissage, les constituants des snRNP U4, U6 et U5 sont recyclés au niveau des CB, afin de reformer une tri-snRNP fonctionnelle pour un nouveau cycle d'épissage (Stanek *et al.*, 2008) (**Fig. 10**).

### **2.3. Des mutations dans le gène codant PRP31 à l'origine de rétinites pigmentaires**

Les rétinites pigmentaires (*retinitis pigmentosa*) constituent un groupe de maladies génétiques touchant la vision. Elles se caractérisent par une dégénérescence graduelle des cellules de la rétine conduisant à une perte progressive de la vision nocturne et de la vision latérale, puis à une perte de la vision centrale. La plupart des mutations recensées touchent des gènes exprimés spécifiquement dans les cellules photoréceptrices. De façon surprenante, des mutations dans 3 gènes liés au processus d'épissage des pré-ARNm ont été reliées à la forme autosomale dominante de la maladie. Celles-ci affectent en effet le gène codant la protéine PRP8 de la snRNP U5, ainsi que ceux codant les protéines PRP3 et PRP31 de la di-snRNP U4/U6 (Chakarova *et al.*, 2002; McKie *et al.*, 2001; Vithana *et al.*, 2001). Ces mutations, qui ont pour effet de perturber l'assemblage du spliceosome (Tanackovic *et al.*, 2011), ont une répercussion drastique sur l'épissage de transcrits spécifiques des cellules photoréceptrices (Mordes *et al.*, 2007). Une quarantaine de mutations différentes ont notamment été identifiées chez des patients au sein du gène PRP31 (Rio Frio *et al.*, 2008). Beaucoup d'entre elles provoquent de larges délétions dans la séquence protéique ou conduisent à une interruption prématurée de la traduction. Néanmoins, certaines de ces mutations conduisent à l'expression

de protéines mutées ponctuellement au niveau d'un seul acide aminé. C'est notamment le cas des mutations A194/E et A216/P qui concernent des résidus situés respectivement dans l'hélice 6, formant le domaine *Coiled-coil*, et dans la boucle connectant les domaines NOP et *Coiled-coil* (Liu *et al.*, 2007b) (**Fig. 30 et 32**). Il a notamment été montré que la mutation A216/P, qui perturbe l'import de la protéine PRP31 dans le noyau, affecte la prolifération cellulaire et altère la structure des CB (Deery *et al.*, 2002; Huranova *et al.*, 2009). L'analyse de la structure du complexe U4-15.5K-PRP31 a révélé que les acides aminés A194 et A216 ne sont pas impliqués dans des contacts directs avec l'ARN ou la protéine 15.5K, en revanche, ils sont tous deux exposés au solvant et pourraient établir des contacts avec d'autres partenaires (Liu *et al.*, 2007b) (**Fig. 32**). Un candidat potentiel pourrait être la protéine PRP6 de la snRNP U5, puisque des expériences de double-hybride ont démontré que l'introduction d'un acide glutamique en position 194 de la protéine PRP31 altère de façon significative son interaction avec PRP6 (Liu *et al.*, 2007b). Néanmoins, étant donné leur nature et leur localisation, une autre possibilité est que ces deux substitutions perturbent le repliement local de la protéine et affectent de façon indirecte sa fonction.

#### **2.4. Des mutations touchant le motif en K-turn du snRNA U4atac à l'origine du syndrome de Taybi-Linder**

Le syndrome de Taybi-Linder (TALS, encore appelé MOPD I pour *Microcephalic Osteodysplastic Primordial Dwarfism type I*) est une maladie autosomale récessive caractérisée par de graves problèmes de développement intra-utérins, associés à une mortalité infantile précoce (Majewski & Spranger, 1976; Majewski *et al.*, 1982; Sigaudy *et al.*, 1998). Décrit pour la première fois en 1967, ce n'est que très récemment que les bases moléculaires de cette pathologie ont été caractérisées (Edery *et al.*, 2011; He *et al.*, 2011). Plusieurs substitutions, localisées dans le snRNA U4atac et associées à un défaut d'épissage par le spliceosome mineur, semblent en effet à l'origine de son développement (Edery *et al.*, 2011; He *et al.*, 2011). L'une de ces mutations est retrouvée dans la tige-boucle III, adjacente au site de fixation des protéines Sm (G111/A), tandis que plusieurs autres sont localisées dans l'hélice I bordant le motif en *K-turn* (G30/A, G50/A/C, G51/A, C53/G, G55/A) (Edery *et al.*, 2011; He *et al.*, 2011). Il a notamment été montré qu'une mutation compensatoire, restaurant un appariement canonique de l'hélice I au sein de l'un des snRNA U4atac variant, permet de

restaurer partiellement la fonction d'épissage (He *et al.*, 2011). En perturbant la formation du motif en *K-turn*, il est donc probable que ces différentes mutations altèrent la fixation de la protéine 15.5K et le recrutement subséquent des protéines PRP31, CYPH, PRP3 et PRP4, conduisant à un défaut d'assemblage de la tri-snRNP U4atac/U6atac.U5 (Edery *et al.*, 2011).

### 3. Les snoRNP à boîtes C/D et H/ACA

Les snoRNP à boîtes C/D et H/ACA sont des ribonucléoparticules dont la spécificité d'action repose sur l'appariement spécifique du snoRNA avec un ARN cible. Le rôle principal de ces RNP est l'apposition de modifications chimiques à des sites précis du pré-ARNr (cf. Introduction, Chapitre 1, 3.2) et de certains autres ARN substrats. De plus, comme nous l'avons déjà abordé, certaines snoRNP constituent des chaperonnes essentielles pour la formation des complexes pré-ribosomiques et pour la réalisation des coupures précoces aux sites A0, A1 et A2 du pré-ARNr (cf. Introduction, Chapitre 1, 3.3 & 3.4). A la différence des UsnRNP spliceosomales, les snoRNP sont conservées chez les archées sous la forme de sRNP, suggérant que l'apparition des systèmes de modifications guidées par des ARN non codants constitue un événement ancien, antérieur à 2 milliards d'années et à la séparation archées/eucaryotes. La composition, la formation et la fonction de ces particules ont fait l'objet de nombreuses recherches. Au cours de la dernière décennie, de nombreuses données structurales ont notamment été obtenues à partir des sRNP d'archées et ont permis de mieux comprendre leur fonctionnement à l'échelle moléculaire.

#### 3.1. Les snoRNP à boîtes C/D

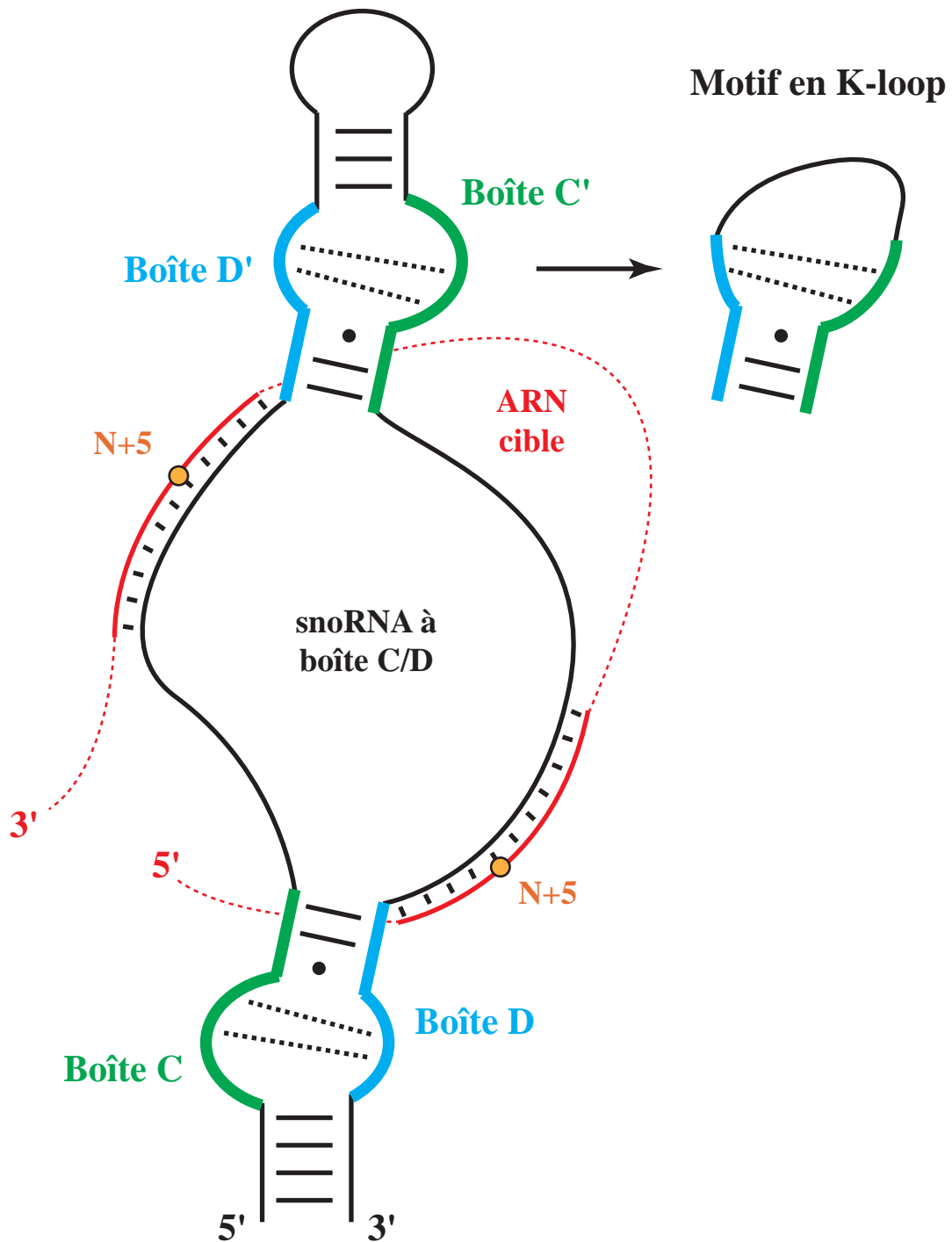
Les snoRNP à boîtes C/D catalysent des réactions de 2'-O-méthylation. Ces particules s'organisent autour d'une base commune avec la snRNP U4, formée par l'association d'un ARN porteur d'un motif en *K-turn* et de la protéine Snu13p/15.5K. Au sein de cette famille, la snoRNP U3 présente des caractéristiques structurales et fonctionnelles singulières. Cette particule essentielle, qui intervient exclusivement dans le cadre de la maturation nucléolytique des ARNr, est retrouvée chez tous les eucaryotes. Elle a fait l'objet de recherches approfondies et constitue un remarquable modèle pour l'étude des complexes

ribonucléoprotéiques et de leur biogenèse. Nous reviendrons en détail sur la snoRNP U3 après avoir présenté les snoRNP à boîtes C/D.

### 3.1.1. Organisation et fonction des snoRNP à boîtes C/D

#### 3.1.1.1. Les snoRNA à boîtes C/D

Les snoRNA à boîtes C/D, identifiés il y a plus de 30 ans par fractionnement d'ARN isolés du nucléole, sont des petites molécules de taille variant entre 70 et 300 nucléotides environ (Lestrade & Weber, 2006; Nakamura *et al.*, 1968; Piekna-Przybylska *et al.*, 2007; Weinberg & Penman, 1968). Comme nous l'avons vu précédemment, ces ARN sont caractérisés par la présence de séquences conservées, appelées boîte C (5'-RUGAUGA-3') et boîte D (5'-CUGA-3'), qui s'assemblent pour former un motif en *K-turn*, responsable du recrutement de la protéine Snu13p/15.5K chez les eucaryotes et de la protéine L7Ae chez les archées (cf. Introduction, Chapitre 2, 1.1) (**Fig. 23E**). Les boîtes C et D sont respectivement localisées en 5' et en 3' de l'ARN et la formation du motif C/D conduit à la mise en place d'une structure en tige-boucle terminale, clampant l'ARN à ces deux extrémités (Bachelier *et al.*, 1995) (**Fig. 34**). La plupart des snoRNA possèdent également deux séquences supplémentaires, les boîtes C' et D', qui présentent un degré de conservation plus faible que les boîtes C et D, mais qui s'assemblent également pour former un motif C'/D' structuré en *K-turn* (Kiss-Laszlo *et al.*, 1998; Tycowski *et al.*, 1996). Ces deux séquences occupent une position interne, la boîte D' est retrouvée dans la moitié 5' de l'ARN, tandis que la boîte C' est située dans la partie 3' (Kiss-Laszlo *et al.*, 1998) (**Fig. 34**). Les boîtes C' et D' sont généralement séparées par 3 à 10 nucléotides non appariés ou par une séquence de plus grande taille formant une tige-boucle interne (Kiss-Laszlo *et al.*, 1998). Chez les archées, les sRNA à boîtes C/D ont une taille plus petite, généralement comprise entre 50 et 70 nucléotides et présentent une meilleure conservation du couple de boîtes C'/D' (Gaspin *et al.*, 2000; Omer *et al.*, 2000). Une étude menée chez les espèces du genre *Pyrococcus* a néanmoins démontré que les boîtes C' et D' sont généralement séparées par une séquence courte de 3 ou 4 nucléotides, conduisant à la formation d'un motif de type *K-loop*, également retrouvé au sein des sRNA à boîtes H/ACA (Nolivos *et al.*, 2005) (**Fig. 23H**). Chez les eucaryotes, la distance entre les motifs C/D et C'/D' varie, en fonction des ARN, d'une à plusieurs dizaines de nucléotides, en revanche chez



**Figure 34 : Représentation schématique d'un snoRNA à boîtes C/D.** Les boîtes C, D et C', D' s'assemblent pour former des motifs structurés en K-turn. Lorsque la séquence entre les boîtes C' et D' est courte (3 à 4 nt), le motif se structure en K-loop. Ce cas de figure est parfois retrouvé chez les eucaryotes, il est en revanche très fréquent chez les archées. Les ARN cibles s'apparient avec les régions guides, situées directement en 5' des boîtes D et D'. La 2'-O-méthylase NOP1/Fibrillarine cible de façon invariante le 5<sup>ème</sup> résidu apparié en amont des boîtes D et D'.

beaucoup d'archées cet espacement présente une longueur moyenne de 12 ou 13 nucléotides (Tran *et al.*, 2005).

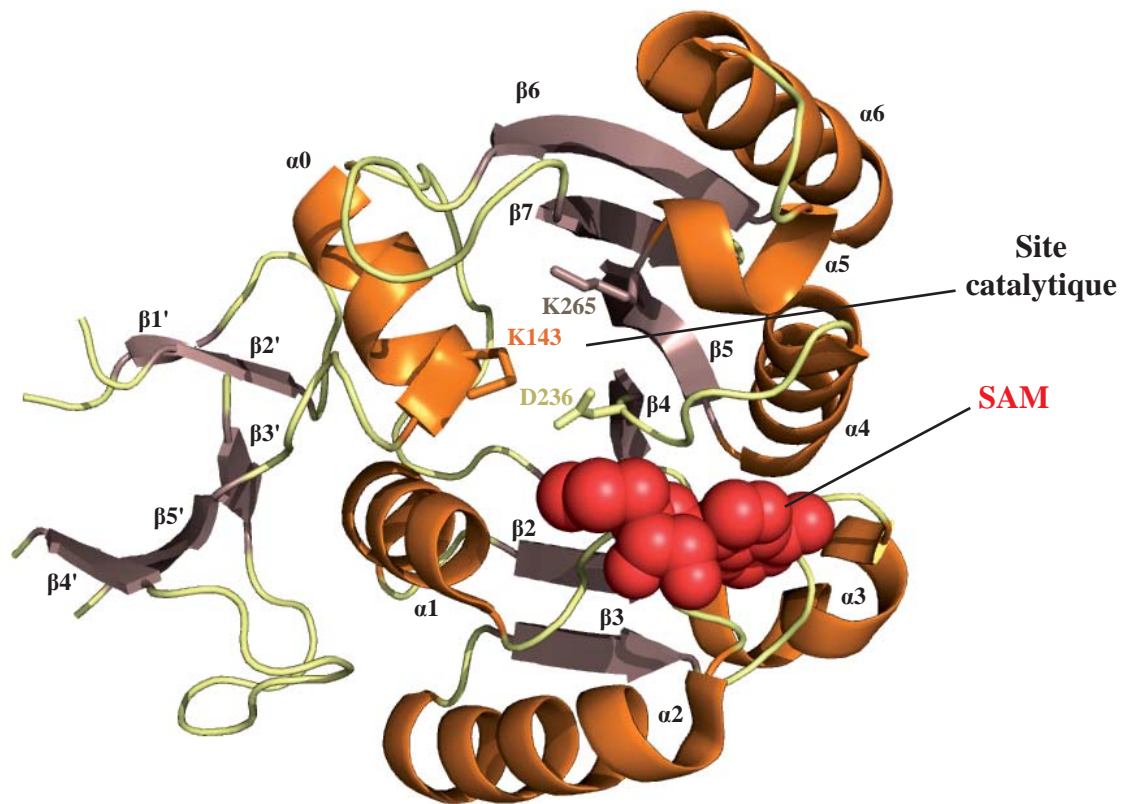
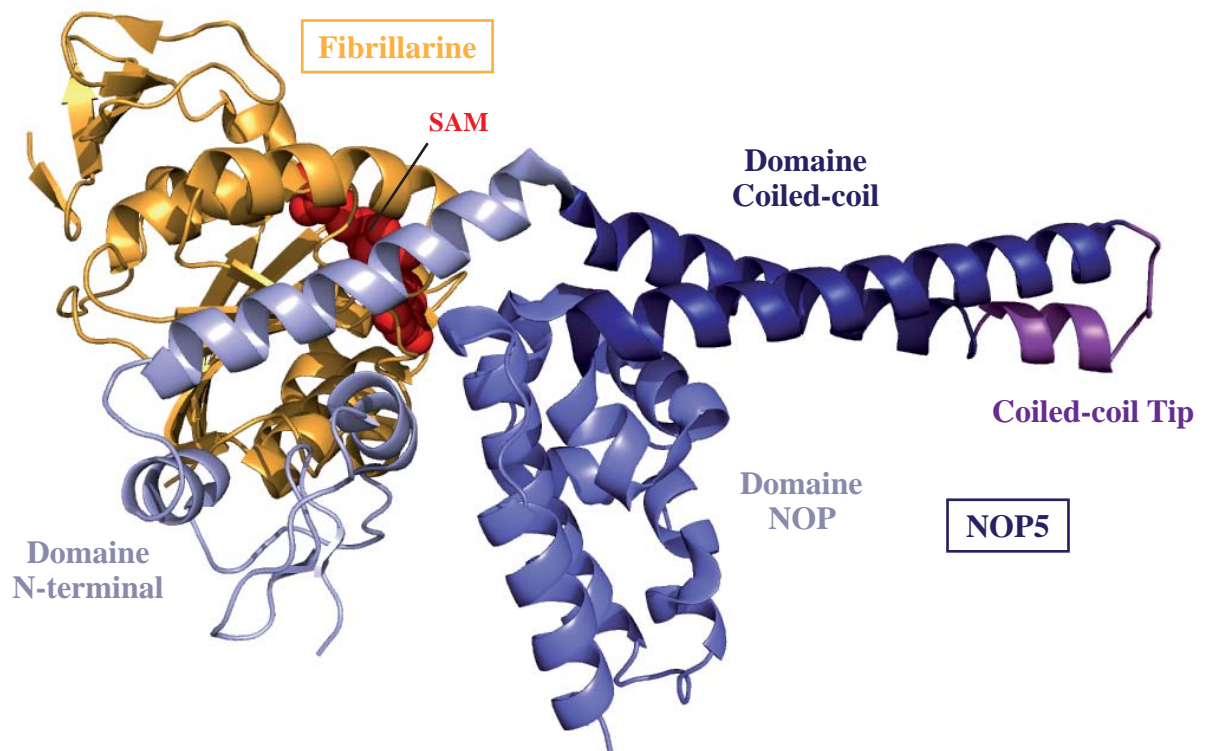
L'appariement d'un snoRNA avec son ARN cible conduit à la formation d'une double-hélice parfaite, d'une longueur de 10 à 20 paires de bases (Cavaille *et al.*, 1996; Kiss-Laszlo *et al.*, 1996). Ces duplex ARN guide/ARN substrat impliquent les régions situées directement en amont des boîtes D et D', et fournissent une information structurale essentielle pour la 2'-O-méthylase Nop1p/Fibrillarine. Cette dernière cible en effet de façon invariante le 5<sup>ème</sup> nucléotide apparié en amont des boîtes D et D', en suivant la règle dite « N plus cinq » (Bachellerie & Cavaille, 1997; Filipowicz *et al.*, 1999; Kiss-Laszlo *et al.*, 1996; Weinstein & Steitz, 1999). Au-delà de leur rôle fonctionnel, de nombreuses études ont également démontré l'importance des boîtes C et D pour la biogenèse des snoRNP. Ces séquences sont essentielles pour la stabilité des snoRNA (Caffarelli *et al.*, 1996; Huang *et al.*, 1992; Li & Fournier, 1992), pour la triméthylation de leur coiffe (Speckmann *et al.*, 2000), pour leur localisation nucléolaire (Lange *et al.*, 1998; Narayanan *et al.*, 1999b; Samarsky *et al.*, 1998; Speckmann *et al.*, 1999) et pour l'assemblage des particules (Caffarelli *et al.*, 1996; Cavaille & Bachellerie, 1996; Watkins *et al.*, 1996; Xia *et al.*, 1997). Enfin, quelques études menées chez les rongeurs et chez *S. cerevisiae* ont été réalisées afin d'identifier la présence de résidus modifiés au sein des snoRNA à boîtes C/D (Massenet *et al.*, 1998). Chez le rat, le snoRNA U3 contient une 2'-O-méthyladénine située en aval de la coiffe, ainsi que deux pseudouridines localisées au niveau de la boîte A' (Myslinski *et al.*, 1990; Reddy *et al.*, 1985). Ces deux dernières modifications, qui ciblent des nucléotides impliqués dans l'appariement entre le snoRNA U3 et le pré-ARNr, pourraient ainsi stabiliser la formation du duplex entre les deux ARN (Massenet *et al.*, 1998). A titre de comparaison, des travaux réalisés au laboratoire ont en revanche démontré l'absence de pseudouridines au sein du snoRNA U3 de levure (Massenet *et al.*, 1998). Chez la souris, une pseudouridine, une 2'-O-méthyladénine et une 2'-O-méthyluridine, ont également été identifiées au sein du snoRNA U8 (Kato & Harada, 1984). En revanche, l'identité des facteurs impliqués dans ces modifications n'a pas été clairement établie. Une hypothèse envisageable est que ces modifications soient catalysées par les scaRNP au cours des étapes de maturation des snoRNP au sein des CB.

### 3.1.1.2. Les protéines des snoRNP à boîtes C/D

La protéine L7Ae chez les archées et la protéine Snu13p/15.5K chez les eucaryotes constituent avec le snoRNA, la plateforme de base pour l'assemblage de la snoRNP à boîtes C/D. Chez la levure, la déplétion de la protéine Snu13p conduit à une forte diminution du taux de snoRNA à boîtes C/D et à un défaut de maturation des ARNr (Watkins *et al.*, 2000). En plus de ce facteur, les snoRNP à boîtes C/D eucaryotes contiennent 3 protéines supplémentaires : la 2'-O-méthylase Nop1p/Fibrillarine et les protéines Nop56p/NOP56 et Nop58p/NOP58. Chez les archées ces deux derniers facteurs sont remplacés par un homologue unique, la protéine NOP5.

#### 3.1.1.2.1. La 2'-O-méthylase Nop1p/Fibrillarine

La Fibrillarine, appelée Nop1p chez la levure, a été identifiée au sein des centres fibrillaire dense du nucléole eucaryote (Ochs *et al.*, 1985; Schimmang *et al.*, 1989) (cf. Introduction, Chapitre 1, 1.4.1) (**Fig. 5C**). Cette protéine, essentielle à la viabilité cellulaire et à la production des ARNr, est conservée chez les eucaryotes et chez les archées (Amiri, 1994; Jansen *et al.*, 1991; Tollervey *et al.*, 1991; Tollervey *et al.*, 1993). Chez la levure *S. cerevisiae*, sa déplétion inhibe le méthylation des ARNr et affecte principalement la production de l'ARNr 18S en inhibant les clivages précoces aux sites A0, A1 et A2 (Tollervey *et al.*, 1991). Elle conduit également à une diminution du taux de nombreux snoRNA, à l'exception notoire du snoRNA U3 (Lafontaine & Tollervey, 2000; Tollervey *et al.*, 1991). Chez l'homme, sa déplétion est également associée à une diminution spécifique du taux de snoRNA à boîtes C/D (Watkins *et al.*, 2004). Les structures de la Fibrillarine et de certains de ses domaines ont été résolues pour plusieurs espèces d'archées et chez l'humain (Aittaleb *et al.*, 2003; Deng *et al.*, 2004; Wang *et al.*, 2000; Wu *et al.*, non publié) (**Fig. 35A**). Cette protéine appartient à la famille des méthyltransférases SAM-dépendantes de classe I (Schubert *et al.*, 2003). Elle contient un domaine C-terminal caractéristique, composé de 7 brins  $\beta$  flanqués par 6 hélices  $\alpha$ , qui porte l'activité catalytique. A son extrémité N-terminale, ce domaine est connecté, par l'intermédiaire d'une hélice  $\alpha_0$ , à un petit domaine conservé, formé de 5 brins  $\beta$ . La Fibrillarine est responsable du transfert d'un groupement méthyle depuis un donneur S-Adénosine-(L)-Méthionine (SAM) vers un ribose cible (Aittaleb *et al.*,

**A****B**

**Figure 35 : Structures tridimensionnelles des protéines Fibrillarine et NOP5.** (A) Structure de la protéine humaine Fibrillarine, obtenue par radiocristallographie (Wu *et al.*, non publié; PDB : 2IPX). Les résidus KDK formant la triade catalytique sont indiqués. La structure du domaine N-terminal GAR, retrouvé spécifiquement chez les eucaryotes, n'a pas été résolue. Le donneur de méthyle, S-Adénosyl-L-Méthionine (SAM), est représenté par des sphères rouges. Le site actif contient la triade catalytique, formée par les résidus K143, D236 et K265. (B) Structure du complexe formé par les protéines Fibrillarine et NOP5 chez l'archaea *Archaeoglobus fulgidus*, obtenue par radiocristallographie (Aittaleb *et al.*, 2003; PDB : 1NT2).



2003; Wang *et al.*, 2000). Son activité semble reposer sur une triade catalytique K-D-K conservée et localisée dans la poche de fixation du substrat (Feder *et al.*, 2003). L'aspartate de cette tétrade joue un rôle primordial dans la catalyse, il pourrait en effet être responsable de la déprotonation du groupement 2'-OH du ribose et/ou de la stabilisation du co-substrat SAM (Aittaleb *et al.*, 2003). De façon intéressante, la Fibrillarine n'est active que lorsqu'elle est intégrée au sein d'une RNP. Aucune activité n'a en effet été observée à l'état isolé, tandis que son fonctionnement a été caractérisé au sein de particules de levure purifiées (Galardi *et al.*, 2002) ou de particules d'archées reconstituées à partir de protéines recombinantes (Galardi *et al.*, 2002; Omer *et al.*, 2002) (V. Senty-Ségault, Communication personnelle). De plus, il a été montré chez les archées que le domaine C-terminal est également responsable d'une interaction avec la protéine NOP5 (Aittaleb *et al.*, 2003) (**Fig. 35B**). *In vitro*, la protéine seule n'interagit pas ou très peu avec les snoRNA (Fatica *et al.*, 2000a; Omer *et al.*, 2002), tandis que des expériences de pontages aux UV ont révélées l'existence de contacts directs avec l'ARN à l'intérieur des RNP (Cahill *et al.*, 2002; Granneman *et al.*, 2009). Plus récemment, la résolution de structures cristallographiques de sRNP à boîtes C/D d'archées a confirmé l'existence d'interactions entre la Fibrillarine et l'hélice intermoléculaire formée par le sRNA guide et son substrat (Lin *et al.*, 2011; Ye *et al.*, 2009). Enfin, chez les eucaryotes, la Fibrillarine contient un domaine N-terminal GAR (*Glycine and Arginine Rich*) qui n'est pas retrouvé dans les protéines d'archées.

### 3.1.1.2.2. Les protéines Nop56p/NOP56 et Nop58p/NOP58

Les protéines Nop56p et Nop58p ont été identifiées chez la levure *S. cerevisiae* suite à un crible de létalité synthétique avec des mutants de la protéine Nop1p (Gautier *et al.*, 1997). Ces deux protéines essentielles partagent une forte homologie de séquence (45% d'identité) et pourraient tirer leur origine dans un événement de duplication de gène qui se serait produit uniquement chez les eucaryotes (Gautier *et al.*, 1997). Néanmoins, des expériences de complémentations croisées, réalisées chez la levure, ont montré que ces deux protéines paralogues ne peuvent se substituer l'une à l'autre, suggérant l'acquisition de fonctions distinctes (Gautier *et al.*, 1997). En effet, malgré leur présence commune au sein des snoRNP à boîtes C/D, leur déplétion conduit à des phénotypes différents (Lafontaine & Tollervey, 1999; Lafontaine & Tollervey, 2000). L'absence de Nop58p est associée à une forte

diminution du taux de snoRNA à boîtes C/D et à un défaut de production de l'ARNr 18S par inhibition des clivages aux sites A0, A1 et A2 (Lafontaine & Tollervey, 1999). La déplétion de Nop56p conduit en revanche à un effet moins prononcé sur la production de l'ARNr 18S et n'affecte pas le taux des snoRNA (Lafontaine & Tollervey, 2000). En revanche, contrairement à la Fibrillarine, la déplétion de Nop56p et Nop58p conduit à un impact moins drastique sur le taux de 2'-O-méthylation des ARNr (Lafontaine & Tollervey, 1999; Lafontaine & Tollervey, 2000). Chez l'homme la déplétion de NOP56 et NOP58 affecte l'accumulation des snoRNA à boîtes C/D avec un effet plus prononcé pour la protéine NOP58 (Watkins *et al.*, 2004). Comme nous l'avons déjà évoqué, ces deux protéines présentent une forte homologie de séquence avec la protéine Prp31p/PRP31 de la snRNP U4 (**Fig. 30**). Cette homologie concerne essentiellement le domaine NOP qui, au sein de la snRNP U4, constitue un motif de reconnaissance ARN-protéine. En revanche, les protéines Nop56p/NOP56 et Nop58p/NOP58 possèdent un domaine C-terminal spécifique, formé de répétition KKE/D. Alors qu'il n'est pas requis pour la fonction et la localisation des deux protéines chez la levure (Gautier *et al.*, 1997), ce domaine semble néanmoins contenir des déterminants d'interaction pour l'hyperméthylase Tgs1p (Mouaikel *et al.*, 2002). De plus, une étude récente a révélé qu'il jouait à la fois le rôle de signal de localisation nucléaire (NLS) et nucléolaire (NoLS) pour les protéines humaines (Pradet-Balade *et al.*, 2011). De façon intéressante, des motifs similaires sont retrouvés dans la pseudouridyl-synthase Cbf5p des snoRNP à boîtes H/ACA et dans les protéines humaines MAP1A et MAP1B chez qui ils permettent l'interaction avec les microtubules (Gautier *et al.*, 1997; Noble *et al.*, 1989). Chez les archées, la structure cristallographique de la protéine NOP5 a été résolue en complexe avec la Fibrillarine ou intégrée au sein de la sRNP (Aittaleb *et al.*, 2003; Lin *et al.*, 2011; Oruganti *et al.*, 2007; Xue *et al.*, 2010; Ye *et al.*, 2009) (**Fig. 35B**). Le domaine NOP, situé dans la partie C-terminale de la protéine, possède également la capacité d'interagir avec le sRNA et avec la protéine L7Ae (Ye *et al.*, 2009). Ce résultat confirme l'analogie avec la protéine Prp31p/PRP31, puisque la protéine NOP5 et ses deux homologues eucaryotes, sont également incapables d'interagir indépendamment avec l'ARN ou avec la protéine L7Ae/Snu13p/15.5K (Omer *et al.*, 2002; Watkins *et al.*, 2002). D'autre part, les analyses structurales ont montré que la protéine NOP5 interagit avec l'hélice II du motif en *K-turn* au sein de la sRNP (Lin *et al.*, 2011; Xue *et al.*, 2010; Ye *et al.*, 2009). Auparavant, l'équipe de R. Lührmann avait déjà démontré chez l'humain que la conservation de la séquence de cette hélice est primordiale pour le recrutement des protéines NOP56 et NOP58 sur l'ARN (Watkins *et al.*, 2002). Comme dans la protéine PRP31, le motif NOP de la protéine NOP5 est précédé par un domaine *Coiled-coil*

coiffé d'un domaine *Tip*, qui permet la dimérisation de la protéine (Ye *et al.*, 2009). Enfin, la partie N-terminale constitue un domaine à part entière, composé de 3 hélices  $\alpha$  et de deux brins  $\beta$ , et impliqué dans l'interaction avec la Fibrillarine (Aittaleb *et al.*, 2003).

### 3.1.1.3. Structure et fonction des snoRNP à boîtes C/D

Les snoRNP à boîtes C/D forment des structures complexes, comprenant généralement deux RNP, respectivement assemblées sur les motifs C/D et C'/D'. Le snoRNA, qui assure la reconnaissance de l'ARN substrat, constitue le module variable du complexe et permet d'adapter la spécificité d'action de la particule en fonction des ARN cibles. Comme nous l'avons abordé dans la partie précédente, ces particules exercent leur fonction au sein du nucléole lors de la maturation des ARNr. Des travaux réalisés chez les mammifères et chez le xénope, ont également démontré leur implication dans la 2'-O-méthylation du snRNA U6 au cours de sa phase de maturation nucléolaire (Ganot *et al.*, 1999; Lange & Gerbi, 2000; Tycowski *et al.*, 1998). Au sein du cerveau humain, il a également été montré que le snoRNA MBII-52 pourrait modifier spécifiquement une adénosine du pré-ARNm du récepteur de la sérotonine 5-HT<sub>2C</sub>, afin de la protéger d'une modification par édition (A  $\rightarrow$  I), catalysée par l'enzyme ADAR2 (Cavaille *et al.*, 2000; Cavaille *et al.*, 2002; Vitali *et al.*, 2005). De plus, chez certains eucaryotes, des particules à boîtes C/D, retrouvées au sein des scaRNP, participent à la maturation des autres UsnRNA au niveau des CB (Darzacq *et al.*, 2002; Jady *et al.*, 2003; Jady & Kiss, 2001). Enfin, chez les archées certains ARNt sont également la cible de sRNP à boîtes C/D (Clouet d'Orval *et al.*, 2001; Joardar *et al.*, 2008; Singh *et al.*, 2004).

#### 3.1.1.3.1. Pré-requis pour l'assemblage des particules

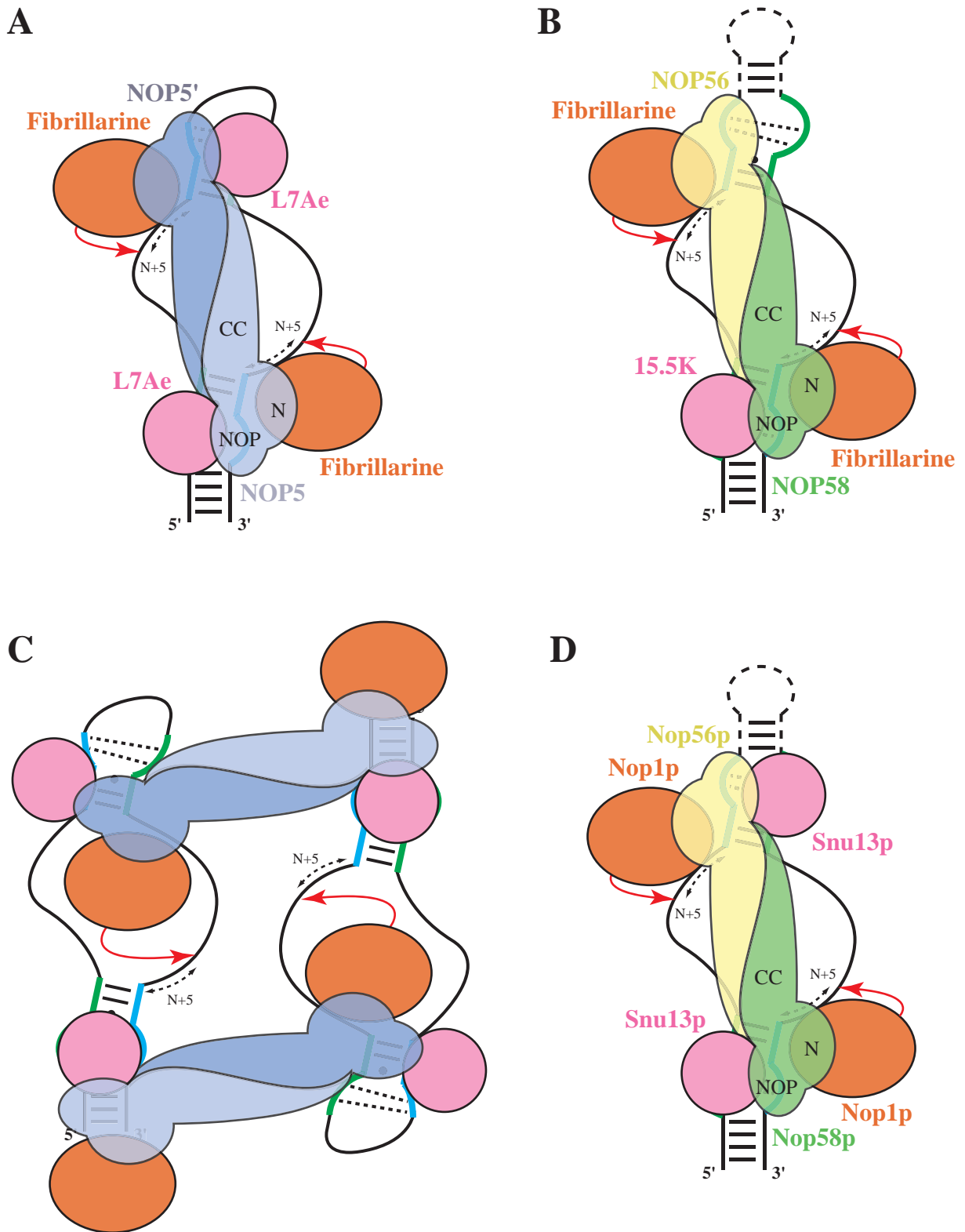
Comme cela a été démontré pour la snRNP U4, l'assemblage des snoRNP à boîtes C/D en extrait cellulaire implique la fixation préalable de la protéine Snu13p/15.5K sur le motif en *K-turn* formé par les boîtes C et D (Watkins *et al.*, 2002). Un résultat similaire a été observé chez les archées, puisque les protéines NOP5 et Fibrillarine ne sont intégrées dans la particule qu'en présence de la protéine L7Ae (Omer *et al.*, 2002). Chez la levure *S. cerevisiae*, des

expériences d'immunoprécipitation ont également permis de montrer que le recrutement de la protéine Nop56p requiert la présence de Nop1p, tandis que les protéines Nop1p et Nop58p peuvent s'assembler de façon indépendante (Lafontaine & Tollervey, 2000). Bien que la formation d'un pré-complexe Nop1p-Nop56p n'a pas été observé chez la levure (Lafontaine & Tollervey, 2000), une étude de microscopie a montré que les protéines humaines Fibrillarine et NOP56 sont associées *in vivo*, indépendamment de la snoRNP (Lechertier *et al.*, 2009). Chez les archées, les protéines Fibrillarine et NOP5 peuvent également interagir sous la forme d'un hétérodimère dont la formation semble faciliter leur incorporation au sein de la sRNP (Aittaleb *et al.*, 2003; Rashid *et al.*, 2003; Tran *et al.*, 2003). A l'inverse des protéines de levure, des expériences réalisées *in vitro* ont montré que l'association de la Fibrillarine avec le sRNA est dépendante de la protéine NOP5 (Omer *et al.*, 2002).

### 3.1.1.3.2. Modèles symétriques et asymétriques

Bien que ces premières données apportent quelques éléments de réponse, la compréhension des événements d'assemblage des particules à boîtes C/D demeure une tâche complexe et nécessite la prise en compte de la présence des motifs C/D et C'/D' sur un même ARN. Chez les archées, plusieurs études ont mis en évidence l'association d'un jeu de protéines L7Ae-NOP5-Fibrillarine par couple de boîtes (Omer *et al.*, 2002; Tran *et al.*, 2003). De plus, les réactions de 2'-O-méthylation, catalysées au niveau des motifs C/D et C'/D', se sont avérées optimales lorsque ces deux particules sont retrouvées au sein du même ARN (Rashid *et al.*, 2003; Tran *et al.*, 2003). Sur la base de ces résultats et de la capacité de dimérisation de la protéine NOP5, un modèle d'assemblage symétrique a ainsi été proposé (Aittaleb *et al.*, 2003) (**Fig. 36A**). Dans ce modèle, les deux couples de boîtes sont occupés chacun par un jeu de 3 protéines et communiquent par l'intermédiaire de la dimérisation des domaines *Coiled-coil* des deux protéines NOP5. Cette configuration repose donc sur un couplage physique et fonctionnel entre les particules et pourrait expliquer la conservation de l'espacement d'une douzaine de nucléotides observé entre les deux motifs (Tran *et al.*, 2005) (cf. paragraphe 3.1.1.1).

Chez les eucaryotes, un modèle d'assemblage asymétrique avait initialement été proposé sur la base d'études de pontages covalents et d'interférences de modifications par des agents



**Figure 36 : Représentation schématique des différents modèles proposés pour l'assemblage des snoRNP et des sRNP à boîtes C/D.** (A) Modèle d'assemblage symétrique proposé chez les archées (Aittaleb *et al.*, 2003). (B) Modèle d'assemblage asymétrique proposé chez *Xenopus laevis* (Cahill *et al.*, 2002). (C) Modèle d'assemblage en di-sRNP asymétrique proposé chez les archées (Bleichert *et al.*, 2009; Xue *et al.*, 2010). Cette configuration implique une réaction de cross-méthylation. Une version symétrique de ce modèle a également été proposée. (D) Modèle d'assemblage symétrique proposé chez la levure *Saccharomyces cerevisiae* (Qu *et al.*, 2011). Les différents domaines des protéines NOP sont mentionnés, N : N-terminal ; CC : Coiled-coil ; NOP : domaine NOP. Les flèches rouges symbolisent les réactions de 2'-O-méthylations catalysées par la Fibrillarine.

chimiques réalisées dans des ovocytes de xénopes (Cahill *et al.*, 2002; Szewczak *et al.*, 2002; Szewczak *et al.*, 2005) (**Fig. 36B**). Dans ce modèle, et conformément à l'absence fréquente d'hélice I au niveau du motif C'/D', la protéine 15.5K interagit uniquement avec le motif C/D. Les protéines NOP56 et NOP58 occupent respectivement les motifs C'/D' et C/D. Enfin, seule la Fibrillarine est associée aux deux couples de boîtes, assurant ainsi la 2'-O-méthylation par chacune des deux particules. De plus, étant donné que la formation d'une RNP au niveau du motif C/D est susceptible de protéger les extrémités 5' et 3' de l'ARN de la dégradation par des exonucléases (Weinstein & Steitz, 1999), cette répartition asymétrique des protéines NOP56 et NOP58 pourrait expliquer le faible impact de NOP56 sur la stabilité des snoRNA.

En 2009, une étude par microscopie électronique a permis de reconstituer en 3 dimensions une sRNP catalytiquement active de l'archée *Methanocaldococcus jannaschii* assemblée à partir de protéines recombinantes (Bleichert *et al.*, 2009). Les résultats obtenus ont révélé la présence de 4 jeux de protéines L7Ae-NOP5-Fibrillarine au sein d'une même RNP de poids moléculaire estimé à 366 kDa, suggérant ainsi l'association de deux sRNP traditionnelles au sein d'une di-sRNP (**Fig. 36C**). Dans le modèle structural déduit de ces résultats, l'interaction des deux protéines NOP5 par leur domaine *Coiled-coil* n'est pas réalisée entre les motifs du même ARN, mais implique un couplage physique entre les motifs de deux ARN distincts. Quatre particules seraient ainsi retrouvées au sein du même macro-complexe et pourraient ainsi assurer la modification de 4 ARN substrats. Par la suite, la résolution de la structure cristallographique d'une demi-sRNP en association avec son substrat a apporté des arguments en faveur de ce modèle (Xue *et al.*, 2010). En effet, la duplication de cette demi-sRNP, associée à une méthode de *docking* moléculaire, a permis de modéliser une di-sRNP et d'en prédire son fonctionnement. Afin de répondre à des problèmes stériques liés à l'accès au substrat par la Fibrillarine, les auteurs de cette étude proposent que la 2'-O-méthylation au niveau d'un motif donné fasse appel à la Fibrillarine située sur le second motif du même sRNA. De façon surprenante, cette configuration implique également le relargage de la protéine L7Ae qui lui est associée, afin de faciliter la fixation du substrat et le positionnement de la sous-unité catalytique.

Cependant, une étude structurale récente a apporté une explication alternative à ce modèle de *cross-methylation*, en cristallisant pour la première fois une sRNP complète, formée autour d'un ARN porteur de deux couples de boîtes et apparié simultanément à deux substrats (Lin *et*

*al.*, 2011) (**Fig. 37**). Dans ce modèle, chaque couple de boîtes est associé à un lot de protéines L7Ae-NOP5-Fibrillarine et assure la 2'-O-méthylation de son propre substrat. A la différence des structures obtenues précédemment (Xue *et al.*, 2010; Ye *et al.*, 2009), la sous-unité catalytique, formée par le domaine N-terminal de la protéine NOP5 et par la Fibrillarine, adopte un positionnement fixe et parfaitement compatible avec l'acte catalytique. De plus, en accord avec le modèle classique, la structure de cette sRNP bipartite confirme l'importance fonctionnelle du couplage physique entre les deux motifs du même ARN par l'intermédiaire des domaines *Coiled-coil* des protéines NOP5. Comme nous le verrons par la suite, si cette dernière étude n'explique pas l'association de deux particules fonctionnelles au sein d'une même di-sRNP (Bleichert *et al.*, 2009), elle répond en revanche aux problèmes d'ordre stérique liés à la catalyse et propose certainement le modèle le plus cohérent avec l'ensemble des données déjà établies.

Pour finir, une autre étude récente, réalisée chez la levure *S. cerevisiae*, a relancé l'idée d'un modèle d'assemblage symétrique pour les particules eucaryotes (Qu *et al.*, 2011) (**Fig. 36D**). A la différence des archées, les grandes distances retrouvées entre les motifs C/D et C'/D' eucaryotes, ainsi que l'apparente distribution asymétrique des protéines, a longtemps suggéré un fonctionnement indépendant des deux motifs. En exprimant différents variants d'un snoRNA à boîtes C/D, porteur de deux couples de boîtes, dans des cellules de levures, l'équipe de E.S. Maxwell a néanmoins montré que si le motif C/D peut fonctionner de façon autonome, la fonctionnalité d'un motif C'/D' structuré en *K-loop* requiert la proximité d'un motif C/D. Ces expériences ont également révélé qu'une distance inter-motifs de 12 à 36 nucléotides est requise pour le fonctionnement correct du motif C'/D'. De plus, lorsque la longueur entre les 2 motifs est supérieure à cette limite, les snoRNA retrouvés chez la levure compensent généralement l'excédent de séquence par la formation de structures en tige-boucle. Enfin, à la différence du modèle établi chez le xénope, des expériences d'immunoprécipitation ont permis de mettre en évidence la présence de la protéine Snu13p sur le motif C'/D'. La faible affinité de la protéine Snu13p observé *in vitro* pour les motifs en *K-loop* pourrait donc être compensée *in vivo* par des contacts entre les différentes protéines de la snoRNP. Ces derniers résultats sont donc en faveur d'un modèle d'assemblage symétrique chez la levure. Les auteurs proposent ainsi qu'un lot de trois protéines Snu13p-Nop1p-Nop56p/Nop58p soit associé par couple de boîtes et que la distance séparant les deux motifs soit fixée par la dimérisation des protéines Nop56p et Nop58p.

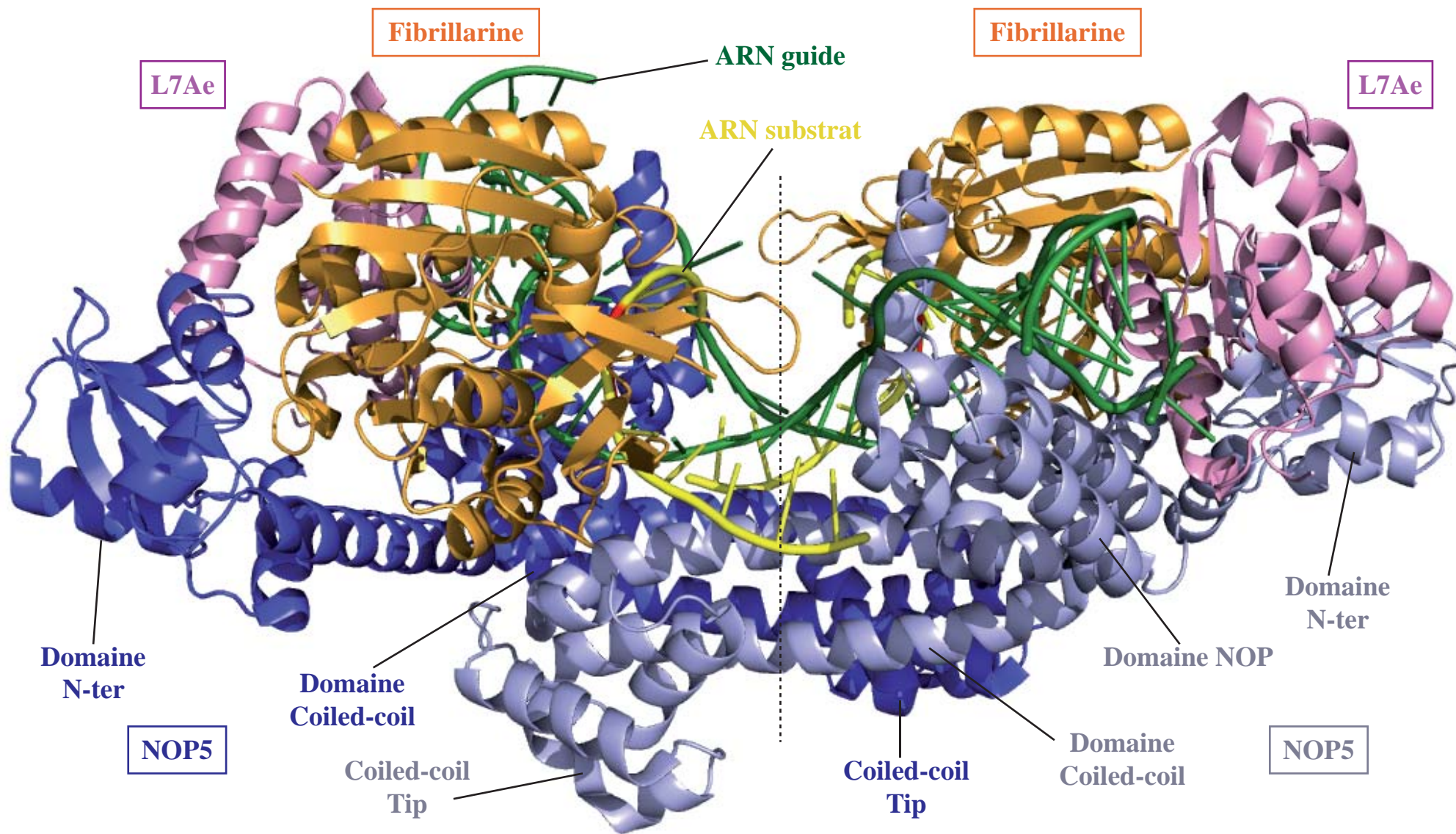


Figure 37 : Structure tridimensionnelle d'une sRNP à boîtes C/D de *Sulfolobus solfataricus* établie par radiocristallographie (d'après Lin *et al.*, 2011; PDB : 3PLA). L'ARN guide porte deux motifs en K-turn et permet l'assemblage de deux particules reliées par la dimérisation des protéines NOP5 (en bleu foncé et bleu clair). Les ARN guides et substrats apparaissent respectivement en vert et jaune. Le trait en pointillés marque l'axe de symétrie entre les deux particules. Un code couleur spécifique est utilisé pour chaque protéine et pour ses domaines.



### 3.1.1.3.3. Bases moléculaires de la 2'-O-méthylation guidée par l'ARN

Comme nous venons de le décrire, trois structures tridimensionnelles de sRNP d'archées sont actuellement disponibles (Lin *et al.*, 2011; Xue *et al.*, 2010; Ye *et al.*, 2009). Dans ces différentes structures, les hélices  $\alpha 8$  et  $\alpha 11$  du domaine NOP de la protéine NOP5 s'associent avec les hélices  $\alpha 2$  et  $\alpha 3$  de la protéine L7Ae pour former un domaine de fixation à l'ARN qui clampe le motif en *K-turn* de part et d'autre (**Fig. 38**). L'interface d'interaction entre les deux protéines implique les résidus K289, A292 et R350 de la protéine NOP5 et les acides aminés N40, E47, E64 et I65 de la protéine L7Ae. De façon intéressante, ces mêmes déterminants sont retrouvés à l'interface d'interaction entre les protéines humaines 15.5K et PRP31 (Liu *et al.*, 2007b), suggérant ainsi que le mode d'interaction des domaines NOP avec les protéines de la famille L7Ae est conservé au cours de l'évolution. D'autre part, les protéines L7Ae et NOP5 établissent de faibles contacts avec l'hélice I du *K-turn*, illustrant ainsi leur tolérance vis-à-vis de l'absence de cette tige dans un motif en *K-loop*. En revanche, ces deux protéines établissent de nombreux contacts avec l'hélice II. De façon remarquable, l'hélice  $\alpha 9'$  du domaine NOP, localisée à la jonction entre les hélices  $\alpha 9$  et  $\alpha 10$ , forme une importante protubérance qui ouvre l'extrémité de l'hélice II au niveau de la paire de bases A-U et assure la séparation physique des brins guide et non guide (**Fig. 39A**). Cette protubérance contient notamment un motif GAEK, strictement conservé au sein des protéines NOP5, Nop56p/NOP56 et Nop58p/NOP58, mais absent des protéines Prp31p/PRP31 (Xue *et al.*, 2010) (**Fig. 30A**). Dans la protéine humaine PRP31, cette région joue néanmoins un rôle comparable puisque l'hélice  $\alpha 10$  constitue une butée permettant son ajustement à une hélice II de 2 paires de bases (Liu *et al.*, 2007b) (cf. Introduction, Chapitre 2, 2.1.3.2) (**Fig. 32**). En revanche, alors que les deux brins de cette hélice sont séparés par la protubérance de NOP5 dans les sRNP, l'hélice II du motif en *K-turn* du snRNA U4 est refermée par une boucle terminale, stabilisée par des contacts avec la boucle située entre les hélices  $\alpha 10$  et  $\alpha 11$  de la protéine PRP31. Cette légère différence d'agencement révèle donc une variation majeure entre les 2 systèmes puisqu'elle assure, dans le cas de la snRNP U4, la stabilisation d'une structure fermée, tandis qu'elle permet dans le cas de la sRNP, le maintien d'une structure ouverte essentielle à la fixation de l'ARN substrat.

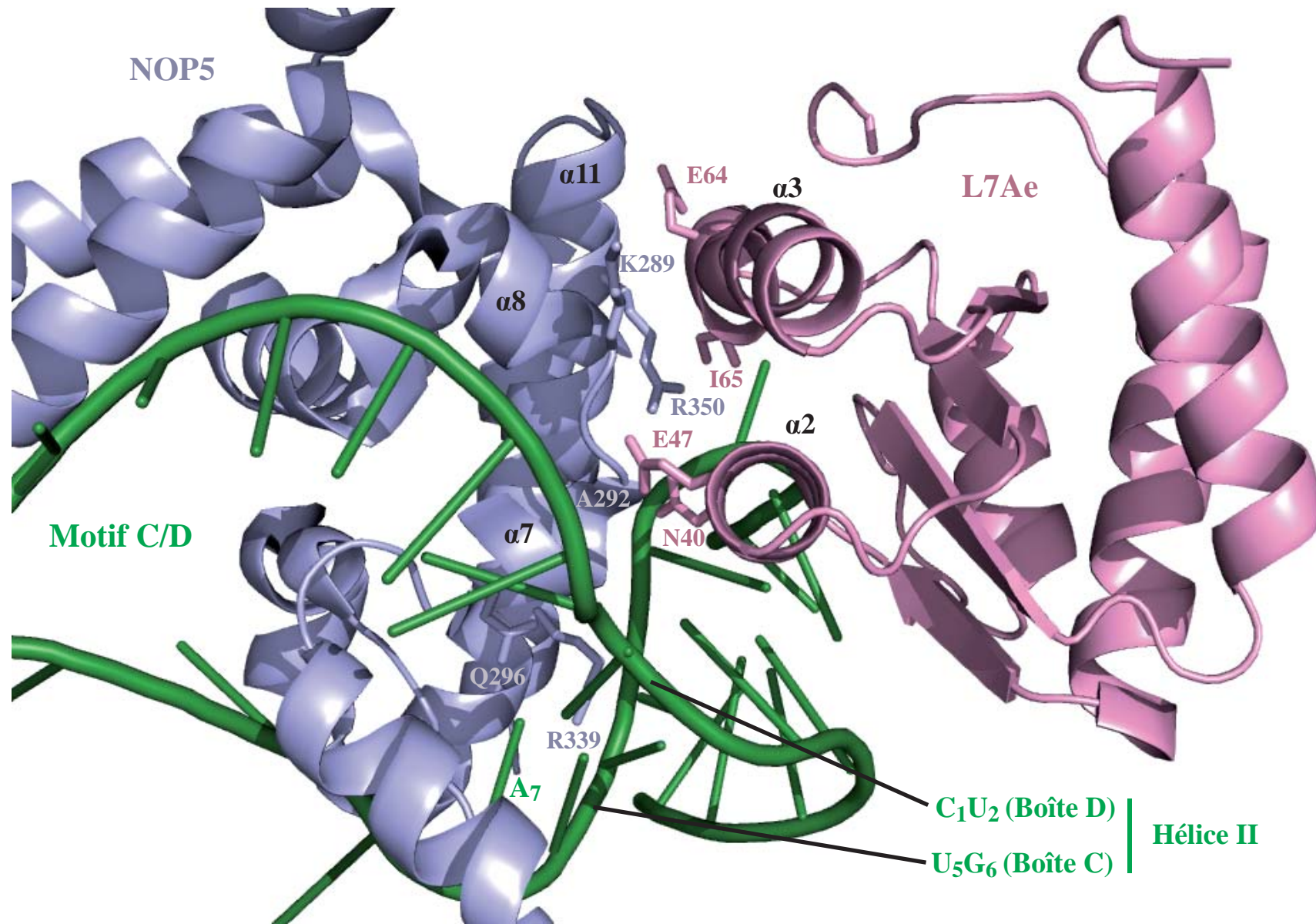
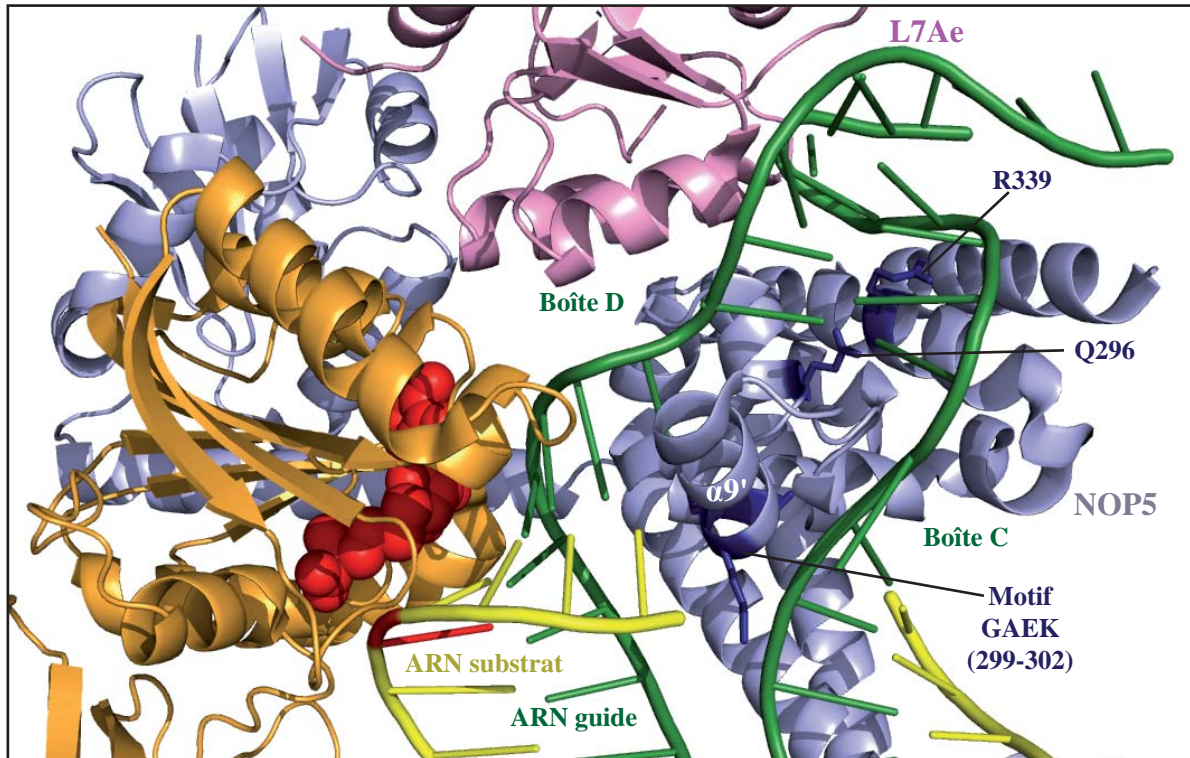
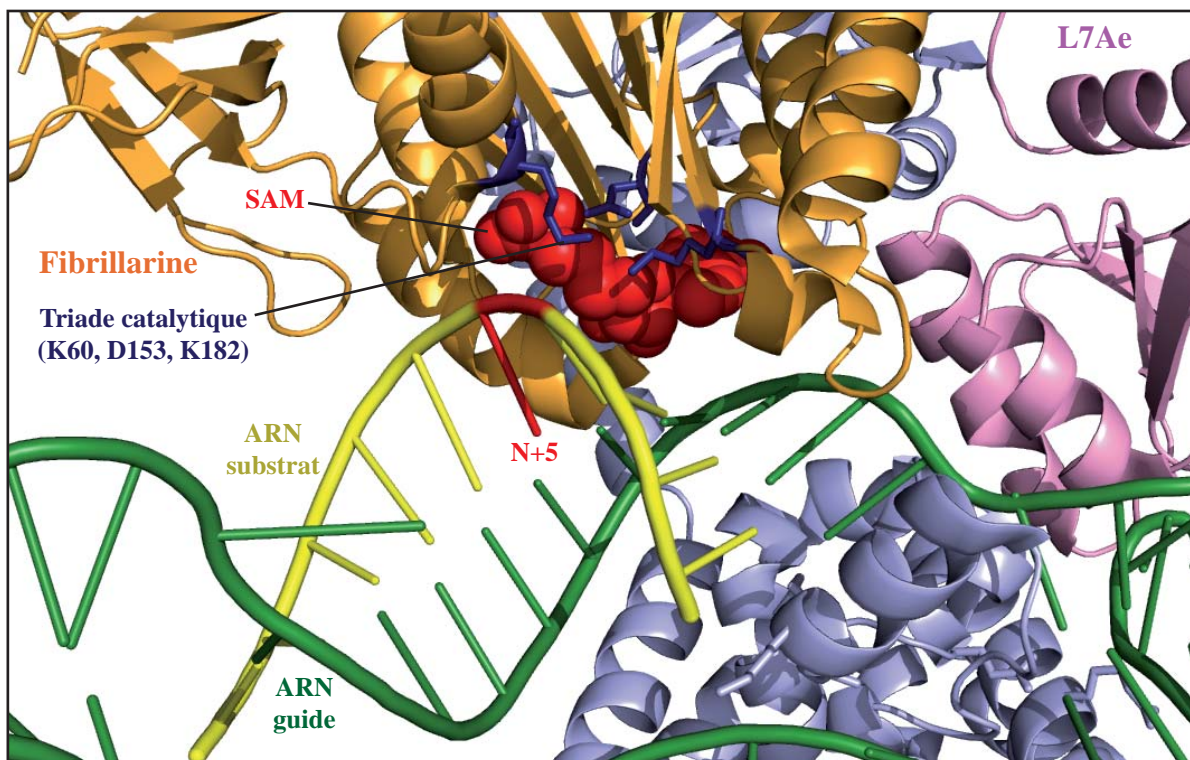


Figure 38 : Complexe formé entre un motif C/D structuré en K-turn et les protéines L7Ae et NOP5 au sein de la sRNP à boîtes C/D de *Sulfolobus solfataricus* (d'après Lin *et al.*, 2011; PDB : 3PLA).

A



B



**Figure 39** : Vue détaillée du domaine catalytique d'une sRNP à boîtes C/D de *Sulfolobus solfataricus* (d'après Lin *et al.*, 2011; PDB : 3PLA). (A) Mise en évidence du rôle joué par NOP5 dans l'ouverture de l'ARN guide. Les résidus Q296 et R339, ainsi que le motif GAEK sont représentés en bleu foncé. (B) Vue du site catalytique de la Fibrillarine et du placement de l'ARN substrat. La SAM (S-Adénosyl-L-Méthionine) est indiquée par des sphères rouges. Le résidu N+5 de l'ARN substrat ciblé par la Fibrillarine est représenté en rouge. La triade catalytique de la Fibrillarine (KDK) est représentée en bleu foncé.

Une autre différence majeure relevée entre l'assemblage de la snRNP U4 et celui des particules à boîtes C/D est la forte dépendance des protéines NOP56 et NOP58 vis-à-vis de la séquence de l'hélice II du motif C/D (Watkins *et al.*, 2002). En effet, si la longueur de cette hélice s'avère déterminante pour le recrutement de PRP31 sur le snRNA U4, la conservation d'une hélice II formée par les appariements U•U, G-C et A-U constitue un élément primordial pour l'assemblage des particules C/D. Les différentes analyses structurales ont ainsi révélé que les résidus Q296 et R339 de la protéine L7Ae de *Sulfolobus solfataricus*, sont impliqués dans la reconnaissance spécifique de la paire G-C terminale et du dernier résidu A de la boîte C, grâce à la formation de plusieurs liaisons hydrogène avec les plateaux de bases (Lin *et al.*, 2011; Ye *et al.*, 2009) (**Fig. 38 et 39A**). Ces deux résidus, qui sont conservés au sein des protéines NOP5, Nop56p/NOP56 et Nop58p/NOP58, se sont par ailleurs révélés importants pour l'assemblage de la particule dans d'autres études (Aittaleb *et al.*, 2003; Hardin *et al.*, 2009). Dans la protéine PRP31, le résidu R est conservé mais se trouve impliqué dans des interactions aspécifiques avec le squelette ribose-phosphate de l'ARN, tandis que le résidu Q, remplacé par un acide aminé variable, est retrouvé dans la région proximale de la boucle terminale de l'ARN (Liu *et al.*, 2007b).

Dans la structure de la sRNP bipartite, le substrat de 10 nucléotides établit un appariement parfait avec la région guide (Lin *et al.*, 2011). Le duplex ARN guide-substrat se trouve alors positionné dans un canal articulé le long du dimère de *Coiled-coil* des protéines NOP5, qui établit des interactions aspécifiques avec les phosphates de l'ARN guide. Ce canal est fermé à ses extrémités par les domaines NOP des deux protéines NOP5 et offre un accès à la Fibrillarine au niveau du sillon mineur du duplex ARN (**Fig. 37**). Le module catalytique, formé par la Fibrillarine et par le domaine N-terminal de la protéine NOP5, est stabilisé par une interaction avec la protéine L7Ae d'une part et avec le domaine *Tip* de la seconde protéine NOP5 d'autre part. Cette configuration permet d'éviter un glissement de la Fibrillarine le long du duplex et assure le positionnement optimal du 5<sup>ème</sup> résidu de l'ARN substrat au sein du site actif de l'enzyme (**Fig. 39B**). Enfin, cette même étude a montré que ce complexe actif est formé au niveau des deux couples de boîtes, suggérant ainsi qu'une même sRNP peut catalyser de façon simultanée la 2'-O-méthylation de deux ARN substrats.

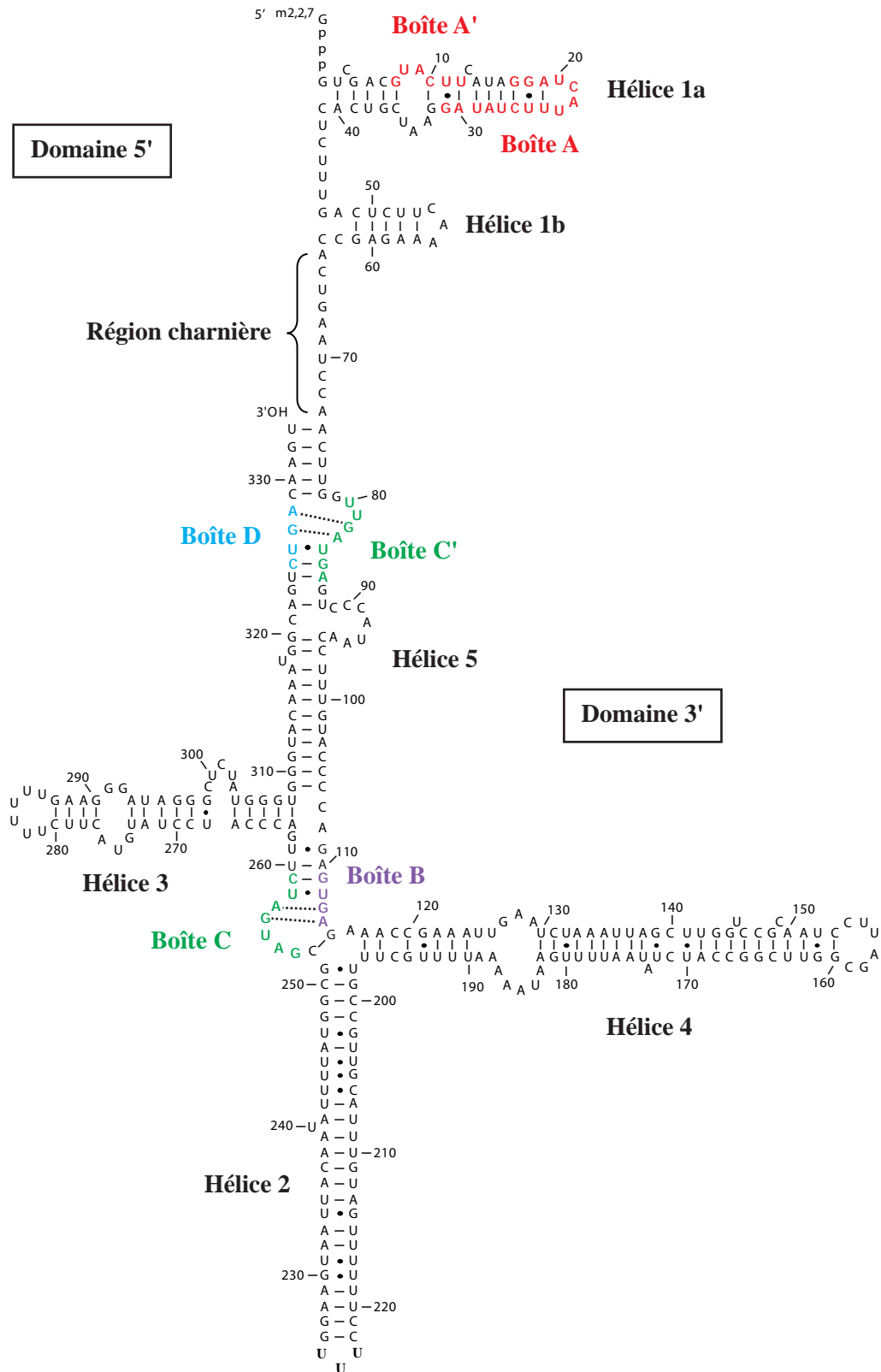
### 3.1.2. La snoRNP U3, une particule singulière impliquée dans la maturation nucléolytique du pré-ARNr

Le snoRNA U3 est essentiel à la viabilité cellulaire, il est conservé chez l'ensemble des eucaryotes et constitue le snoRNA le plus abondant de la cellule (Epstein *et al.*, 1984; Prestayko *et al.*, 1970; Prestayko *et al.*, 1971; Zieve & Penman, 1976). Il se distingue des autres snoRNA à boîtes C/D par des caractéristiques structurales et fonctionnelles qui lui sont propres. Comme nous l'avons abordé dans le chapitre précédent, le snoRNA U3 n'est pas impliqué dans des réactions de 2'-O-méthylation, en revanche, il joue un rôle majeur dans la formation des complexes pré-ribosomiques et dans le déroulement des clivages précoces du pré-ARNr (cf. Introduction, Chapitre 1, 3.4.1). Ce rôle particulier repose notamment sur une architecture singulière. Le snoRNA U3 possède en effet deux motifs structurés en *K-turn* : un motif C'/D qui permet le recrutement des protéines constitutives des snoRNP à boîtes C/D, ainsi qu'un motif B/C, qui est reconnu par la protéine Snu13p/15.5K et par la protéine Rrp9p/h55K, spécifique de cet ARN.

#### 3.1.2.1. Le snoRNA U3

Chez de nombreux eucaryotes, le snoRNA U3 est transcrit par l'ARN polymérase II et possède une coiffe 2,2,7-triméthylguanosine (Reddy & Busch, 1983). Chez les plantes et les algues unicellulaires, il est en revanche transcrit par l'ARN pol III et possède une coiffe 5'- $\gamma$ -monométhyl-phosphate (Antal *et al.*, 2000; Shimba *et al.*, 1992). Dans tous ces organismes, le snoRNA U3 possède deux domaines principaux : une région 5' très conservée, qui est impliquée dans la reconnaissance du pré-ARNr, ainsi qu'une région 3', plus variable, qui contient les motifs C'/D et B/C (Fournier *et al.*, 1998; Jeppesen *et al.*, 1988; Parker & Steitz, 1987; Segault *et al.*, 1992) (**Fig. 40**). Ces deux domaines sont séparés par une région charnière, également appelée *3'-hinge*.

En plus des boîtes C', D, B et C (Hughes *et al.*, 1987; Tycowski *et al.*, 1993; Wise & Weiner, 1980) (**Fig. 23F et G**), le snoRNA U3 possède deux autres séquences conservées, les boîtes A et A', qui sont situées dans le domaine 5', également appelé domaine fonctionnel (Myslinski *et al.*, 1990; Wise & Weiner, 1980) (**Fig. 40**). Ces boîtes sont essentielles à la fonction de l'ARN, puisqu'elles sont impliquées dans la formation d'hélices intermoléculaires avec les



**Figure 40 :** Structure secondaire du snoRNA U3 (snR17A) de *Saccharomyces cerevisiae* obtenue par modifications chimiques et enzymatiques en solution (d'après Ségault *et al.*, 1992; Marmier-Gourrier *et al.*, 2003). Les différentes séquences conservées sont indiquées en couleur. Le domaine 5' (ou domaine fonctionnel) est séparé du domaine 3' par une région charnière.

régions cibles du pré-ARNr (Hughes, 1996; Mereau *et al.*, 1997). A l'état libre, ces deux séquences sont piégées au sein de structures en tige-boucle qui se forment, au nombre de 1 ou 2, au sein du domaine 5' (Fournier *et al.*, 1998; Jeppesen *et al.*, 1988; Parker & Steitz, 1987; Segault *et al.*, 1992).

Le domaine 3' assure quant à lui l'ancrage des protéines liées de manière stable à la snoRNP U3 (Hartshorne & Agabian, 1994; Jeppesen *et al.*, 1988; Mereau *et al.*, 1997; Parker & Steitz, 1987). Ce domaine fait l'objet d'importantes variations de taille et de structure entre les espèces. Ces variations se traduisent essentiellement par l'insertion de structures en tige-boucle dans la région proximale du motif B/C. Chez l'homme, le motif B/C est encadré par deux motifs en tige-boucle (Parker & Steitz, 1987), tandis que chez la levure *S. cerevisiae*, ce motif se situe au centre d'une structure cruciforme comportant 3 tiges-boucles (Segault *et al.*, 1992) (**Fig. 40**).

L'étude détaillée de ces différents éléments de séquence et de structure a permis de mettre en évidence leur importance pour la biogenèse et la fonction de la particule chez la levure (Clery *et al.*, 2007b; Hughes, 1996; Marmier-Gourrier *et al.*, 2003; Mereau *et al.*, 1997; Samarsky & Fournier, 1998). Comme c'est le cas pour les autres snoRNA, les boîtes C' et D sont requises pour l'accumulation du snoRNA U3, pour l'hyperméthylation de sa coiffe et pour son adressage nucléolaire. Les boîtes B et C ne sont pas nécessaires pour la production du snoRNA U3 et pour sa localisation nucléolaire, elles sont en revanche importantes pour sa fonction. De même, le domaine 5' fonctionnel, n'est pas requis pour la biogenèse de l'ARN. D'autre part, la délétion des hélices 2, 3 et 4 encadrant le motif B/C du snoRNA U3 chez la levure (*i.e.* ARN U3 $\Delta$ 2,3,4), n'altère ni la biogenèse, ni la fonction de l'ARN. Chez l'homme le snoRNA U3 présente des caractéristiques similaires, à la différence de la boîte C qui s'avère requise pour maintenir le taux de snoRNA U3 dans la cellule (Granneman *et al.*, 2004; Knox *et al.*, 2011). Cette observation suggère donc que la biogenèse de la snoRNP U3 humaine est plus sensible à l'assemblage de la particule B/C, que ne l'est la snoRNP U3 de levure. D'autre part, il a été montré que le snoRNA U3 transite de façon séquentielle au sein des différentes composantes du nucléole (Lazdins 1997 ; Gerbi 1997). Son appariement avec le pré-ARNr et les étapes de clivages précoces s'effectuent en effet au niveau des DFC, puis en association avec les structures pré-ribosomiques il migrerait vers les GC, avant d'être relargué, vraisemblablement par l'action de l'hélicase Has1p (Liang & Fournier, 2006) (**Fig. 5**). Des expériences de localisation réalisées dans des cellules HeLa ont également permis de

montrer l'importance de la partie 3' de la région charnière (*3'-hinge*) et du motif B/C pour la localisation subnucléolaire du snoRNA U3 humain au niveau des GC, illustrant ainsi leur importance pour le déroulement correct des étapes initiales de la biogenèse des ribosomes (Granneman *et al.*, 2004). Enfin, chez les vertébrés, le snoRNA U3 fait l'objet de réactions de pseudouridylation, il contient notamment une pseudouridine en position 8 de la boîte A' (Reddy *et al.*, 1980). Au laboratoire, S. Massenet a en revanche montré que le snoRNA U3 de *S. cerevisiae* ne possède aucun résidu pseudouridylé (Massenet *et al.*, 1998).

### 3.1.2.2. Les protéines de la snoRNP U3

La purification de la snoRNP U3, réalisée chez l'humain puis chez la levure, a permis de caractériser les protéines stablement associées au snoRNA U3 (Lubben *et al.*, 1993; Watkins *et al.*, 2000). En plus des protéines Snu13p/15.5K, Nop1p/Fibrillarine, NOP56 et NOP58, caractéristiques des snoRNP à boîtes C/D, ces travaux ont conduit à l'identification d'un nouveau composant spécifique de cette particule : la protéine humaine U3-55K et son homologue chez la levure, la protéine Rrp9p (Lubben *et al.*, 1993; Venema *et al.*, 2000; Watkins *et al.*, 2000). Ces deux protéines, qui sont requises pour la viabilité cellulaire, se sont avérées essentielles pour la fonction de la snoRNP U3 (Granneman *et al.*, 2004; Lukowiak *et al.*, 2000; Venema *et al.*, 2000). En revanche, alors que la déplétion de Rrp9p chez la levure affecte peu l'accumulation du snoRNA U3, l'expression de la protéine humaine U3-55K est essentielle à la stabilité de l'ARN (Knox *et al.*, 2011; Venema *et al.*, 2000). Enfin, comme la protéine PRP4 de la snRNP U4, les protéines Rrp9p et U3-55K sont caractérisées par la présence de 7 domaines WD40 qui pourraient se structurer en un large domaine en *helix-proppeler* ((Granneman *et al.*, 2002; Pluk *et al.*, 1998); C. Charron & N. Rolland, Communication personnelle).

Comme cela a été démontré pour la protéine PRP31 et pour les composants des snoRNP à boîtes C/D, le recrutement du facteur Rrp9p/U3-55K sur l'ARN implique la fixation préalable de la protéine Snu13p/15.5K sur le motif B/C (Clery *et al.*, 2007b; Granneman *et al.*, 2002). Son incorporation au sein de la particule repose donc également sur des déterminants protéiques et ARN. Des expériences réalisées chez l'humain ont notamment permis de montrer que les répétitions de motifs WD40, présentes à l'extrémité C-terminale de la protéine, sont nécessaires et suffisantes pour permettre l'association de U3-55K avec un



complexe U3B/C-15.5K *in vitro* (Granneman *et al.*, 2002). De plus, l'équipe de R. Lührmann a mis en évidence l'importance de résidus conservés en surface de la protéine 15.5K pour la formation de ce complexe en extrait cellulaire et pour l'interaction directe entre les deux protéines *in vitro* (Schultz *et al.*, 2006b). De façon intéressante, ces acides aminés (A88 et R91) sont situés sur l'hélice  $\alpha 4$  de la protéine 15.5K, qui ne semble pas impliquée dans une interaction avec les domaines NOP. Cette région de la protéine 15.5K, également importante pour le recrutement du complexe CYPH-hPRP4-hPRP3 sur le snRNA U4 (Schultz *et al.*, 2006b), pourrait donc constituer un site spécifique d'interaction avec des protéines structurées en *helix-proppeler*, comme PRP4 et Rrp9p/U3-55K.

Concernant la protéine humaine U3-55K, une étude a montré qu'une boucle interne, en amont de l'hélice I du motif B/C, constitue un déterminant essentiel pour le recrutement de cette protéine sur l'ARN (Granneman *et al.*, 2002). Chez la levure, des travaux réalisés au laboratoire ont permis de mettre en évidence 3 paramètres importants pour la fixation de Rrp9p sur l'ARN : la présence de la paire de bases G-C inversée dans l'hélice II du *K-turn*, l'hélice I réduite à une unique paire de bases, ainsi que la présence d'un résidu G en position 1 du *K-turn* (Clery *et al.*, 2007b) (cf. Introduction, Chapitre 2, 1.2.4) (**Fig. 23**). *In vivo*, la mutation de ces déterminants abolit en effet le recrutement de Rrp9p sur l'ARN U3 $\Delta$ 2,3,4 et affecte la croissance cellulaire, sans affecter de façon drastique sa stabilité. Ces données confirment ainsi l'importance du motif B/C pour la fonctionnalité du snoRNA U3 et suggèrent que la protéine Rrp9p joue un rôle clé dans le processus de biogenèse des ribosomes. Plus récemment, l'utilisation d'une approche de pontages aux UV, couplée à une analyse par séquençage à haut-débit (CRAC, *Cross-linking and Analysis of cDNA*,) a permis de mettre en évidence deux sites de fixation de la protéine Rrp9p au niveau des hélices 2 et 4 du snoRNA U3 (Granneman *et al.*, 2009). Ces interactions contribuent sans doute à stabiliser la protéine au sein de la snoRNP, cependant, en accord avec la grande variabilité inter-espèces de cette région, ces deux hélices ne sont requises ni pour la viabilité cellulaire, ni pour le recrutement de Rrp9p *in vivo* (Clery *et al.*, 2007b; Samarsky & Fournier, 1998; Venema *et al.*, 2000).

Une fois incorporée au sein des complexes pré-ribosomiques, d'autres facteurs comme le complexe MPP10 (Mpp10p-Imp3p-Imp4p) ou les protéines Bms1p, Rcl1p, Sof1p et Dhr1p, interagissent avec la snoRNP U3 (Bax *et al.*, 2006; Colley *et al.*, 2000; Granneman *et al.*, 2003; Karbstein *et al.*, 2005; Lee & Baserga, 1999). Des travaux récents ont conduit à la

caractérisation d'un nouveau complexe qui constituerait un état intermédiaire entre la forme stable de la particule U3 (12S) et sa forme intégrée au SSU-processome (80S) (Turner *et al.*, 2009). Cette particule 50S, identifiée chez l'humain, s'accumule en effet lorsque la transcription du pré-ARNr est bloquée ou lorsque des protéines du complexe UTP-t, appartenant au SSU-processome, sont déplétées. Elle ne contient pas les sous-complexes UTP-t, UTP-B, MPP10 et BMS1/RCL1, en revanche, elle renferme, en plus de la snoRNP U3, certaines protéines retrouvées dans les complexes pré-ribosomiques, comme l'hélicase DBP4, la protéine RRP5, ou encore la Nucléoline.

### 3.1.2.3. Structure de la snoRNP U3

A l'heure actuelle, aucune structure tridimensionnelle de la snoRNP U3 n'est disponible, cependant des données récentes, appuyées par le modèle d'organisation des sRNP d'archées, ont apporté un éclairage nouveau sur l'architecture probable de la particule (Knox *et al.*, 2011). De plus, ces travaux ont révélé un lien inattendu entre l'assemblage de la snoRNP U3 et le processus de différenciation cellulaire.

Comme nous l'avons évoqué précédemment, le motif C'/D du snoRNA U3 correspond à un motif C/D non canonique (cf. Introduction, Chapitre 2, 1.2.4). Chez les mammifères et chez les amphibiens, le résidu U canonique retrouvé en position 3 du *K-turn*, est remplacé par un G dans le motif C'/D (**Fig. 23**). Différentes études ont montré que cette configuration, très défavorable pour le recrutement de la protéine Snu13p/15.5K *in vitro*, pourrait affecter de façon drastique la formation de la particule (Marmier-Gourrier *et al.*, 2003; Nottrott *et al.*, 1999; Watkins *et al.*, 2002). Néanmoins, dans le cas du snoRNA U3, ce motif est parfaitement fonctionnel *in vivo* et permet le recrutement des protéines spécifiques des snoRNP à boîtes C/D. Afin de mieux appréhender ce paradoxe, l'équipe de N.J. Watkins a étudié *in vivo* les effets associés à la mutation cumulée des boîtes C' et C (Knox *et al.*, 2011). Ainsi, lorsque la boîte C' est remplacée par une séquence consensus 5'-AUGAUGA-3', une mutation de la boîte C, délétère pour la fixation de la protéine U3-55K, n'engendre plus d'effet négatif sur l'accumulation du snoRNA U3. De même, la diminution du taux de snoRNA U3 provoqué par la déplétion de la protéine U3-55K est également compensé par la présence d'une boîte C' optimisée. Ces résultats suggèrent donc qu'il existerait, au sein de la particule U3, une forme de coopérativité entre les motifs B/C et C'/D. Afin de confirmer cette hypothèse, les auteurs

ont étudié l'association de la protéine NOP58 avec l'ARN muté au niveau de sa boîte C. Ils ont ainsi montré que le recrutement de la protéine NOP58 sur l'ARN est altéré par la mutation de la boîte C, excepté lorsqu'une la boîte C' est remplacée par la séquence consensus. L'ensemble de ces résultats indique donc que l'incorporation de la protéine U3-55K au niveau de la particule B/C pourrait favoriser l'assemblage de la particule C'/D et ainsi compenser la présence d'une séquence C' sous-optimale. Etant donné que le recrutement de la protéine NOP58 est essentiel pour la stabilité des snoRNA, cette observation pourrait donc expliquer l'importance de la boîte C du snoRNA U3 humain pour sa stabilité *in vivo*.

En utilisant l'approche CRAC, qui consiste à analyser par séquençage à haut débit, des fragments d'ARN pontés covalamment à des protéines d'intérêt, l'équipe de D. Tollervey avait permis de cartographier des interactions ARN-protéines au sein de la snoRNP U3 de levure (Granneman *et al.*, 2009). Alors que la protéine Nop58p s'associe à l'hélice II du motif C'/D, les protéines Nop56p et Fibrillarine sont pontées à la base de l'hélice 3 qui est située à proximité du motif B/C. Ce résultat est d'autant plus intéressant que des interactions physiques ont été détectées entre U3-55K et les protéines NOP56 et NOP58 (Knox *et al.*, 2011). En tenant compte des avancées récentes sur les sRNP d'archées, l'équipe de N.J. Watkins a donc proposé un modèle d'assemblage pour la snoRNP U3 (**Fig. 41**). Dans ce modèle pseudo-symétrique, le motif C'/D serait occupé par les protéines 15.5K et NOP58, tandis que le motif B/C serait associé aux protéines 15.5K, U3-55K et NOP56. Les deux particules contiendraient également une protéine Fibrillarine en association avec les protéines NOP56 et NOP58. Enfin, comme dans les modèles classiques de snoRNP (**Fig. 36**), les deux particules seraient en contact par le biais de la dimérisation des domaines *Coiled-coil* de ces deux protéines.

Un tel modèle pourrait être appliqué aux autres snoRNA U3 eucaryotes, qui possèdent tous des boîtes C' sous optimales. Des travaux réalisés au laboratoire sur le snoRNA U3 de levure ont notamment permis de montrer que la longueur de l'hélice 5 séparant les motifs C'/D et B/C est un paramètre déterminant pour la stabilité de l'ARN *in vivo* (N. Marmier-Gourrier, N. Rolland, Communication personnelle). Dans l'hypothèse de l'établissement d'une interaction entre les deux particules, cette contrainte de distance pourrait ainsi correspondre à l'espacement nécessaire à la dimérisation des protéines Nop56p et Nop58p ancrées sur chacun des motifs.

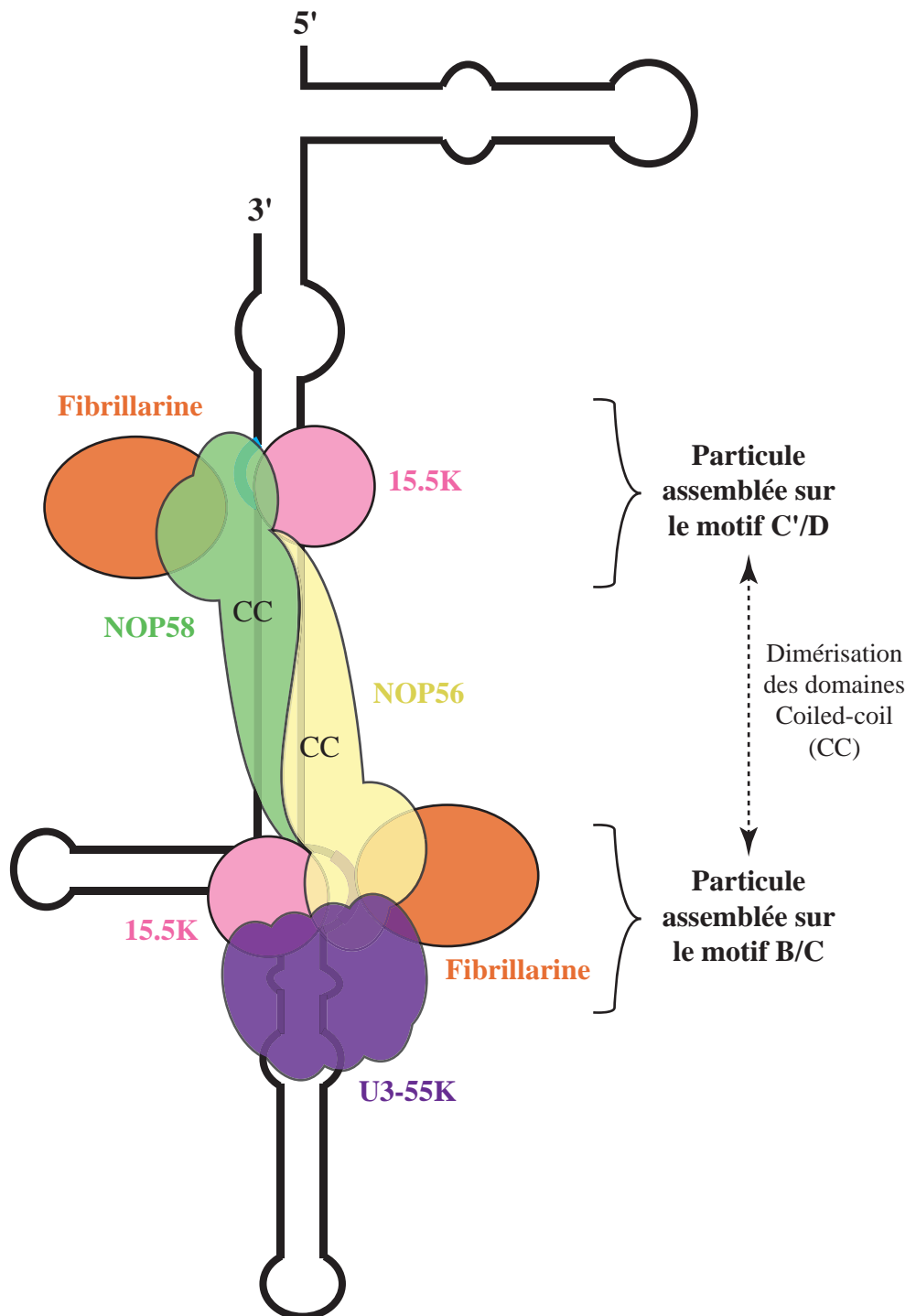


Figure 41 : Modèle d'assemblage de la snoRNP U3 humaine (d'après Knox *et al.*, 2011).

### 3.1.2.4. Assemblage de la snoRNP U3 et différenciation cellulaire

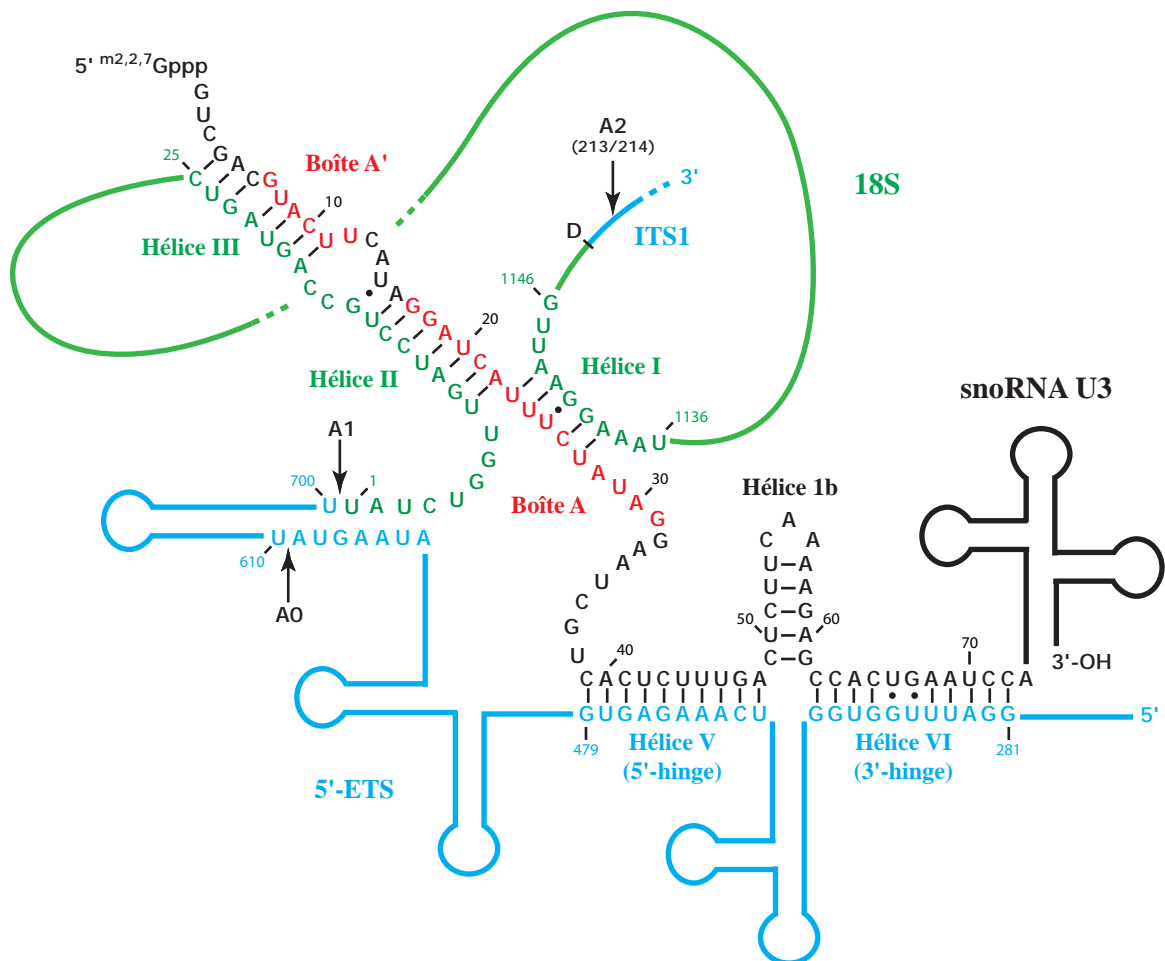
Outre les aspects liés à l'organisation structurale de la particule, l'équipe de N.J. Watkins a également montré l'existence d'une co-régulation du taux de protéine U3-55K et de snoRNA U3 au cours de la différenciation cellulaire (Knox *et al.*, 2011). En effet, il est clairement établi que le taux de production des ribosomes est ajusté en fonction des besoins cellulaires. Ainsi leur biogenèse diminue lors de la différenciation terminale des cellules, tandis qu'elle augmente dans la plupart des cas de cancers (Ruggero & Pandolfi, 2003). Au niveau du nucléole, cette régulation s'exerce sur le niveau de transcription du pré-ARNr et sur le taux de maturation des ARNr (Mayer & Grummt, 2006). Parmi les composants stablement associés au snoRNA U3, seule la protéine U3-55K fait l'objet d'une telle régulation, puisqu'à la différence de la Fibrillarine et des protéines NOP56 et NOP58, son taux chute au cours de la différenciation des cellules épithéliales (Knox *et al.*, 2011). Cette diminution s'accompagne donc vraisemblablement d'un défaut d'assemblage de la protéine NOP58 sur le motif C'/D, conduisant à la destabilisation spécifique du snoRNA U3. Ces dernières données apportent donc des éléments de réflexion quant à la pérennité évolutive des boîtes C' sous optimales. Chez l'homme, le caractère non consensuel de la boîte C' pourrait en effet donner prise à un contrôle du taux de snoRNA exercé par la protéine U3-55K, dont la propre régulation au cours de la différenciation cellulaire pourrait permettre un ajustement de la production de ribosomes. Cette hypothèse est d'autant plus intéressante que la protéine U3-55K, qui est fréquemment surproduite durant la tumorigenèse, a été associée à la survenue de cancers du sein (Luo *et al.*, 2001; Sjoblom *et al.*, 2006; Sperger *et al.*, 2003). Enfin, ces phénomènes de régulation impliquent également d'autres composants du SSU-processome, puisque le taux de certaines protéines comme MPP10, UTP12, PRP43 ou la Nucléoline, est altéré durant la différenciation des cellules épithéliales (Knox *et al.*, 2011).

### 3.1.2.5. Fonction de la snoRNP U3

Comme nous l'avons évoqué dans le premier chapitre, la snoRNP U3 est un composant essentiel du SSU-processome, dont la formation à l'extrémité 5' du pré-ARNr naissant, constitue une étape clé de la synthèse des ribosomes. Ce macro-complexe est à l'origine d'un long processus conduisant à la maturation nucléolytique de l'ARNr 18S et à l'assemblage de

la sous-unité 40S. De nombreuses études ont montré l'importance de la snoRNP U3 dans l'initiation de ce processus (cf. Introduction, Chapitre 1, 3.4.1) (**Fig. 19**). Cette dernière est en effet essentielle pour la mise en place du SSU-processome (Osheim *et al.*, 2004) et pour le déroulement des clivages précoces aux sites A0, A1 et A2 du pré-ARNr (Hughes & Ares, 1991; Kass *et al.*, 1990; Li *et al.*, 1990; Zagorski *et al.*, 1988).

Bien qu'elle n'ait jamais été impliquée de façon directe dans la catalyse, la snoRNP U3 semble jouer un rôle majeur dans l'architecture du SSU-processome et dans la structuration du pré-ARNr. Ce rôle de chaperonne repose notamment sur la capacité du snoRNA U3 à interagir de façon spécifique avec certaines séquences du pré-ARNr (**Fig. 42**). Ces interactions, relativement bien conservées chez les eucaryotes, impliquent le domaine 5' du snoRNA U3, ainsi que la région charnière le reliant au domaine 3' (Antal *et al.*, 2000; Beltrame & Tollervey, 1992; Borovjagin & Gerbi, 2000; Hughes, 1996; Mereau *et al.*, 1997). Chez la levure et chez le xénope, plusieurs hélices intermoléculaires ont été caractérisées entre le pré-ARNr et le snoRNA U3. Les boîtes A' et A du snoRNA forment respectivement un et deux appariements conservés avec la séquence de l'ARNr 18S (Borovjagin & Gerbi, 2000; Hughes, 1996; Hughes *et al.*, 1987; Mereau *et al.*, 1997; Myslinski *et al.*, 1990). De façon intéressante, ces duplex, baptisés hélices I, II et III, impliquent une région de l'ARNr engagée dans la formation du pseudo-nœud central, un élément de structure universellement conservé au sein de la petite sous-unité ribosomique (Gutell *et al.*, 1994; Neefs *et al.*, 1993). Cette observation suggère donc que, par le biais de ces appariements, le snoRNA U3 pourrait retarder la formation du pseudo-nœud central et ainsi contribuer au repliement séquentiel de l'ARNr au cours de l'assemblage de la future sous-unité ribosomique (Hughes, 1996). En plus de ces 3 hélices, un autre appariement conservé, appelé hélice V (ou appariement 5'-*hinge*), a été décrit entre la région charnière du snoRNA U3 et la séquence 5'-ETS du pré-ARNr (Beltrame & Tollervey, 1995; Borovjagin & Gerbi, 2000; Brule *et al.*, 1996). Initialement, la formation d'une hélice IV entre la région charnière du snoRNA et l'ARN 18S, avait également été prédite chez la levure (Mereau *et al.*, 1997). Cependant, de récentes données ont permis d'infirmer cette hypothèse en montrant que la séquence charnière du snoRNA U3 participe en réalité à la formation d'une hélice VI conservée (ou appariement 3'-*hinge*) avec la région 5'-ETS (Borovjagin & Gerbi, 2000; Dutca *et al.*, 2011; Marmier-Gourrier *et al.*, 2011). De plus, différents travaux réalisés chez la levure ont permis de montrer l'importance de ces hélices pour la production de l'ARNr 18S (Beltrame & Tollervey, 1995; Dutca *et al.*, 2011; Marmier-Gourrier *et al.*, 2011; Sharma & Tollervey, 1999). En effet, à l'exception de



**Figure 42 : Représentation schématique des appariement formés entre le snoRNA U3 et le pré-ARNr 35S chez la levure *Saccharomyces cerevisiae* (d'après Hughes *et al.*, 1996; Méreau *et al.*, 1997; Dutca *et al.*, 2011; Marmier-Gourrier *et al.*, 2011). L'ARNr 18S est représenté en vert, les séquences 5'-ETS et ITS1 sont représentées en bleu. Le snoRNA U3 est représenté en noir, les boîtes A et A' sont indiquées en rouge. Les sites de coupures endonucléolytiques A0, A1 et A2 du pré-ARNr sont indiqués. 5'-ETS : 5'- External Transcribed Spacer ; ITS1 : Internal Transcribed Spacer.**

l'hélice III qui n'est requise que pour les coupures aux sites A1 et A2, ces différents appariements sont importants pour les clivages aux trois sites A0, A1 et A2.

En parallèle de ces études, d'autres travaux ont établi l'importance de ces appariements pour la formation complète du SSU-processome au niveau des DFC (Dutca *et al.*, 2011; Granneman *et al.*, 2004). La formation de l'hélice VI (*3'-hinge*) semble jouer un rôle nucléateur dans ce processus, en permettant le recrutement du complexe MPP10 (Dutca *et al.*, 2011; Marmier-Gourrier *et al.*, 2011). Une fois recrutées, il semblerait que les protéines Mpp10p, Imp3p et Imp4p facilitent la formation des autres appariements intermoléculaires en interagissant avec certaines régions du snoRNA U3 (Dutca *et al.*, 2011; Gerczei & Correll, 2004; Gerczei *et al.*, 2009; Marmier-Gourrier *et al.*, 2011; Wormsley *et al.*, 2001). *In vitro*, il a notamment été montré que les protéines Imp3 et Ump4p, contribuent à la formation et à la stabilisation de l'hélice V (*5'-hinge*), tandis que la formation des hélices II et III serait facilitée par la protéine Imp4p (Gerczei & Correll, 2004). Ainsi, la formation séquentielle et assistée de ces appariements constitue une étape clé, dont le déroulement correct conduit à la réalisation des clivages précoces du pré-ARNr et à la formation de la particule pré-40S.

### 3.2. Les snoRNP à boîtes H/ACA

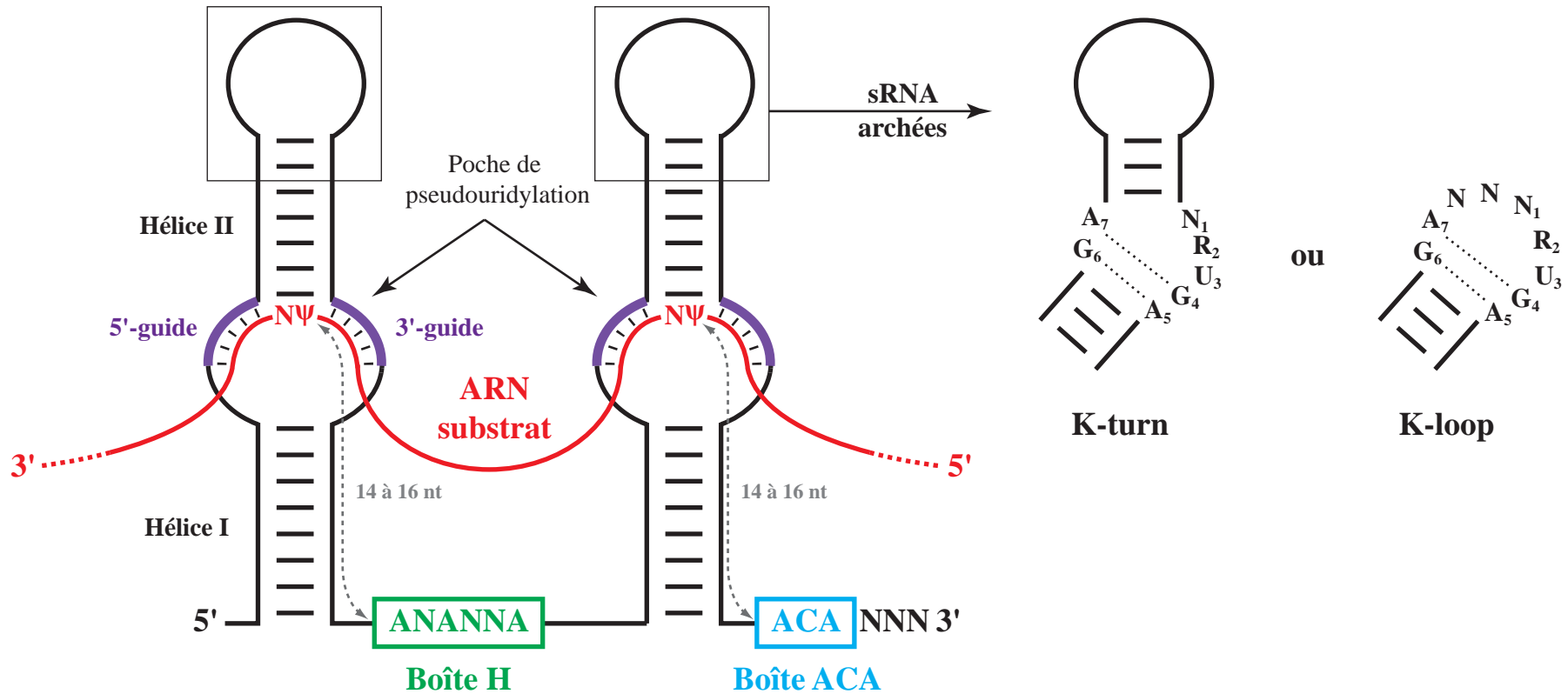
Les snoRNP à boîtes H/ACA catalysent des réactions de pseudouridylation. Comme nous l'avons évoqué, certaines d'entre elles sont également impliquées dans la maturation nucléolytique du pré-ARNr. Bien qu'elles partagent des éléments communs avec les snoRNP à boîtes C/D, ces particules présentent néanmoins des différences architecturales majeures, qui leur confèrent un statut particulier au sein de la famille des RNP L7Ae.

#### 3.2.1. Les snoRNA à boîtes H/ACA

Chez les eucaryotes, les snoRNA à boîtes H/ACA sont caractérisés par la présence de deux structures en tige-boucle, séparées par une région charnière simple brin contenant une séquence conservée 5'-ANANNA-3', appelée boîte H (*Hinge*) (Ganot *et al.*, 1997b) (**Fig. 43**). La seconde structure tige-boucle est également suivie d'une autre séquence conservée, la boîte



**snoRNA  
à boîtes H/ACA**



**Figure 43 : Représentation schématique d'un snoRNA à boîtes H/ACA.** L'ARN cible, représenté en rouge, s'apparie avec les séquences guides 5' et 3' de la poche de pseudouridylation. Le premier résidu non apparié du substrat, représenté par un Ψ, est ciblé par l'enzyme CBF5/Dyskérine. Chez les archées, les sRNA à boîtes H/ACA, qui comptent 1 à 3 structures tige-boucle et qui ne présentent pas de boîte H, sont caractérisés par la présence d'un motif en K-turn ou en K-loop à l'apex de chaque structure tige-boucle.

ACA, formée par le tri-nucléotide 5'-ACA-3' et située 3 résidus en amont de l'extrémité 3' du snoRNA (Balakin *et al.*, 1996). A l'exception de quelques organismes (ex : *Euglena gracilis* ou *Trypanosoma*) qui peuvent présenter des structures alternatives avec 1 ou 3 tiges-boucles (Myslyuk *et al.*, 2008; Russell *et al.*, 2004), l'organisation générale des snoRNA à boîtes H/ACA est relativement bien conservée chez les eucaryotes. Les boîtes H et ACA sont nécessaires à la stabilité du snoRNA, à sa maturation, ainsi qu'à l'activité de pseudouridylation (Balakin *et al.*, 1996; Bortolin *et al.*, 1999; Ganot *et al.*, 1997b). Chaque structure en tige-boucle contient une boucle interne appelée « poche de pseudouridylation », dont les brins 5' et 3' sont impliqués dans la reconnaissance spécifique de l'ARN cible. La formation de deux hélices intermoléculaires entre l'ARN guide et l'ARN substrat conduit à l'exposition d'un di-nucléotide 5'-UN-3' non apparié au sommet de la poche (Ganot *et al.*, 1997a; Ni *et al.*, 1997). Le résidu 5' du di-nucléotide est l'uridine ciblée par la modification. De façon conservée, ce dernier est toujours localisé 14 à 16 nucléotides en amont des boîtes H ou ACA (Ganot *et al.*, 1997a; Ni *et al.*, 1997).

Chez les archées, les sRNA à boîtes H/ACA ont été découverts postérieurement chez *Archaeoglobus fulgidus* (Tang *et al.*, 2002). A la différence des snoRNA eucaryotes, les sRNA H/ACA d'archées peuvent être constitués de 1 à 3 structures en tige-boucle (Muller *et al.*, 2008; Rozhdestvensky *et al.*, 2003). Ils sont également caractérisés par l'absence de la boîte H et par la présence d'une boîte ACA de type 5'-ANA-3' située directement en aval de chaque tige-boucle. Comme nous l'avons déjà évoqué, ils ont pour particularité de posséder un motif en *K-turn* ou en *K-loop*. Ce dernier, retrouvé au sein de l'hélice II de chaque tige-boucle, est situé à une distance conservée de 5 à 6 paires de bases de la poche de pseudouridylation (Fourmann, 2009; Muller *et al.*, 2008; Rozhdestvensky *et al.*, 2003). De plus, des travaux réalisés au laboratoire par S. Muller et J.B. Fourmann ont également permis de mettre en évidence que la réaction de pseudouridylation tolère la présence d'un tri-nucléotide 5'-UNN-3' non apparié au niveau de la poche de pseudouridylation (Fourmann, 2009; Muller *et al.*, 2008).

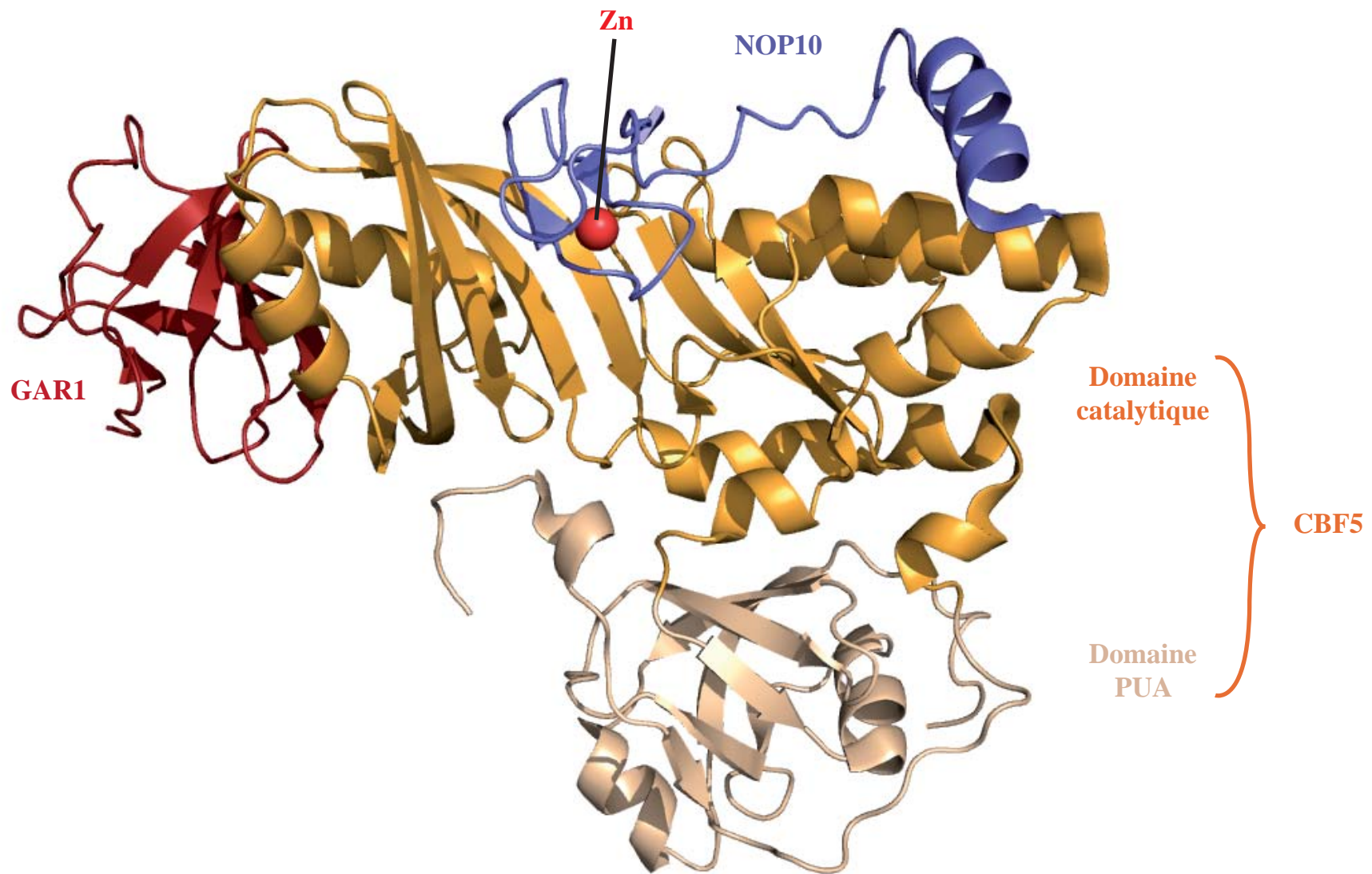
### 3.2.2. Les protéines des snoRNP à boîtes H/ACA

En plus de la protéine Nhp2p/NHP2 de la famille L7Ae, qui est essentielle à la stabilité de l'ARN et à l'activité de la particule (Henras *et al.*, 1998; Watkins *et al.*, 1998), les snoRNP à

boîtes H/ACA eucaryotes contiennent 3 autres protéines constitutives : les protéines Gar1p/GAR1, Nop10p/NOP10 et Cbf5p/CBF5. Cette dernière, également appelée NAP57 ou Dyskérine, en référence à son lien avec la dyskératose congénitale, porte l'activité pseudouridyl-synthase. Ces différentes protéines sont conservées chez les archées à l'exception de Nhp2p/NHP2 qui est remplacée par la protéine L7Ae.

### 3.2.2.1. La protéine Cbf5p/Dyskérine

Identifiée chez la levure pour son interaction avec les centromères, puis pour sa présence au sein des particules H/ACA, Cbf5p (*Centromere Binding Factor 5*) est une protéine essentielle à la viabilité cellulaire (Jiang *et al.*, 1993; Watkins *et al.*, 1998). Des expériences de déplétion réalisées chez la levure ont montré que la protéine Cbf5p est requise pour la fonction des particules H/ACA, mais également pour la stabilité des snoRNA (Lafontaine *et al.*, 1998). Comme les protéines Nop56p et Nop58p, Cbf5p contient une extrémité C-terminale riche en motifs répétés KKE/D, qui interagit avec l'hyperméthylase Tgs1p (Mouaikel *et al.*, 2002). Néanmoins, ces motifs qui sont absents chez les archées, ne sont pas requis pour la fonctionnalité de la protéine (Watkins *et al.*, 1998; Winkler *et al.*, 1998). Par ailleurs, elle est également caractérisée par une forte homologie de séquence avec l'enzyme TruB, impliquée dans la pseudouridylation de la position 55 des ARNt chez *E. coli* (Aravind & Koonin, 1999; Koonin, 1996; Lafontaine & Tollervey, 1998). Chez les archées, plusieurs structures tridimensionnelles de la protéine CBF5, libre ou intégrée au sein de différents complexes, ont été résolues, confirmant ainsi l'homologie structurale avec la protéine TruB (Hamma *et al.*, 2005; Li & Ye, 2006; Manival *et al.*, 2006; Rashid *et al.*, 2006). Ces deux protéines partagent en effet deux domaines caractéristiques : un domaine catalytique spécifique de toutes les pseudouridyl-synthases (Lafontaine *et al.*, 1998), ainsi qu'un domaine PUA (*PseudoUridine synthase and Archaeosine transglycosylase*) situé à l'extrémité C-terminale de la protéine et fréquemment muté dans les cas de dyskératose congénitale (Heiss *et al.*, 1998) (**Fig. 44**). Le domaine PUA est un motif d'interaction avec l'ARN retrouvé dans de nombreuses enzymes de modifications (Aravind & Koonin, 1999). La résolution de la structure tridimensionnelle d'une sRNP à boîtes H/ACA chez les archées a notamment permis de mettre en évidence son interaction avec la boîte ACA avec l'ARN (Li & Ye, 2006; Zhou *et al.*, 2011). En plus de son rôle dans la réaction de pseudouridylation, le domaine catalytique de CBF5 est impliqué dans des interactions avec les protéines NOP10 et GAR1 (Li & Ye, 2006; Manival *et al.*, 2006;



**Figure 44 : Structure tridimensionnelle du complexe formé par les protéines GARI1, CBF5 et NOP10 chez *Pyrococcus furiosus* (Rashid *et al.*, 2006; PDB : 2EY4).** La structure a été obtenue par radiocristallographie. Le domaine PUA de CBF5 est formé par la partie C-terminale de la protéine et par un brin  $\beta$  situé à l'extrémité N-terminale. PUA : PseudoUridine synthase and Archaeosine transglycosylase.

Rashid *et al.*, 2006). Enfin, une étude récente, réalisée chez l'homme, a mis en évidence l'existence d'une isoforme cytoplasmique de la dyskérine (Angrisani *et al.*, 2011). Sa production résulte d'événements d'épissage alternatif, mais sa fonction précise, vraisemblablement indépendante des snoRNP, n'a pas encore été caractérisée.

### 3.2.2.2. La protéine Gar1p/GAR1

Bien qu'essentielle à l'activité de la particule, la protéine Gar1p/GAR1 est la seule protéine à ne pas être requise pour la stabilité du snoRNA (Bousquet-Antonelli *et al.*, 1997; Girard *et al.*, 1992). La résolution de sa structure tridimensionnelle chez les archées a révélé une architecture en tonneau formé de 6 brins  $\beta$  (Rashid *et al.*, 2006) (**Fig. 44**). Chez les eucaryotes, cette protéine est caractérisée par la présence supplémentaire d'un domaine GAR (*Glycin-Arginin Rich*) à chacune de ses extrémités. Ces domaines interagissent spécifiquement avec les ARN et contribuent à la localisation nucléolaire de nombreuses protéines comme la Nucléoline (Ginisty *et al.*, 1999). De plus, les résidus arginine de ce domaine C-terminal font l'objet de méthylations asymétriques (Frankel & Clarke, 1999; Wang & Meier, 2004; Xu *et al.*, 2003).

### 3.2.2.3. La protéine NOP10

La protéine Nop10p/NOP10 est essentielle pour l'activité de la particule et pour la stabilité des snoRNA à boîtes H/ACA (Henras *et al.*, 1998). Il s'agit d'une protéine de petite taille (une soixantaine de résidus) dont la structure a été résolue chez les archées et chez les eucaryotes. La protéine Nop10p/NOP10 eucaryote est formée de 2 brins  $\beta$ , reliés à une unique hélice  $\alpha$  C-terminale par une région charnière flexible (Hamma *et al.*, 2005; Khanna *et al.*, 2006; Reichow & Varani, 2008). Chez les archées, sa région N-terminale est en revanche composée de 4 brins  $\beta$ , qui adoptent une structure de type  $\beta$  *ribbon*, chélateur d'un atome de zinc (Hamma *et al.*, 2005; Manival *et al.*, 2006; Rashid *et al.*, 2006) (**Fig. 44**). Bien que ce domaine N-terminal soit impliqué dans l'interaction avec la protéine CBF5, la mutation des 4 cystéines responsables de la chélation du zinc, n'altère ni la formation de la particule, ni son

activité (Manival *et al.*, 2006). Enfin, au sein de la sRNP, NOP10 établit également des contacts avec le sRNA et avec la protéine L7Ae (Li & Ye, 2006).

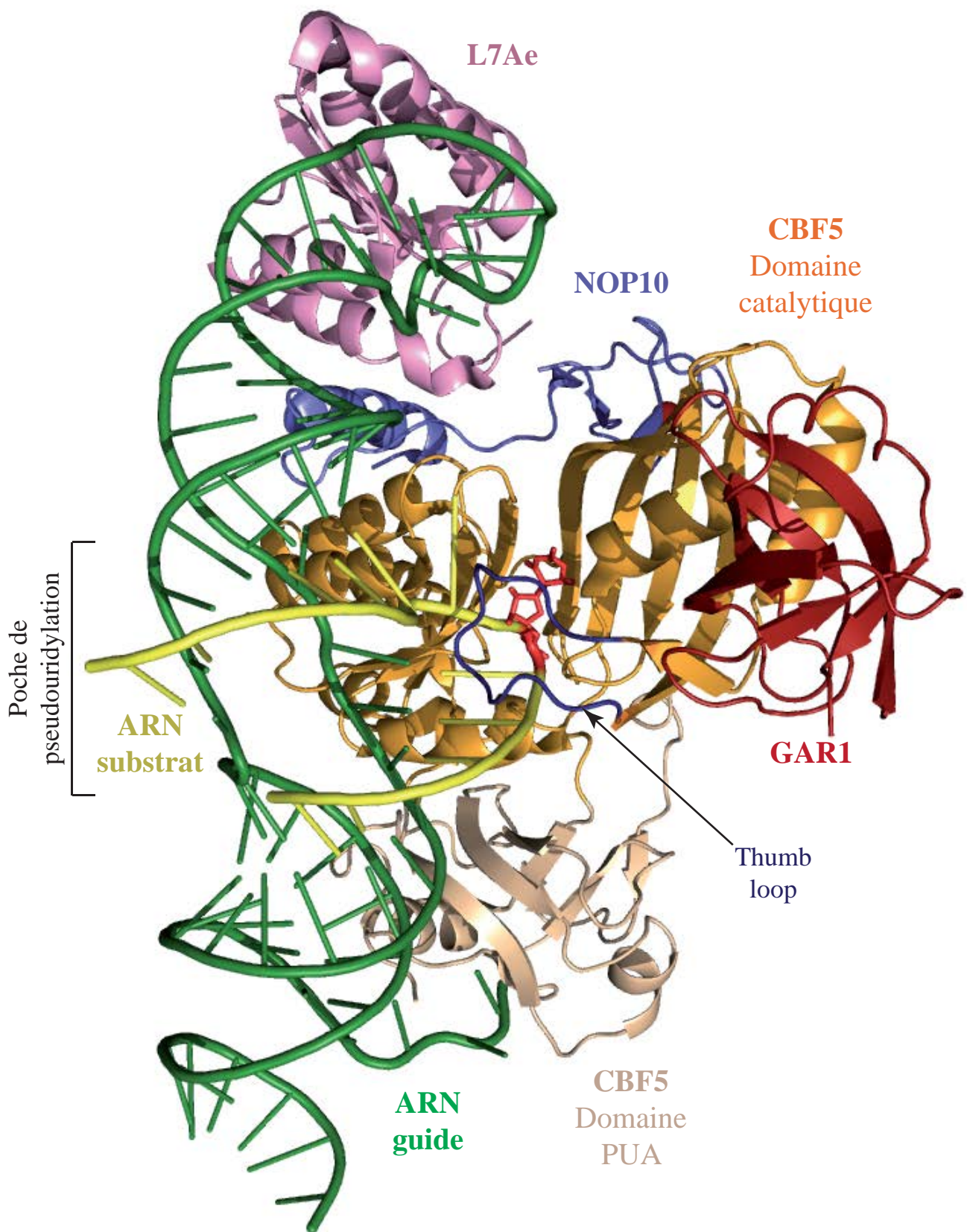
### 3.2.3. Structure et fonction des snoRNP à boîtes H/ACA

En plus de leur rôle dans la maturation des ARNr, les particules à boîtes H/ACA sont impliquées dans la modification des UsnRNA. Chez les mammifères, le snRNA U6 est modifié par des snoRNP à boîtes H/ACA au sein du nucléole (Ganot *et al.*, 1999; Kiss *et al.*, 2004), tandis que des réactions de pseudouridylation catalysées par des scaRNP, ciblent les autres UsnRNA spliceosomaux au niveau des CB (Darzacq *et al.*, 2002; Jady *et al.*, 2003; Jady & Kiss, 2001; Kiss *et al.*, 2004; Kiss *et al.*, 2002). Enfin, de façon plus surprenante, des travaux récents réalisés chez les archées ont permis de mettre en évidence que la protéine CBF5, seule ou associée aux protéines NOP10 et GAR1, peut catalyser une réaction de pseudouridylation non guidée au niveau de la position 55 des ARNt (Fourmann, 2009; Gurha *et al.*, 2007; Muller *et al.*, 2007; Roovers *et al.*, 2006).

Comme c'est le cas pour les particules à boîtes C/D, l'étude *in vitro* des sRNP à boîtes H/ACA d'archées a fourni de nombreuses informations sur le fonctionnement de ce système de modification. Au laboratoire, B. Charpentier a mis au point une approche de reconstitution de sRNP actives *in vitro*, qui a permis de définir les éléments clés conduisant à l'assemblage d'une particule fonctionnelle (Charpentier *et al.*, 2007; Charpentier *et al.*, 2005). Il a notamment démontré qu'un complexe contenant le sRNA, ainsi que les protéines CBF5 et NOP10 constitue l'entité minimale permettant la réaction de pseudouridylation. La présence supplémentaire des protéines L7Ae et GAR1 s'est néanmoins révélée nécessaire pour assurer l'activité optimale de la particule. En comparaison avec les résultats décrits précédemment, ces données révèlent donc une différence fondamentale entre les sRNP à boîtes C/D et H/ACA. En effet, alors que l'assemblage de la particule C/D obéit à un processus séquentiel hautement dépendant du recrutement de la protéine L7Ae sur l'ARN, la formation de la particule à boîtes H/ACA ne semble pas reposer sur les mêmes contraintes. A la différence de la Fibrillarine, l'enzyme CBF5 peut en effet interagir de façon spécifique avec l'ARN, de plus c'est sa fixation et non celle de L7Ae, qui est à l'origine du recrutement des protéines NOP10 et GAR1 (Baker *et al.*, 2005; Charpentier *et al.*, 2005).

Chez les eucaryotes, la caractérisation du fonctionnement des snoRNP à boîtes H/ACA a initialement été entrepris en extraits cellulaires (Dragon *et al.*, 2000; Wang & Meier, 2004; Wang *et al.*, 2002) ou à partir de particules purifiées (Karijolich *et al.*, 2007; Ma *et al.*, 2005a; Xiao *et al.*, 2009). Des expériences réalisées à partir de protéines produites en lysat de réticulocytes ont notamment permis d'établir les interactions entre les composants de la snoRNP (Wang & Meier, 2004). Comme chez les archées, CBF5 peut interagir indépendamment avec NOP10 et GAR1, de plus il est possible de former un complexe trimérique CBF5-NOP10-NHP2 qui reconnaît de façon spécifique les snoRNA à boîtes H/ACA. Plus récemment, l'équipe de K. Ye, a réussi à reconstituer une particule active *in vitro*, à partir de protéines de *S. cerevisiae* co-purifiées sous forme recombinantes, après leur co-expression chez *E. coli* (Li *et al.*, 2011c). A la différence de la protéine L7Ae d'archée, qui joue un rôle critique pour l'activité des sRNP à boîtes H/ACA archées (Baker *et al.*, 2005; Charpentier *et al.*, 2005), les résultats obtenus par cette approche ont révélé que Nhp2p n'est pas nécessaire pour assurer la pseudouridylation, mais qu'elle est néanmoins requise pour assurer un turnover réactionnel optimal (Li *et al.*, 2011c).

Ces dernières années, plusieurs structures tridimensionnelles de sRNP à boîtes H/ACA, contenant ou non l'ARN substrat, ont été résolues (Duan *et al.*, 2009; Li & Ye, 2006; Liang *et al.*, 2007a; Liang *et al.*, 2009a; Zhou *et al.*, 2011) (**Fig. 45**). Ces études ont notamment permis de confirmer le rôle central joué par la protéine CBF5 qui établit de nombreux contacts avec l'ARN et les protéines NOP10 et GAR1. Au sein de la particule, l'ARN guide interagit de façon linéaire avec une surface protéique formée par les protéines CBF5, NOP10 et L7Ae. Les extrémités du sRNA sont stabilisées d'une part grâce à l'interaction de CBF5 avec la boîte ACA et avec la base de la tige, d'autre part grâce à la fixation de L7Ae sur le motif apical en *K-loop*. De façon intéressante, l'appariement 3' entre l'ARN guide et l'ARN substrat dans la poche de pseudouridylation, mime la tige T $\Psi$ C reconnue par l'enzyme TruB au sein des ARNt. De plus, la boucle  $\beta$ 7- $\beta$ 10 de la protéine CBF5 forme une structure « en pouce », appelée *Thumb loop*, qui est conservée sous une forme plus courte au sein de la protéine TruB et qui semble jouer un rôle clé dans l'accessibilité de la poche de pseudouridylation pour le substrat (Duan *et al.*, 2009; Liang *et al.*, 2007a; Liang *et al.*, 2009a). En effet, en absence de substrat, cette boucle, stabilisée par une interaction avec la protéine GAR1, est dans une configuration ouverte favorisant l'accès à un ARN cible. En revanche, en présence de ce dernier, la boucle se referme sur la poche de pseudouridylation et participe au maintien du substrat dans le site actif de l'enzyme. De plus, il a été montré que



**Figure 45 : Structure tridimensionnelle d'une sRNP à boîte H/ACA de *Pyrococcus furiosus* obtenue par radiocristallographie (d'après Duan *et al.*, 2009; PDB : 3HAY). La structure a été obtenue par radiocristallographie. L'ARN substrat (en jaune) est positionné dans la poche de pseudouridylation. Le nucléotide ciblé par la modification, ici une fluorouridine, est représenté en rouge. Le site actif de l'enzyme CBF5 est refermé par la boucle  $\beta 7$ -  $\beta 10$  (Thumb loop).**



des changements de conformation de l'ARN guide en présence de la protéine L7Ae favorisent également son interaction avec le substrat (Liang *et al.*, 2007a). Des résultats obtenus au laboratoire par J.B. Fourmann et M. Blaud, indiquent que ce rôle de L7Ae repose sur des contraintes induites par son interaction avec la protéine NOP10 au sein de la particule (Fourmann, Blaud *et al.*, non publié). Chez *S. cerevisiae*, la résolution de la structure cristallographique d'un complexe Nop10p-Cbf5p-Gar1p partiel, obtenu par co-expression chez *E. coli*, a révélé la présence au niveau de la protéine Gar1p, d'une extension C-terminale, baptisée domaine CTE (*C-Terminal Extension*), absente chez les archées (Li *et al.*, 2011c). Cette dernière est responsable de contacts hydrophobes supplémentaires avec le pouce de Cbf5p et pourrait permettre, dans des conditions de températures moins élevées, un meilleur contrôle du turnover réactionnel.

Enfin, l'analyse de ces données structurales a également révélé une grande similarité entre les mécanismes d'action de l'enzyme CBF5 et de la pseudouridyl-synthase bactérienne TruB, responsable de la modification des ARNt. Il a notamment été proposé que ces deux protéines appartenant à la même famille, dérivent d'une même ARNt-pseudouridyl-synthase ancestrale (Lafontaine & Tollervey, 1998). Chez les eucaryotes, une duplication du gène codant ce précurseur pourrait ainsi être à l'origine de l'apparition de deux enzymes distinctes : la protéine PUS4, responsable, comme la protéine TruB bactérienne, de la modification des ARNt, et la protéine CBF5 (Lafontaine & Tollervey, 1998). Ces données, ainsi que la capacité de CBF5 à fonctionner, seule ou en association avec les protéines NOP10 et GAR1, dans le cadre de la modification non guidée des ARNt (Fourmann, 2009; Gurha *et al.*, 2007; Muller *et al.*, 2007; Roovers *et al.*, 2006), suggère que l'acquisition de sa fonction guidée, constitue un événement évolutif secondaire (Lafontaine & Tollervey, 1998). Cette hypothèse pourrait notamment expliquer les différences majeures observées entre les sRNP à boîtes C/D et H/ACA d'archées. En effet, alors que la protéine L7Ae est essentielle pour initier la formation d'une particule C/D active, la possibilité d'assembler un module catalytique CBF5-GAR1-NOP10 actif semble totalement modifier les règles d'assemblage des particules H/ACA, ainsi que l'importance fonctionnelle de ses différents composants. A la différence des autres RNP de la famille L7Ae, la construction des particules H/ACA autour de l'enzyme CBF5 et non autour du complexe *K-turn*-L7Ae/Nhp2p/NHP2, pourrait donc être responsable d'une modification de la pression de sélection, à l'origine de l'apparition du *phylum* NHP2 et de la perte des motifs en *K-turn* dans les snoRNA à boîtes H/ACA eucaryotes.

### 3.2.4. Le complexe télomérase chez les vertébrés

De façon surprenante, un motif H/ACA a également été découvert dans la région 3' de l'ARN télomérique (TERC) chez les vertébrés (Mitchell *et al.*, 1999a). Ce motif possède une organisation classique, il présente en effet deux structures en tige-boucle et permet le recrutement des protéines constitutives des snoRNP à boîtes H/ACA (Antal *et al.*, 2002; Dez *et al.*, 2001; Dragon *et al.*, 2000; Mitchell *et al.*, 1999b). Cette particule H/ACA n'a pour l'instant été associée à aucune réaction de pseudouridylation, en revanche, elle est essentielle pour la maturation de l'ARN TERC et pour l'activité du complexe télomérase (Mitchell *et al.*, 1999a). Chez la levure, TLC1, le composant ARN de la télomérase, ne possède pas de motif H/ACA, en revanche, comme c'est le cas pour les UsnRNA, sa région 3' s'associe avec les protéines Sm (Seto *et al.*, 1999).

### 3.2.5. Des mutations à l'origine de la Dyskératose Congénitale touchent des composants des particules à boîtes H/ACA

La dyskératose congénitale (DC) ou syndrome de Zinsser-Cole-Engman est une maladie génétique rare dont la prévalence est estimée à une personne sur un million (Walne & Dokal, 2009). Cette maladie polymorphe se caractérise par des altérations des tissus à fort taux de renouvellement, comme la peau et la moelle osseuse, et peut dans certains cas aboutir à la survenue de cancers. Il existe trois formes distinctes de DC : une forme liée au chromosome X, une forme autosomale dominante et une forme autosomale récessive. Bien que dans plus de 60% des cas, les facteurs génétiques impliqués dans le développement de la maladie soient inconnus, des mutations touchant certains composants des particules à boîtes H/ACA ont été identifiées pour chacune de ces formes. Ainsi, la forme liée au chromosome X peut impliquer des mutations au niveau de l'enzyme CBF5, la forme autosomale dominante a été associée à des mutations dans les gènes codant l'ARN TERC du complexe de la télomérase et sa transcriptase inverse TERT (*Telomerase Reverse Transcriptase*), enfin des mutations relevées dans les gènes codant NHP2, NOP10 et TERT ont également été identifiées dans des cas de forme autosomale récessive.

Etant donné que ces mutations touchent à la fois les snoRNP à boîtes H/ACA, les scaRNP et le complexe télomérase, il est relativement difficile d'établir la part de chacun des défauts engendré (biogenèse des ribosomes, épissage et réplication des télomères) dans développement de la pathologie. Néanmoins, les nouvelles données structurales obtenues à partir du modèle archée apportent des informations précieuses pour essayer de comprendre les dysfonctionnements engendrés à l'échelle moléculaire par certaines de ces mutations (Vulliamy & Dokal, 2008). Une étude récente, basée sur la reconstitution de particules à partir de protéines produites en lysat de réticulocytes, a permis de mieux comprendre l'impact de certaines mutations sur la formation des RNP (Trahan *et al.*, 2010). De façon intéressante, les mutations V126M et Y139H de la protéine NHP2 inhibent son interaction avec la protéine NOP10 qui constitue une interface clé pour la formation de la particule. La mutation R34W de NOP10 affecte quant à elle la capacité de la protéine à interagir avec l'ARN à boîtes H/ACA. En revanche, de façon surprenante, la plupart des mutations de la protéine CBF5, qui se localisent au niveau du domaine PUA, n'ont pas d'effet majeurs sur la formation des particules *in vitro*. Comme nous l'aborderons par la suite, d'autres données suggèrent que ces mutations n'affectent pas directement les contacts entre les constituants de la RNP, mais altèrent certaines étapes de la biogenèse des particules (Grozdanov *et al.*, 2009a).



## CHAPITRE 3

### Biogenèse des RNP de la famille L7Ae

Le système archée constitue un excellent modèle, qui a fourni de nombreuses informations sur la structure et sur le fonctionnement des particules à boîtes C/D et H/ACA. La production des différentes protéines sous forme recombinante, la reconstitution de particules actives *in vitro* et la résolution de leurs structures tridimensionnelles a permis de mettre à jour les bases moléculaires de leur assemblage et de l'acquisition d'une activité catalytique spécifique. Chez les eucaryotes, la reconstitution de snoRNP fonctionnelles *in vitro* s'avère en revanche beaucoup plus complexe et semble nécessiter le recours à des techniques de co-expression de protéines recombinantes chez *E. coli* (Li *et al.*, 2011c). De telles approches permettent en effet de parer à la faible solubilité individuelle des protéines, en les plaçant dans un contexte proche de leur environnement au sein des RNP matures. Cette différence caractéristique semble inhérente à la complexité des processus de biogenèse se déroulant dans les cellules eucaryotes. A l'image de la production des ribosomes, la formation des RNP eucaryotes repose en effet sur l'intervention de facteurs qui dirigent et coordonnent en *trans* les grandes étapes de biogenèse. Ainsi, dès leur synthèse, les différentes sous-unités de chaque RNP sont prises en charge par une multitude de facteurs dont elles sont devenues intrinsèquement dépendantes et qui assurent des fonctions de modification, de chaperonnage, de transport et d'assemblage. Parmi les mécanismes les mieux caractérisés, la production des UsnRNP spliceosomales et celle des snoRNP à boîtes H/ACA, illustrent de façon remarquable, différents aspects de ces phénomènes. La formation des UsnRNP implique notamment le complexe SMN, qui assure un contrôle de la spécificité d'assemblage des particules. La biogenèse des snoRNP à boîtes H/ACA fournit quant à elle un modèle d'assemblage co-transcriptionnel et fait appel à une stratégie de mimétisme moléculaire, qui consiste au remplacement transitoire de certains composants de la RNP par des facteurs d'assemblage structurellement proches. Après avoir défini les principes gouvernant l'assemblage des complexes macromoléculaires, nous passerons donc en revue ces deux processus, qui constituent des exemples détaillés concernant certaines RNP de la famille L7Ae. Puis, nous

terminerons en dressant l'état des connaissances sur la biogenèse des snoRNP à boîtes C/D au début de mon travail de thèse.

## 1. Principes généraux régissant l'assemblage des macro-complexes *in vivo*

D'un point de vue thermodynamique, l'assemblage d'un complexe macromoléculaire, comme une RNP, dépend de la probabilité de rencontre de ses différents constituants et de la stabilité de leur association (Berg & von Hippel, 1985). De façon paradoxale, le milieu intracellulaire constitue un environnement relativement défavorable, puisque la densité moléculaire élevée qui y règne (20 à 30% du volume total), contraint la libre diffusion des molécules et favorise les interactions aspécifiques (Ellis, 2001). De plus, les contacts entre les différentes sous-unités protéiques d'un macro-complexe impliquent généralement des régions hydrophobes, dont l'exposition au solvant favorise l'agrégation des protéines à l'état isolé (Ellis, 2001). Enfin, il est fréquent que le repliement complet de certaines sous-unités ne s'effectue que par le biais de contacts intermoléculaires, s'établissant au sein des complexes matures (Luger *et al.*, 1997; Pomeranz Krummel *et al.*, 2009). L'ensemble de ces caractéristiques biologiques constitue donc autant d'obstacles aux processus d'assemblage et plusieurs stratégies ont émergées, au cours de l'évolution, afin d'en maintenir l'efficacité et la spécificité (Chari & Fischer, 2010) : (i) la synthèse des différents constituants d'un complexe macromoléculaire fait généralement l'objet d'un contrôle précis et coordonné, visant à assurer leur production en quantité adéquate, (ii) l'interférence entre la biogenèse des particules et leur implication fonctionnelle dans le métabolisme cellulaire est souvent limitée par la séparation spatiale de ces processus, (iii) les cellules ont également recours à des protéines chaperonnes qui assurent un pré-repliement de certaines sous-unités et évitent la mise en place de structures ou d'interactions trop précoces, (iv) enfin, d'autres facteurs agissant en *trans* dans différents processus (*i.e.* maturation, modifications, transport), interviennent de façon interdépendante et contribuent à la hiérarchisation des différentes étapes de biogenèse.

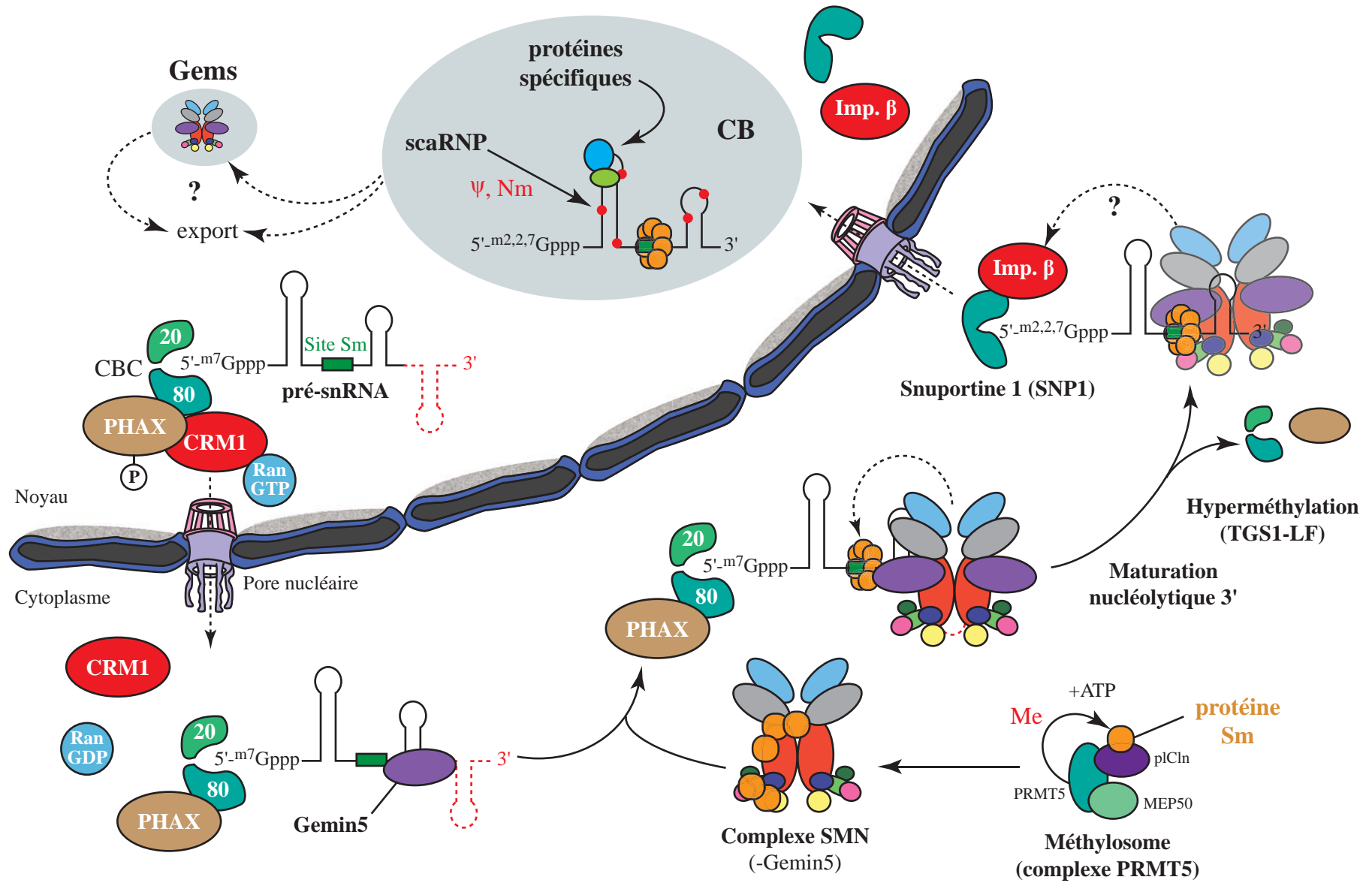
## 2. Biogenèse des UsnRNP, un modèle pour le contrôle de la spécificité d'assemblage

Chez les animaux, chez les protistes et chez les plantes, la biogenèse des snRNP U1, U2, U4 et U5 est un processus dynamique qui implique le transit des UsnRNA par le cytoplasme (Patel & Bellini, 2008; Will & Luhrmann, 2001) (**Fig. 46**). Durant cette phase, les protéines Sm sont assemblées sur l'ARN au cours d'un processus dirigé par le complexe chaperon SMN. Chez la levure *S. cerevisiae*, qui constitue un modèle privilégié pour l'étude des UsnRNP, aucune étape d'assemblage cytoplasmique n'a en revanche été observée. De plus, alors que la fonction de la protéine SMN semble conservée chez *Schizosaccharomyces pombe* (Owen *et al.*, 2000; Paushkin *et al.*, 2000), aucun facteur équivalent n'a été identifié chez *S. cerevisiae*. Pour cette raison, nous nous attacherons à décrire principalement dans cette partie, les processus de biogenèse des snRNP chez les animaux. La snRNP U6, dont l'ARN est transcrit par l'ARN polymérase III, suit quant à elle une voie spécifique, exclusivement nucléaire et indépendante du complexe SMN.

### 2.1. Voies de biogenèse des UsnRNP

#### 2.1.1. Biogenèse nucléocytoplasmique des snRNP U1, U2, U4 et U5

Comme nous l'avons abordé au cours du premier chapitre, les pré-UsnRNA, qui sont transcrits par l'ARN polymérase II, acquièrent de façon spécifique une coiffe  $m^7G$  et possèdent une extension au niveau de leur extrémité 3' (**Fig. 46**). A la suite de la transcription, le complexe CBC interagit spécifiquement avec la coiffe  $m^7G$  et assure le recrutement de l'adaptateur PHAX (Ohno *et al.*, 2000; Segref *et al.*, 2001) (**Fig. 3**). Le facteur d'export CRM1, en association avec le complexe Ran-GTP, reconnaît alors l'adaptateur PHAX sous sa forme phosphorylée et exporte l'ARN vers le cytoplasme (Askjaer *et al.*, 1999; Ohno *et al.*, 2000; Segref *et al.*, 2001). Les étapes initiales de la biogenèse du snRNA U2 pourraient néanmoins présenter certaines spécificités, puisqu'un précurseur de cet ARN a été détecté au sein des CB, suggérant l'existence d'un premier transit par ces corps nucléaires avant la phase d'export cytoplasmique (Smith & Lawrence, 2000).



**Figure 46 : Représentation schématique de la voie de biogenèse des snRNP U1, U2, U4 et U5 chez les animaux.** Les membres du complexe SMN sont représentés par le code couleur suivant : SMN (rouge); Gemin2 (jaune); Gemin3 (gris); Gemin4 (bleu clair); Gemin5 (mauve); Gemin6 (rose); Gemin7 (vert clair); Gemin8 (bleu foncé); Unrip (vert foncé). Les protéines Sm sont représentées en orange. CB : Cajal Bodies; Ψ : pseudouridine; Nm : 2'-O-méthylation.



Suite à l'export, l'hydrolyse du GTP et la déphosphorylation de PHAX conduisent à la dissociation partielle du complexe d'export (Kitao *et al.*, 2008; Ohno *et al.*, 2000; Segref *et al.*, 2001). PHAX demeure en effet associée au complexe pré-UsnRNA/CBC jusqu'à l'hyperméthylation de la coiffe, vraisemblablement afin d'éviter toute interaction avec le facteur d'initiation de la traduction eIF4E. Au sein du cytoplasme et en coopération avec le complexe PRMT5, le complexe SMN facilite l'assemblage spécifique des 7 protéines Sm sur l'extrémité 3' de l'ARN et reste associé à la particule (Massenet *et al.*, 2002; Meister *et al.*, 2001a; Meister *et al.*, 2001c; Pellizzoni *et al.*, 2002b). Les protéines Sm et le complexe SMN coopèrent ensuite pour recruter l'hyperméthylase TGS1 cytoplasmique (Mouaikel *et al.*, 2003) et pour permettre la maturation nucléolytique finale du pré-UsnRNA (Seipelt *et al.*, 1999; van Hoof *et al.*, 2000b). On remarquera que chez la levure *S. cerevisiae*, l'hyperméthylation de la coiffe, qui requiert également l'assemblage préalable des protéines Sm, se déroule au sein du nucléole (Mouaikel *et al.*, 2002). La coiffe <sup>m<sup>2,2,7</sup></sup>G nouvellement synthétisée et l'heptamère de protéines Sm constituent alors un signal de localisation nucléaire bipartite. La coiffe triméthylée est en effet reconnue par la protéine Snurportine 1 (SNP1), qui interagit avec l'importine β et qui assure le ré-import de la particule vers le noyau (Huber *et al.*, 1998; Palacios *et al.*, 1997). Le mode de contribution des protéines Sm à ce processus demeure incertain. Il a été proposé qu'elles assurent le recrutement d'un facteur d'import non identifié (Palacios *et al.*, 1997), néanmoins les protéines Sm pourraient également interagir avec l'importine β, afin de favoriser son recrutement (Fischer *et al.*, 1993; Narayanan *et al.*, 2002).

Après l'entrée dans le noyau, les facteurs d'import sont dissociés et les pré-UsnRNP sont acheminées transitoirement vers les CB (Sleeman & Lamond, 1999b). Au sein de ces corpuscules, les UsnRNA subissent des réactions de 2'-O-méthylation et de pseudouridylation spécifiques, guidées par des scaRNA (Darzacq *et al.*, 2002; Jady *et al.*, 2003). Les UsnRNA contiennent en effet de nombreux nucléotides modifiés par 2'-O-méthylation, pseudouridylation et méthylation de bases (Massenet *et al.*, 1998). Comme c'est le cas pour les ARNr, l'abondance de ces modifications semble corrélée au degré de complexité de l'organisme, puisque le nombre de résidus modifiés est largement plus élevé chez les vertébrés que chez les organismes unicellulaires (Massenet *et al.*, 1998). De façon intéressante, ces modifications se concentrent essentiellement dans les régions impliquées dans des appariements intermoléculaires lors de l'assemblage du spliceosome et pourraient contribuer à stabiliser son architecture. L'importance des modifications a notamment été

démontrée pour le snRNA U2, puisque l'abolition de certaines d'entre elles perturbent l'assemblage final de la snRNP, ainsi que la mise en place et l'activité du spliceosome (Donmez *et al.*, 2004; Lin & Kielkopf, 2008; Newby & Greenbaum, 2001; Yu *et al.*, 1998; Zhao & Yu, 2004). Chez les vertébrés, le snRNA U2 contient notamment 2 résidus 2'-O-méthylés et 6 pseudouridines au niveau de sa région d'interaction avec la boîte de branchement (Massenet *et al.*, 1999). Chez la levure seuls trois d'entre eux, les résidus Ψ35, Ψ42 et Ψ44, sont conservés (Massenet *et al.*, 1999). Alors que les positions 35 et 44 sont respectivement modifiées par les enzymes Pus7p et Pus1p (Ma *et al.*, 2003; Massenet *et al.*, 1999), une étude plus récente a démontré que la position 42 est ciblée par le snoRNA à boîtes H/ACA snR81 (Ma *et al.*, 2005a). Ce résultat est d'autant plus intéressant que cet ARN guide également la pseudouridylation du résidu U1051 de l'ARNr 25S (Schattner *et al.*, 2004). Malgré l'absence de scaRNP chez la levure, l'existence de modifications guidées au niveau des UsnRNA suggère qu'il existe une phase nucléolaire au cours de la biogenèse de la snRNP U2. Plus précisément, ce type de modification pourrait s'effectuer au sein des NB, qui partagent un certain nombre de similarités fonctionnelles avec les CB (Verheggen *et al.*, 2001) (cf. Introduction, Chapitre 1, 1.5.2) (**Fig. 6**). Enfin, de façon surprenante, une étude récente a révélé l'existence de deux pseudouridylations supplémentaires induites en conditions de stress sur le snRNA U2 de levure (Wu *et al.*, 2011). Celles-ci concernent les positions 56 et 93, qui sont localisées au niveau d'une région de l'ARN sujette à de larges remaniements conformationnels au cours de l'épissage. La position 56 est ciblée par l'enzyme Pus7p, tandis que la position 93 est ciblée également par la snoRNP à boîtes H/ACA snR81. Les auteurs ont démontré que le caractère inductible de ces modifications repose sur la reconnaissance d'une séquence cible sous-optimale. De plus, ils proposent que leur incidence en conditions de stress contribue à stabiliser certaines structures intramoléculaires essentielles à la réaction d'épissage.

En plus des étapes de modifications et d'après l'exemple de la biogenèse de la snRNP U2, il a été proposé que l'assemblage des protéines spécifiques de chaque UsnRNP se déroule au sein des CB (Nesic *et al.*, 2004). Le relargage du complexe SMN de la particule pourrait également s'effectuer au niveau de ces régions. En effet, la Coiline semble faciliter cette dissociation en rentrant en compétition avec le complexe SMN pour la fixation de la protéine SmB' (Hebert *et al.*, 2001). Une étude récente a notamment montré que l'état de phosphorylation de la Coiline pourrait réguler ces événements (Toyota *et al.*, 2010). Sous une forme initialement déphosphorylée, cette dernière faciliterait le recrutement des UsnRNP en

interagissant avec la protéine SMN, puis sa phosphorylation favoriserait son interaction avec la protéine SmB' au détriment de la protéine SMN qui serait alors dissociée du complexe. Enfin, à l'issue de ces étapes, les UsnRNP matures s'accumulent dans les régions inter-chromatiniennes, au niveau des speckles, qui constituent des sites de stockage dédiés à l'approvisionnement des sites de transcriptions (Lamond & Spector, 2003) (cf. Introduction, Chapitre 1, 1.3) (**Fig. 4**).

En parallèle de ces données, des travaux réalisés chez le xénope ont mis en évidence un trafic particulier des UsnRNP (Gerbi *et al.*, 2003a). En effet, des expériences de micro-injections dans des oocytes ont révélé la présence des snRNA U2, U4 et U5 au sein du nucléole (Gerbi & Lange, 2002; Yu *et al.*, 2001). Concernant le snRNA U2, cette localisation transitoire dépend de la présence des protéines Sm et de l'hyperméthylation de la coiffe, elle semblerait donc postérieure au passage de l'ARN dans le cytoplasme (Yu *et al.*, 2001). Concernant les snRNA U4 et U5, l'adressage nucléolaire est observé très rapidement après l'injection, il ne dépend pas de la fixation des protéines Sm, ni de l'hyperméthylation de la coiffe, ni de la présence du snRNA U6, et pourrait donc se dérouler avant l'export vers le cytoplasme (Gerbi *et al.*, 2003b). De plus, il a été montré que la localisation nucléolaire du snRNA U4 requiert la présence de la protéine 15.5K, suggérant que le recrutement de ce facteur sur l'ARN pourrait se dérouler avant l'export cytoplasmique (Gerbi *et al.*, 2003b). Quelle soit précoce (U4 et U5), ou plus tardive (U2), la présence transitoire des UsnRNA au sein du nucléole, suggère donc l'existence d'étapes de maturation spécifiques au sein de ce domaine nucléaire. Les études réalisées sur le snRNA U2 ont notamment démontré que certaines réactions de 2'-O-méthylation et de pseudouridylation s'effectuent au sein du nucléole (Yu *et al.*, 2001). Ceci suggère donc, qu'en plus de l'intervention des scaRNP dans les CB, les UsnRNA pourraient également subir des modifications catalysées par les snoRNP à boîtes C/D et H/ACA au sein du nucléole.

### **2.1.2. Biogenèse nucléaire de la snRNP U6**

Le snRNA U6, transcrit par l'ARN polymérase III et porteur d'une coiffe  $\gamma$ -monométhylphosphate (mpppG) (Singh & Reddy, 1989), suit une voie de biogenèse spécifique. Après sa synthèse, le UsnRNA est stabilisé par la fixation de la protéine La

(Lhp1p chez la levure) au niveau d'une séquence poly(U) terminale (Kufel *et al.*, 2003; Wolin & Cedervall, 2002). Dans un second temps, les protéines Lsm (*Sm-like*), qui sont phylogénétiquement apparentées aux protéines Sm et qui sont conservées chez les eucaryotes, sont recrutées à la place de la protéine La (Achsel *et al.*, 1999; Pannone *et al.*, 2001; Seraphin, 1995). Leur assemblage, sous la forme d'un hexamère au niveau de la séquence polyU, constitue un élément essentiel pour l'adressage successif de la particule, vers le nucléole, puis vers les CB (Gerbi & Lange, 2002). Au sein du nucléole, l'ARN subit ensuite des réactions de 2'-O-méthylation et de pseudouridylation spécifiques catalysées par les snoRNP à boîtes C/D et H/ACA (Ganot *et al.*, 1999; Lange & Gerbi, 2000; Tycowski *et al.*, 1998). Avant son accumulation au niveau des speckles, la particule transite par les CB. Le snRNA U6 s'y apparie avec le snRNA U4 pour former la di-snRNP U4/U6, puis la tri-snRNP U4/U6.U5 qui constitue une entité fonctionnelle pour la réaction d'épissage (Liu *et al.*, 2006; Novotny *et al.*, 2010; Schaffert *et al.*, 2004; Stanek *et al.*, 2003) (cf. Introduction, Chapitre 2, 2.2) (**Fig. 33**).

## 2.2. Rôle du complexe SMN dans l'assemblage spécifique des UsnRNP

Comme nous l'avons évoqué précédemment, le domaine 3' (*core domain*) des snRNP U1, U2, U4 et U5, est formé par l'oligomérisation des protéines Sm B/B', D1, D2, D3, E, F et G au niveau d'une séquence de type [RA(U)<sub>5-6</sub>GR] située dans la région 3' de chaque UsnRNA (Branlant *et al.*, 1982; Kambach *et al.*, 1999; Leung *et al.*, 2011; Liautard *et al.*, 1982; Pomeranz Krummel *et al.*, 2009; Raker *et al.*, 1999; Stark *et al.*, 2001) (cf. Introduction, chapitre 2, 2.1.3.1) (**Fig. 29 & 31**). Plusieurs études ont montré la possibilité d'assembler directement les protéines Sm sur un UsnRNA *in vitro*, suggérant que ces différents constituants contiennent l'information nécessaire pour l'assemblage du domaine *core* des UsnRNP (Hermann *et al.*, 1995; Leung *et al.*, 2011; Pomeranz Krummel *et al.*, 2009; Raker *et al.*, 1999). Néanmoins, l'étude détaillée de ce processus d'assemblage dans le contexte cellulaire, a révélé que la formation du domaine 3', constitue un processus séquentiel complexe, qui nécessite une source d'énergie et l'intervention de nombreux facteurs protéiques (Gubitz *et al.*, 2004).

La phase cytoplasmique d'assemblage des UsnRNP implique en effet l'intervention de deux complexes protéiques majeurs (**Fig. 46**). D'une part, le complexe SMN, composé des

protéines SMN, Unrip et Gemin 2 à 8 (Baccon *et al.*, 2002; Carissimi *et al.*, 2005; Carissimi *et al.*, 2006; Charroux *et al.*, 1999; Charroux *et al.*, 2000; Grimmier *et al.*, 2005b; Gubitzi *et al.*, 2002; Liu & Dreyfuss, 1996; Pellizzoni *et al.*, 2002a). D'autre part, le méthylosome 20S, formé par l'association de l'arginine-méthyltransférase PRMT5/JBP1 et des protéines pICln et MEP50 (Friesen *et al.*, 2001b; Friesen *et al.*, 2002; Meister *et al.*, 2001b; Pu *et al.*, 1999). Au cours de ce processus, le complexe SMN agit en tant que complexe chaperon en facilitant l'assemblage des protéines Sm sur le UsnRNA (Buhler *et al.*, 1999; Fischer *et al.*, 1997; Liu *et al.*, 1997; Meister *et al.*, 2001a; Pellizzoni *et al.*, 1998). Cette assistance, assure notamment un contrôle de la spécificité d'assemblage, puisque des expériences réalisées en extrait cellulaire ont démontré que l'absence du complexe SMN conduit à la formation de complexes illégitimes entre les protéines Sm et de nombreux ARN (Pellizzoni *et al.*, 2002b). Cette fonction de contrôle repose sur la capacité du complexe SMN à interagir à la fois avec les protéines Sm et avec les UsnRNA. Ces derniers contiennent en effet des déterminants permettant le recrutement spécifique du complexe SMN (Golembe *et al.*, 2005; Yong *et al.*, 2004; Yong *et al.*, 2002). A l'exception du snRNA U1, le domaine reconnu par le complexe SMN comprend généralement les structures en tige-boucle de la région 3' terminale des UsnRNA, ainsi que le site de fixation des protéines Sm (Yong *et al.*, 2004; Yong *et al.*, 2002). Des expériences récentes ont par ailleurs démontré que la protéine Gemin5 est responsable de cette reconnaissance (Battle *et al.*, 2006; Yong *et al.*, 2010). Elle se fixerait sur les pré-UsnRNA porteurs d'une extension au niveau de l'extrémité 3' et favoriserait de façon secondaire le recrutement des autres sous-unités du complexe SMN pré-associées aux protéines Sm (Yong *et al.*, 2010). En effet, les différents constituants du complexe SMN établissent des interactions spécifiques avec plusieurs sous-groupes de protéines Sm. Les protéines Gemin6 et Gemin7 possèdent notamment des structures proches de celles des protéines Sm, qui pourraient faciliter le recrutement de ces dernières sur le complexe SMN (Ma *et al.*, 2005b). De plus, il a été montré que les protéines SmB/B', SmD1 et SmD3, possèdent des domaines GAR, riches en glycine et en arginine, qui font l'objet de méthylation et qui sont importants pour l'assemblage des particules (Brahms *et al.*, 2001; Friesen *et al.*, 2001a). Plusieurs études ont mis en évidence que ces résidus sont ciblés par le méthylosome avant leur interaction avec le complexe SMN (Brahms *et al.*, 2001; Friesen *et al.*, 2001b; Meister *et al.*, 2001b; Meister & Fischer, 2002; Pu *et al.*, 1999). La protéine pICln possède une forte affinité pour les domaines Sm et permettrait la séquestration des protéines Sm nouvellement synthétisées (Pu *et al.*, 1999). Au cours d'un mécanisme ATP-dépendant, l'enzyme PRMT5 catalyserait ensuite la diméthylation symétrique des résidus arginines afin

« d'activer » les protéines Sm, en leur conférant une affinité accrue pour le complexe SMN (Brahms *et al.*, 2001; Friesen *et al.*, 2001a; Friesen *et al.*, 2001b; Meister *et al.*, 2001b). Enfin, l'association physique du méthylosome avec le complexe SMN permettrait le transfert des protéines Sm activées vers le complexe chaperon afin de procéder aux étapes finales d'assemblage (Meister & Fischer, 2002). En plus de ces événements précoces de méthylation, une autre étude a révélé que la fonction du complexe SMN pourrait être régulée par la phosphorylation de la protéine SMN (Grimmler *et al.*, 2005a). Le complexe SMN isolé de la fraction cytoplasmique, qui s'avère compétent pour l'assemblage des UsnRNP en extrait cellulaire, est associé à une forme phosphorylée de la protéine SMN. En revanche, le complexe purifié à partir de la fraction nucléaire, qui ne permet plus d'assurer la fonction d'assemblage, renferme une forme hypophosphorylée de la protéine. De plus, les auteurs ont montré que la mutation des résidus sérines ciblés par les phosphorylations, conduit à la perte de l'activité chaperonne du complexe SMN, suggérant ainsi un rôle clé de ces modifications dans l'activation du complexe chaperon.

A titre de comparaison, l'assemblage des protéines Sm (Smd1p, Smd2p, Smd3p, Sme1p, Smx3p, Smx2p) chez la levure *S. cerevisiae* constitue vraisemblablement un processus nucléaire. A défaut de la présence du complexe SMN, il a été proposé que la protéine Lhp1p contribue à la formation du domaine *core* des UsnRNP (Xue *et al.*, 2000a). Son interaction avec l'extension 3' des pré-UsnRNA est en effet importante pour promouvoir l'assemblage de la protéine Smd1p sur les UsnRNA.

Enfin, en marge de son rôle dans l'assemblage des UsnRNP, le complexe SMN a été impliqué dans d'autres processus cellulaires, démontrant ou suggérant son rôle dans l'assemblage d'autres catégories de RNP :

(i) L'assemblage de la snRNP U7, impliquée dans la maturation de l'extrémité 3' des ARNm d'histones, est notamment dépendant du complexe SMN (Pillai *et al.*, 2003). Au sein de son domaine 3', les protéines SmB, Smd3, SmE, SmF et SmG, sont associées aux protéines Lsm10 et Lsm11 qui constituent également des substrats du complexe SMN.

(ii) Au laboratoire, N. Piazzon a également démontré l'association du complexe SMN avec la protéine FMRP, qui a été impliquée dans les processus de transport et de traduction des ARNm au sein des cellules neuronales (Piazzon *et al.*, 2008).

(iii) Enfin, plusieurs études ont révélé l'interaction de la protéine SMN avec les protéines Fibrillarine et GAR1, constituants respectifs des snoRNP à boîtes C/D et H/ACA (Jones *et al.*, 2001; Liu & Dreyfuss, 1996; Pellizzoni *et al.*, 2001). A l'image des protéines Sm, ces deux facteurs sont caractérisés par la présence de séquences riches en dipeptides glycine-arginine. De façon intéressante, la mutation du domaine GAR de la Fibrillarine conduit à une perte d'interaction avec la protéine SMN (Jones *et al.*, 2001; Pellizzoni *et al.*, 2001). De plus, l'expression d'un variant de la protéine SMN, déficient pour l'assemblage des UsnRNP, est associée à une localisation aberrante de la Fibrillarine et du snoRNA U3 au niveau des structures périnucléolaires (Pellizzoni *et al.*, 2001). Enfin, la protéine SMN est requise pour l'accumulation dans les CB, du facteur NOPP140, une chaperonne impliquée dans la biogenèse des snoRNP (Renvoise *et al.*, 2009).

### 2.3. Assemblage des UsnRNP et syndrome d'amyotrophie spinale

De nombreuses maladies génétiques sont liées à des défauts du processus d'épissage. Le plus fréquemment, les mutations incriminées dans la survenue de ces pathologies impliquent des éléments de séquences agissant en *cis* dans la régulation de l'épissage et affectent spécifiquement la biogenèse de certains ARNm (Cartegni *et al.*, 2002). Néanmoins, dans certains cas, les mutations touchent des éléments agissant en *trans* et conduisent à des effets plus larges. Parmi cette seconde catégorie, l'une des pathologies les plus étudiée est l'amyotrophie spinale ou SMA (*Spinal Muscular Atrophy*), un désordre neuromusculaire monogénique, causé par la réduction du taux de la protéine SMN (Chari *et al.*, 2009).

La SMA est une maladie autosomale récessive, caractérisée par une dégénérescence des motoneurons  $\alpha$  de la colonne vertébrale (Kostova *et al.*, 2007). Cette pathologie, qui touche une naissance sur 6000, conduit à une diminution progressive des fonctions musculaires et provoque une forte mortalité infantile. Sa survenue a été corrélée à des mutations au niveau du gène codant la protéine SMN, baptisée *Survival of Motor Neurons* en référence à la maladie (Lefebvre *et al.*, 1995). Chez l'homme, ce gène est présent au niveau du chromosome 5, sous la forme de deux copies issues d'un événement de duplication : une copie télomérique (SMN1) et une copie centromérique (SMN2) (Lefebvre *et al.*, 1995). Pour plus de 98% des patients atteints de SMA, la survenue de la pathologie est liée à des mutations ou des délétions

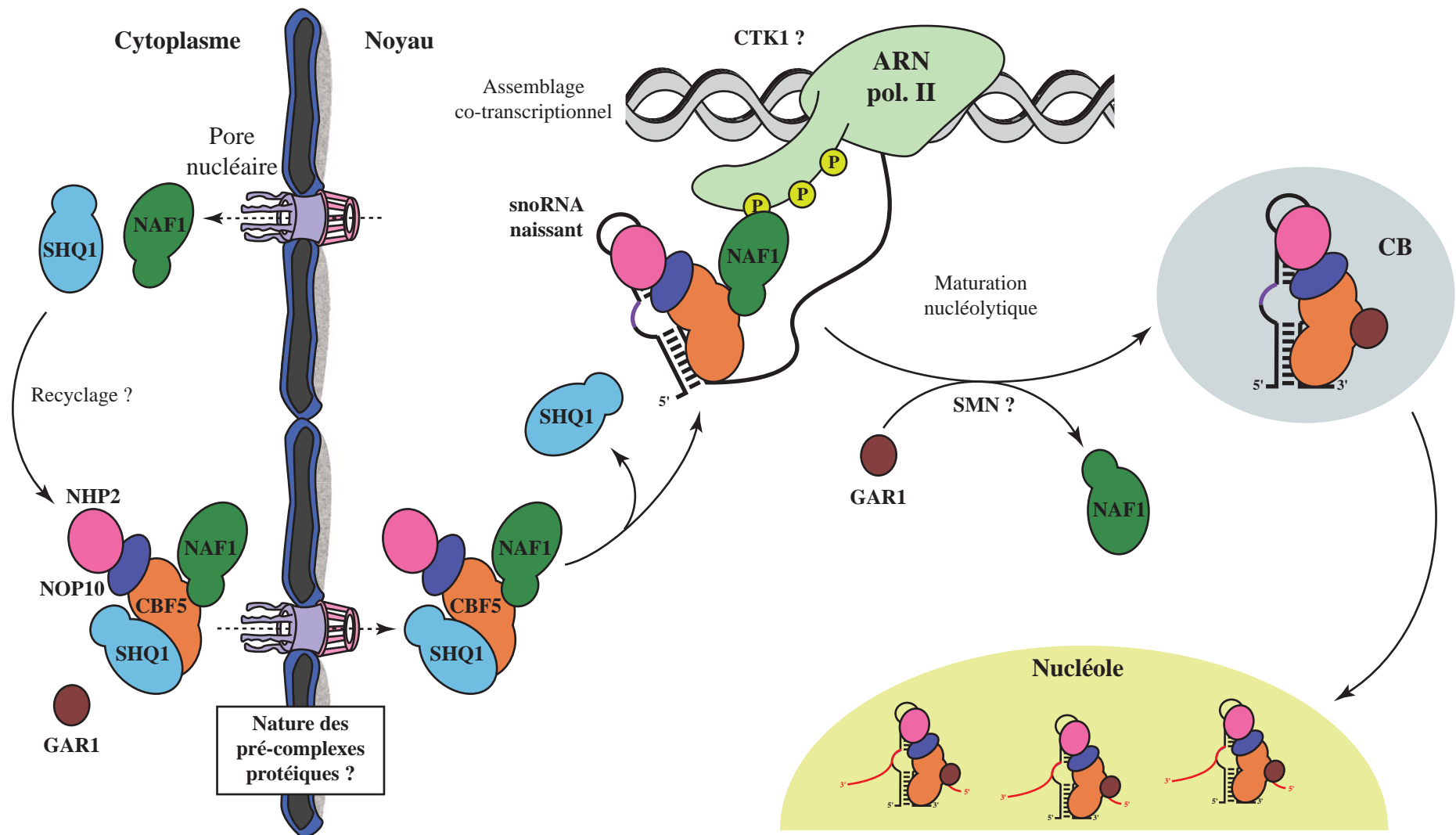
au niveau du gène SMN1 (Lefebvre *et al.*, 1995). Or, un défaut de fonctionnement du gène SMN1 ne peut pas être totalement compensé par le gène SMN2, puisqu'une altération d'épissage inhérente à ce *locus* (substitution conservée dans l'exon 7) conduit à la production majoritaire d'une forme tronquée instable de la protéine (Lorson *et al.*, 1999).

Comme toutes les maladies touchant des facteurs généraux, les mutations du gène SMN1 ont un caractère pléiotrope qui complique l'étude étiologique de la SMA. Une question qui se pose à l'heure actuelle est de comprendre pourquoi les mutations du gène SMN1 affectent spécifiquement les motoneurons ? A ce titre, le lien mis à jour entre SMN et la protéine FMRP, qui joue un rôle important dans les cellules neuronales, constitue un nouvel élément de réflexion (Piazzon *et al.*, 2008). Néanmoins, une étude réalisée chez le Zébrafish a montré que l'injection de UsnRNP matures exogènes dans des motoneurons déficients en protéine SMN permet de rétablir un phénotype normal (Winkler *et al.*, 2005). Ce résultat souligne donc l'importance d'un défaut d'assemblage des UsnRNP dans la pathologie. Une explication fréquemment proposée est que le déclenchement de la SMA soit lié à un effet de seuil. En effet, un taux réduit de UsnRNP fonctionnelles pourrait être toléré par certains tissus, tout en affectant plus spécifiquement certaines cellules, plus tributaires de l'efficacité et de la fidélité de la machinerie d'épissage (Chari *et al.*, 2009).

### **3. Biogenèse des snoRNP à boîtes H/ACA, un modèle d'assemblage co-transcriptionnel**

Au cours de la dernière décennie, de nombreuses études ont permis de mettre à jour les mécanismes de biogenèse des snoRNP à boîtes H/ACA. A la différence des UsnRNA, les snoRNA à boîtes H/ACA ne font pas l'objet d'un export cytoplasmique et l'assemblage des protéines sur l'ARN est un processus exclusivement nucléaire (Pradet-Balade *et al.*, 2011). La biogenèse de ces particules implique notamment l'intervention des facteurs auxiliaires Naf1p/NAF1 et Shq1p/SHQ1, qui coordonnent le transport des protéines vers le noyau et leur assemblage sur l'ARN en cours de transcription (Kiss *et al.*, 2010) (**Fig. 47**).





**Figure 47 : Représentation schématique du processus d'assemblage des snoRNP à boîtes H/ACA.** Ce mécanisme semble conservé entre l'homme et la levure. Le transit par les Cajal-Bodies (CB) et l'intervention éventuelle du complexe SMN constituent des spécificités humaines. La nature des complexes protéiques antérieurs à la formation de la pré-snoRNP demeure incertaine (cf. Texte Introduction, Chapitre 3, 3.3.2).

### 3.1. Transcription et maturation des snoRNA à boîtes H/ACA

Chez la levure, les ARN à boîtes H/ACA sont majoritairement codés au sein d'unités transcriptionnelles mono- ou polycistroniques, alors que chez les vertébrés, tous ces ARN, à l'exception de l'ARN de la télomérase, sont hébergés dans des introns de pré-ARNm (Lestrade & Weber, 2006; Piekna-Przybylska *et al.*, 2007). Comme nous l'avons déjà abordé, cette configuration génétique détermine le mode de production et de maturation de l'ARN (cf. Introduction, Chapitre 1, 2.3.2 et 2.4.2). Les snoRNA cistroniques acquièrent de façon co-transcriptionnelle, une coiffe  $m^7G$  au niveau de leur extrémité 5', tandis que leur extrémité 3' est initialement définie grâce à l'intervention du complexe NRD1 (Steinmetz *et al.*, 2001). Alors que la majorité de ces transcrits subit de façon secondaire une hyperméthylation de la coiffe par l'enzyme Tgs1p (Mouaikel *et al.*, 2002), certains snoRNA monocistroniques peuvent subir une excision de leur extrémité 5'. Cet événement est catalysé par l'action de l'endoribonucléase III Rnt1p, qui intervient également dans le cadre de la maturation des transcrits polycistroniques (Chanfreau *et al.*, 1998a; Chanfreau *et al.*, 1998b). Enfin, comme cela a été démontré pour la protéine Nop1p dans le cas des snoRNA à boîtes C/D (Giorgi *et al.*, 2001; Morlando *et al.*, 2004), la protéine Cbf5p s'est avérée requise pour la terminaison de la transcription des snoRNA à boîtes H/ACA cistroniques, suggérant ainsi l'existence de couplage entre ce processus et les événements d'assemblage des snoRNP (Ballarino *et al.*, 2005). Les snoRNA introniques sont quant à eux générés à partir du débranchement des structures en lasso issues de l'épissage (Kiss & Filipowicz, 1995; Ooi *et al.*, 1998). Dans tous les cas, la définition de l'extrémité 3' mature implique une digestion 3'-5' par l'exosome (Allmang *et al.*, 1999; van Hoof *et al.*, 2000b), tandis que la définition des extrémités 5' non coiffées, pourrait impliquer, comme dans le cas des snoRNA à boîtes C/D, l'activité exonucléase 5'-3' des enzymes Rat1p ou Xrn1p (Petfalski *et al.*, 1998; Qu *et al.*, 1999).

### 3.2. Transport et localisation des snoRNP à boîtes H/ACA

Des études de localisation réalisées chez le xénope avaient initialement montré l'importance des boîtes H et ACA pour la localisation nucléolaire des snoRNA, mais également de l'ARN de la télomérase, qui transite par le nucléole au cours du cycle cellulaire (Lange *et al.*, 1999; Narayanan *et al.*, 1999a). Par la suite, l'équipe de T. Kiss a mis en évidence l'existence d'un

signal de localisation spécifique des CB, retrouvé chez les mammifères, chez la drosophile et chez les plantes (Richard *et al.*, 2003). Ce signal, baptisé CAB box (*Cajal Bodies box*), est une courte séquence UGAG, localisée dans la boucle terminale des tige-boucle 5' et/ou 3' des motifs H/ACA. Il est retrouvé au sein des scaRNA et de l'ARN de la télomérase des vertébrés, et sa mutation conduit à l'accumulation de ces ARN dans le nucléole (Richard *et al.*, 2003; Theimer *et al.*, 2007). Récemment, un nouveau facteur, appelé WDR79/TCAB1 a été impliqué dans l'adressage vers les CB des ARN porteurs de la CAB box (Tycowski *et al.*, 2009). Son recrutement sur les particules, qui requiert à la fois la présence de la CAB box et de la boîte ACA, est en effet nécessaire pour leur accumulation dans les CB. De plus, comme cela a été montré pour les snoRNP à boîtes C/D, les facteurs PHAX et CRM1, pourraient également être impliqués dans certaines étapes de transport intranucléaire des snoRNP à boîtes H/ACA et du complexe télomérase (Boulon *et al.*, 2004; Pradet-Balade *et al.*, 2011). Chez la levure en revanche, le transport des snoRNP à boîtes H/ACA est assez peu connu et ces particules ne semblent pas s'accumuler de façon spécifique dans les NB, comme c'est le cas pour les snoRNP à boîtes C/D (Verheggen *et al.*, 2001).

### 3.3. Assemblage des snoRNP à boîtes H/ACA

Parmi les facteurs intervenant en *trans* au cours de la biogenèse des RNP, les protéines Naf1p/NAF1 et Shq1p/SHQ1 semblent spécifiques des snoRNP boîtes H/ACA. Ces deux protéines, dont la fonction semble conservée chez les eucaryotes, ont été caractérisées chez la levure *S. cerevisiae* (Dez *et al.*, 2002; Fatica *et al.*, 2002b; Yang *et al.*, 2002), puis chez l'humain (Grozdanov *et al.*, 2009b; Hoareau-Aveilla *et al.*, 2006). Elles présentent toutes deux la capacité de transiter entre le cytoplasme et le noyau, mais sont en revanche exclues du nucléole et des CB humains (Dez *et al.*, 2002; Fatica *et al.*, 2002b; Grozdanov *et al.*, 2009b; Kittur *et al.*, 2006; Yang *et al.*, 2002). De plus, les protéines Naf1p/NAF1 et Shq1p/SHQ1, qui ne sont pas associées aux snoRNP matures (Dez *et al.*, 2002; Kittur *et al.*, 2006; Yang *et al.*, 2002), peuvent interagir avec la pseudouridyl-synthase CBF5 *in vivo* et *in vitro* (Darzacq *et al.*, 2006; Dez *et al.*, 2002; Fatica *et al.*, 2002b; Grozdanov *et al.*, 2009b; Yang *et al.*, 2002). La déplétion de ces deux protéines essentielles provoque une diminution drastique du taux de snoRNA à boîtes H/ACA, ainsi que de l'ARN de la télomérase humaine, et affecte le processus de maturation des ARNr (Dez *et al.*, 2002; Fatica *et al.*, 2002b; Grozdanov *et al.*,

2009b; Hoareau-Aveilla *et al.*, 2006; Yang *et al.*, 2002). Enfin, il a été montré que la déplétion de la protéine Naf1p/NAF1 chez la levure et chez l'homme conduit également à un défaut d'accumulation des protéines Cbf5p/Dyskérine, Nop10p/NOP10 et Gar1p/GAR1 (Darzacq *et al.*, 2006; Dez *et al.*, 2002; Hoareau-Aveilla *et al.*, 2006).

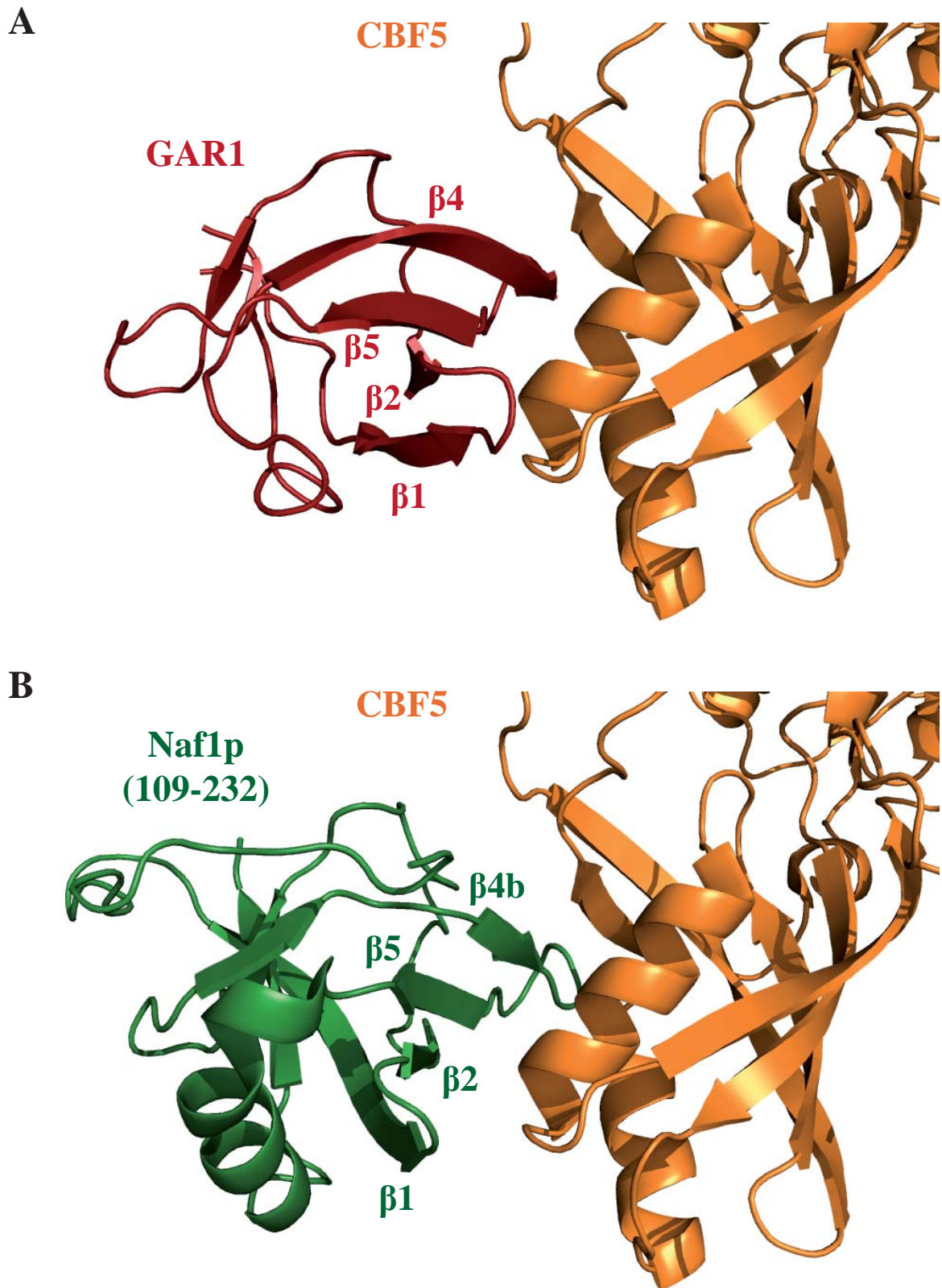
### 3.3.1. Rôle de la protéine Naf1p/NAF1

Chez la levure, des expériences d'immunoprécipitation de la chromatine (ChIP, *Chromatin ImmunoPrecipitation*) ont démontré l'association co-transcriptionnelle des protéines Naf1p et Cbf5p avec les snoRNA cistroniques et introniques (Ballarino *et al.*, 2005; Yang *et al.*, 2005). Ce recrutement, qui est dépendant de la transcription et qui requiert l'intégrité de la séquence du snoRNA, s'effectue après l'initiation de la transcription, puisque les deux facteurs ne sont pas détectés au niveau des régions promotrices (Yang *et al.*, 2005). De plus, des tests d'interaction par double-hybride et des expériences de co-immunosélection ont montré l'association de la protéine Naf1p avec le domaine CTD de l'ARN polymérase II sous sa forme phosphorylée (Fatica *et al.*, 2002). En revanche, l'implication de la kinase Ctk1p, responsable de la phosphorylation de la Serine 2 du domaine CTD, demeure incertaine. En effet, alors que l'équipe de I. Bozzoni suggère son implication dans le mode de recrutement de Naf1p et Cbf5p, les travaux des groupes de G. Chanfreau et Y. Henry démontrent que sa délétion n'influe pas sur l'assemblage co-transcriptionnel des particules H/ACA (Ballarino *et al.*, 2005; Yang *et al.*, 2005). Le couplage physique entre les machineries de transcription et d'assemblage des RNP a été néanmoins confirmé par co-immunosélection. *In vivo*, Naf1p est en effet associée avec les facteurs de transcription Spt16p, Sub1p et Tfg1p, qui sont également détectés sur les sites de transcription des snoRNA à boîtes H/ACA (Yang *et al.*, 2005), tandis que Cbf5p a été co-purifiée avec la protéine Rtf1p, qui appartient au complexe PAF1 et qui interagit également avec Spt16p (Krogan *et al.*, 2004).

Chez l'homme, la nature co-transcriptionnelle de l'assemblage des snoRNP à boîtes H/ACA a été confirmée par des approches de ChIP et de microscopie (Darzacq *et al.*, 2006). En effet, l'équipe de T. Meier a montré que NAF1, ainsi que les protéines CBF5, NOP10 et NHP2 sont détectées au niveau des sites de transcription des snoRNA (Darzacq *et al.*, 2006). L'incorporation de la protéine GAR1 au sein de la particule apparaît en revanche comme un

événement plus tardif, qui pourrait se dérouler au niveau des CB. Des expériences réalisées *in vitro* ont également permis de montrer que la protéine NAF1 peut interagir avec CBF5 seule ou intégrée au sein d'un complexe trimérique formé par les protéines CBF5, NOP10 et NHP2 (Darzacq *et al.*, 2006). D'autre part, les interactions établies par CBF5 avec NAF1 et GAR1 se sont révélées exclusives, puisqu'en présence de NAF1, GAR1 n'est plus incorporée au sein du complexe protéique. L'ensemble de ces données suggère donc qu'un pré-complexe formé par les protéines NAF1, CBF5, NOP10 et NHP2 pourrait être recruté sur l'ARN en cours de transcription, puis dans un second temps, le facteur NAF1, serait échangé avec la protéine GAR1 afin de constituer la RNP définitive (**Fig. 47**). De façon intéressante, des données structurales obtenues à partir de la protéine Naf1p de levure sont venues conforter ce modèle (**Fig. 48**). Une analyse de séquence avait initialement révélée la présence de plusieurs motifs caractéristiques au sein de la protéine Naf1p (Fatica *et al.*, 2002). Cette dernière contient notamment un domaine présentant une forte similitude avec la protéine Gar1p (32% d'identité et 49% de similarité), un motif NLS, ainsi que deux régions d'homologie (277-365 et 391-448) avec le facteur de terminaison de la transcription Nrd1p. De façon intéressante, la résolution de la structure tridimensionnelle du fragment (109-232) de la protéine Naf1p a permis de mettre en évidence une forte homologie structurale avec la protéine GAR1 (Leulliot *et al.*, 2007), dont la structure a été résolue chez les archées (cf. Introduction, Chapitre 2, 3.2.2.2) (**Fig. 44 et 45**) et plus récemment chez *S. cerevisiae* (Li *et al.*, 2011b; Li *et al.*, 2011c). Ce domaine, qui est requis pour la fonctionnalité de la protéine *in vivo*, est en effet constitué d'un tonneau  $\beta$ , formé par 6 brins  $\beta$  antiparallèles, et contient une hélice  $\alpha$  supplémentaire localisée entre les brins  $\beta 5$  et  $\beta 6$ . La surface d'interaction avec CBF5, définie au sein de la protéine GAR1 par les brins  $\beta 1$ ,  $\beta 2$ ,  $\beta 4$  et  $\beta 5$  (Rashid *et al.*, 2006) (**Fig. 48A**), apparaît relativement bien conservée dans la structure de Naf1p (**Fig. 48B**). De plus, cette surface, composée des brins  $\beta 1$ ,  $\beta 2$ ,  $\beta 4b$  et  $\beta 5$  dans la protéine Naf1p, est également impliquée dans la dimérisation du facteur d'assemblage, suggérant que son interaction avec la pseudouridine-synthase pourrait être régulée négativement par la formation d'un homodimère (Leulliot *et al.*, 2007).

Comme nous l'avons déjà abordé, la protéine GAR1 semble impliquée dans la régulation de l'activité de la particule en interagissant avec la boucle  $\beta 7$ - $\beta 10$  de CBF5, qui module l'accessibilité de la poche de pseudouridylation (Duan *et al.*, 2009; Liang *et al.*, 2007; Liang *et al.*, 2009) (cf. Introduction, Chapitre 2, 3.2.3) (**Fig. 45**). Etant donné les similarités structurales observées entre Naf1p et Gar1p, il a été proposé que la prise en charge de



**Figure 48 : Comparaison structurale du domaine (109-232) de la protéine Naf1p de *Saccharomyces cerevisiae* et de la protéine GAR1 de *Pyrococcus furiosus*.** (A) La structure de la protéine GAR1 en interaction avec la protéine CBF5 provient de la structure du complexe GAR1-CBF5-NOP10 (Rashid *et al.*, 2006; PDB : 2EY4). (B) La structure tridimensionnelle du domaine (109-232) de Naf1p a été obtenue par radiocristallographie (Leulliot *et al.*, 2007; PDB : 2V3M). Le remplacement de la protéine GAR1 par le domaine homologue du facteur Naf1p a été réalisé *in silico* à l'aide du logiciel PyMol.

Cbf5p/CBF5 par Naf1p/NAF1 contribue à inhiber son activité enzymatique, jusqu'à l'issue des étapes de transport et d'assemblage des snoRNP (Leulliot *et al.*, 2007). Récemment, l'équipe de K. Ye a en effet mis en évidence, chez la levure, que le remplacement de Gar1p par Naf1p au sein d'une snoRNP à boîtes H/ACA reconstituée *in vitro*, est associée à une très faible activité (Li *et al.*, 2011c). De plus, l'échange entre Naf1p et Gar1p semble résulter d'une meilleure affinité de Gar1p pour Cbf5p (Li *et al.*, 2011c). Enfin, le domaine CTE de Gar1p (cf. Introduction, Chapitre 2, 3.2.3), absent dans la protéine Naf1p, pourrait être à l'origine de ces deux différences majeures, puisque sa délétion diminue fortement son affinité pour la protéine Cbf5p et altère l'activité de la particule (Li *et al.*, 2011c).

Ces résultats illustrent donc la nature du mécanisme d'échange entre Naf1p et Gar1p. De façon intéressante, ce processus repose sur un mimétisme structural entre les deux protéines et pourrait permettre de passer d'un état non fonctionnel à une configuration active de la particule. Chez l'homme, cet événement semble se dérouler au niveau des CB et pourrait faire intervenir le complexe SMN qui est enrichi dans ces corpuscules nucléaires (Darzacq *et al.*, 2006). En effet, par l'intermédiaire de leur domaine commun, également présent chez la Fibrillarine, les protéines NAF1 et GAR1 présentent une certaine homologie structurale avec les protéines Sm (Kambach *et al.*, 1999). De plus, comme nous l'avons déjà mentionné, la protéine SMN interagit également avec la protéine GAR1 qui pourrait constituer un substrat pour le complexe chaperon (Pellizzoni *et al.*, 2001). Enfin, comme c'est le cas pour les protéines Sm, la protéine GAR1 contient des résidus arginines méthylés, qui pourraient améliorer significativement son affinité pour le complexe SMN (Wang & Meier, 2004).

### 3.3.2. Rôle de la protéine Shq1p/SHQ1

Contrairement à la protéine Naf1p/NAF1, Shq1p/SHQ1 n'est pas retrouvée au niveau des sites de transcription des snoRNA à boîtes H/ACA (Grozdanov *et al.*, 2009b; Yang *et al.*, 2005). De plus, bien qu'elle réalise la navette entre le cytoplasme et le noyau, elle ne possède pas de signal de localisation nucléaire et sa translocation vers le noyau semble dépendre de son association avec Cbf5p/CBF5, qui possède son propre NLS (Grozdanov *et al.*, 2009b). L'analyse de séquence de la protéine Shq1p/SHQ1 a révélé la présence d'un domaine CS (*CHORD-containing proteins and Sgt1*) au niveau de sa région N-terminale (Singh *et al.*,

2009). Cet élément, également appelé *HSP20-like domain*, est retrouvé dans de nombreuses co-chaperonnes et constitue un domaine d'interaction avec les chaperonnes Hp82p/HSP90. La structure tridimensionnelle du domaine CS de la protéine Shq1p a été résolue chez la levure (Godin *et al.*, 2009; Singh *et al.*, 2009). Bien que requis pour la fonction de la protéine, ce dernier n'interagit ni avec HSP90, ni avec la protéine Cbf5p (Godin *et al.*, 2009; Singh *et al.*, 2009). Plus récemment, deux équipes ont résolu, par radiocristallographie, la structure tridimensionnelle du domaine C-terminal de la protéine Shq1p de levure, libre ou en complexe avec Cbf5p (Li *et al.*, 2011b; Walbott *et al.*, 2011). Celui-ci, baptisé SSD (*Shq1p Specific Domain*), adopte une structure globulaire, formée de 17 hélices  $\alpha$  et 2 brins  $\beta$ . Son interaction avec la protéine Cbf5p implique le domaine PUA de la pseudouridyl-synthase, ainsi qu'une extension C-terminale, retrouvée spécifiquement chez les eucaryotes. De façon intéressante cette extension flexible, qui présente une forte tendance à la dégradation (Li *et al.*, 2011c), adopte une structure stable au sein du complexe avec Shq1p (Li *et al.*, 2011b; Walbott *et al.*, 2011). Ce résultat illustre de façon remarquable la dépendance de certaines protéines eucaryotes vis-à-vis des facteurs d'assemblage, puisque la faible solubilité de Cbf5p pourrait être compensée par sa prise en charge par la protéine Shq1p. De plus, cette interaction s'est également avérée incompatible avec la fixation de Cbf5p sur le snoRNA (Walbott *et al.*, 2011). L'interface d'interaction entre Cbf5p et Shq1p mime en effet l'interaction entre l'enzyme et le snoRNA au sein de la particule mature. Les contacts électrostatiques entre Cbf5p et le squelette phosphate de l'ARN sont notamment remplacés par des interactions de même nature avec des résidus aspartates ou glutamates du facteur d'assemblage (Li *et al.*, 2011b; Walbott *et al.*, 2011).

L'ensemble de ces résultats suggère donc que Shq1p/SHQ1 pourrait être impliquée dans la prise en charge précoce de la protéine Cbf5p/CBF5, afin de prévenir des interactions aspécifiques avec les ARN, jusqu'à son incorporation au sein des pré-snoRNP (Godin *et al.*, 2009; Grozdanov *et al.*, 2009b; Li *et al.*, 2011b; Walbott *et al.*, 2011) (**Fig. 47**). Toutefois, la nature des complexes protéiques précédant la formation de ces pré-snoRNP demeure incertaine. Initialement, une étude réalisée à partir de la protéine CBF5 de rat (alias NAP57) traduite *in vitro* indiquait que son interaction avec SHQ1 était inhibée en présence des protéines NAF1, NOP10 et NHP2 (Grozdanov *et al.*, 2009b). En revanche, les résultats obtenus récemment à partir des protéines de levure recombinantes indiquent que l'interaction entre Cbf5p et Shq1p peut s'établir en présence des autres protéines de la particule et du facteur Naf1p (Li *et al.*, 2011b).



Enfin, plusieurs études ont révélé que l'interaction entre Cbf5p et Shq1p pourrait jouer un rôle critique dans la survenue de certaines formes de Dyskératose Congénitale (DC) (Grozdanov *et al.*, 2009a; Li *et al.*, 2011b; Walbott *et al.*, 2011) (cf. Introduction, Chapitre 2, 3.2.5). Les mutations touchant la protéine CBF5 ne sont pas localisées au niveau du domaine catalytique central, mais se concentrent au niveau des régions N- et C-terminales de la protéine. En se basant sur les structures de sRNP d'archées et en utilisant une approche de modélisation moléculaire, l'équipe de T. Meier a mis en évidence que le cluster formé par ces mutations ne touche vraisemblablement pas les régions impliquées dans des contacts entre les constituants de la RNP. En revanche, les récentes données structurales, couplées à des tests d'interactions *in vitro*, indiquent que certaines de ces mutations affectent la capacité de la protéine à interagir avec le facteur Shq1p/SHQ1 (Grozdanov *et al.*, 2009a; Li *et al.*, 2011b; Walbott *et al.*, 2011). A titre d'exemple, la mutation A353V, qui est retrouvée dans 40% des cas de DC liée au chromosome X, renforce l'interaction avec le facteur d'assemblage, tandis d'autres mutations présentent un effet inverse en altérant l'interaction (Grozdanov *et al.*, 2009a). Dans le contexte de la biogenèse des particules, ces deux cas de figure pourraient induire des effets similaires, puisque des expériences réalisées en extrait cellulaire ont montré que la déstabilisation de la protéine CBF5, tout comme sa rétention par SHQ1, conduisent à un défaut d'assemblage de particules fonctionnelles (Grozdanov *et al.*, 2009a; Grozdanov *et al.*, 2009b).

#### 4. Biogenèse des snoRNP à boîtes C/D

Comme nous l'avons abordé précédemment, les mécanismes de transcription et de maturation des snoRNA à boîtes C/D ont fait l'objet de nombreuses investigations (cf. Introduction, Chapitre 1). De même, comme nous allons le voir, le transport et la localisation des snoRNP à boîtes C/D ont largement été étudiés et constituent à l'heure actuelle un modèle de choix pour l'étude des processus de trafic intranucléaire. Plusieurs travaux, menés chez la levure et chez l'homme, ont notamment abouti à la proposition d'un modèle détaillé pour la biogenèse de la snoRNP U3. Cependant, à la différence des UsnRNP et des snoRNP à boîtes H/ACA, les événements qui régissent *en trans* l'assemblage de ces particules demeurent encore relativement peu connus. Depuis le début des années 2000, plusieurs équipes ont néanmoins

identifié de nouveaux facteurs requis pour l'accumulation des snoRNP à boîtes C/D dans les cellules, suggérant l'existence d'une infrastructure potentiellement dédiée à leur assemblage.

#### 4.1. Transcription et maturation des snoRNA à boîtes C/D

Comme c'est le cas pour les snoRNA à boîtes H/ACA, il existe, pour les gènes de snoRNA à boîtes C/D, plusieurs configurations génomiques différentes, donnant lieu à plusieurs voies de maturation possibles. Chez la levure *S. cerevisiae*, la majorité des snoRNA sont monocistroniques (ex : U3) ou polycistroniques (ex : U14-snR190) (Piekna-Przybylska *et al.*, 2007). Chez les vertébrés, hormis quelques cas de snoRNA monocistroniques (ex : U3, U8, U13), la plupart d'entre eux sont en revanche hébergés dans des introns (ex : U14) (Lestrade & Weber, 2006). Que ce soit chez l'homme ou chez la levure, la libération des snoRNA à boîtes C/D introniques a fait l'objet d'études détaillées (cf. Introduction, Chapitre 1, 2.1.2) (**Fig. 11**). Deux principaux mécanismes ont ainsi été identifiés : une voie majoritaire impliquant le débranchement de la structure en lasso (Cavaille *et al.*, 1996; Kiss & Filipowicz, 1993; Ooi *et al.*, 1998; Petfalski *et al.*, 1998) et une voie minoritaire impliquant un clivage endonucléolytique indépendant de l'épissage (Cavaille *et al.*, 1996; Ooi *et al.*, 1998; Petfalski *et al.*, 1998; Villa *et al.*, 1998). La libération des snoRNA codés au sein d'unités transcriptionnelles indépendantes nécessite quant à elle l'intervention de nombreux facteurs, au cours d'un processus complexe coordonné par le système NRD1 (cf. Introduction, Chapitre 1, 2.4.2.3) (Steinmetz *et al.*, 2001) (**Fig. 12**). Les transcrits polycistroniques sont dissociés par des clivages endonucléolytiques catalysés par l'action de l'endoribonucléase Rnt1p (Chanfreau *et al.*, 1998a; Chanfreau *et al.*, 1998b). De façon générale, la définition des extrémités 3' et 5' matures implique des étapes de digestion exonucléolytiques catalysées respectivement par l'exosome (Allmang *et al.*, 1999; van Hoof *et al.*, 2000) et par les enzymes Rat1p ou Xrn1p (Petfalski *et al.*, 1998; Qu *et al.*, 1999) (**Fig. 11 & 12**). Enfin, tandis que les snoRNA cistroniques porteurs d'une coiffe <sup>m7</sup>G subissent généralement une hyperméthylation catalysée par l'enzyme Tgs1p/TGS1 (ex : U3) (Mouaikel *et al.*, 2002), certains snoRNA identifiés chez la levure, sont excisés de leur extrémité 5' par l'action de l'endoribonucléase Rnt1p (ex : snR190) (Chanfreau *et al.*, 1998b; Lee *et al.*, 2003).

## 4.2. Transport et localisation des snoRNP à boîtes C/D

De nombreuses études réalisées chez différents organismes, ont démontré que le motif C/D constitue un élément nécessaire et suffisant pour l'adressage nucléolaire des snoRNA (Lange *et al.*, 1998; Narayanan *et al.*, 1999b; Samarsky *et al.*, 1998; Verheggen *et al.*, 2001). Au cours de leur biogenèse, les snoRNP à boîtes C/D, transitent également par des domaines spécialisés : les CB chez les animaux et les NB chez la levure *S. cerevisiae* (cf. Introduction, Chapitre 1, 1.5) (**Fig. 5**). Ces corpuscules renferment notamment l'enzyme Tgs1p/TGS1 et sont le lieu de l'hyperméthylation de la coiffe <sup>m7</sup>G portée par certains snoRNA. Ces processus de transport impliquent des événements de régulations spécifiques que nous allons détailler. De plus, une étude récente réalisée chez l'homme suggère que la snoRNP U8 présente une phase de maturation cytoplasmique au cours de sa biogenèse.

### 4.2.1. Chez la levure

Chez la levure, l'expression d'un snoRNA U14 stabilisé artificiellement par l'allongement de sa tige terminale (Huang, 2002), a permis de mettre en évidence l'importance des 4 protéines constitutives Snu13p, Nop1p, Nop56p et Nop58p pour la localisation nucléolaire (Verheggen *et al.*, 2001). En effet, la déplétion de ces protéines inhibe l'accumulation de l'ARN dans le nucléole, avec un effet plus marqué pour les protéines Snu13p et Nop1p (Verheggen *et al.*, 2001). De même, la localisation nucléolaire des protéines Nop1p et Nop58p est également dépendante de l'expression des autres protéines de la particule, suggérant que seules les snoRNP complètement assemblées sont adressées au niveau du nucléole (Verheggen *et al.*, 2001). D'autre part, l'accumulation de snoRNA au sein des NB a été observée pour le snoRNA U3 et pour l'ARN U14 placé dans un contexte intronique (Verheggen *et al.*, 2001). Comme c'est le cas pour l'adressage nucléolaire, la présence de l'ARN U14 au sein de ces corpuscules requiert la présence du motif C/D et des 4 protéines constitutives, indiquant que les particules sont assemblées au cours d'un processus antérieur. De plus, le transit des snoRNP par les NB pourrait être régulé par les facteurs Srp40p et Nsr1p, homologues respectifs des protéines humaines NOPP140 et Nucléoline (Verheggen *et al.*, 2001). En effet, la présence de la protéine Srp40p est essentielle pour la rétention du snoRNA U14 dans les NB, tandis que la protéine Nsr1p est requise pour sa dispersion dans le nucléole. Cet effet de

dispersion pourrait notamment être favorisé par l'appariement du snoRNA avec le pré-ARNr, qui s'est avéré dépendant de la Nucléoline (Ginisty *et al.*, 1998; Turner *et al.*, 2009). Enfin, récemment, la recherche systématique de mutants délétères pour la localisation nucléolaire des snoRNA a été réalisée chez *S. cerevisiae* (Qiu *et al.*, 2008). Parmi les facteurs caractérisés, les protéines Tao3p et Nop4p sont notamment requises pour l'accumulation de certains snoRNA dans les NB, tandis que les protéines Sec14p et Htl1p sont impliquées dans leur dispersion depuis les NB, vers le reste du nucléole. Le rôle précis joué par ces facteurs, de même que leur implication directe dans ces processus, restent cependant à déterminer.

#### 4.2.2. Chez les vertébrés

Chez les vertébrés, le transit des snoRNA par les CB requiert également la présence du motif C/D (Narayanan *et al.*, 1999b; Samarsky *et al.*, 1998). En revanche, parmi les quatre protéines constitutives des particules C/D, seul le recrutement de la protéine 15.5K est nécessaire pour l'accumulation du snoRNA U3 au niveau des CB (Boulon *et al.*, 2004). Les protéines NOP56 et NOP58 sembleraient quant à elles jouer un rôle plus important dans l'adressage nucléolaire, puisque leur domaine C-terminal basique est nécessaire et suffisant pour assurer leur propre localisation au sein du noyau, mais également au niveau du nucléole (Pradet-Balade *et al.*, 2011). Comme chez la levure, la protéine NOPP140 semble également impliquée dans ces événements de transport. Cette dernière, qui s'associe avec les snoRNA à boîtes C/D et H/ACA et qui transite entre les CB et le nucléole (Watkins *et al.*, 2004; Yang *et al.*, 2000), pourrait en effet réguler l'accumulation des particules dans ces deux compartiments. De plus, comme cela a été montré chez la levure, la Nucléoline pourrait également jouer un rôle clé dans la rétention nucléolaire des snoRNA, notamment en favorisant leur incorporation au sein des complexes pré-ribosomiques (Ginisty *et al.*, 1998; Turner *et al.*, 2009).

D'autre part, les facteurs PHAX et CRM1 ont également été impliqués dans les processus de trafic intranucléaire des snoRNP (Boulon *et al.*, 2004) (**Fig. 49**). Comme nous l'avons décrit précédemment, ces facteurs jouent un rôle clé dans le processus de biogenèse des UsnRNP spliceosomales, en assurant l'export des UsnRNA porteurs d'une coiffe <sup>m7</sup>G vers le cytoplasme. De façon paradoxale, le snoRNA U3, qui présente cette même caractéristique, ne fait pas l'objet d'un export cytoplasmique (Boulon *et al.*, 2004; Terns & Dahlberg, 1994;

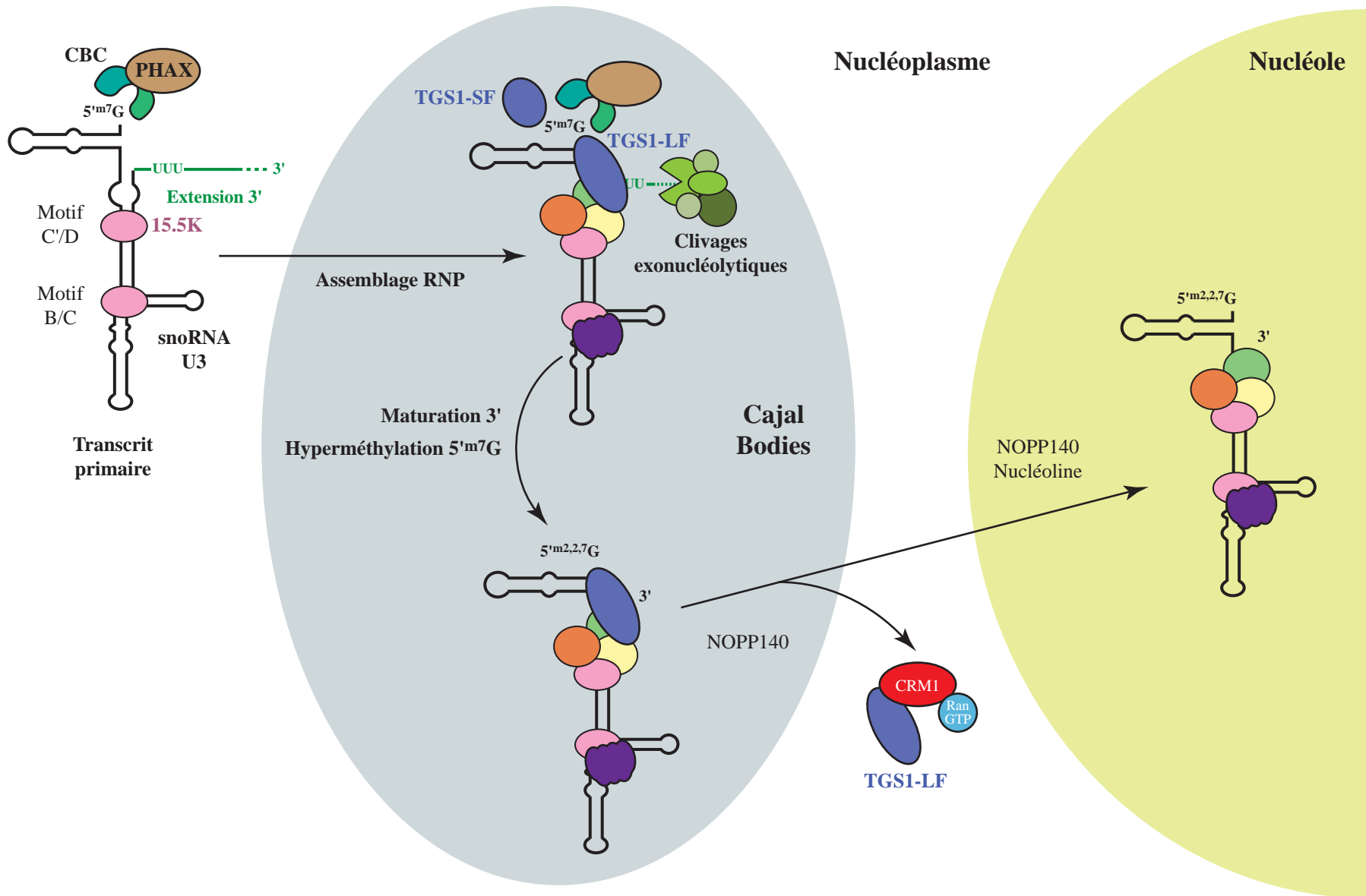


Figure 49 : Modèle proposé pour la voie de biogenèse de la snoRNP U3 chez les vertébrés (Verheggen *et al.*, 2002; Boulon *et al.*, 2004; Pradet-Ballade *et al.*, 2011).

Terns *et al.*, 1995), de plus, sa coiffe  $m^7G$  est importante pour promouvoir son adressage vers les CB (Boulon *et al.*, 2004). En utilisant le modèle du snoRNA U3, l'équipe d'E. Bertrand a mis en évidence un rôle nouveau des facteurs PHAX et CRM1 dans le transport des snoRNA vers les CB, puis vers le nucléole (Boulon *et al.*, 2004). En effet, la protéine PHAX, qui est requise pour la production du snoRNA U3, est initialement recrutée sur le snoRNA U3 porteur d'une coiffe  $m^7G$  et favorise son accumulation dans les CB. Puis, dans un second temps, le complexe CRM1/Ran-GTP est recruté sur l'ARN porteur d'une coiffe  $m^{2,2,7}G$  et induit son adressage vers le nucléole. L'association de la protéine PHAX avec d'autres ARN porteurs d'une coiffe  $m^7G$ , comme les snoRNA U8 et U13, ainsi que l'ARN du complexe télomérase, a également été mis en évidence (Boulon *et al.*, 2004; Watkins *et al.*, 2004; Watkins *et al.*, 2007). De plus, il semblerait que ce rôle puisse être étendu aux ARN non coiffés, puisque la protéine PHAX contrôle également la production du snoRNA intronique U14 (Watkins *et al.*, 2004) et que son rôle dans le transport du snoRNA U3 a été confirmé en absence de coiffe  $m^7G$  (Boulon *et al.*, 2004).

De façon intéressante, une étude menée récemment par l'équipe d'E. Bertrand, a permis de montrer que l'implication du facteur CRM1 dans ce processus est dépendante de l'hyperméthylase cytoplasmique TGS1-LF et de la séquence C-terminale basique de la protéine NOP58 (Pradet-Balade *et al.*, 2011). La protéine TGS1-LF, qui possède une localisation principalement cytoplasmique, est en effet détectée en quantité plus réduite au sein des CB (Pradet-Balade *et al.*, 2011; Verheggen *et al.*, 2002). Les auteurs ont ainsi démontré que cette isoforme de l'enzyme possède la capacité d'interagir avec le domaine C-terminal de la protéine NOP58 et pourrait de ce fait, inhiber le départ des snoRNP vers le nucléole (Pradet-Balade *et al.*, 2011). A l'inverse, en présence du complexe Ran-GTP, le facteur CRM1 reconnaît un signal d'export (NES) spécifique de la protéine TGS1-LF, dissocie son interaction avec NOP58, et favorise à la fois le retour de l'hyperméthylase vers le cytoplasme et l'adressage nucléolaire des snoRNP (Pradet-Balade *et al.*, 2011). De façon intéressante, ce système semble indépendant de la coiffe, puisque ce rôle de CRM1 a été confirmé pour de nombreux snoRNA à boîtes C/D introniques. De plus, les snoRNA à boîtes H/ACA et l'ARN du complexe télomérase semblent également concernés, suggérant qu'en plus de son rôle dans le transport nucléocytoplasmique, le système PHAX/CRM1 pourrait jouer un rôle général dans le trafic intranucléaire de divers ARN non codants. Cette double fonction pourrait ainsi coordonner les événements de biogenèse des snoRNP et de production des ribosomes. En cas de saturation des systèmes d'export, le titrage de la protéine CRM1 par

les sous-unités pré-60S et pré-40S, pourrait en effet, induire un phénomène de rétro-contrôle, conduisant au ralentissement du processus de synthèse des ribosomes, par la limitation de l'adressage nucléolaire des snoRNP (Pradet-Balade *et al.*, 2011).

### 4.2.3. L'export cytoplasmique de la pré-snoRNP U8

Malgré l'implication du système PHAX/CRM1/Ran-GTP, les snoRNA ne semblent pas faire l'objet d'un export vers le cytoplasme comme c'est le cas pour les UsnRNA spliceosomaux. Plusieurs études réalisées dans des oocytes de xénope, avaient initialement exclu cette possibilité pour le snoRNA U3 (Terns & Dahlberg, 1994; Terns *et al.*, 1995). Plus récemment, des expériences réalisées chez l'homme à partir du snoRNA U3, d'un snoRNA à boîtes C/D intronique artificiel, du snoRNA à boîtes H/ACA U64, et de l'ARN du complexe télomérase, ont également démontré l'absence de phase cytoplasmique au cours de la biogenèse de ces ARN (Pradet-Balade *et al.*, 2011). Cependant, en dépit de ces nombreuses observations, certains snoRNA pourraient échapper à cette règle, puisqu'une étude réalisée par N.J. Watkins a permis de mettre en évidence l'existence de complexes cytoplasmiques contenant des précurseurs du snoRNA U8 humain (Watkins *et al.*, 2007). En plus de sa forme mature, quatre précurseurs (I à IV) du snoRNA U8 sont détectés dans les cellules. La caractérisation des pré-snoRNP contenant ces différentes espèces a permis de mettre en évidence des complexes de taille, de composition et de localisation variable (Watkins *et al.*, 2007). Dans le noyau, les intermédiaires longs I et II sont enrichis au sein de complexes 12S, contenant notamment les protéines NOP56, NOP58, Fibrillarine et PHAX. En revanche, les précurseurs courts III et IV sont abondamment retrouvés au sein d'un complexe 17S, caractérisé par la présence additionnelle des facteurs CRM1 et Ran-GTP. Dans les extraits cytoplasmiques, deux particules ont été majoritairement détectées : un complexe 17S contenant les intermédiaires III et IV, ainsi qu'un complexe 25S enrichi en précurseur IV porteur d'une coiffe <sup>m2,2,7</sup>G. Les protéines CRM1 et Ran-GTP ne sont pas associés à ces deux complexes, la protéine PHAX est détectée uniquement au sein du complexe 17S et la Snurportine est enrichie au sein du complexe 25S. L'ensemble de ces données révèle donc une voie de biogenèse similaire à celle des UsnRNP spliceosomales. Le pré-snoRNA U8 serait exporté vers le cytoplasme par le système PHAX/CRM1/Ran-GTP, puis l'hyperméthylation de la coiffe, catalysée par l'enzyme cytoplasmique TGS1-LF, permettrait ensuite le remplacement de la protéine PHAX par la

Snurportine, afin de favoriser le ré-import de la RNP dans le noyau. La détection des protéines constitutives des snoRNP au sein des particules 12S, suggère en revanche que leur assemblage s'effectue au cours d'un processus nucléaire précoce. En comparaison avec les événements d'assemblage cytoplasmique des UsnRNP, l'intérêt fonctionnel de l'export du snoRNA U8 reste donc à déterminer.

### **4.3. Importance des modifications post-traductionnelles pour la biogenèse des snoRNP à boîtes C/D ?**

Récemment, plusieurs études ont identifié la présence de modifications post-traductionnelles au niveau de certaines protéines constitutives des snoRNP humaines (Matic *et al.*, 2010; Westman & Lamond, 2011; Westman *et al.*, 2010). C'est notamment le cas de la protéine NOP58 qui est modifiée, dans sa région C-terminale riche en motifs KKE/D, par SUMOylation des résidus K467 et K497 (Westman *et al.*, 2010), et par phosphorylation du résidu S502 (Matic *et al.*, 2010). L'ajout des groupements SUMO de NOP58 pourrait être catalysé par les enzymes SUMO2/3 et SUMO1 au niveau du nucléoplasme ou des CB, tandis que les isopeptidases SEN3/5 pourraient assurer leur déconjugaison au niveau du nucléole (Westman *et al.*, 2010). La Sérine en position 502 pourrait quant à elle être ciblée par la kinase CK2 (Matic *et al.*, 2010). De plus, ces différentes modifications pourraient être liées par des événements de co-régulation puisque la Sérine en position 502 est requise pour la SUMOylation optimale de la Lysine 497 (Matic *et al.*, 2010).

L'inhibition de la SUMOylation de la protéine NOP58 par mutagenèse dirigée n'a pas révélé d'impact majeur sur la localisation de la protéine dans les CB ou dans le nucléole, en revanche, l'absence des groupements SUMO conduit à une diminution de l'affinité de NOP58 pour les snoRNA à boîtes C/D *in vivo* (Westman *et al.*, 2010). Etant donné l'importance de la protéine NOP58 pour l'adressage nucléolaire des ARN, l'absence de SUMOylation conduit également à un défaut d'accumulation des snoRNA au niveau du nucléole et à leur rétention au niveau de *foci* nucléoplasmiques correspondant aux sites de transcription (Westman *et al.*, 2010). Les auteurs proposent que l'addition des groupements SUMO au niveau de la région C-terminale de la protéine NOP58 conduit à des changements conformationnels et/ou à la formation de nouvelles surfaces de contacts (Westman *et al.*, 2010). L'acquisition de ces



nouvelles propriétés pourrait ainsi permettre d'augmenter l'affinité de la protéine pour les snoRNA ou faciliter son interaction avec un facteur intervenant en *trans* au cours de la biogenèse des particules (Westman *et al.*, 2010). D'autre part, les protéines NOP58 nucléolaires, qui correspondent à la fraction majoritairement SUMOylée, se sont révélées plus stables que les protéines NOP58 nucléoplasmiques. Etant donné que la protéine NOP58 est également la cible d'ubiquitinations et de dégradation par le protéasome, il a été proposé que les modifications post-traductionnelles contribuent au contrôle qualité de la biogenèse des particules (Westman & Lamond, 2011). La SUMOylation pourrait en effet constituer une étiquette spécifique des snoRNP correctement assemblées, inversement l'accumulation nucléoplasmique de protéines NOP58 non SUMOylées pourrait constituer un signal de « défaut d'assemblage » et conduire à la dégradation ciblée de ces protéines. Enfin, la déconjugaison des groupements SUMO au niveau du nucléole par les enzymes SEN 3/5 pourrait également contribuer au recyclage ou au traitement des particules à l'issue de leur intervention dans la biogenèse des ribosomes.

De façon intéressante, les résidus K467 et K497 sont conservés dans la protéine Nop58p de levure. En revanche, la protéine NOP56 humaine, qui contient également le domaine C-terminal riche en lysines, est relativement peu SUMOylée, suggérant que ce mode de régulation est spécifique de NOP58 (Westman *et al.*, 2010). Enfin, la recherche exhaustive des protéines humaines SUMOylées a révélé que NOP58, mais également les protéines NHP2 et CBF5, figurent parmi les principaux facteurs SUMOylés du nucléole (Westman *et al.*, 2010). Ces nouvelles données suggèrent donc que les modifications post-traductionnelles, et en particulier les SUMOylations, pourraient constituer un système général de régulation de la biogenèse des snoRNP.

#### **4.4. La snoRNP U3, un modèle d'étude pour la biogenèse des snoRNP**

Etant donné son caractère essentiel et son implication majeure dans la production des ribosomes, la snoRNP U3 a fait l'objet de nombreuses recherches et la caractérisation des processus gouvernant sa propre biogenèse a été entreprise dans différents organismes.

#### 4.4.1. Biogenèse de la snoRNP U3 chez la levure

Chez la levure *S. cerevisiae*, le snoRNA U3 est codé au sein de deux unités transcriptionnelles distinctes, appelées *SNR17A* et *SNR17B* (Hughes *et al.*, 1987). Le gène *SNR17A*, localisé au niveau du chromosome 15, code le snoRNA U3A (333 nt), qui constitue la forme majoritairement exprimée dans la cellule (environ 90%). Le gène *SNR17B*, localisé sur le chromosome 16, est quant à lui, responsable de l'expression d'une forme minoritaire U3B, plus courte de 1 nucléotide (332 nt) et présentant 92% d'identité avec l'ARN U3A. Cette situation est retrouvée fréquemment chez les eucaryotes, puisque de nombreux organismes, dont l'homme, le xénope, le rat ou encore la souris, comptent plusieurs copies du gène du snoRNA U3 au sein de leur génome (Bernstein *et al.*, 1983; Mazan & Bachellerie, 1988; Reddy *et al.*, 1985; Savino *et al.*, 1992; Suh *et al.*, 1986; Suh *et al.*, 1991). Les gènes codant le snoRNA U3 chez les levures des genres *Saccharomyces*, *Kluyveromyces* et *Pichia*, présentent néanmoins une particularité de taille, puisqu'ils contiennent un intron épissé par la machinerie spliceosomale (Brule *et al.*, 1995; Brule *et al.*, 1996; Fournier *et al.*, 1998; Myslinski *et al.*, 1990). Dans le snoRNA U3A de *S. cerevisiae*, cet intron est situé entre les boîtes A et A' (Myslinski *et al.*, 1990). Il forme une longue structure en tige-boucle dont la formation ne semble pas influencer la structuration du domaine 3' (Mougin *et al.*, 1996). D'autre part, malgré un grand degré de divergence, les introns des pré-snoRNA U3A et U3B présentent deux caractéristiques communes : un site 5' d'épissage séquestré au sein d'une hélice et un site de branchement sous-optimal (Brule *et al.*, 1995; Mougin *et al.*, 1996; Myslinski *et al.*, 1990). Compte tenu de l'efficacité élevée du processus d'épissage du snoRNA U3, cette configuration défavorable suggère donc l'intervention de facteurs influençant positivement l'excision de l'intron *in vivo* (Mougin *et al.*, 1996).

Le processus de maturation nucléolytique du snoRNA U3 a été caractérisé chez la levure *S. cerevisiae* par l'équipe de D. Tollervy (Kufel *et al.*, 2000; Kufel *et al.*, 2003) (**Fig. 50**). Dans le modèle proposé, un long transcrit primaire pourrait être clivé de façon précoce par l'endonucléase Rnt1p. Cette dernière reconnaît en effet une structure en tige-boucle située dans l'extension 3' de l'ARN et génère une double restriction aux sites +21 et +58. Alors que les sous-produits de cette réaction sont dégradés par l'exonucléase 5'-3' Rat1p, la nouvelle extrémité du snoRNA U3 est rapidement dégradée jusqu'aux nucléotides +12 ou +18. Cette dégradation pourrait être catalysée par l'exosome et/ou par les exonucléases de la famille Rex1-3p, tandis que les deux nouveaux intermédiaires U3-I (+12) et U3-II (+18) sont

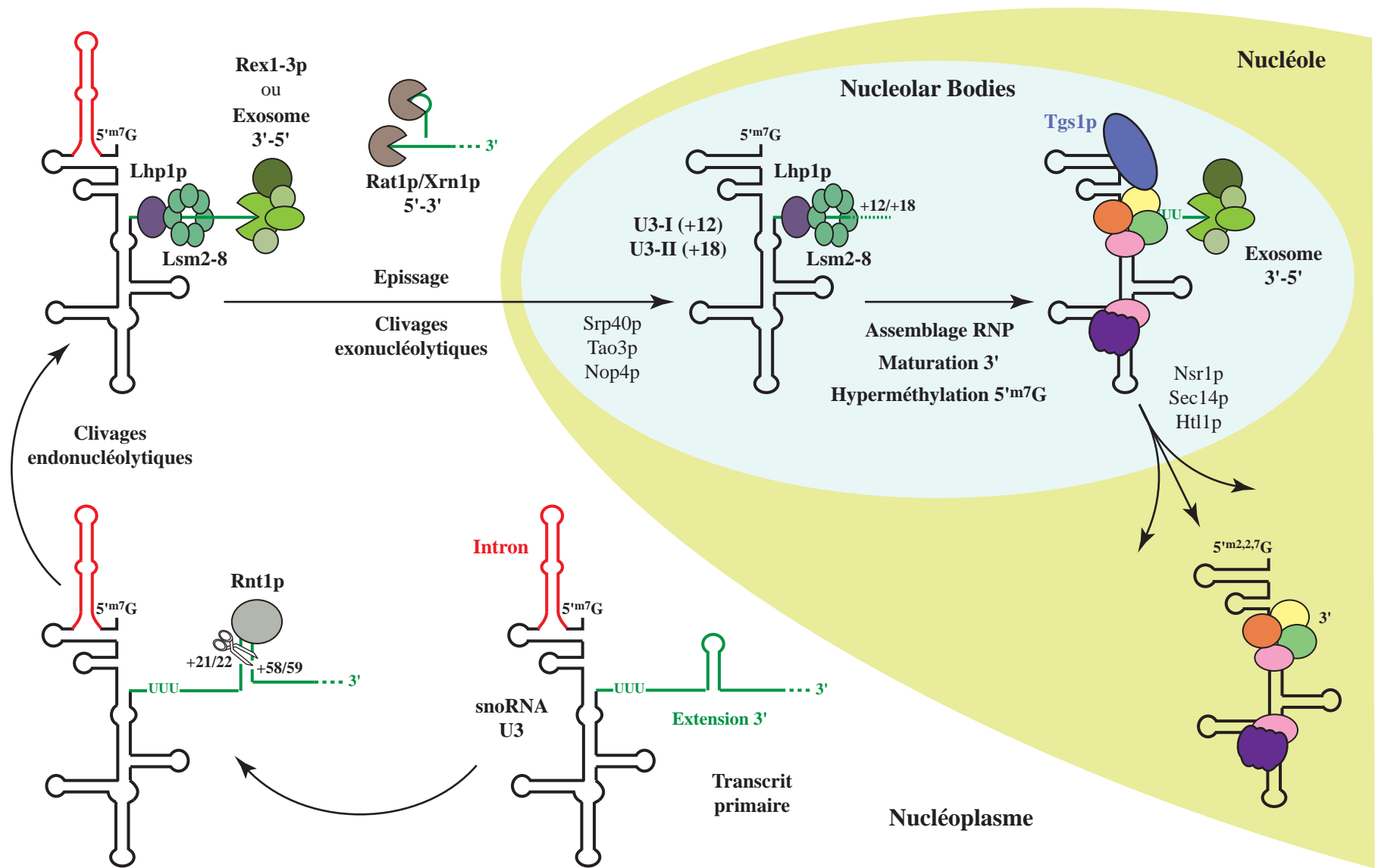


Figure 50 : Modèle proposé pour la voie de biogenèse de la snoRNP U3 chez *Saccharomyces cerevisiae* (d'après Kufel *et al.*, 2000; Verheggen *et al.*, 2002; Qiu *et al.*, 2008).

stabilisés par la présence du facteur Lhp1p et des protéines Lsm. Ces facteurs sont en effet recrutés de façon synergique, au niveau de séquences riches en U, situées en amont des sites de coupures. Leur absence conduit à la disparition des intermédiaires +12 et +18, sans affecter de façon majeure le taux de snoRNA U3 mature. Etant donné que les protéines Nop1p, Nop56p et Nop58p ne semblent pas associées à ces deux intermédiaires, il a été proposé que leur recrutement sur le motif C'/D déplace les protéines Lhp1p et Lsm, afin de permettre la formation de l'extrémité 3' mature. Cette dernière étape est catalysée par l'exosome, puisque la déplétion de sa sous-unité catalytique provoque l'accumulation des deux espèces +12 et +18. D'autre part, seule la forme mature du snoRNA U3 est porteuse d'une coiffe <sup>m2,2,7</sup>G, ce qui indique que l'étape d'hyperméthylation est postérieure à la maturation nucléolytique. Enfin, dans un contexte sauvage, la séquence intronique est excisée à partir d'un long précurseur dont l'extrémité 3' n'a pas été déterminée. Ce dernier étant associé avec les protéines Lhp1p et Lsm, il a été proposé que l'étape d'épissage intervienne après la formation des extrémités +12 et +18. Néanmoins, l'épissage pourrait se dérouler plus précocement, puisqu'il n'est pas exclu que ces facteurs soient recrutés sur les intermédiaires non maturés ou qu'ils se fixent au niveau d'autres séquences poly(U), situées plus en aval.

Par la suite, des expériences de microscopie électronique ont permis de mettre en évidence la localisation nucléaire des différents intermédiaires du snoRNA U3 (Verheggen *et al.*, 2002). Alors que les espèces non épissées sont uniquement détectées dans le nucléoplasme, les intermédiaires +12 et +18 sont détectés à la fois dans le nucléoplasme et dans les NB. Ce résultat indique que les étapes finales de biogenèse de la snoRNP pourraient se dérouler dans ces corpuscules. En plus de l'hyperméthylation de la coiffe catalysée par l'enzyme Tgs1p, l'élimination des extensions +12 et +18 de l'ARN pourrait donc s'effectuer dans les NB. Ces événements semblent par ailleurs faire l'objet d'un couplage avec l'assemblage de la particule C'/D, puisque le recrutement de l'enzyme Tgs1p pourrait impliquer son interaction avec le domaine C-terminal basique (KKE/D) de la protéine Nop58p (Mouaikel *et al.*, 2002). En revanche, l'absence d'association des protéines Nop1p, Nop56p et Nop58p avec les précurseurs +12 et +18 semble paradoxale, puisque ces dernières sont requises pour l'accumulation du snoRNA U14 dans les NB (Verheggen *et al.*, 2001). Une possibilité est que d'autres facteurs se substituent aux protéines Nop1p, Nop56p et Nop58p pour l'adressage et/ou la rétention des intermédiaires +12 et +18 du snoRNA U3 dans les NB. Enfin, une autre éventualité est que les protéines constitutives de la particule soient associées de façon labile avec les pré-snoRNA, jusqu'aux étapes finales de la biogenèse.

#### 4.4.2. Biogenèse de la snoRNP U3 chez les vertébrés

Chez les vertébrés, les processus de maturation nucléolytique du snoRNA U3 sont moins connus. Chez le rat, deux précurseurs, porteurs d'extensions de 8 et 14 nucléotides à leur extrémité 3', ont été identifiés (Stroke & Weiner, 1985), tandis que chez l'homme, un intermédiaire unique, porteur d'une extension de 8 nucléotides, est détecté majoritairement (Watkins *et al.*, 2004). L'étude de la localisation du snoRNA U3 de rat, exprimé de façon hétérologue dans des cellules humaines, a permis d'établir des comparaisons avec les données obtenues chez la levure (Verheggen *et al.*, 2002) (**Fig. 49**). Tout d'abord, l'inactivation de la boîte C, qui est délétère pour le recrutement des protéines de la particule, conduit à une perte de stabilité du snoRNA U3 mature, sans affecter le taux des deux intermédiaires +8 et +14. De plus, des expériences d'immunoprécipitation ont démontré que seule la protéine 15.5K est associée à ces précurseurs, suggérant qu'en absence des protéines NOP58 et Fibrillarine, ces derniers doivent être stabilisés par des facteurs supplémentaires, comme cela a été montré pour les intermédiaires +12 et +18 chez la levure. Des études de localisation ont également mis en évidence que les précurseurs du snoRNA U3 de rat sont détectés essentiellement au niveau des sites de transcription et des CB, tandis que l'ARN mature est retrouvé dans les CB et dans le nucléole. En plus de la Fibrillarine (Gerbi & Borovjagin, 1997; Raska *et al.*, 1990), l'équipe d'E. Bertrand a mis en évidence que l'hyperméthylase TGS1 et les protéines 15.5K, NOP56 et NOP58 s'accumulent dans les CB (Verheggen *et al.*, 2002). De façon similaire à la biogenèse de la snoRNP U3 de levure, ces données suggèrent donc que la formation de la coiffe  $m^{2,2,7}G$ , la maturation nucléolytique terminale de l'extrémité 3', ainsi que l'assemblage des protéines constitutives, pourraient se dérouler au sein de ces corpuscules. Seule la protéine 15.5K, déjà présente sur les pré-snoRNAU3 et requise pour l'adressage vers les CB, s'associerait avec l'ARN au cours d'événements plus précoces (Boulon *et al.*, 2004; Verheggen *et al.*, 2002). Par ailleurs, la co-localisation fréquente des CB avec les gènes codant les snoRNA à boîtes C/D, suggère l'existence de relations entre ces étapes de la biogenèse des particules et le processus de transcription (Gao *et al.*, 1997; Verheggen *et al.*, 2002).

Enfin, en utilisant le modèle humain, l'équipe de R. Lührmann, a entrepris la caractérisation des complexes contenant les formes immatures et mature du snoRNA U3 (Watkins *et al.*, 2004). Au cours de cette étude, la composition en ARN et en protéines de particules isolées sur gradients de glycérol a été analysée de façon ciblée par les approches de Northern-blot et

de Western-blot. Le précurseur du snoRNA U3, ainsi que sa forme mature, ont été détectés dans des extraits nucléoplasmiques (CB compris), et seule la forme mature s'est révélée porteuse d'une coiffe <sup>m2,2,7</sup>G. De plus, alors que le pré-snoRNA U3 est enrichi dans des particules d'environ 16S, la forme mature est préférentiellement associée à des complexes 12S, suggérant le départ d'un certain nombre de partenaires suite aux événements de maturation. Des expériences de co-immunoprécipitation ont en effet révélé l'association du pré-snoRNA U3 avec les protéines LSM4, RRP46, La, NOPP140, TIP48, TIP49, PHAX, CRM1, tandis que le snoRNA U3 mature est co-purifié avec les protéines TIP49, CRM1, TGS1, NOPP140. Ces résultats sont cohérents avec les données présentées précédemment. En effet, alors que la protéine PHAX, impliquée dans l'adressage vers les CB, est associée uniquement au pré-U3, le facteur CRM1, responsable de l'adressage nucléolaire, demeure associé au snoRNA U3 mature (Boulon *et al.*, 2004; Pradet-Balade *et al.*, 2011). De même, l'hyperméthylase TGS1 est enrichie dans les complexes contenant le snoRNA U3 mature, qui constitue vraisemblablement son substrat privilégié (Kufel *et al.*, 2000; Verheggen *et al.*, 2002). En revanche, l'analyse des fractions nucléolaires a révélé qu'aucun de ces facteurs ne reste associé au snoRNA U3 mature après son adressage vers son site d'action. De façon intéressante, les protéines Fibrillarine, NOP56, NOP58 et U3-55K, associées au snoRNA U3 mature, sont également co-purifiées avec le pré-snoRNA U3 au sein des complexes 16S. A la différence des travaux réalisés à partir des snoRNA U3 de levure et de rat, ce résultat indique donc que le recrutement des protéines constitutives de la particule pourrait s'effectuer de façon précoce, sur l'ARN non maturé. L'analyse de la stabilité des pré-snoRNP U3 caractérisées au cours de cette étude, a en revanche révélé une différence notable entre les complexes nucléoplasmiques et les complexes isolés du nucléole. Ces derniers peuvent en effet être purifiés en utilisant des concentrations élevées en NaCl (750 mM), tandis que les particules issues du nucléoplasme ne supportent pas des concentrations salines excédant 150 à 250 mM. Une explication à ce phénomène pourrait être que l'intégration de la snoRNP U3 nucléolaire au sein du processome 90S fournisse un environnement protecteur. Néanmoins, les auteurs proposent que la particule puisse gagner en stabilité suite à des remaniements conformationnels majeurs se produisant de façon tardive au sein des complexes nucléoplasmiques. Les hélicases hexamériques TIP48 et TIP49 pourraient notamment participer à ces événements, puisque des protéines appartenant à la même famille sont impliquées dans le remodelage ATP-dépendant de complexes multiprotéiques (Hanson & Whiteheart, 2005) (cf. Introduction, Chapitre 3, 4.6.1).

#### 4.5. Vers un modèle d'assemblage pour les snoRNP à boîtes C/D ?

Comme nous l'avons abordé précédemment, le système d'assemblage co-transcriptionnel des snoRNP à boîtes H/ACA est relativement bien conservé chez les eucaryotes et ne présente pas de variations majeures en fonction de la configuration génétique du snoRNA. En revanche, concernant les snoRNP à boîtes C/D, la situation apparaît différente, puisque le recrutement des protéines constitutives semble varier en fonction du mode de production et de maturation de l'ARN.

Parmi les snoRNA codés par des séquences introniques, deux voies de maturation ont été décrites (cf. Introduction, Chapitre 1, 2.3.2) : (i) une voie majeure dépendante de l'épissage, qui implique un positionnement strict du snoRNA au sein de l'intron, (ii) une voie mineure indépendante de l'épissage, qui nécessite la formation d'une hélice stable flanquant le snoRNA à ses extrémités. Dans ces deux cas de figures, l'intégrité du motif C/D est requis pour la maturation nucléolytique de l'ARN et l'assemblage de la snoRNP semble précéder sa libération du pré-ARNm (Giorgi *et al.*, 2001; Hirose *et al.*, 2003; Hirose & Steitz, 2001). Dans la voie majeure, l'assemblage de la particule interviendrait lors de la formation du spliceosome C1, puisque le recrutement des protéines humaines 15.5K et Fibrillarine a été observé à ce stade du processus d'épissage (Hirose *et al.*, 2003) (**Fig. 11A**). Comme la libération du snoRNA, cet assemblage est dépendant de la réaction d'épissage et du positionnement du snoRNA au sein de l'intron (Hirose *et al.*, 2003). Chez l'homme, la protéine IBP160, identifiée comme facteur de couplage, pourrait en effet coordonner les événements d'épissage, d'assemblage et de maturation du snoRNA (Hirose *et al.*, 2006). Dans la voie minoritaire, qui concerne seulement 5 snoRNA chez l'homme, l'assemblage de la particule s'effectuerait avant l'épissage (**Fig. 11B**) (Hirose *et al.*, 2003; Hirose & Steitz, 2001). Le recrutement des protéines humaines 15.5K et Fibrillarine a en effet été constaté au niveau du précurseur non mûri de ces snoRNA et ce, même lorsque la réaction d'épissage est inhibée (Hirose *et al.*, 2003). Des expériences réalisées chez la levure semblent confirmer ces observations, puisque le recrutement des protéines Nop1p et Nop58p a également été constaté au niveau des pré-ARNm contenant le snoRNA U18 (Giorgi *et al.*, 2001) (cf. Introduction, Chapitre 1, 2.3.2) (**Fig. 11B**). De plus, par le biais d'une interaction physique avec l'endonucléase Rnt1p, la protéine Nop1p pourrait faciliter la maturation nucléolytique du snoRNA (Giorgi *et al.*, 2001). En revanche, la protéine Nop56p n'a pas été détectée à ce stade

du processus de biogenèse, suggérant une association plus labile ou un recrutement plus tardif (Giorgi *et al.*, 2001).

Dans ces différents travaux, l'assemblage de toutes les protéines constitutives des snoRNP n'a pas été testé de façon systématique, néanmoins, l'association des protéines Nop1p/Fibrillarine et Nop58p avec les pré-ARNm suggère le recrutement préalable de la protéine Snu13p/15.5K. Comme nous l'avons abordé précédemment, sa fixation sur les motifs en *K-turn*, constitue un événement nucléateur, essentiel au recrutement des autres protéines. Cependant, à la différence des snoRNA à boîtes H/ACA et à l'exception notoire du snoRNA U3, les snoRNA à boîtes C/D sont des ARN relativement peu structurés et la formation d'un motif C/D structuré en *K-turn* est soumis à d'importantes contraintes de flexibilité. Pour parer à cet inconvénient, il a notamment été proposé que la présence d'une tige flanquant le snoRNA, ou que l'intervention d'un facteur de couplage, constituent des adaptations visant à faciliter le rapprochement des boîtes C et D, ainsi que le recrutement subséquent des protéines constitutives de la snoRNP (Hirose *et al.*, 2006).

D'autre part, la mise en évidence d'événements d'assemblage antérieurs ou concomitants à la réaction d'épissage, suggère fortement que les snoRNP sont formées de façon co-transcriptionnelle. Des expériences d'immunoprécipitation de la chromatine (ChIP) réalisées chez la levure ont notamment permis de mettre en évidence le recrutement co-transcriptionnel des protéines Nop1p, Nop56p et Nop58p sur le snoRNA U18, qui emprunte les voies de maturation majeure et mineure, ainsi que sur le snoRNA U24 maturé par la voie majeure (Vincenti *et al.*, 2007). Ces associations, uniquement détectables dans la portion du gène codant le snoRNA, sont en effet dépendantes de la transcription et impliquent la présence d'un motif C/D intègre. Chez les vertébrés, d'autres études corroborent l'idée d'un assemblage précoce, se déroulant au niveau des sites de transcription. Des données de microscopie indiquent en effet que la protéine 15.5K néosynthétisée, s'accumule initialement au niveau des speckles et progresse de façon unidirectionnelle et dépendante de la transcription, en direction des CB, puis du nucléole (Leung & Lamond, 2002). De plus, les CB, qui pourraient centraliser une partie des événements d'assemblage, présentent une dynamique importante qui se matérialise notamment par une localisation préférentielle au niveau des gènes hébergeant des snoRNA introniques (Schul *et al.*, 1999a).



Concernant les snoRNA codés au sein d'unités transcriptionnelles indépendantes, une étude également basée sur l'approche de ChIP, a permis de détecter la protéine Nop1p au niveau du gène codant le snoRNA monocistronique snR13 de levure (Morlando *et al.*, 2004). Cette association précoce pourrait notamment faire l'objet d'un couplage avec les événements de terminaison de la transcription, puisque Nop1p facilite le recrutement de la protéine Ref2p, impliquée dans le découplage du processus de clivage/polyadénylation (Morlando *et al.*, 2004) (cf. Introduction, Chapitre 1, 2.4.2.3) (**Fig. 12**). Bien que les autres protéines de la particule n'aient pas été intégrées à cette étude, ce résultat suggère néanmoins que l'association de la protéine Nop1p au niveau des sites de transcription du snoRNA snR13 repose sur l'assemblage tout ou partiel de la particule C/D. Cependant, en dépit de cette donnée, relativement peu de données ont été collectées au sujet de l'assemblage des snoRNA monocistroniques et la majorité des informations disponibles concernent le snoRNA U3.

Chez la levure, comme chez l'homme, il a notamment été proposé que l'assemblage de la snoRNP U3 constitue un événement tardif, se déroulant de façon postérieur à la maturation nucléolytique de l'ARN (Kufel *et al.*, 2000; Verheggen *et al.*, 2002) (**Fig. 50**). Ce mode d'assemblage, relativement différent de ceux proposés pour les snoRNA introniques, pourrait être spécifique de la snoRNP U3, dont la formation semble reposer sur des contraintes particulières, liées notamment à la présence d'un motif C'/D sous-optimal et à une dépendance vis-à-vis de la protéine Rrp9p/U3-55K (Knox *et al.*, 2011) (N. Rolland, V. Senty, Communication personnelle). Néanmoins, l'analyse des pré-snoRNP U3 purifiées à partir d'extraits nucléaires humains, a révélé des perspectives nouvelles en suggérant l'existence d'un processus de remodelage, assurant la transition de pré-snoRNP peu stables vers des particules matures correctement assemblées (Watkins *et al.*, 2004). En effet, malgré des interactions labiles, les protéines Fibrillarine, NOP56 et NOP58 ont été détectées en association avec le pré-snoRNA U3 dépourvu de coiffé triméthylée et non mûré en 3'. Cette association s'établit au sein de larges complexes nucléoplasmiques, qui renferment également de nombreux facteurs intervenant en *trans* dans la biogenèse de la snoRNP. Parmi ces facteurs, l'identification des protéines La et LSM suggère que leur présence n'est pas strictement exclusive avec celle des protéines constitutives de la snoRNP et que ces différentes protéines pourraient être associées de façon précoce avec les mêmes ARN. Ces différences avec les résultats obtenus lors des précédentes études, semblent s'expliquer par l'utilisation d'approches de purification moins stringentes, privilégiant des concentrations salines faibles et l'absence de détergents (Kufel *et al.*, 2000; Verheggen *et al.*, 2002; Watkins

*et al.*, 2004). La mise en perspective de l'ensemble de ces observations révèle néanmoins que le processus d'assemblage des snoRNP pourrait se décomposer en plusieurs étapes successives, conduisant de façon séquentielle à la formation des particules matures. La transition entre ces différents états semble se dérouler au sein de complexes macromoléculaires, contenant des facteurs capables de diriger et de coordonner les événements d'assemblage, de maturation, de modification et de transport. Parmi les facteurs identifiés à ce jour, les AAA<sup>+</sup>-ATPases TIP48 et TIP49 constituent des candidats d'intérêt, néanmoins la caractérisation plus systématique des protéines contenues au sein de ces complexes sera nécessaire pour améliorer la compréhension des mécanismes de biogenèse des snoRNP à boîtes C/D.

#### **4.6. Mise en évidence de nouveaux facteurs impliqués dans l'assemblage des snoRNP à boîtes C/D**

Différents facteurs requis pour la biogenèse des snoRNP à boîtes C/D ont été identifiés au cours de ces dernières années. Parmi ces facteurs, certains ne semblent pas impliqués directement dans les processus de maturation des snoRNA et sont des candidats potentiels pour participer aux événements d'assemblage des particules. C'est notamment le cas des hélicases Rvb1p/TIP49 et Rvb2p/TIP48, des protéines Pih1p, Bcd1p et Cwc24p récemment caractérisées chez la levure, ainsi que des facteurs Rsa1p et NUFIP, respectivement identifiés par notre équipe et par celle d'E. Bertrand.

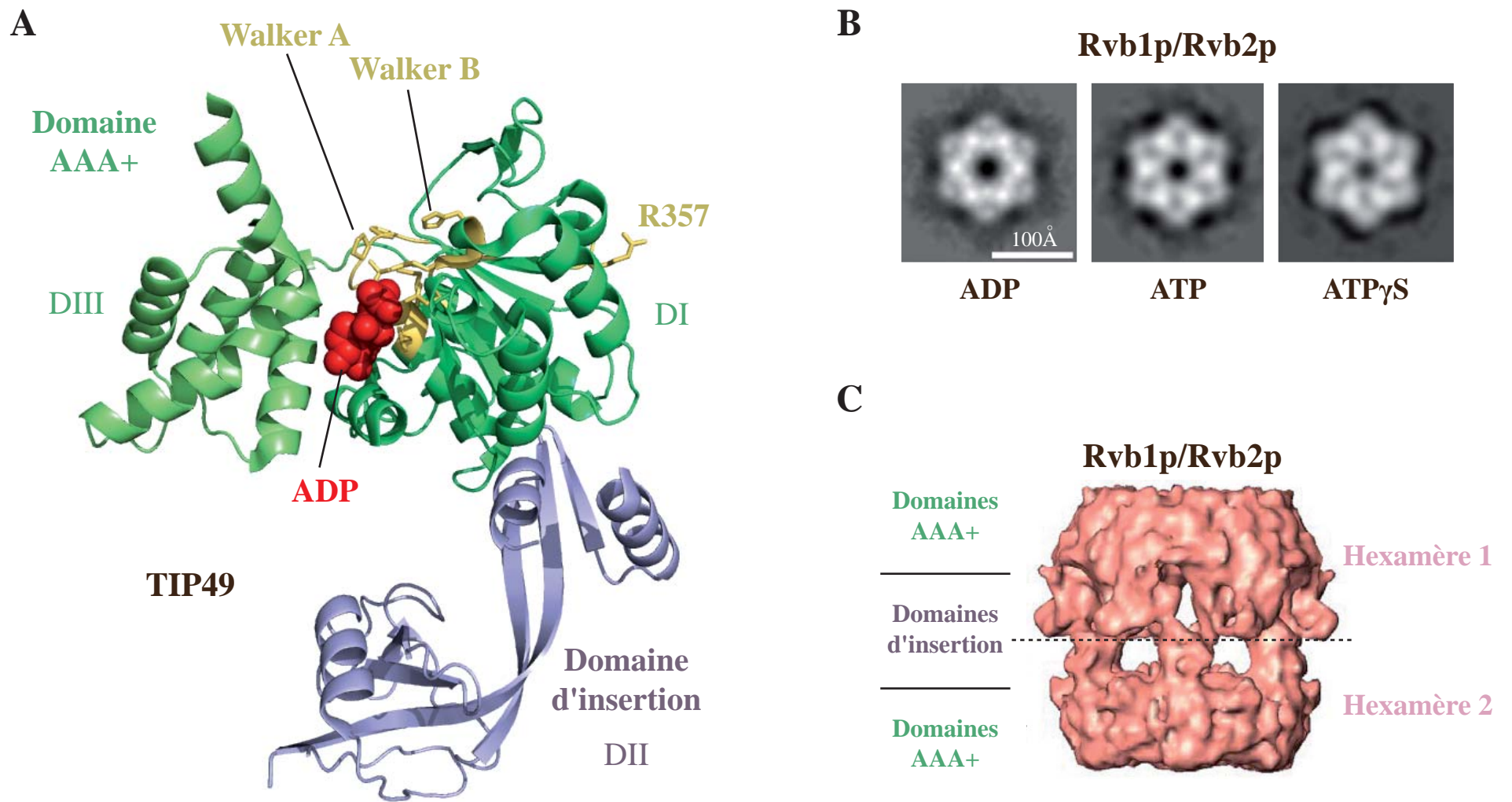
##### **4.6.1. Les AAA<sup>+</sup>-ATPase Rvb1p/TIP49 et Rvb2p/TIP48**

L'association des protéines Rvb1p/TIP49 et Rvb2p/TIP48 avec un snoRNA à boîtes C/D a initialement été mise en évidence chez la souris (Newman *et al.*, 2000; Watkins *et al.*, 1998), avant d'être confirmée chez l'homme (Watkins *et al.*, 2002; Watkins *et al.*, 2004). En raison de leur caractérisation dans plusieurs organismes eucaryotes et de leur implication dans de nombreux processus cellulaires, ces deux protéines possèdent une terminologie complexe. Rvb1p/TIP49 a été initialement identifiée au sein d'un complexe contenant la protéine TBP (*TATA box Binding Protein*) et baptisée TIP49 pour *TBP-Interacting Protein 49* (Kanemaki *et*

*al.*, 1997). Appelée Rvb1p chez la levure, on la retrouve également sous les termes TIP49a, RuvBL1, Pontin52, TAP54a, Tih1p, p50, NMP238 ou ECP54 (Bauer *et al.*, 1998; Fuchs *et al.*, 2001; Holzmann *et al.*, 1998; Lim *et al.*, 2000; Makino *et al.*, 1999; Qiu *et al.*, 1998; Salzer *et al.*, 1999; Shen *et al.*, 2000). Rvb2p/TIP48 est un paralogue de Rvb1p/TIP49 que l'on retrouve également sous les termes de Tip48, Tip49b, RuvBL2, Reptin52, TAP54b, Tih2p, p47, and ECP51 (Bauer *et al.*, 1998; Fuchs *et al.*, 2001; Gohshi *et al.*, 1999; Kanemaki *et al.*, 1999; Lim *et al.*, 2000; Qiu *et al.*, 1998; Salzer *et al.*, 1999; Wood *et al.*, 2000). Comme c'est le cas pour NOP56 et NOP58, ces deux protéines partagent une forte homologie et une bonne conservation au sein du *phylum* eucaryote, et sont présentes sous la forme d'un homologue unique chez les archées (Kanemaki *et al.*, 1999; Makino *et al.*, 1999; Newman *et al.*, 2000).

#### 4.6.1.1. Données structurales

La structure tridimensionnelle de la protéine humaine TIP49 a été résolue par radiocristallographie (Matias *et al.*, 2006) et plusieurs études par microscopie électronique ont été réalisées à partir des protéines Rvb1p/TIP49 et Rvb2p/TIP48 chez l'homme (Puri *et al.*, 2007) et chez la levure (Gribun *et al.*, 2008; Torreira *et al.*, 2008) (**Fig. 51**). Ces deux protéines contiennent trois domaines distincts (DI, DII et DIII). Les domaines I et III forment un ensemble caractéristique des protéines ATPase AAA+ (*ATPases associated with diverse cellular processes*) (Matias *et al.*, 2006) (**Fig. 51A**). Le domaine I, qui assure la fixation du nucléotide, adopte une structure de type  $\alpha\beta\alpha$ , formée par un feuillet de 5 brins  $\beta$ , pris en sandwich par des hélices  $\alpha$ . Le domaine III est en revanche formé exclusivement d'hélices  $\alpha$ . A l'intérieur du domaine I, on distingue également deux éléments structuraux caractéristiques des protéines NTPases : les motifs Walker A et B. Le motif Walker A (GPPGTGKT dans TIP49) est important pour la fixation et l'orientation du nucléotide, tandis que le motif Walker B (DEVH dans TIP49) est nécessaire pour la réaction d'hydrolyse (Cheung *et al.*, 2010). Les protéines Rvb1p/TIP49 et Rvb2p/TIP48 diffèrent en revanche des autres protéines AAA+ par l'insertion, entre les motifs Walker A et B, du domaine II, principalement composé de brins  $\beta$  (Matias *et al.*, 2006). De part sa structure et la répartition de ses charges de surfaces, ce dernier pourrait notamment constituer un domaine d'interaction avec l'ADN (Matias *et al.*, 2006).



**Figure 51 : Structures des protéines AAA+ Rvb1p/TIP49 et Rvb2p/TIP48.** (A) Structure tridimensionnelle de la protéine humaine TIP49 obtenue par radiocristallographie (Matias *et al.*, 2006). DI, DII, DIII : Domaines I, II et III. (B) Visualisation par microscopie électronique des structures hexamériques formées par les protéines Rvb1p et Rvb2p de *Saccharomyces cerevisiae*, en présence d'ADP, d'ATP ou d'ATP $\gamma$ S (un analogue d'ATP non hydrolysable) (Gribun *et al.*, 2008). (C) Modèle de structure en double-anneau obtenu par cryo-microscopie électronique à partir des protéines Rvb1p et Rvb2p de *Saccharomyces cerevisiae* (Torreira *et al.*, 2008).

La structure cristallographique de la protéine TIP49 humaine et les différentes données de microscopie électronique ont révélé que les protéines Rvb1p/TIP49 et Rvb2p/TIP48 recombinantes s'associent sous la forme d'un anneau hexamérique (Gribun *et al.*, 2008; Matias *et al.*, 2006; Puri *et al.*, 2007; Torreira *et al.*, 2008) (**Fig. 51B et C**). Cette particularité est également une caractéristique des protéines de type AAA+ (Neuwald *et al.*, 1999). Les contacts entre chaque sous-unité sont réalisés par l'intermédiaire des domaines AAA+, de telle sorte que chaque nucléotide est pris en sandwich entre deux monomères (Matias *et al.*, 2006). Cette configuration particulière est importante pour l'hydrolyse du nucléotide puisque, pour chaque site actif, une arginine (R357 dans TIP49 humaine) de la sous-unité adjacente contribue à la catalyse (Davey *et al.*, 2003; Hishida *et al.*, 2004; Johnson & O'Donnell, 2003; Karata *et al.*, 1999) (**Fig. 51A**). La nature du nucléotide contenu dans le site actif est également déterminante pour la conformation de l'hexamère, puisque de larges changements ont été constatés entre les formes associées à l'ADP et à l'ATP (Gribun *et al.*, 2008) (**Fig. 51B**). La présence d'ATP se traduit notamment par la diminution de la taille du pore central, révélant ainsi que l'activité ATPase constitue le moteur de la flexibilité structurale du complexe. De plus, et même si la question fait toujours débat et qu'il n'existe pas de preuve absolue, différents résultats suggèrent que les protéines Rvb1p/TIP49 et Rvb2p/TIP48 forment des anneaux hétérohexamériques (Cheung *et al.*, 2010). En effet, une étude réalisée à partir des protéines de levure a révélé que le mélange des deux paralogues Rvb1p et Rvb2p améliore de façon significative leur capacité d'oligomérisation, ainsi que leur activité ATPase (Gribun *et al.*, 2008; Ikura *et al.*, 2000). Enfin, alors que certaines études concluent à la formation d'anneaux hexamériques simples (Gribun *et al.*, 2008), d'autres travaux proposent l'association des hexamères deux à deux pour former des structures en double anneaux, reliées par l'intermédiaire des domaines d'insertion (Puri *et al.*, 2007; Torreira *et al.*, 2008) (**Fig. 51C**).

#### 4.6.1.2. Données fonctionnelles

Le domaine AAA+ des protéines Rvb1p/TIP49 et Rvb2p/TIP48 présente une homologie de séquence avec l'hélicase bactérienne RuvB, qui dissocie les doubles brins d'ADN dans le sens 5'-3' (Hishida *et al.*, 2004; Kanemaki *et al.*, 1997; Tsaneva *et al.*, 1993). Une telle fonction a donc été proposée pour les protéines Rvb1p/TIP49 et Rvb2p/TIP48 (Cheung *et al.*, 2010;

Matias *et al.*, 2006). L'hexamère pourrait en effet interagir avec les fourches d'ADN, de telle sorte qu'un des brins d'ADN passe à travers le pore central, tandis que le second soit stabilisé à l'extérieur de l'anneau par l'intermédiaire d'une interaction avec les domaines II. La progression de la fourche le long du duplex d'ADN pourrait ainsi être stimulée par l'hydrolyse de l'ATP. En effet, la protéine TIP49 humaine est capable d'interagir directement avec des molécules d'ADN double brin et simple brin, ainsi qu'avec des molécules d'ARN simple brin (Matias *et al.*, 2006). En revanche les données obtenues *in vitro* n'ont pas encore permis de caractériser clairement cette activité hélicase et la littérature révèle une certaine controverse à ce sujet. Parmi les résultats marquants, et à l'image de l'activité ATPase, il semblerait néanmoins que l'activité hélicase soit améliorée par la présence des deux protéines Rvb1p et Rvb2p chez la levure (Gribun *et al.*, 2008; Ikura *et al.*, 2000). Celle-ci semble pouvoir s'exercer de façon bidirectionnelle avec une large préférence pour le sens 5'-3' (Gribun *et al.*, 2008; Kanemaki *et al.*, 1999). La présence d'une extrémité 5' ou 3' sortante au niveau du duplex d'ADN favorise également l'activité du complexe Rvb1p-Rvb2p (Gribun *et al.*, 2008). Enfin, la difficulté à caractériser cette activité hélicase pourrait reposer sur l'implication de cofacteurs protéiques permettant d'augmenter cette activité *in vivo*.

Outre cette fonction d'hélicase, de nombreux facteurs de la famille des protéines AAA+ exercent leur activité sur des polypeptides (Hanson & Whiteheart, 2005; Neuwald *et al.*, 1999). Il existe en effet de nombreux exemples indiquant que les changements conformationnels des protéines AAA+ induit par le cycle d'hydrolyse des nucléotides sont mis à profit pour assurer la déstructuration de certaines protéines, ou la dissociation d'agrégats et de complexes protéiques (Hanson & Whiteheart, 2005; Neuwald *et al.*, 1999). L'étude des protéines bactériennes CplA et CplX, qui adressent des polypeptides déstructurés à la protéase CplP, suggère notamment que le transit des protéines clientes par le pore central de l'hexamère pourrait contribuer à ce processus de dénaturation (Kenniston *et al.*, 2003; Sauer *et al.*, 2004).

Comme les autres membres de la famille des protéines AAA+, les hélicases Rvb1p/TIP49 et Rvb2p/TIP48 ont été impliquées dans de nombreux processus cellulaires (Huen *et al.*, 2010; Jha & Dutta, 2009). Elles interviennent notamment dans les mécanismes de régulation de la transcription et de réparation de l'ADN, de part leur appartenance au complexe d'acétylation des histones TIP60 (Ikura *et al.*, 2000; Jha *et al.*, 2008) et aux complexes de remodelage de la chromatine INO80 (Jonsson *et al.*, 2004; Ohdate *et al.*, 2003; Shen *et al.*, 2000) et SWR-C

(Cai *et al.*, 2005; Kobor *et al.*, 2004; Krogan *et al.*, 2003b; Mizuguchi *et al.*, 2004). Elles interagissent également avec différents facteurs de transcription nécessaires pour la progression du cycle cellulaire, comme c-Myc (Bellosta *et al.*, 2005; Etard *et al.*, 2005; Wood *et al.*, 2000) et la  $\beta$ -Caténine (Bauer *et al.*, 2000; Feng *et al.*, 2003; Kim *et al.*, 2005). Par l'intermédiaire de ces différentes fonctions, généralement dépendantes de l'activité ATPase, il a ainsi été montré que les protéines Rvb1p et Rvb2p sont impliquées dans la transcription de plus de 5% des gènes de levure (Jonsson *et al.*, 2001).

Concernant la biogenèse des snoRNP, il a notamment été montré que la déplétion des protéines humaines TIP49 et TIP48 affecte le taux de snoRNA à boîtes C/D (Watkins *et al.*, 2004). Chez la levure, la déplétion de la protéine Rvb2p perturbe la production des ARNr et affecte le taux des snoRNA à boîtes C/D, à l'exception notoire du snoRNA U3 (King *et al.*, 2001). La protéine Rvb2p est également requise pour maintenir le taux de snoRNA à boîtes H/ACA, suggérant un rôle élargi à la biogenèse des deux classes de snoRNP (King *et al.*, 2001). Son expression est également importante pour la localisation nucléolaire de certains composants des snoRNP, puisque la déplétion de Rvb2p conduit à une délocalisation des protéines Nop1p et Gar1p vers le nucléoplasme et le cytoplasme (King *et al.*, 2001). Une délocalisation similaire de la protéine Nop1p avait précédemment été constatée suite à la déplétion de Nop58p (Wu *et al.*, 1998). Ce résultat constitue une donnée intéressante puisque des données obtenues chez l'homme indiquent que le recrutement des deux hélicases et de la protéine NOP58 sur les snoRNA pourrait s'effectuer de façon commune (Watkins *et al.*, 2002). En effet, l'association des protéines humaines TIP48 et TIP49 sur un snoRNA à boîtes C/D est dépendante de la séquence de l'hélice II et semble reposer sur les mêmes déterminants que ceux nécessaires à la fixation de NOP58 (Watkins *et al.*, 2002). D'autre part, de nombreux résultats indiquent que les deux hélicases interviennent au cours des étapes nucléoplasmiques de la biogenèse des particules. Chez la levure, la protéine Rvb2p qui est enrichie dans le nucléoplasme, apparaît exclue du nucléole (King *et al.*, 2001), tandis que les protéines p50 et p55 de souris présentent également une localisation nucléoplasmique (Bauer *et al.*, 1998; Holzmann *et al.*, 1998; Newman *et al.*, 2000). De façon intéressante, des expériences d'immunoprécipitation réalisées chez la souris ont montré que ces deux protéines s'associent avec le snoRNA U14 incubé *de novo* dans un extrait cellulaire, tandis qu'elles n'interagissent pas avec les snoRNP U14 préformées, présentes dans l'extrait (Bauer *et al.*, 1998; Holzmann *et al.*, 1998; Newman *et al.*, 2000). Ce résultat, ainsi que l'identification des protéines TIP49 et TIP48 au sein des pré-snoRNP U3 humaines (Watkins *et al.*, 2004),

indiquent donc que les deux hélicases ne présentent pas de réelle affinité pour les snoRNP matures, mais qu'elles interagissent néanmoins avec les composants des snoRNP au cours de leur biogenèse.

#### 4.6.2. La protéine Nop17p/NOP17 (alias Pih1p)

La protéine Nop17p (*Nucleolar protein 17*), alias Pih1p (*Protein interacting with Hsp90-1*), a initialement été identifiée par double-hybride comme partenaire de la sous-unité Rrp43p de l'exosome (Gonzales *et al.*, 2005; Oliveira *et al.*, 2002), de la protéine Nop58p (Gonzales *et al.*, 2005), et de la protéine Nop53p, un facteur impliqué dans la biogenèse des ARNr 25S et 5.8S (Granato *et al.*, 2005). Ces interactions ont été confirmées par des expériences de co-immunosélection, en revanche, aucune interaction n'a été détectée entre Nop17p et les autres protéines constitutives des snoRNP à boîtes C/D (Gonzales *et al.*, 2005; Granato *et al.*, 2005). D'autre part, les tests de double-hybride ont permis de montrer que la région 217-512 de Nop58p, qui comprend le domaine NOP et une portion du domaine *Coiled-coil* (**Fig. 30**), est suffisante pour permettre l'interaction avec Nop17p (Gonzales *et al.*, 2005). Cette dernière n'est pas essentielle à la viabilité cellulaire, mais sa déplétion conduit à un défaut de croissance qui est accentué à 37°C (Gonzales *et al.*, 2005). Ce phénotype est associé à une réduction du nombre de polysomes et de la quantité totale de ribosomes. Cet effet repose notamment sur une diminution du taux d'ARNr 18S et 25S, et sur l'accumulation des précurseurs 35S et 23S (Gonzales *et al.*, 2005). De plus, des données d'analyse à haut-débit ont révélé la présence des protéines Nop58p et Nop17p au sein des mêmes complexes pré-ribosomiques (Krogan *et al.*, 2004). Ces différents résultats suggèrent donc un rôle de la protéine Nop17p dans la biogenèse des ribosomes, néanmoins d'autres données indiquent que cette protéine pourrait intervenir en amont, dans le cadre de la biogenèse des snoRNP à boîtes C/D. En effet, la protéine Nop17p, qui s'accumule au sein du nucléole et du nucléoplasme, s'est avérée importante pour la localisation correcte des protéines Snu13p, Nop1p, Nop56p et Nop58p au niveau du nucléole (Gonzales *et al.*, 2005). Sa déplétion conduit notamment à une accumulation aberrante de ces protéines au niveau du nucléoplasme et du cytoplasme. De plus, la protéine Nop17p ne semble pas associée de façon stable aux snoRNA à boîtes C/D, mais son absence conduit à une augmentation drastique de l'association de la protéine Nop58p avec les snoRNA U3 et U14 (Gonzales *et al.*, 2005). Ce résultat surprenant suggère



donc que l'interaction entre Nop17p et Nop58p pourrait contribuer à réguler le processus d'assemblage des particules. En revanche, la déplétion de Nop17p n'a pas révélé d'effet marqué sur le taux de snoRNA U3 (Gonzales *et al.*, 2005).

### **4.6.3. Le complexe R2TP, cofacteur des chaperonnes HSP90**

#### **4.6.3.1. Les chaperonnes HSP90**

Les protéines de type HSP90 sont des chaperonnes ubiquitaires, retrouvées chez les bactéries et chez les eucaryotes. Elles jouent un rôle majeur dans les processus de signalisation cellulaire, puisqu'elles sont requises pour maintenir l'activité de nombreux facteurs de signalisation, comme les protéines kinases et les récepteurs aux hormones stéroïdiennes (Caplan *et al.*, 2007). Cette fonction s'inscrit dans un rôle global de maintien de l'homéostasie cellulaire. De nombreuses études ont en effet révélé que les protéines HSP90 peuvent compenser des variations phénotypiques en protégeant certaines protéines de la dégradation et en facilitant leur repliement dans un état actif (Rutherford, 2003). Cette fonction constitue un système de sécurité visant à maintenir de nombreuses fonctions essentielles en dépit des mutations ou des variations environnementales. Néanmoins, il est clairement établi que les chaperonnes HSP90 jouent un rôle important dans le développement tumoral (Trepel *et al.*, 2010). Ces chaperonnes agissent en effet sur de nombreuses oncoprotéines en les protégeant de la dégradation, en dépit de leur surexpression ou des mutations dont elles sont la cible. La surexpression des HSP90 constitue donc un marqueur de référence pour la recherche de cellules cancéreuses et l'inhibition de leur activité représente une stratégie de choix dans le cadre des thérapies anticancéreuses. Actuellement, 13 inhibiteurs sont en phase d'essai clinique (Kim *et al.*, 2009b).

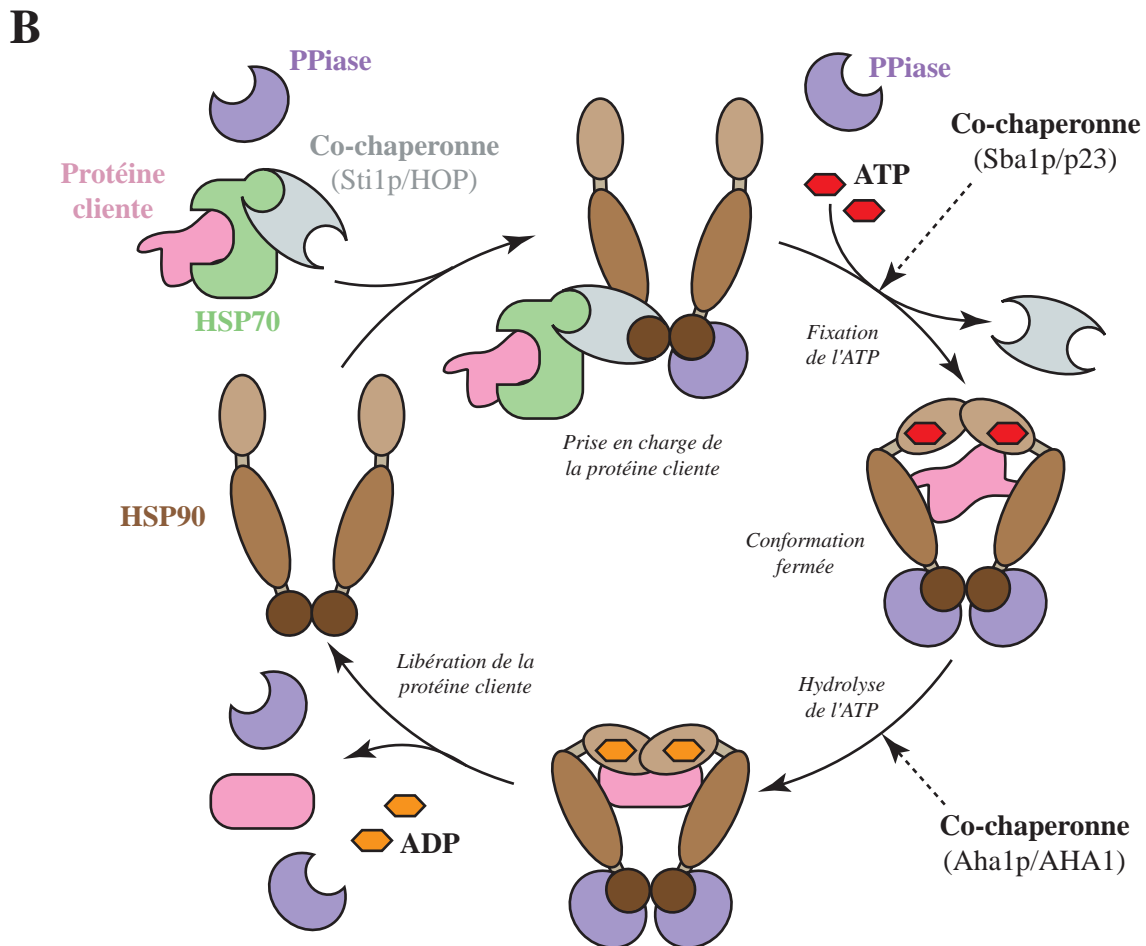
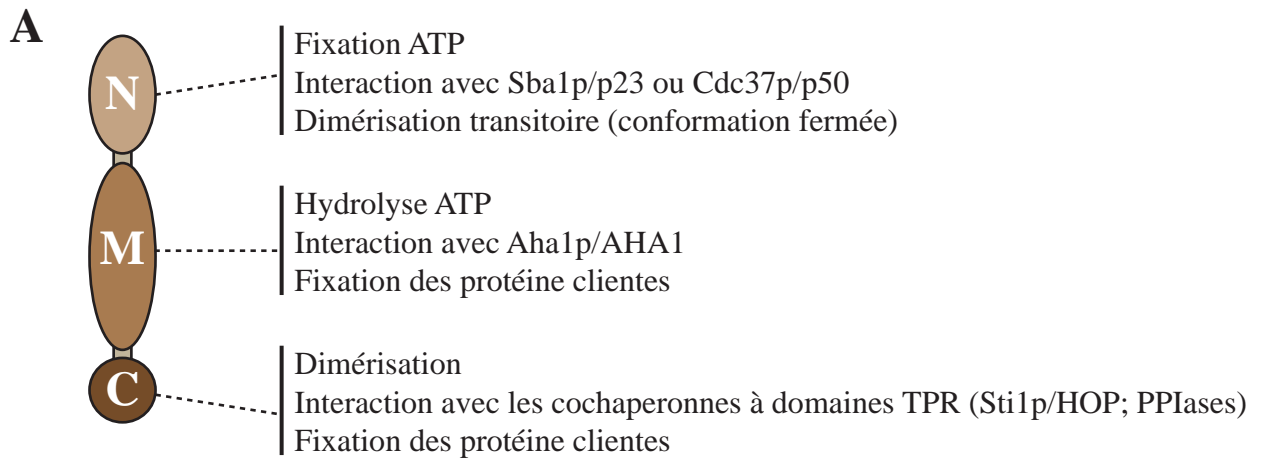
Comme dans beaucoup d'autres organismes, il existe chez la levure *S. cerevisiae* deux isoformes de la chaperonne : la protéine Hsc82p, qui est exprimée de façon constitutive, et la protéine Hsp82p, qui est surexprimée en condition de stress thermique ou en phase stationnaire de croissance (Borkovich *et al.*, 1989; Donalies & Stahl, 2001). Ces isoformes, principalement cytoplasmiques, semblent faire l'objet de passages transitoires dans le noyau durant la phase exponentielle de croissance (Huh *et al.*, 2003; Tapia & Morano, 2010). Une

étude récente a en revanche montré que la protéine Hsp82p a tendance à s'accumuler au sein du noyau en phase de quiescence (Tapia & Morano, 2010). Cet import est dépendante de l'importine Srp1p/Kap95p et pourrait permettre de cibler spécifiquement certaines protéines au niveau du noyau (Tapia & Morano, 2010).

Parmi les nombreuses chaperonnes qui facilitent le repliement des protéines cellulaires, les protéines HSP90 présentent certaines particularités (Zhao & Houry, 2005). Elles interviennent en effet exclusivement sur les protéines dans un état de repliement avancé, proche de l'état natif. De plus, elles s'associent avec des cofacteurs protéiques, appelés cochaperonnes, qui régulent leur cycle d'activité. Les HSP90 agissent en effet au sein de larges complexes macromoléculaires contenant différents co-facteurs, mais également d'autres chaperonnes, comme les HSP70 et les HSP40, ainsi que les protéines substrats, appelées « protéines clientes ». Il a notamment été proposé que la prise en charge des protéines clientes par ces macro-complexes chaperons implique l'intervention successive des protéines HSP40, HSP70, puis HSP90 (Frydman & Hohfeld, 1997). C'est ainsi HSP70 qui délivrerait la protéine cliente à la chaperonne HSP90 afin de procéder aux étapes finales de son repliement.

Les chaperonnes HSP90 sont composées de 3 principaux domaines : un domaine N-terminal qui contient le site de fixation à l'ATP (Prodromou *et al.*, 1997), un domaine central, essentiel pour l'hydrolyse de l'ATP et pour l'association avec les protéines clientes (Meyer *et al.*, 2003), et un domaine C-terminal, responsable de la dimérisation de la protéine et du recrutement des cochaperonnes contenant des domaines TPR (*Tetratricopeptid Repeat*) (Scheufler *et al.*, 2000) (**Fig. 52A**). Les domaines TPR sont composés de plusieurs motifs hélices-boucle-hélices de 34 acides aminés, appelés motifs TPR, et sont fréquemment retrouvés parmi les cofacteurs des protéines HSP (D'Andrea & Regan, 2003). La protéine Sti1p/HOP, une cochaperonne majeure de HSP90, contient notamment plusieurs domaines TPR, lui permettant d'interagir avec les motifs C-terminaux EEVD des chaperonnes HSP90 et HSP70 (Scheufler *et al.*, 2000). Cette double interaction permet ainsi de former le complexe HSP90-HOP-HSP70, qui constitue l'entité minimale pour le fonctionnement de la chaperonne (**Fig. 52B**).

Au cours de son cycle d'activité, HSP90 adopte une forme dimérique afin de former une pince moléculaire, dont la conformation ouverte/fermée est régulée par l'hydrolyse d'ATP (**Fig. 52B**) (Mayer, 2010; Pearl & Prodromou, 2006). Au début du cycle, HSP90 est associée



**Figure 52 : Structure et fonction des chaperonnes HSP90 (adapté de Li *et al.*, 2011; Mayer *et al.*, 2010).** (A) Représentation schématique des 3 domaines de la protéine chaperonne HSP90 et de leurs fonctions associées. N : "N-terminal domain"; M : "Middle domain"; C : "C-terminal domain". (B) Cycle d'activité de la chaperonne HSP90 et régulation par différents cofacteurs. PPIase : Peptidyl-Prolyl Isomérase (ex: FKBP51, FKBP52, FKBP37, Cpr6-7p/CYP40).

à des cochaperonnes comme Sti1p/HOP ou Cdc37p/p50 qui facilitent la prise en charge de la protéine cliente (Felts *et al.*, 2007; Hessling *et al.*, 2009; Roe *et al.*, 2004). Le passage à une conformation fermée stable requiert ensuite la fixation d'ATP et la présence de cochaperonnes comme Sba1p/p23 (Ali *et al.*, 2006). En fin de cycle, certaines cochaperonnes comme Aha1p/AHA1 vont ensuite favoriser l'hydrolyse de l'ATP et le relargage de la protéine cliente correctement structurée (Hessling *et al.*, 2009; Meyer *et al.*, 2004). Plusieurs PPIases (*Peptidyl-Prolyl isomérases*) interviennent également au cours de ce cycle. Ces dernières contiennent en effet des domaines TPR, qui leur permettraient notamment de déplacer la protéine Sti1p/HOP de son site de fixation (Li *et al.*, 2011a) (**Fig. 52B**).

#### 4.6.3.2. Identification du complexe R2TP

Afin d'exercer sa fonction sur un large spectre de protéines clientes, il semble que les chaperonnes HSP90 puissent s'adjoindre la fonction de différents jeux de cochaperonnes (Zhao & Houry, 2007). Ces dernières, spécifiques de certains types de substrats, lui permettraient en effet d'agir de façon ciblée dans certaines voies du métabolisme cellulaire, tout en conservant son statut de chaperonne ubiquitaire. En 2005, l'équipe de W.A. Houry a notamment réalisé une série de cribles à grande échelle visant à identifier de nouveaux partenaires pour les isoformes Hsp82p et Hsc82p de levure (Zhao *et al.*, 2005). Cette démarche a permis d'identifier un total de 627 partenaires, substrats ou cofacteurs, soit environ 10% du protéome de levure. Parmi ces candidats potentiels, deux nouveaux cofacteurs ont été caractérisés : la protéine Nop17p, rebaptisée Pih1p, et une petite protéine de 111 acides aminés, baptisée Tah1p pour *TPR-containing protein associated with Hsp90*. Cette dernière contient en effet un domaine TPR similaire au domaine TPR2B de la cochaperonne Sti1p/HOP (Scheufler *et al.*, 2000).

Les interactions des protéines Pih1p et Tah1p entre elles et avec la chaperonne ont initialement été identifiées par la technique de double-hybride (Ito *et al.*, 2001; Zhao *et al.*, 2005). Cette approche a notamment permis de mettre en évidence l'interaction de la protéine Tah1p avec l'extrémité C-terminale (599-709) de la chaperonne, suggérant un mode d'interaction similaire à celui de Sti1p/HOP. Des expériences de pull-down réalisées à partir de protéines recombinantes, ont ensuite permis de préciser la nature de ces interactions (Zhao

*et al.*, 2005). En effet, la protéine Tah1p interagit à la fois avec les protéines Pih1p et Hsp82p, tandis que l'association de Hsp82p avec Pih1p n'est possible qu'en présence de Tah1p. Ceci suggère donc que l'interaction entre Hsp82p et Pih1p observée par double-hybride pourrait être médiée par la protéine Tah1p endogène, et que ces trois facteurs pourraient être associés au sein d'un même complexe. En utilisant un système rapporteur, les auteurs ont également démontré l'importance des protéines Pih1p et Tah1p pour la structuration optimale du récepteur aux glucocorticoïdes et de la tyrosine kinase v-Src, deux protéines clientes de la chaperonne. Enfin, dans le but d'identifier d'autres substrats spécifiques des cofacteurs Pih1p et Tah1p, l'équipe de W.A. Houry a également entrepris leur purification par l'approche TAP-tag (Zhao *et al.*, 2005). De façon surprenante, cette démarche n'a pas permis d'observer leur association avec les protéines Hsp82p ou Hsc82p, suggérant la nature transitoire de ces interactions et/ou la formation d'une quantité restreinte de complexes *in vivo*. Les auteurs ont en revanche montré que les protéines Pih1p et Tah1p sont abondamment co-purifiées l'une avec l'autre, ainsi qu'avec les AAA+-ATPases Rvb1p et Rvb2p. Le complexe formé par ces 4 protéines a été baptisé R2TP pour « Rvb1p-Rvb2p-Tah1p-Pih1p ». Des expériences réalisées à partir des souches de levure invalidées ont notamment révélé que l'absence des cofacteurs Pih1p ou Tah1p conduit à une incorporation accrue des deux hélicases au sein du complexe de remodelage de la chromatine INO80 (Zhao *et al.*, 2005).

Bien que les expériences de co-immunosélection n'aient pas permis de mettre en évidence une association stable avec la chaperonne *in vivo*, l'ensemble de ces données suggère que Pih1p et Tah1p constituent de nouveaux cofacteurs des protéines HSP90 (Zhao *et al.*, 2005). D'autre part, la localisation nucléaire des protéines Pih1p et Tah1p indique que leur interaction avec la chaperonne pourrait s'effectuer soit de façon transitoire au niveau du cytoplasme, soit dans des conditions de croissance particulières, au niveau du noyau (Tapia & Morano, 2010; Zhao *et al.*, 2005). Le rôle fonctionnel du complexe R2TP reste cependant à déterminer. Comme cela a été proposé, ce dernier pourrait en effet constituer un intermédiaire, conduisant à la formation des complexes de remodelage de la chromatine INO80 et SWR-C. Néanmoins, l'implication des protéines Rvb1p, Rvb2p et Pih1p dans les processus de biogenèse des snoRNP, suggère que ce complexe pourrait constituer une entité à part entière, exerçant sa fonction dans le cadre d'autres processus cellulaires. Enfin, en utilisant une approche de protéomique, une autre équipe a récemment mis en évidence que les protéines TIP49, TIP48 et PIH1 sont co-immunosélectionnées avec HSP90 humaine, suggérant la conservation du complexe R2TP chez les eucaryotes (Te *et al.*, 2007).

#### 4.6.4. La protéine Cwc24p

Un crible d'interactions par double-hybride, réalisé à partir de la protéine Nop17p (alias Pih1p), a permis de mettre en évidence un autre facteur potentiellement impliqué dans la biogenèse des snoRNP à boîtes C/D : la protéine Cwc24p (Gonzales *et al.*, 2005). Cette protéine essentielle à la viabilité cellulaire, contient notamment un domaine RING, ainsi qu'un domaine à doigt de zinc, contenu dans le fragment minimum d'interaction avec Pih1p (Cwc24p<sub>86-164</sub>) (Goldfeder & Oliveira, 2008). L'interaction entre les deux partenaires a notamment été confirmée *in vitro* par GST-pull down (Goldfeder & Oliveira, 2008).

La protéine Cwc24p est enrichie dans le nucléoplasme, mais ne s'accumule pas dans le nucléole (Goldfeder & Oliveira, 2008). Sa déplétion ralentit le processus de maturation des ARNr en affectant le clivage au site A0, et à moindre mesure, les clivages en A1 et A2 (Goldfeder & Oliveira, 2008). Cet effet ne semble pas direct et pourrait résulter d'un défaut de biogenèse des snoRNP à boîtes C/D, en particulier de la snoRNP U3. En effet, l'utilisation de puces permettant l'analyse du profil d'épissage a permis de corrélérer la déplétion de Cwc24p à un défaut d'épissage du pré-snoRNA U3 et de plusieurs pré-ARNm contenant des snoRNA à boîtes C/D (ex : snR38, snR54) (Goldfeder & Oliveira, 2008). Cet effet est spécifique, puisque peu de pré-ARNm sont affectés par cette déplétion. De plus, il pourrait reposer sur un lien entre la protéine Cwc24p et le spliceosome, puisque cette dernière interagit physiquement avec le facteur d'épissage Cef1p (Goldfeder & Oliveira, 2008; Ohi *et al.*, 2002). En revanche, des expériences d'immunoprécipitation de la chromatine et de co-immunoprécipitation n'ont pas permis de démontrer le recrutement de la protéine Cwc24p sur l'ARN et/ou l'ADN *in vivo* (Goldfeder & Oliveira, 2008).

L'ensemble de ces données suggère donc que Cwc24p pourrait intervenir de façon transitoire au cours de la biogenèse des pré-snoRNP. De par sa capacité à interagir avec Pih1p et Cef1p, elle pourrait notamment contribuer au couplage entre les processus d'assemblage des RNP et la maturation du snoRNA par le spliceosome (Goldfeder & Oliveira, 2008). De façon intéressante, cette étude révèle également que l'excision de l'intron du snoRNA U3 de levure et la maturation des snoRNA intronique pourrait impliquer des mécanismes communs.

#### 4.6.5. La protéine Bcd1p

En 2003, l'équipe de T.R. Hughes a réalisé un crible à grande échelle de mutants de levures associés à des défauts de biogenèse de certains ARN non codants (Peng *et al.*, 2003). Cette approche, basée sur l'utilisation de puces à ADN, a notamment permis d'identifier une nouvelle protéine, baptisée Bcd1p (*Box C/D snoRNA 1*), qui s'est avérée essentielle à la viabilité cellulaire et à l'accumulation de l'ensemble des snoRNA à boîtes C/D (Hiley *et al.*, 2005; Peng *et al.*, 2003). Bcd1p possède un domaine à doigt de zinc, ainsi qu'un signal de localisation nucléaire. L'analyse sur puce des ARN co-immunosélectionnés avec la protéine TAP-Bcd1p a révélé un enrichissement spécifique en snoRNA à boîtes C/D, en revanche la faible quantité de matériel obtenue n'a pas permis d'identifier les partenaires protéiques associés à Bcd1p. A l'image du rôle joué par la protéine Naf1p dans le cadre de la biogenèse des snoRNP à boîtes H/ACA, il a notamment été proposé que Bcd1p puisse servir de facteur d'échange avec la protéine Nop56p (Matera *et al.*, 2007). Comme la protéine Gar1p, cette dernière est en effet la seule protéine de la particule qui n'est pas essentielle pour la stabilité des snoRNA et son recrutement pourrait s'effectuer de façon secondaire. Néanmoins, cette hypothèse n'a jamais été vérifiée expérimentalement et le rôle précis joué par Bcd1p demeure indéterminé.

#### 4.6.6. Les protéines Rsa1p et NUFIP

##### 4.6.6.1. Rsa1p, un nouveau partenaire de la snoRNP U3

Au laboratoire, Nathalie Marmier-Gourrier et Bruno Charpentier ont entrepris l'identification de nouvelles protéines associées au snoRNA U3, en utilisant l'approche de triple-hybride (Marmier-Gourrier, 2003). Le crible, réalisé à partir d'une banque génomique et d'un ARN rapporteur contenant le motif B/C du snoRNA U3, a permis d'identifier deux clones de la protéine Rsa1p.

Précédemment, l'équipe de P. Linder avait proposé l'implication de ce facteur dans le processus d'assemblage de la sous-unité ribosomique 60S, suite à la découverte de son lien synthétique létal avec l'hélicase Dbp6p (Kressler *et al.*, 1999). Cette protéine, baptisée Rsa1

pour *Ribosome assembly 1*, n'est pas essentielle à la viabilité cellulaire, mais sa délétion confère un défaut de croissance particulièrement marqué à 37°C (Kressler *et al.*, 1999). Elle contient un NLS potentiel au niveau de son extrémité C-terminale (363-PNYKNRK-369) et présente une localisation principalement nucléoplasmique (Kressler *et al.*, 1999). Elle est en revanche exclue du nucléole.

La délétion de cette protéine conduit à une diminution du taux de sous-unités 60S et à l'accumulation de polysomes incomplets (Kressler *et al.*, 1999). Le suivi de la production des ARNr par *pulse-chase labeling* a également révélé une diminution du taux d'ARNr 25S et du ratio 25S/18S. Cette diminution n'est cependant pas le résultat d'une perte de stabilité de l'ARNr 25S, mais repose sur un ralentissement de sa production. L'ensemble de ces effets sont détectables à 30°C, mais apparaissent beaucoup plus marqués à 37°C. D'autre part, la voie de maturation de l'ARNr 18S semble également affectée par l'absence de la protéine Rsa1p, puisque le taux d'ARN mature diminue légèrement et de façon concomitante à l'accumulation de l'intermédiaire 23S. Les auteurs suggèrent néanmoins que cette observation constitue un effet secondaire, résultant vraisemblablement d'un rétrocontrôle négatif déclenché par le défaut de biogenèse de la sous-unité 60S (Kressler *et al.*, 1999).

De façon intéressante, l'équipe de P. Linder a également montré que la délétion du gène *RSA1* conduit à l'accumulation de sous-unité 60S dépourvues de la protéine ribosomique Rpl10p (alias Qsr1p) (Kressler *et al.*, 1999). Cette protéine, essentielle pour l'association des deux sous-unités ribosomiques, serait incorporée au sein de la particule 60S, suite à un échange avec le facteur d'export Nmd3p au niveau du cytoplasme (Dick *et al.*, 1997a; Eisinger *et al.*, 1997; Hedges *et al.*, 2005) (cf. Introduction, Chapitre 1, 3.4.2) (**Fig. 20**). Ces résultats suggèrent donc que la protéine Rsa1p pourrait être impliquée dans les étapes tardives de l'assemblage de la sous-unité 60S, en particulier dans le cadre de l'échange Nmd3p-Rpl10p (Kressler *et al.*, 1999). Néanmoins, son intervention pourrait avoir lieu en amont de ce processus, puisque Rsa1p, qui est essentiellement nucléoplasmique, est également requise pour l'export de la particule 60S vers le cytoplasme (Gadal *et al.*, 2001). L'accumulation de polysomes incomplets pourrait donc résulter d'une rétention nucléaire de la sous-unité 60S, associée à un ralentissement coordonné de la production des ARNr. Par ailleurs, la protéine Rsa1p présente également différents liens génétiques avec les protéines de la grande sous-unité Rpl3p, Rpl10p et Rpl29p, ainsi qu'avec les protéines Nop8p et Rsa3p, deux autres



facteurs impliqués dans la biogenèse des ribosomes (de la Cruz *et al.*, 2004; DeLabre *et al.*, 2002; Kressler *et al.*, 1999).

Même si son rôle précis reste à déterminer, l'ensemble de ces données suggèrent donc l'implication de Rsa1p dans la biogenèse des ribosomes, en particulier dans les étapes tardives d'assemblage de la sous-unité 60S. En plus de cette fonction, la découverte de son association avec le snoRNA U3 constitue un autre lien inattendu avec le processus de production des ribosomes. Des études préliminaires menées au laboratoire par Bruno Charpentier, ont notamment permis de montrer que la portion C-terminale de la protéine (acides aminés 230 à 381 ; alias Rsa1p<sub>230-381</sub> ou N3C1) est suffisante pour compléter le défaut de croissance d'une souche  $\Delta RSA1$  et pour assurer son interaction par triple-hybride avec le snoRNA U3. En revanche, des tests d'interaction réalisés par retard sur gel, ont démontré l'absence d'interaction directe entre les deux partenaires et suggèrent l'existence d'un partenaire intermédiaire, responsable du recrutement de la protéine Rsa1p sur le snoRNA U3 *in cellulo*. En collaboration avec notre groupe, l'équipe d'E. Bertrand a entrepris la recherche des protéines associées à Rsa1p par une approche de double-hybride ciblée. Ces expériences ont notamment révélé l'interaction de Rsa1p et de son fragment C-terminal avec la protéine Snu13p, ainsi qu'avec l'hyperméthylase Tgs1p et la protéine Lsm4p (Boulon, 2005). Parmi ces facteurs, la protéine Snu13p, qui interagit spécifiquement avec le motif B/C du snoRNA U3, constitue le meilleur candidat pour médier le recrutement de Rsa1p sur l'ARN, d'autant plus que la destruction de ce motif par mutation de la boîte C, abolit l'interaction entre Rsa1p et l'ARN par l'approche triple-hybride. La protéine Rsa1p, qui n'est pas retrouvée au sein de la particule mature (Watkins *et al.*, 2000) et qui ne s'accumule pas dans le nucléole (Kressler *et al.*, 1999), pourrait donc interagir de façon transitoire avec la snoRNP U3 au cours de son transit par le nucléoplasme.

#### **4.6.6.2. NUFIP, un nouveau partenaire de la protéine 15.5K**

En parallèle de ces travaux, l'équipe d'E. Bertrand a réalisé un crible d'interaction par double-hybride afin d'identifier de nouveaux partenaires protéiques pour la protéine humaine 15.5K (Boulon, 2005). Cette démarche a permis de mettre en évidence son interaction avec un nouveau facteur d'intérêt : la protéine NUFIP (*Nuclear FMRP Interacting Protein*). Cette

dernière possède un domaine à doigt de zinc de type C2H2 (174-201), un NLS bipartite (243-260), un NES dépendant de CRM1 dans sa région C-terminale et un site de phosphorylation par les kinases cdk (292-295) (Bardoni *et al.*, 2003; Bardoni *et al.*, 1999; Cabart *et al.*, 2004). Comme Rsa1p, elle interagit par double-hybride avec les protéines LSM4 et TGS1 (Boulon, 2005). Faiblement détectée au niveau du cytoplasme et du nucléole, NUFIP présente une localisation principalement nucléoplasmique (Bardoni *et al.*, 1999; Cabart *et al.*, 2004). De plus, son expression, qui a été constatée dans de nombreux types cellulaires, est constante au cours du cycle cellulaire (Bardoni *et al.*, 1999; Cabart *et al.*, 2004).

Initialement, la protéine NUFIP a été caractérisée comme partenaire d'interaction de FMRP, une protéine dont l'expression est dérégulée au cours du développement du syndrome du X-fragile (Bardoni *et al.*, 1999). FMRP est impliquée dans la stabilité, le transport et le contrôle traductionnel de certaines catégories d'ARNm (De Rubeis & Bagni, 2010). Elle exerce une action essentielle au niveau des cellules neuronales, où elle participe au sein de larges RNP, à l'export des ARNm vers le cytoplasme et à leur traduction localisée au niveau des dendrites. Les travaux de B. Bardoni ont notamment permis de mettre en évidence une co-localisation des deux protéines au niveau du noyau, ainsi qu'un profil d'expression similaire et prédominant au niveau du cortex, de l'hippocampe et des cellules de Purkinje (Bardoni *et al.*, 1999). Comme FMRP, la protéine NUFIP a également la capacité d'interagir physiquement avec l'ARN, puisqu'elle interagit *in vitro* avec des ARN poly(G) ou poly(U) (Bardoni *et al.*, 1999). Cette propriété est cependant indépendante du domaine à doigt de zinc et pourrait reposer sur un domaine structural encore non décrit. L'ensemble de ces données suggère donc l'implication des protéines NUFIP et FMRP dans des processus communs, vraisemblablement liés au métabolisme des ARN au sein du noyau.

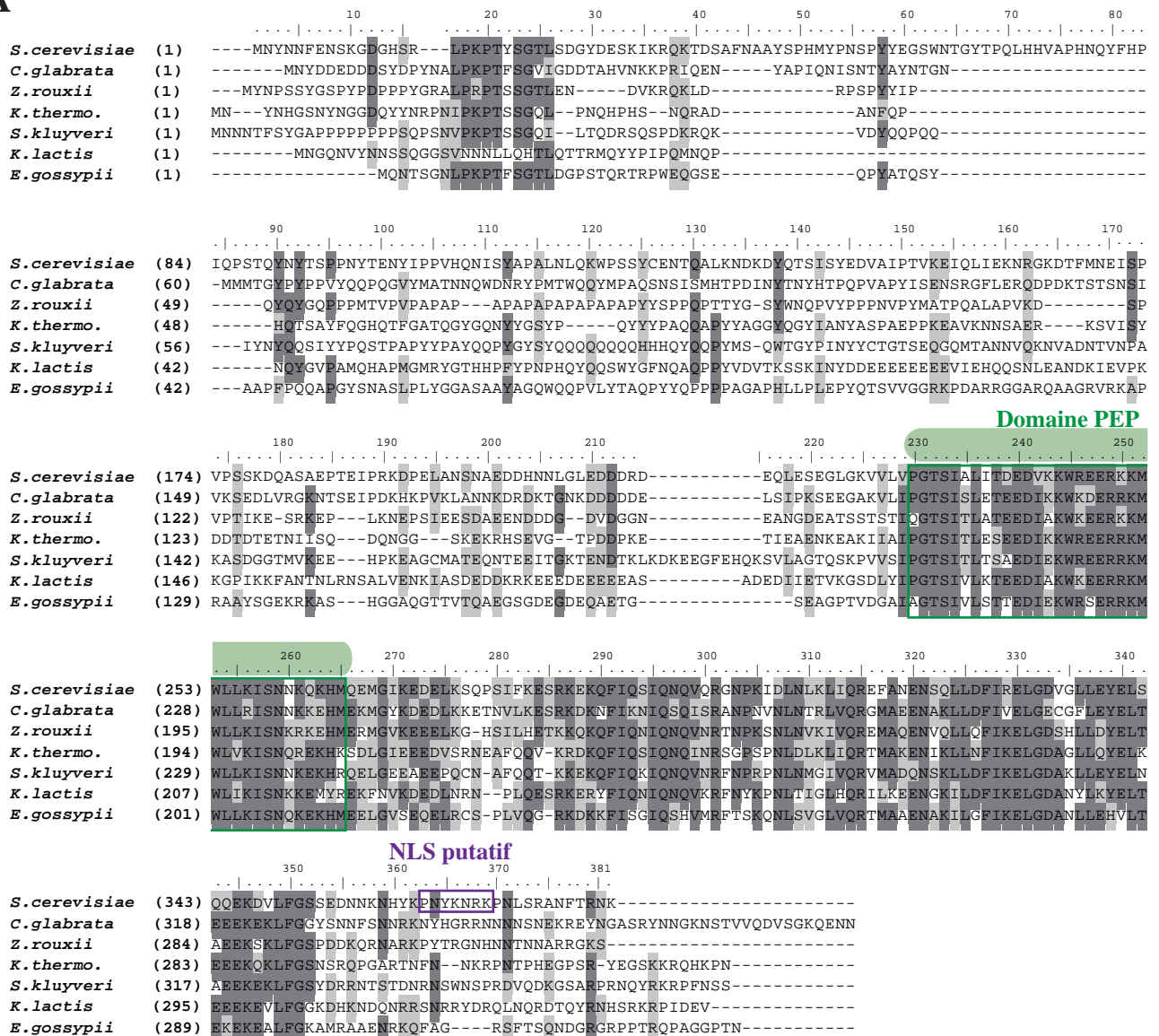
Cette fonction pourrait notamment concerner les étapes précoces de la biogenèse des ARNm, puisqu'un lien entre NUFIP et la protéine BRCA1, a été mis à jour (Cabart *et al.*, 2004). BRCA1 est un suppresseur de tumeur bien caractérisé, dont la mutation confère un risque élevé de développement de cancer du sein ou des ovaires (Linger & Kruk, 2010). Ce facteur possède de multiples rôles puisqu'il a été impliqué dans la réponse aux dommages à l'ADN, dans la progression du cycle cellulaire, dans l'apoptose, ou encore dans la réponse aux hormones stéroïdiennes (Rosen *et al.*, 2006). L'ensemble de ces fonctions semble reposer sur la capacité de BRCA1 à réguler la transcription. Cette dernière est en effet associée à différents complexes de remodelage de la chromatine, ainsi qu'à l'ARN polymérase II, dont

elle peut stimuler ou inhiber l'activité en fonction de son association avec différents partenaires protéiques. Des expériences de co-immunosélection et de GST-pull down ont permis de mettre en évidence une interaction entre les protéines NUFIP et BRCA1 (Cabart *et al.*, 2004). Cette interaction semble par ailleurs intervenir dans le contexte transcriptionnel. NUFIP présente en effet la capacité de stimuler spécifiquement l'activité ARN polymérase II à la fois *in vitro* et *in vivo*. De plus, cette propriété qui requiert l'intégrité du domaine à doigt de zinc, s'est révélée optimale en présence de BRCA1. D'autres données indiquent que la protéine NUFIP s'associe également avec le complexe PTEF-b, en particulier par l'intermédiaire d'une interaction avec la sous-unité cycline T1 (Cabart *et al.*, 2004). Comme nous l'avons abordé précédemment, le complexe PTEF-b intervient dans la stimulation de l'élongation de la transcription par phosphorylation de la Sérine 2 du domaine CTD (cf. Introduction, Chapitre 1, 2.1) (Komarnitsky *et al.*, 2000). Dans son étude, l'équipe de S. Murphy a ainsi montré que l'effet activateur de la transcription par la protéine NUFIP requiert la présence de la cycline T1 (Cabart *et al.*, 2004). Comme le complexe PTEF-b, NUFIP est retrouvée au sein du complexe de pré-initiation de la transcription (PIC, *Pré-Initiation Complex*) et reste associée avec la polymérase au cours de l'élongation. Enfin, les auteurs ont également démontré qu'en présence d'ATP, NUFIP favorise la transition du PIC vers un complexe d'élongation (Cabart *et al.*, 2004). L'ensemble de ces résultats suggère donc que la protéine NUFIP pourrait faciliter le recrutement et/ou l'intervention du complexe PTEF-b afin de promouvoir l'activation de la transcription dépendante du facteur BRCA1.

#### 4.6.6.3. Rsa1p et NUFIP, homologues fonctionnels ?

La recherche des homologues de la protéine NUFIP par alignements de séquence permis de mettre en évidence sa conservation chez de nombreux eucaryotes, à l'exception de certains organismes unicellulaires comme *S. cerevisiae*. Inversement, la protéine Rsa1p, et en particulier sa moitié C-terminale, présente une bonne conservation chez les levures (**Fig. 53A**), mais ne possède pas de réels homologues chez les autres organismes du *phylum* eucaryote. De façon inattendue, l'utilisation de l'algorithme de recherche PSI-BLAST a cependant révélé une conservation de séquence partielle entre la protéine Rsa1p de levure et la protéine humaine NUFIP (Boulon, 2005). Cette similitude se concentre essentiellement au niveau d'une région centrale de 37 résidus, baptisée domaine PEP et pour laquelle les deux

**A**



**B**



**Figure 53 : Conservation de séquence des protéines Rsa1p et NUFIP. (A)** Alignement de séquences des protéines Rsa1p retrouvées chez différentes espèces de levures hémiascomycètes. Les acides aminés identiques dans plus de 50% des séquences alignées sont représentés sur fond gris foncé. Les acides aminés similaires dans plus de 50% des séquences alignées sont représentés sur fond gris clair. *K. thermo.* : *K. thermotolerans*. **(B)** Alignement de séquences des domaines PEP retrouvés dans la protéine Rsa1p de *S. cerevisiae* et dans les protéines NUFIP animales. Seuls les acides aminés identiques ou similaires à 100% sont représentés respectivement sur fond gris foncé et gris clair. Ces alignements ont été réalisés par ClustalW Multiple Alignment à l'aide du logiciel BioEdit (Hall, T.A. 1999).

protéines partagent 34% d'identité et 60% de similarité de séquence (**Fig. 53B**). De plus, la bonne conservation de ce motif parmi les homologues des protéines Rsa1p et NUFIP, indique que l'ensemble de ces facteurs pourraient avoir une fonction conservée au cours de l'évolution. De façon intéressante, la mise en évidence de leur association respective avec les protéines Snu13p et 15.5K suggère que le motif PEP pourrait former un domaine d'interaction dédié à la protéine de la famille L7Ae. Etant donné le grand degré de conservation entre les protéines Snu13p et 15.5K (**Fig. 21B**), il est en effet probable que cette interaction repose sur des éléments structuraux communs. Le domaine PEP constituerait alors l'élément de base d'une nouvelle fonction, impliquant les protéines Rsa1p et NUFIP dans le métabolisme des snoRNP à boîtes C/D et/ou de la snRNP U4.



# OBJECTIFS





## OBJECTIFS DE LA THÈSE

Les RNP contenant des protéines de la famille L7Ae/L30 constituent un remarquable exemple d'évolution et de diversification des fonctions biologiques à l'échelle moléculaire. D'origine ancienne, vraisemblablement antérieures à la séparation des *phyla* archées/eucaryotes, ces particules ont semble-t-il dérivé du ribosome ancestral, pour acquérir de nouvelles potentialités liées au processus d'expression génique. Deux grandes classes de RNP de la famille L7Ae/L30 sont retrouvées chez les archées : les sRNP à boîtes C/D et H/ACA. Chez les eucaryotes, la famille L7Ae/L30 semble, en revanche, avoir fait l'objet d'une diversification accrue, puisqu'en plus des snoRNP à boîtes C/D et H/ACA, elle compte notamment la snRNP U4 spliceosomale, ainsi que le complexe télomérase et les mRNP SECIS chez les vertébrés.

Comme nous l'avons abordé dans la partie introductive, la comparaison structurale des protéines de la famille L7Ae/L30 révèle un grand degré de conservation, qui se matérialise notamment par un même mode d'interaction avec l'ARN. Sur cette base commune, la variation locale de certains déterminants permet néanmoins des ajustements fins et confère à chaque protéine son propre spectre de spécificité. Ainsi, les complexes formés entre les protéines de la famille L7Ae/L30 et leurs partenaires ARN constituent souvent des entités à part entière, présentant de nouvelles caractéristiques structurales et servant de socle à la formation de différentes classes de RNP. Dans le cas des particules à boîtes C/D et de la snRNP U4, le recrutement des autres protéines de la RNP implique notamment un module structural doué d'une spécificité de reconnaissance mixte protéine/ARN : le domaine NOP. Ce dernier, présent dans la protéine d'archée NOP5 et dans les protéines eucaryotes NOP56, NOP58 et PRP31, interagit de façon stable avec le complexe ARN-L7Ae/Snu13p/15.5K et contribue de ce fait à la hiérarchisation des événements d'assemblage (cf. Introduction, Chapitre 2). De plus, ce domaine favorise également la spécificité du processus, puisque la variation de certains déterminants entre les protéines NOP5, NOP56, NOP58 et la protéine PRP31 assure la discrimination entre les ARN à boîtes C/D et le snRNA U4.

L'analyse de ces caractéristiques intrinsèques suggère donc que les différents constituants des RNP pourraient contenir, dans leur structure, l'information nécessaire et suffisante pour assurer leur assemblage spécifique. Cette propriété a notamment été vérifiée chez les archées, puisque des sRNP à boîtes C/D et H/ACA actives *in vitro* peuvent être assemblées à partir de protéines recombinantes, et que de nombreuses structures tridimensionnelles de sRNP ont déjà été résolues. Chez les eucaryotes en revanche, la situation apparaît relativement plus compliquée, les approches de reconstitution ne sont toujours pas maîtrisées et relativement peu de données structurales sont actuellement disponibles. Comme nous l'avons déjà discuté, cette différence notoire entre les deux systèmes, repose probablement sur la complexification des cellules eucaryotes. Au cours de l'évolution, ces dernières ont en effet subi d'importants changements d'ordre structurel. Elles ont notamment fait l'objet d'une large augmentation de volume, couplée à une compartimentalisation de l'espace cellulaire. Sous l'effet de l'accroissement du protéome et du transcriptome, la densité moléculaire a également augmentée chez ces organismes. Dans un tel contexte, la mise en place des machineries moléculaires fondamentales, en particulier celles impliquées dans le processus d'expression des gènes, est devenue dépendante de facteurs spécialisés contrôlant l'efficacité et la spécificité des processus de biogenèse.

Ainsi, la mise à jour des événements de production des RNP de la famille L7Ae/L30 a révélé l'existence de nombreuses étapes, se déroulant de façon séquentielle et ordonnée, à travers plusieurs compartiments cellulaires. Les ARN, transcrits par l'ARN polymérase II, font l'objet d'événements de maturation nécessitant l'intervention d'une cohorte de facteurs protéiques. Parmi eux, certains sont directement impliqués dans les processus de clivages et de modifications, d'autres assurent en revanche le découplage des événements de maturation (e.g. découplage clivage/polyadénylation), afin d'assurer un processus spécifique pour chaque classe d'ARN. Les protéines, qui font également l'objet de modifications post-traductionnelles, nécessitent quant à elles une prise en charge par des facteurs de chaperonnage et de transport, avant d'être assemblées sur les ARN. Enfin, l'ensemble de ces facteurs et les constituants des particules matures, établissent de nombreuses interactions, qui contribuent à assurer une coordination étroite entre les processus de maturation des ARN, d'assemblage des particules et de transport.

Parmi ces différentes étapes, les mécanismes moléculaires régissant l'assemblage des RNP demeurent relativement peu connus et leur caractérisation constitue aujourd'hui un objectif

majeur. A ce titre, et comme nous l'avons abordé dans l'introduction, les processus d'assemblage des UsnRNP et des snoRNP à boîtes H/ACA, ont fait l'objet d'attentions particulières au cours des dernières années. Leur étude a notamment permis de mettre à jour deux systèmes modèles : (i) le complexe SMN qui contrôle la spécificité d'assemblage des UsnRNP spliceosomales, (ii) le système SHQ1/NAF1, qui procède par mimétisme moléculaire, afin d'assurer le caractère séquentiel de l'assemblage des snoRNP à boîtes H/ACA.

A mon arrivée au laboratoire, B. Charpentier avait débuté la caractérisation de la protéine Rsa1p, récemment identifiée comme partenaire de la snoRNP U3 et de la protéine Snu13p chez *S. cerevisiae*. L'hypothèse formulée alors, était que le domaine PEP, partagé par Rsa1p et la protéine humaine NUFIP, constitue un domaine d'interaction dédié à la protéine Snu13p. En premier lieu, nous nous sommes donc attachés à vérifier cette hypothèse en étudiant l'interaction entre les deux partenaires *in vitro*. D'autre part, aucun modèle concernant les mécanismes d'assemblage et de contrôle qualité des snoRNP à boîtes C/D n'avait encore été proposé. Seuls quelques facteurs, comme les hélicases Rvb1p/TIP49 et Rvb2p/TIP48, ou les protéines Pih1p, Cwc24p et Bcd1p, identifiés au cours d'études indépendantes, étaient supposés pouvoir jouer un rôle dans ces phénomènes. Comme cela avait été réalisé auparavant pour les UsnRNP et pour les snoRNP à boîtes H/ACA, nous avons donc entrepris la caractérisation de l'infrastructure protéique impliquée dans les mécanismes d'assemblage des snoRNP à boîtes C/D chez *S. cerevisiae*. Afin de mener à bien ce travail, nous avons défini quatre grandes questions qui nous ont guidés dans nos expérimentations et dans notre réflexion :

- (i) Quels liens physiques s'établissent entre les facteurs d'assemblage présumés et les constituants des snoRNP ?
- (ii) Existe-t-il des liens physiques et fonctionnels entre les différents facteurs d'assemblage présumés ?
- (iii) Quelle est l'importance individuelle et collective de ces facteurs pour la biogenèse des snoRNP et plus particulièrement pour les étapes d'assemblage ?
- (iv) Etant donné les nombreuses homologies relevées entre les différentes RNP de la famille L7Ae, existe-t-il des mécanismes d'assemblage communs, en particulier dans le but d'assurer un contrôle de la spécificité ?

---

Au cours de notre démarche, nous avons essentiellement focalisé notre attention sur le nouveau facteur Rsa1p, sur la protéine Bcd1p, sur les membres du complexe R2TP et, par extension, sur la chaperonne Hsp82p. L'utilisation de l'organisme modèle *S. cerevisiae* comme système d'étude, nous a notamment permis de coupler les approches de biochimie et de biologie moléculaire à des techniques de génétique de la levure et d'analyses globales :

1- Nous avons notamment réalisé des cribles d'interaction entre les différentes protéines d'intérêt, en utilisant les approches complémentaires de double-hybride, de pull-down et de co-expression chez *E. coli*. Pour la majorité des nouveaux facteurs étudiés, nous avons entrepris la cartographie de leurs domaines d'interaction et, dans certains cas, des acides aminés impliqués. Cette démarche a été particulièrement utilisée dans le cadre de l'étude de l'interaction entre Rsa1p et la protéine Snu13p, puisqu'elle a notamment abouti à la proposition d'un modèle d'interaction entre les deux partenaires.

2- En plus des interactions entre protéines, nous avons également étudié la capacité des différents facteurs à interagir de façon directe ou indirecte avec les snoRNA. Des expériences ont ainsi été réalisées *in vivo*, par des approches de co-immunoprécipitation et de triple-hybride, et *in vitro*, par des approches de retard sur gel, de pull-down et d'empreintes.

3- L'importance des différents facteurs pour la biogenèse des snoRNP à boîtes C/D a également été étudiée *in vivo*. Pour cela nous avons entrepris la caractérisation phénotypique des souches invalidées pour les gènes d'intérêt. Les effets associés à ces délétions ont notamment été vérifiés par des tests de croissance, par des expériences de Northern-blot et par des tests de Western-blot. Nous avons ainsi pu constater l'importance individuelle de chaque facteur pour la production des snoRNA, pour la biogenèse des ARNr et pour la stabilité des différentes protéines d'intérêt. Le rôle collectif de ces facteurs a également été testé par la création et la caractérisation de souches de levures doublement invalidées. Enfin, l'utilisation de ces souches de levure modifiées nous a également permis de tester les capacités de complémentations des domaines ou des variants d'interaction générés pour les différents facteurs.

4- Afin de mieux comprendre la mécanistique du processus d'assemblage, nous avons cherché à reproduire la formation de snoRNP dans différents contextes, cellulaires ou acellulaires, en présence ou en absence des différents facteurs. Nous avons notamment utilisé

des approches de triple-hybride, de double-hybride ou de co-immunoprécipitation dans des souches invalidées pour certains gènes d'intérêt. Nous avons également utilisés des approches *in vitro*, basées sur les techniques de pull-down et de retard sur-gel, en utilisant des protéines recombinantes.

5- Enfin, outre les démarches ciblées, citées ci-dessus, nous avons également cherché à valider nos résultats par des approches plus globales. En utilisant l'approche TAP-tag, couplée à une analyse systématique par spectrométrie de masse, nous avons ainsi entrepris la purification et la caractérisation des complexes macromoléculaires contenant les protéines d'intérêt. Cette démarche nous a notamment permis d'étudier l'association des différents facteurs *in vivo*, mais également d'identifier de nouvelles protéines d'intérêt. Enfin, la réalisation de cribles d'interactions génétiques, nous a permis d'établir pour certains facteurs, une cartographie exhaustive des liens fonctionnels établis avec de nouveaux partenaires protéiques.

L'ensemble de ce travail a été réalisé en parallèle de la thèse de Régis Back, dont l'objectif était d'obtenir des données structurales sur les composants des snoRNP et sur les facteurs requis pour leur biogenèse. Au cours de notre démarche, nous avons donc contribué à la validation biologique des données structurales obtenues par RMN, tout en bénéficiant de leur apport précieux pour la compréhension des mécanismes biologiques étudiés. De plus, notre travail a également été mené dans le cadre de plusieurs collaborations : avec les équipes d'E. Bertrand (IGM, Montpellier), d'A. Krol (IBMC, Strasbourg) et de S. Moréra (LEBS, Gif-sur-Yvette) au sujet de l'étude des facteurs d'assemblage, avec l'équipe d'A. Van Dorssealer (IPHC, Strasbourg) pour les analyses par spectrométrie de masse, et avec C. Saveanu (Institut Pasteur, Paris) pour la réalisation des cribles d'interactions génétiques.



# MATERIEL & METHODES





# MATERIEL & METHODES

## 1. Matériel

### 1.1. Les souches de cellules

#### 1.1.1. Les souches de la bactérie *Escherichia coli*

- Pour l'amplification et l'extraction de plasmides :

##### *E. coli* DH5a

(*supE44*,  $\Delta$ *lacU169* ( $\Phi$ 80 *lacZ* $\Delta$ M15), *hsdR17*, *recA1*, *endA1*, *gyrA96*, *thi-1*, *relA1*)

- Pour la surproduction de protéines :

##### *E. coli* BL21-Codon Plus (DE3)

(F<sup>-</sup> *ompT hsdS* (rb<sup>-</sup>mb<sup>-</sup>) *dcm* + Tet<sup>r</sup> *gal*  $\lambda$  (DE3) *endA Hte* [*argU ileY leuW Cam<sup>r</sup>*])

#### 1.1.2. Les souches de la levure *Saccharomyces cerevisiae*

- Pour les études de caractérisation phénotypique :

*S. cerevisiae* **BY4741** (*Mat a*) et **BY4742** (*Mat  $\alpha$* ) (*his3* $\Delta$ 1; *leu2* $\Delta$ 0; *met15* $\Delta$ 0; *ura3* $\Delta$ 0) représentent les souches parentales, de *mating-type* opposé, à partir desquelles des invalidations de gènes ont été effectuées par recombinaison homologue, pour générer les souches suivantes (Euroscarf):

BY4741 $\Delta$ RSA1 et BY4742 $\Delta$ RSA1: YPL193w::kanMX4;

BY4741 $\Delta$ HIT1 et BY4742 $\Delta$ HIT1: YJR055w::kanMX4;

BY4741 $\Delta$ PIH1 et BY4742 $\Delta$ PIH1: YHR034c::kanMX4;

BY4741 $\Delta$ TAH1 et BY4742 $\Delta$ TAH1: YCR060w::kanMX4;

By4741 $\Delta$ RTT106 : YCR060w::kanMX4.

*S. cerevisiae* **W303-GAL::SNU13** (*Mat a*; *trp1*- $\Delta$ 63; *his3*- $\Delta$ 200; *ura3*-52; *lys2*-801; *ade2*-101; *leu2*- $\Delta$ 1; *HIS3*::*GAL1*::YEL026w) ((Watkins *et al.*, 2000); aimablement fournie par P. Fabrizio)

*S. cerevisiae* **JG540-GAL::NHP2** (*Mat a*, *ade2*, *ade3*, *leu2*, *lys2.1*, *trp1.1*, *tyr7.1*, *ura3*, *can1*; *URA3*::*GAL1*::YDL208w) ((Henras *et al.*, 1998); aimablement fournie par Y. Henry)

*S. cerevisiae* **tetO<sub>7</sub>::BCD1** (*Mat a*; pYHR040w::kanR-tet07-TATA; *URA3*::CMV-tTA ; *his3*-1 ; *leu2*-0 ; *met15*-0) ((Peng *et al.*, 2003); aimablement fournie par T.R. Hughes)

*S. cerevisiae* **Y3068** (*MAT  $\alpha$* ; *can1* $\Delta$ ::*MFA1pr-HIS3 his3* $\Delta$ 1 *leu2* $\Delta$ 0 *ura3* $\Delta$ 0 *met15* $\Delta$ 0 *lys2* $\Delta$ 0) ((Zhao *et al.*, 2005); aimablement fournie par W.A. Houry)

*S. cerevisiae*  $\Delta$ HSP82 +TAP-HSC82 (R0009) (*MAT*  $\alpha$ ; *can1* $\Delta$ ::*MFA1pr-HIS3 his3* $\Delta$ 1 *leu2* $\Delta$ 0 *ura3* $\Delta$ 0 *met15* $\Delta$ 0 *lys2* $\Delta$ 0 *hsp82* $\Delta$ ::*LEU2*; TAP-HSC82) (construite à partir de la souche Y3068 (Zhao *et al.*, 2005); aimablement fournie par W.A. Houry)

- Pour les tests de double-hybride :

*S. cerevisiae* Y187 (*Mat*  $\alpha$ ) et Y190 (*Mat*  $a$ )

(*gal4* $\Delta$  *gal80* $\Delta$  *ade2-101 his3-200 leu2-3,112 lys2-801 trp1-901 ura3-52*, *URA3*::*Gal1UAS GAL1TATA-LacZ*, *LYS2*::*GAL1UASHisTATA-HIS3*)

Ces souches ont été invalidées par recombinaison homologue pour le gène *RSAl* :

Y187- $\Delta$ *RSAl* et Y190- $\Delta$ *RSAl* : YPL193w::kanMX6 ((Boulon *et al.*, 2008) ; construites par B. Charpentier).

*S. cerevisiae* CG1945 (*Mat*  $a$ )

(*ura3-52, his3-200, ade2-101, lys2-801, trp1-901, leu2-3, 112, gal4-542, gal80-538, cyh*<sup>r2</sup>, *LYS2*::*GAL1UAS-GAL1TATA-HIS3, URA3*::*GAL4*<sub>17-mers(x3)</sub>-*CYC1TATA-LacZ*)

Cette souche est utilisée à la place de la souche Y190 afin de détecter des interactions par double-hybride, sans utilisation de 3-AT. Le système rapporteur de cette souche génère en effet moins de bruit de fond (expression basale du gène rapporteur en absence du facteur Gal4) que celui de la souche Y190. Son utilisation est privilégiée pour détecter les interactions de faible intensité.

- Pour les tests de triple-hybride :

*S. cerevisiae* R40-coat (*Mat*  $\alpha$ ) et L40-coat (*Mat*  $a$ )

(*ura3-52, leu2-3, 112, his3D200, trp1D1, ade2, LYS2*::(*LexAop*)4-*HIS3, ura3*::(*LexAop*)8-*lacZ, LexA-MS2 coat (TRP1)*)

Ces souches ont été invalidées par recombinaison homologue pour le gène *RSAl* :

R40-coat- $\Delta$ *RSAl* et L40-coat- $\Delta$ *RSAl* : YPL193w::kanMX6 ((Boulon *et al.*, 2008) ; construites par B. Charpentier).

- Pour les expériences d'immunosélection :

*S. cerevisiae* TAP-tag (*Mat*  $a$ )

Toutes ces souches partagent le génotype suivant : *ade2; arg4; leu2-3,112; trp1-289; ura3-52*. Les différentes souches, commercialisées par Euroscarf expriment une protéine étiquetée à son extrémité C-terminale par l'étiquette TAP. Ces souches génétiquement modifiées ont été obtenues après recombinaison au *locus* codant la protéine d'une construction portant le marqueur d'auxotrophie *URA3* :

TAP-*RSAl* : *URA3*::TAP-YPL193w;

TAP-*HIT1* : *URA3*::TAP-YJR055w;

TAP-*PIH1* : *URA3*::TAP-YHR034c;

TAP-*TAH1* : *URA3*::TAP-YCR060w;

TAP-*RVB1* : *URA3*::TAP- YDR190c;

TAP-*RVB2* : *URA3*::TAP-YPL235w;

TAP-*BCD1* : *URA3*::TAP-YHR040w;

TAP-*NOP1* : *URA3*::TAP-YDL014w;

TAP-*NOP56* : *URA3*::TAP-YLR197w;

TAP-*NOP58* : *URA3*::TAP-YOR310c;

TAP-*NAF1* : *URA3*::TAP-YNL124w.

## 1.2. Les plasmides

### Vecteurs bactériens

**Vecteur pCR2.1** (Invitrogen). Le clonage de fragments d'ADN obtenus par amplification par la technique de PCR a été réalisé dans le plasmide pCR2.1 fourni avec le kit *PCR T/A cloning*. Le plasmide est commercialisé sous forme linéarisée et porte un résidu thymine protubérant sur chaque extrémité 3'. Il contient des gènes de résistance à l'ampicilline et à la kanamycine. Il possède également le gène *lacZ* codant la  $\beta$ -galactosidase, placé sous le contrôle du promoteur *Plac* et interrompu par la cassette de clonage. Ce plasmide a été utilisé pour cloner et séquencer certaines ORF isolées à partir d'ADN génomique, ainsi que pour vérifier la nature des produits amplifiés par RT-PCR.

**Vecteur pUC18** (Invitrogen). Ce vecteur a été utilisé pour cloner certaines séquences codant des snoRNA et pour générer, par PCR, des matrices ADN de transcription *in vitro*. Il contient un gène de résistance à l'ampicilline et l'origine de répllication du vecteur pMB1. Il possède également le gène *lacZ* codant la  $\beta$ -galactosidase, placé sous le contrôle du promoteur *Plac* et interrompu par la cassette de clonage.

**Vecteur pGEX-6P-1** (GE Healthcare). Ce vecteur a été utilisé pour la surproduction et la purification de protéines recombinantes chez *E. coli*. Il permet l'expression, par induction à l'IPTG, de protéines fusionnées à l'étiquette GST au niveau de leur extrémité N-terminale. Il contient un gène de résistance à l'ampicilline, le gène *lacI*, l'origine de répllication du vecteur pBR322 et un promoteur Ptac en amont de l'ORF de la GST. Cette dernière est séparée de la cassette de clonage par une séquence codant le site de clivage par la protéase PreScission<sup>TM</sup>. Les clonages dans ce vecteur ont été réalisés entre les sites de restriction *SmaI* et *Sall*.

**Vecteur pET15b** (Novagen). Ce vecteur a été utilisé pour la surproduction et la purification de protéines recombinantes chez *E. coli*. Il permet l'expression, par induction à l'IPTG, de protéines fusionnées une étiquette 6xHis au niveau de leur extrémité N-terminale. Il contient un gène de résistance à l'ampicilline, le gène *lacI*, l'origine de répllication du vecteur pBR322 et un promoteur T7 en amont de la cassette de clonage.

**Vecteur pET28a** (Novagen). Ce vecteur a été utilisé pour la surproduction et la purification de protéines recombinantes chez *E. coli*. Il permet l'expression, par induction à l'IPTG, de protéines fusionnées à une étiquette 6xHis au niveau de leur extrémité N-terminale. Il contient un gène de résistance à la kanamycine, le gène *lacI*, l'origine de répllication du vecteur pBR322 et un promoteur T7 en amont de la cassette de clonage.

Pour ces deux vecteurs, l'étiquette 6xHis peut être éliminée par clivage au site spécifiquement reconnu par la protéase Thrombine, codé en amont de la cassette de clonage.

**Vecteurs pnCS et pnEA-3CH** (Diebold *et al*, 2011 ; aimablement fournis par C. Romier). Ces deux vecteurs ont été utilisés pour les expériences de co-expression et pour la production de protéines radiomarquées lors de synthèse par traduction *in vitro* en extrait cellulaire S30 d'*E. coli*. Le Vecteur pnCS permet l'expression, par induction à l'IPTG, d'une protéine non étiquetée. Construit à partir du vecteur pCDF-DuET, il contient un gène de résistance à la spectinomycine, le gène *lacI*, et un promoteur T7. Les clonages dans ce vecteur ont été réalisés entre les sites de restriction *NdeI* et *BamHI*. Le vecteur pnEA-3CH permet l'expression, par induction à l'IPTG, de protéines fusionnées à une étiquette 6xHis au niveau

de leur extrémité N-terminale. Construit à partir du vecteur pet15b, il contient un gène de résistance à l'ampicilline, le gène *lacI* et un promoteur T7. Originellement pourvu d'une séquence codant un site de coupure à la protéase Thrombine entre l'étiquette et l'ORF d'intérêt (*i.e.* pnEA-3tH), ce vecteur a été modifié au laboratoire par R. Back, afin de faire apparaître un site de coupure à la protéase PreScission<sup>TM</sup>. Les clonages dans ce vecteur ont été réalisés entre les sites de restriction *NdeI* et *BamHI*.

### Vecteurs navettes *E. coli* – *S. cerevisiae*

Les plasmides décrits ci-dessous possèdent l'origine de répllication du plasmide pUC (dérivant de colE1 de pBR322) permettant la répllication chez *E. coli*, et une origine soit 2 $\mu$  ou soit CEN/ARS permettant leur répllication dans le noyau des levures.

**Vecteur pG1** (Addgene). Ce vecteur a été utilisé pour l'expression ectopique de différentes protéines dans les cellules de levure. Il possède l'origine CEN/ARS, permettant sa répllication en copie unique chez la levure. Il contient un gène de résistance à l'ampicilline et le marqueur *TRP1*. La cassette de clonage est située en aval du promoteur du gène *GLD1* (alias *GPD*) codant la GAPDH. Les clonages dans ce vecteur ont été réalisés entre les sites de restriction *BamHI* et *Sall*.

**Vecteur p413-TEF** (ATCC). Ce vecteur a été utilisé pour l'expression ectopique de différentes protéines dans les cellules de levure. Il possède l'origine de répllication CEN/ARS. Il contient un gène de résistance à l'ampicilline et le marqueur d'auxotrophie *HIS3*. La cassette de clonage est située en aval du promoteur TEF. Les clonages dans ce vecteur ont été réalisés entre les sites de restriction *BamHI* et *XhoI*.

**Vecteur pASZ11** (Stotz & Linder, 1990). Ce vecteur a été utilisé pour l'expression ectopique de l'ARN U3 $\Delta$ 2,3,4 dans les cellules de levure et pour générer, par PCR, des matrices ADN de transcription *in vitro*. Il possède l'origine CEN/ARS. Il contient un gène de résistance à l'ampicilline et le marqueur d'auxotrophie *ADE2*. Dans le vecteur pASZ11::U3 $\Delta$ 2,3,4 recombinant, la séquence codant l'ARN U3 $\Delta$ 2,3,4 (dépourvue de la séquence intronique) est placée sous contrôle du promoteur naturel du gène *SNR17A* (Marmier-Gourrier *et al.*, 2003; Mereau *et al.*, 1997; Segault *et al.*, 1992).

**Vecteur p422** (aimablement fourni par E. Bertrand). Ce vecteur a été utilisé pour l'expression ectopique des protéines NUFIP et NUFIP- $\Delta$ PEP dans les cellules de levure, au cours des expériences de triple-hybride *bridged*. Il possède l'origine de répllication 2 $\mu$  permettant sa répllication en copies multiples chez la levure. La cassette de clonage est située en aval du promoteur TEF. Il contient un marqueur d'auxotrophie *ADE2*.

**Vecteur pGFP** (aimablement fourni par E. Bertrand). Ce vecteur a été utilisé pour l'expression ectopique d'une protéine de fusion GFP-Rsa1p dans les cellules de levure. Il contient le marqueur d'auxotrophie *URA3*.

**Vecteur pACT2** (Clontech). Ce vecteur a été utilisé pour les tests de double-hybride et de triple-hybride. Il possède l'origine de répllication 2 $\mu$ , un gène de résistance à l'ampicilline et le marqueur d'auxotrophie *LEU2*. Il permet l'expression d'une protéine d'intérêt fusionnée, à son extrémité N-terminale, au domaine d'activation de la transcription du facteur Gal4 (Gal4-

AD). Cette expression est sous le contrôle du promoteur *ADHI*. Les clonages dans ce vecteur ont été réalisés entre les sites de restriction *BamHI* et *XhoI*.

**Vecteur pGBKT7** (Clontech). Ce vecteur a été utilisé pour les tests de double-hybride. Il possède l'origine  $2\mu$  permettant sa répllication en copies multiples chez la levure. Il contient un gène de résistance à la kanamycine et le marqueur d'auxotrophie *TRP1*. Il permet l'expression de la protéine d'intérêt fusionnée, à son extrémité N-terminale, au domaine de fixation à l'ADN du facteur Gal4 (Gal4-BD). Cette expression est sous contrôle du promoteur *ADHI*. Les clonages dans ce vecteur ont été réalisés entre les sites de restriction *BamHI* et *SalI*.

**Vecteur pIIIAMS2.2** (Université du Wisconsin). Ce vecteur a été utilisé pour les tests de triple-hybride et pour générer, par PCR, des matrices ADN de transcription *in vitro*. Il possède l'origine  $2\mu$  permettant sa répllication en copies multiples chez la levure. Il contient un gène de résistance à l'ampicilline et les marqueurs d'auxotrophie *URA3* et *ADE2*. Il permet l'expression de l'ARN d'intérêt fusionné, au niveau de son extrémité 5', à un double motif tige-boucle de l'ARN du phage MS2. Cette expression est sous le contrôle d'un promoteur spécifique de l'ARN polymérase III. Les clonages dans ce vecteur ont été réalisés entre les sites de restriction *SmaI* et *SphI*.

Les principaux plasmides recombinants utilisés au cours des différentes études sont listés dans le **Tableau 3**.

### 1.3. Les oligonucléotides

Les principaux oligonucléotides utilisés sont décrits dans le **Tableau 4**.

### 1.4. Les milieux de culture

Les différents milieux de culture sont autoclavés à 121°C, sous 1 bar de pression, pendant 20 minutes. Pour la préparation des milieux solides, de l'agar est additionné à raison de 20 g/L avant le passage à l'autoclave.

#### 1.4.1. Cultures de bactéries

**Milieu LB (Luria Bertani)** : peptone pancréatique animale (5 g/l); peptone tryptique de caséine (5 g/l); extrait de levure (5 g/l); NaCl (10 g/l). Le pH est ajusté à 7,5 avec une solution de NaOH 10 N. Les antibiotiques sont ajoutés après l'étape de stérilisation, à une concentration finale de 100  $\mu\text{g/mL}$  pour l'ampicilline, 50  $\mu\text{g/mL}$  pour la kanamycine, 30  $\mu\text{g/mL}$  pour la spectinomycine et 25  $\mu\text{g/mL}$  pour le chloramphénicol.

#### 1.4.2. Cultures de levures

**Milieu riche YPG** : extrait de levure (10 g/L) ; bactopectone (20 g/l) ; galactose (20 g/l).

**Milieu riche YPD** : extrait de levure (10 g/L) ; bactopectone (20 g/l) ; glucose (20 g/l).

Vecteur	ORF/Séquence	Provenance/Référence	
Vecteurs bactériens	<b>pUC18</b>	U14 ( <i>SNR128</i> )	Laboratoire AREMS (B. Charpentier)
		U3Δ2,3,4 WT/mut	(Marmier-Gourrier <i>et al.</i> , 2003)
		miniU3B/C	(Marmier-Gourrier <i>et al.</i> , 2003)
	<b>pGEX-6P1</b>	<i>SNU13</i> WT/mut	(Boulon <i>et al.</i> , 2008)
		<i>RSA1</i>	(Boulon <i>et al.</i> , 2008)
		<i>RSA1</i> <sub>230-381</sub>	(Boulon <i>et al.</i> , 2008)
	<b>pET15b</b>	<i>BCD1</i>	fourni par E. Bertrand
		<i>BCD1</i> -ZF	Laboratoire AREMS (R. Back)
		<i>BCD1</i> -ΔZF	Laboratoire AREMS (R. Back)
	<b>pET28a</b>	yPEP	(Boulon <i>et al.</i> , 2008)
	<b>pnCS</b>	<i>NOP58</i>	Laboratoire AREMS (R. Back)
		<i>NOP58</i> <sub>1-377</sub>	Laboratoire AREMS (R. Back)
		<i>RSA1</i> WT/mut	(Rothé, Back <i>et al.</i> , Publication n°2)
		<i>RSA1</i> <sub>230-381</sub> WT/mut	(Rothé, Back <i>et al.</i> , Publication n°2)
		<i>RSA1</i> <sub>230-375</sub> WT/mut	(Rothé, Back <i>et al.</i> , Publication n°3)
		<i>HIT1</i>	Laboratoire AREMS (R. Back)
		<i>PIH1</i>	Laboratoire AREMS (R. Back)
	<b>pnEA-3CH</b>	<i>SNU13</i> WT/mut	(Rothé, Back <i>et al.</i> , Publication n°2)
		<i>RSA1</i>	(Rothé <i>et al.</i> , Publication n°3)
		<i>RSA1</i> <sub>230-375</sub>	(Rothé <i>et al.</i> , Publication n°3)
	<i>TAH1</i>	Laboratoire AREMS (R. Back)	
Vecteurs navettes	<b>pG1</b>	<i>SNU13</i> WT/mut	(Rothé, Back <i>et al.</i> , Publication n°2)
		<i>RSA1</i> WT/mut	(Rothé, Back <i>et al.</i> , Publication n°2)
		<i>RSA1</i> <sub>230-381</sub> WT/mut	(Rothé, Back <i>et al.</i> , Publication n°2)
		<i>NHP2</i> WT/mut	Laboratoire AREMS (B. Rothé)
		<i>PIH1</i> WT/fragments	Laboratoire AREMS (B. Rothé)
	<b>p413-TEF</b>	flag- <i>NOPI</i>	Laboratoire AREMS (V. Senty-Ségault)
		flag- <i>NOP56</i>	Laboratoire AREMS (V. Senty-Ségault)
		flag- <i>NOP58</i>	Laboratoire AREMS (V. Senty-Ségault)
		<i>BCD1</i> WT/fragments	Laboratoire AREMS (D. Tiotiu)
		<i>HIT1</i> WT/fragments	Laboratoire AREMS (D. Tiotiu)
		<i>TAH1</i> WT/fragments	Laboratoire AREMS (B. Rothé)
	<b>pASZ11</b>	U3 ( <i>SNR17A</i> )	(Marmier-Gourrier <i>et al.</i> , 2003)
		U3Δ2,3,5	(Marmier-Gourrier <i>et al.</i> , 2003)
	<b>p422</b>	<i>NUFIP</i>	(Boulon <i>et al.</i> , 2008)
		<i>NUFIP</i> ΔPEP	(Boulon <i>et al.</i> , 2008)
	<b>pGFP</b>	<i>RSA1</i>	fourni par E. Bertrand
	<b>pACTII</b>	<i>RSA1</i> WT/mut	(Rothé, Back <i>et al.</i> , Publication n°2)
		<i>RSA1</i> <sub>230-381</sub> WT/mut	(Rothé, Back <i>et al.</i> , Publication n°2)
		<i>PIH1</i> WT/fragments	(Back <i>et al.</i> , Publication n°4)
		<i>TAH1</i> WT/fragments	(Back <i>et al.</i> , Publication n°4)
		<i>HSP82</i> WT/mut	(Back <i>et al.</i> , Publication n°4)
		<i>SNU13</i>	(Boulon <i>et al.</i> , 2008)
		<i>NOPI</i>	(Boulon <i>et al.</i> , 2008)
		<i>NOP56</i>	(Boulon <i>et al.</i> , 2008)
		<i>NOP58</i>	(Boulon <i>et al.</i> , 2008)
		<i>RRP9</i>	Laboratoire AREMS (V. Senty-Ségault)
		<i>RTT106</i>	Laboratoire AREMS (D. Tiotiu)
	<b>pGBKT7</b>	<i>SNU13</i> WT/mut	(Rothé, Back <i>et al.</i> , Publication n°2)
		<i>HIT1</i>	(Rothé <i>et al.</i> , Publication n°3)
		<i>PIH1</i> FL/fragments	(Back <i>et al.</i> , Publication n°4)
		<i>TAH1</i> FL/fragments	(Back <i>et al.</i> , Publication n°4)
		<i>TAH1</i> WT/mut	(Back <i>et al.</i> , Publication n°4)
		<i>NOPI</i>	(Boulon <i>et al.</i> , 2008)
		<i>NOP56</i>	(Boulon <i>et al.</i> , 2008)
		<i>NOP58</i>	(Boulon <i>et al.</i> , 2008)
		<i>RRP9</i>	Laboratoire AREMS (V. Senty-Ségault)
		<i>BCD1</i>	Laboratoire AREMS (D. Tiotiu)
	<b>pIIAMS2.2</b>	U3BC	(Boulon <i>et al.</i> , 2008)
		U4/U6	Laboratoire AREMS (A. Mougin)

Tableau 3 : Plasmides recombinants utilisés au cours de la thèse.

N°	Séquence 5'-3'	Description
6687	GATACTCTGCTCTGTGCTATCGTC	ChIP Intron EFB1-snoRNA C/D U18 (3')
6686	CCTATGGTTACATGTAAGGGAAG	ChIP Intron EFB1-snoRNA C/D U18 (5')
6685	CGGTTGGACTCTTCACTGCTGAAA	ChIP Promoteur EFB1 (3')
6684	AACCTGCAGGTGAAGAAGCGCTTT	ChIP Promoteur EFB1 (5')
5260	GTATTTGCGCTCACGTTGTC	ChIP snoRNA C/D snR45 (3')
5259	TTGTGTCCTTTAAACCGTGC	ChIP snoRNA C/D snR45 (5')
5255	GCTTTGCCGTTGCAATTTGT	ChIP snoRNA C/D U3-exon 2 (5') (utilisé avec 3'-41)
5266	GGGTCAAGCTTTGCAATTCT	ChIP snoRNA H/ACA snR32 (3')
5265	GGTTGCCAACGCGAAGC	ChIP snoRNA H/ACA snR32 (5')
7731	CCCGGGATCCCGATGGCGGTGTTGTGG	Clonage Bcd1p (1-62) +BamHI (5')
7732	ACCGGTCGACCGGTACTTTCGTCGTCGCGC	Clonage Bcd1p (1-62) +Sall/STOP (3')
7736	ACCGGTCGACCGGTACACCACTTTTTTTGAATCCA	Clonage Bcd1p (169-260) +Sall/STOP (3')
7735	CCCGGGATCCCGATGCCTATGCAGGAAAAGGGT	Clonage Bcd1p (169-260) +START/BamHI (5')
7738	ACCGGTCGACCGGTGAGCGGGTTCCTGCTGT	Clonage Bcd1p (261-305) +Sall/STOP (3')
7737	CCCGGGATCCCGATGGAGTTGGCTATACATGAGA	Clonage Bcd1p (261-305) +START/BamHI (5')
7740	ACCGGTCGACCGGTATGCAGTGAGGAAATCC	Clonage Bcd1p (306-366) +Sall (3')
7739	CCCGGGATCCCGATGCCACTTGAGCATACGAGT	Clonage Bcd1p (306-366) +START/BamHI (5')
7734	ACCGGTCGACCGGTCAACATAGAATCCACTCCAG	Clonage Bcd1p (63-168) +Sall/STOP (3')
7733	CCCGGGATCCCGATGCACGAAACAAATGCTTATG	Clonage Bcd1p (63-168) +START/BamHI (5')
8290	GTTTCGTACAAGTATTCCTGTACTAGGATTCCTCATATA	Clonage Bcd1p avec ZF-Hit1p (3')
8289	CAAAGCATGTTTCAAGGAACATGATCCTAAGGAGTATAT	Clonage Bcd1p avec ZF-Hitp (5')
8283	CCCGGGATCCCGATGATCTAGTGCAGTTAA	Clonage Hit1p (1-43) +BamHI (5')
8284	ACCGGTCGACCGGTATTCTTATGAACATGCTTTG	Clonage Hit1p (1-43) +Sall/STOP (3')
8285	CCCGGGATCCCGATGAGCGAACAACCCAGGG	Clonage Hit1p (44-69) +BamHI (5')
8287	CCCGGGATCCCGATGAATAAACATTAAGACG	Clonage Hit1p (70-164) +BamHI (5')
8288	ACCGGTCGACCGGTTATTTCTCACCGCATTTAA	Clonage Hit1p (70-164) +Sall/STOP (3')
8286	ACCGGTCGACCGGTATGCGAGTGAAGTGTATG	Clonage Hit1p (74-69) +Sall/STOP (3')
8292	GTTAACGAGTCCAGTTTGTTCGCTGTTGGGTCC	Clonage Hit1p avec ZF-Bcd1p (3')
8291	CAATTGCTCAGGTCAAACAAGCGAACAACCCAGGG	Clonage Hit1p avec ZF-Bcd1p (5')
7562	CCCGGGATCCCGATGGCTAGTGAACCTTTTGAA	Clonage Hsp82p +BamHI (5')
7563	CCGCTCGAGCGCTAATCTACCTCTTCCATTT	Clonage Hsp82p +XhoI (3')
6964	GGATCCGATGGGTAAGACAAACAAGGAAC	Clonage Nhp2p +BamHI (5')
6488	GGATCCATGGGTAAGACAACAAGGAAC	Clonage Nhp2p +BamHI (5')
6489	CTCGAGTCATAAAGCTTGAACCTCTTTG	Clonage Nhp2p +Sall (3')
7308	CCGCTCGAGCGGTTACGAATCTCGGATCTCACA	Clonage Pih1p (1-148) +STOP/XhoI (3')
7309	CCGCTCGAGCGGTTAATTATCATCGCCATCGTC	Clonage Pih1p (1-211) +STOP/XhoI (3')
7310	CCGCTCGAGCGGTTATGCAAGTGCCTGTTTTCAG	Clonage Pih1p (1-252) +STOP/XhoI (3')
7558	CCCGGGATCCCGATGAATGATCCCGACGATG	Clonage Pih1p (212-344) +BamHI/START (5')
7559	CCCGGGATCCCGATGCCTGCACCTGCACCC	Clonage Pih1p (252-344) +BamHI/START (5')
7560	CCCGGGATCCCGATGGAGTGCAGGAAATAC	Clonage Pih1p (263-344) +BamHI/START (5')
7307	CCCGGGATCCCGATGGCCGATTTCTTATTG	Clonage Pih1p +BamHI (5')
7561	CCGCTCGAGCGGTTATATATATATATAGTGTGC	Clonage Pih1p +XhoI (3')
7949	CCCGGGATCCCGATGCAAAATTGTTCTTAGAT	Clonage Rtt106p +BamHI (5')
7950	ACCGGTCGACCGGTCAATCGTATTCTACTCCG	Clonage Rtt106p +Sall (3')
7314	CCGCTCGAGCGGTCATCAGGACCGGTGCTATCC	Clonage Tah1p (1-112) +XhoI (3')
7313	CCGCTCGAGCGGTCACACCTACCGCTCCCTGTGC	Clonage Tah1p (1-92) +STOP/XhoI (3')
8182	CCCGGGATCCCGATGTCAGTACAGATCCCTGTT	Clonage Tah1p (93-111) +BamHI/START (5')
7312	CCCGGGATCCCGATGAGCCAAATTTGAAAAG	Clonage Tah1p +BamHI (5')
4744	GGATCCTTCTGGAACCTCCATAG	Clonage yPEP +START/BamHI (5')
4745	CTCGAGCTATTGCATATGTTCTGTGTTGTTA	Clonage yPEP +STOP/XhoI (3')
4214	GCAGAAAGAGTGCCACAG	KO/vérification promoteur HIT1 (utilisé avec 3'KanMX-1287)
5380	CTAGGAATTTGTAATTCATA	KO/vérification promoteur PIH1 (utilisé avec 3'KanMX-1287)
5383	CAGCGTATATGACTGCACA	KO/vérification promoteur TAH1 (utilisé avec 3'KanMX-1287)
7966	CGTATTCTACTCCGGAACCATCCTCATCATCTTC	Mutagenèse élimination BamHI dans ORF RTT106 (3')
7965	GAAGATGATGAGGATGGTCCGGAGTAGAATACG	Mutagenèse élimination BamHI dans ORF RTT106 (5')
6515	GAAGATGATGGGACAGATGCTGCTGCACATAAGACTGGGATGTG	Mutagenèse Nhp2p E95D96H97/AAA (3')
6514	CACATCCAGTCTTATGTGCGAGCATCTGTCCCATACATCTTC	Mutagenèse Nhp2p E95D96H97/AAA (5')
6839	GTGATCTTACATAAGAGTGGGATGTGGG	Mutagenèse Nhp2p V92/L (3')
6838	CCCACATCCCACTCTTATGTGAAGATCAC	Mutagenèse Nhp2p V92/L (5')
6830	GGTATGAAGATGTATGGGACATCTTATCTTACATAAGAGTGGGATGTGGGAAATAAC	Mutagenèse remplacement $\alpha$ 3-Nhp2p par $\alpha$ 3-Snu13p (3')
6829	GTTATTTCCACATCCCACTCTTATGTGAAGATAAAGATGTCCCATACATCTTCATACC	Mutagenèse remplacement $\alpha$ 3-Nhp2p par $\alpha$ 3-Snu13p (5')
6832	GGAAACAAGACATATGGAACAGAGTGGTCTTACATAGTACTGGTAAATGCAATAGG	Mutagenèse remplacement $\alpha$ 3-Snu13p par $\alpha$ 3-Nhp2p (3')
6831	CCTATTGCAATTTACCAGTACTATGTGAAGACCACTCTGTTCCATATGCTTTGTTCC	Mutagenèse remplacement $\alpha$ 3-Snu13p par $\alpha$ 3-Nhp2p (5')
5920	TAGCCACATTTTTTTCTTCTGCTCTCCACTTTTTAACATCTTC	Mutagenèse Rsa1p E247/A (3')
5919	GAAGATGTTAAAAAGTGGAGAGCAGAAGAAAAAATGTGGCTA	Mutagenèse Rsa1p E247/A (5')
7111	CCACATTTTTTTCTTCTGCTGCCCACCTTTTTAACATCTTC	Mutagenèse Rsa1p R246/A (3')
7110	GAAGATGTTAAAAAGTGGGAGCAGAAGAAAAAATGTGG	Mutagenèse Rsa1p R246/A (5')
5922	TAGCCACATTTTTTTGCTTCTTCTGCTCCACTTTTTAACATCTTC	Mutagenèse Rsa1p R249/A (3')
5921	GAAGATGTTAAAAAGTGGAGAGAAGCAAAAAAATGTGGCTA	Mutagenèse Rsa1p R249/A (5')
5083	CCACATTTTTTTCTTCTTCTGCTGCTTTAACATCTTCATCAGTAAT	Mutagenèse Rsa1p W245/A (3')
5082	ATTACTGATGAAGATGTTAAAAAGGCTAGAGAAGAAAAAATGTGG	Mutagenèse Rsa1p W245/A (5')
5079	AACAAAGACATATGGAACATCTTGTGACACATAGTAATGGTAAATGCAA	Mutagenèse Snu13p E72/A (3')
5078	TTGCATTTACCATTACTATGTGCTGACAAGAATGTTCCATATGCTTTGTT	Mutagenèse Snu13p E72/A (5')
5008	AACAAAGACATATGGAACATTAGCAGCAGACATAGTAATGGTAAATGCAA	Mutagenèse Snu13p E72D73K74/AAA (3')
5007	TTGCATTTACCATTACTATGTGCTGCTGCTAATGTTCCATATGCTTTGTT	Mutagenèse Snu13p E72D73K74/AAA (5')

**Tableau 4a : Oligonucléotides utilisés au cours de la thèse.**  
(cf. suite et légende, page suivante)

N°	Séquence 5'-3'	Description
2132	CATGGCTTAATCTTTGAGAC	Reverse/Northern détection 18S (ARNr)
2131	CTCCGCTTATTGATATGC	Reverse/Northern détection 25S (ARNr)
2905	CTTGCGACCGGCTATTC	Reverse/Northern détection 5'-ETS (ARNr)
2134	CCAGATAACTATCTTAAAAG	Reverse/Northern détection 5'-ETS (ARNr)
4441	AAAGCTCTCATGCTCTTGCC	Reverse/Northern détection ITS1 (ARNr)
2130	TGTTACCTCTGGGCC	Reverse/Northern détection ITS1 (ARNr)
2133	GGCCAGCAATTTCAAGTTA	Reverse/Northern détection ITS2 (ARNr)
8167	AGCTTGAGTTTTCCACACC	Reverse/Northern détection snoRNA C/D snR13
2431	2431 CGAGGAAAGAGACACCATTATC	Reverse/Northern détection snoRNA C/D snR190
6952	GAAATACAAATATCAACATA	Reverse/Northern détection snoRNA C/D snR38
8166	ACGCCCCGGATAACTAGAGT	Reverse/Northern détection snoRNA C/D snR4
8168	TACTTAATGCTTCATAGGAC	Reverse/Northern détection snoRNA C/D snR40
8169	TTATGTTGGCCACTCATCAC	Reverse/Northern détection snoRNA C/D snR63
8232	GCGCCCTAAGCTCTACTA	Reverse/Northern détection snoRNA C/D snR64
8170	TGACCAATCATCAATTCTCC	Reverse/Northern détection snoRNA C/D snR70
1968	TCACTCAGACATCCTAG	Reverse/Northern détection snoRNA C/D U14
8172	TGTGATAGTCAACCACATCTC	Reverse/Northern détection snoRNA C/D U18
6951	GATCTTGGTGATAATTGGTA	Reverse/Northern détection snoRNA C/D U24
41	3' CCGTCAGACTGTTCA 5'	Reverse/Northern détection snoRNA U3 (+ 3' Transcription T7)
8233	CACTCGCGCAGTACCGAT	Reverse/Northern détection snoRNA H/ACA snR8
8171	TACTCGATCCAGCACGGGAA	Reverse/Northern détection snoRNA MRP
4481	GACCAAGGAGTTTGATCAATGA	Reverse/Northern détection snRNA U1
2053	TAAATTTCAACCAGGGGAAACAAATCTCGGACGAA	Reverse/Northern détection snRNA U4
7159	CAGTGCAGGGTCCGAGGTATT	RT-PCR amorce 3' exogène pour PCR (alias 3'PCR)
7517	CAGTGCAGGGTCCGAGGTATTCTGAAAACAAACCTTTGG	RT-PCR amorce rev pré-U3 longs (alias rev) (+ séq 5' = 7159)
7514	GTCGTATCCAGTGCAGGGTCCGAGGTATTTCGCACTGGATACGACAAAAAGTGG	RT-PCR amorce tige-boucle rev pré-U3 +12 (+ séq 5' = 7159)
7515	GTCGTATCCAGTGCAGGGTCCGAGGTATTTCGCACTGGATACGACAAAAGGAA	RT-PCR amorce tige-boucle rev pré-U3 +18 (+ séq 5' = 7159)
7424	GTCGTATCCAGTGCAGGGTCCGAGGTATTTCGCACTGGATACGACACTTGTC	RT-PCR amorce tige-boucle rev U3 (+ séq 3'-7159)
5252	TCTGTGTGCAGCTACTTCA	RT-PCR amorce U3 exon 1 (5') (alias 5'PCRa)
5254	AGTCTTAGGTAAGGTTT	RT-PCR amorce U3 intron (5') (alias 5'PCRb)
967	CTAATACGACTCACTATAGGG	Transcription matrice T7 amorce universelle 5'-T7
2174	TAATACGACTCACTATAGGACCTTTGTACCCAG	Transcription matrice T7 U3B/C et mini-U3B/C (5'-T7)
2176	CTGCCATTTGTACCCA	Transcription matrice T7 U3B/C et mini-U3BC (3')
6980	TAATACGACTCACTATAGGGTTCGACGCTACTTCA	Transcription matrice T7 snoRNA U3 (5'-T7)
6085	TAATACGACTCACTATAGGGCTTGGTTGATGAG	Transcription matrice T7 mini-U3Δ2,3,4 (5'-T7)
6086	GGGCTTGTGACTGCCATTTGTACCC	Transcription matrice T7 mini-U3Δ2,3,4 (3')
5256	CTTACACCAACTAATCCA	Transcription matrice T7 pré-U3 long (3')
1967	TAATACGACTCACTATAGGGTTCGAGGTGATGAAAG	Transcription matrice T7 snoRNA U14 (5'-T7) (3'=1968)

**Tableau 4b : Oligonucléotides utilisés au cours de la thèse (suite).** Le numéro, correspondant à la nomenclature utilisée au laboratoire, la séquence et la fonction de chaque oligonucléotide sont précisés. (5') : Amorce s'hybridant en 5' de la séquence à amplifier; (5'-T7) : Amorce (5') contenant le promoteur de transcription de l'ARN polymérase T7; (3') : Amorce s'hybridant en 3' de la séquence à amplifier.



**Milieu sélectif YNB (Yeast Nitrogen Base) :** Base azotée de levure sans acide aminé (10 g/l); galactose ou glucose (20 g/l); mélange d'acides aminés (20 g/l de chaque acide aminé).

Les antibiotiques sont ajoutés après l'étape de stérilisation, à une concentration finale de 200 µg/mL pour la généticine (G418), 60 µg/mL pour la noureséothricine et 50 µg/mL pour la doxycycline. Les inhibiteurs des protéines chaperonnes HSP90 et HSP70, respectivement la geldanamycine et la pifithrine  $\gamma$ , sont ajoutés à une concentration de 20 µg/mL. Enfin, le 3-aminotriazol, utilisé pour les expériences de double-hybride ou de triple-hybride, est ajouté selon une gamme de concentration variant de 0,1 mM à plus de 50 mM, en fonction des applications.

### 1.5. Les tampons et solutions d'usage courant

**TE :** Tris-HCl 10 mM, pH 7,4 ; EDTA 1 mM.

**TBE :** Tris borate 89 mM, pH 8,3 ; EDTA 2 mM.

**PBS :** Na<sub>2</sub>HPO<sub>4</sub> 10 mM; KH<sub>2</sub>PO<sub>4</sub> 1,8 mM, pH 7,3; NaCl 140 mM; KCl 2,7 mM.

**Tampon D :** HEPES-KOH 20 mM, pH 7,9 ; KCl 150 mM ; MgCl<sub>2</sub> 1,5 mM ; EDTA 0,2 mM.

**Tampon SSPE 20X :** Na<sub>2</sub>HPO<sub>4</sub> 150 mM, pH 7,4 ; NaCl 3 M; EDTA 25 mM

**Tampon SSC 20X :** Citrate de sodium 340 mM, pH 7,2 ; NaCl 150 mM

**Tampon MOPS 10X :** MOPS 200 mM ; Acétate de sodium 50mM ; EDTA 10 mM

**Tampon de migration SDS-PAGE :** Glycine 14,25 g/l ; Tris 3 g/l ; SDS 1 g/l.

**Réactif de Denhardt's 100X :** Ficoll 20 g/L ; Polyvinylpyrrolidone 20 g/L ; BSA 20 g/L

**Solution de polyacrylamide dénaturante :** Acrylamide/Bisacrylamide (23,75/1,25) 25 % (p/v) ; TBE 1X ; urée 8 M.

**Solution de polyacrylamide native :** Acrylamide/Bisacrylamide (38/2) 40 % (p/v) ; TBE 1X.

**Bleu de dépôt pour gel SDS-PAGE :** Tris-HCl 80 mM pH 6.8 ; SDS 3% ; DTT 100 mM ; Glycérol 10% ; Bleu de bromophénol 0.1%.

**Bleu de formamide :** EDTA 20 mM ; bleu de bromophénol 0,05 % ; bleu de xylène cyanol 0,05 % dans de la formamide désionisée.

**Bleu natif pour les expériences de retard sur gel :** HEPES/KOH 20 mM, pH 8,3 ; glycérol 40 % ; bleu de bromophénol 0,05 % ; bleu de xylène cyanol 0,03 %.

**Bleu natif pour dépôt sur gel d'agarose :** EDTA 5 mM ; SDS 0,4 % ; glycérol 5 % ; bleu de bromophénol 0,2 %.

**Solution de coloration pour gel SDS-PAGE :** Ethanol 20% ; Acide Acétique 10% ; 2 g/l Bleu de Coomassie R250.

**Solution de décoloration de gel SDS-PAGE :** Ethanol 20% ; Acide Acétique 10%.

**Solution de coloration au bleu colloïdale :** Bleu de Coomassie G250 0,1% ; sulfate d'ammonium 10% ; acide orthophosphorique 1,6% ; éthanol 20%.

### 1.6. Les anticorps et les sérums

Les sérums et anticorps suivants ont été employés pour les analyses par Western-blot :

- **anti-Snu13p :** Sérum de lapin (aimablement fourni par E. Bertrand). Utilisé à une dilution de 1/5000<sup>ème</sup>.
- **anti-Pih1p-Tah1p :** Sérum de lapin (aimablement fourni par S. Moreira). Utilisé à une dilution de 1/5000<sup>ème</sup>.

- **anti-Trm4p** : Sérum de lapin (aimablement fourni par I. Motorine). Utilisé à une dilution de 1/5000<sup>ème</sup>.
- **anti-Pma1p** : Sérum de lapin (aimablement fourni par B. André). Utilisé à une dilution de 1/10000<sup>ème</sup>.
- **anti-GFP** : Anticorps monoclonal produit chez le lapin (SIGMA). Utilisé à une dilution de 1/4000<sup>ème</sup>.
- **anti-flag** : Anticorps monoclonal produit chez la souris (SIGMA). Utilisé à une dilution de 1/600<sup>ème</sup>.
- **anti-HA** : Anticorps monoclonal produit chez le lapin (SIGMA). Utilisé à une dilution de 1/1000<sup>ème</sup>.

## 2. Méthodes

### 2.1. Techniques de génie génétique

#### 2.1.1. Clonages

Les fragments d'ADN d'intérêts sont amplifiés par PCR en utilisant l'ADN polymérase Phusion® (Finnzymes) et un couple d'oligonucléotides amorces apportant, respectivement en 5' et en 3', des sites de restriction compatibles avec le vecteur de destination. Les ADN double-brin obtenus sont traités par l'endonucléase *DpnI* (MBI-Fermentas), purifiés sur colonnes NucleoSpin® (Machery-Nagel) et quantifiés par mesure de l'absorbance à 260 nm. 2 µg d'ADN sont ensuite digérés par double-restriction enzymatique, dans les conditions décrites par le fournisseur (MBI-Fermentas). Les produits de digestion sont à nouveau purifiés sur colonnes NucleoSpin®, dont la limite d'exclusion permet d'éliminer les produits de clivage provenant des deux extrémités du fragment d'ADN. Le vecteur de destination est préparé de la même façon et subit une étape supplémentaire de déphosphorylation de ses extrémités 5' par action de la phosphatase alcaline CIAP (MBI Fermentas). La ligature par la T4-DNA ligase (MBI-Fermentas) est ensuite réalisée à partir de 50 ng de vecteur et d'un excès molaire de 3 fois de fragment. Après transformation de cellules *E. coli* DH5α super-compétentes et sélection des clones recombinants sur milieu sélectif, les plasmides sont amplifiés par mini-préparation et analysés par séquençage.

#### 2.1.2. Mutagenèse dirigée par PCR

Au cours de ma thèse, l'approche de mutagenèse dirigée par PCR a été utilisée en employant l'ADN polymérase Phusion® (Finnzymes). La réaction est réalisée dans un volume total de 50 µL, en présence de 0,02 unité d'enzyme Phusion® dans son tampon, de 0,2 µM de chaque dNTP (MBI Fermentas), de 250 ng de plasmide recombinant et de 40 nM de chaque amorce complémentaire apportant les séquences mutées. La température et la durée de chaque étape réactionnelle sont ajustées en fonction du Tm des amorces et de la taille du plasmide. Après 30 cycles réactionnels, 10 µl de mélange réactionnel sont traités par l'endonucléase *DpnI* (MBI Fermentas) dans les conditions recommandées par le fournisseur. Après transformation de cellules *E. coli* DH5α super-compétentes et sélection des clones recombinants sur milieu sélectif, les plasmides sont amplifiés par mini-préparation et analysés par séquençage.

### 2.1.3. Production et purification de protéines recombinantes

Les conditions décrites ci-dessous correspondent aux conditions standards, les plus couramment utilisées pour nos protéines d'intérêt. Cependant, dans certains cas, elles peuvent faire l'objet d'adaptations spécifiques (températures de culture, tampon, ...) permettant de répondre aux propriétés de certaines protéines (solubilité, stabilité, ...).

#### 2.1.3.1. Surproduction de protéines chez *E. coli*

Les cellules BL21-DE3 transformées par le (ou les) plasmide(s) d'intérêt (*e.g.* pnEA-3TH + pnCS pour les tests de co-expression) sont utilisées pour ensemercer 1 L de milieu de culture (LB + antibiotiques) à 0,1 unités  $A_{600nm}/mL$ . Les cultures sont placées sous agitation à une température de 30°C. A une concentration cellulaire équivalente à 0,8 unités  $A_{600nm}/mL$ , la surproduction de protéines est induite par ajout d'IPTG à la concentration de 1 mM. Les cultures sont alors incubées sous agitation à une température de 20°C, jusqu'à saturation. A l'issue de la culture, les cellules sont reprises dans un tampon (HEPES-KOH 20 mM, pH 7,5 ; NaCl 400 mM ;  $\beta$ -mercaptoéthanol 5 mM ou DTT 1 mM ; benzamidine 1 mg/mL) à raison de 100 unités  $A_{600nm}/mL$ . Elles sont ensuite lysées, à 4°C, par traitement aux ultra-sons à l'aide d'un sonicateur Sonifer 250 (Branson) (duty-cycle 50% ; Output 6). Le lysat est centrifugé 1 h à 30000g et 4°C. Le surnageant de centrifugation obtenu est alors additionné de PEI à une concentration de 0,0125%. Après agitation manuelle et une nouvelle étape de centrifugation (30 min, 20000g, 4°C), le surnageant obtenu peut être utilisé pour procéder à l'étape de purification.

#### 2.1.3.2. Purification de protéines fusionnées à la GST

La purification des protéines fusionnées à la GST est réalisée sur la résine Glutathione Sépharose (GE Healthcare). Pour l'équivalent de 1L de culture, 500  $\mu L$  de résine sont équilibrés dans le tampon de lyse, puis incubés avec l'extrait total de protéines, sur roue, pendant 2 h, à 4°C. La résine est ensuite lavée par deux lavages successifs de 20 min sur roue, à 4°C, avec 15mL de tampon (HEPES-KOH 50mM, pH 7,5 ; NaCl 1M ; DTT 1mM), puis à nouveau 2 fois 10 min avec 15 mL de tampon (HEPES-KOH 20 mM, pH 7,5 ; NaCl 300 mM ; DTT 1 mM). Si la démarche expérimentale nécessite de conserver la protéine d'intérêt sous forme de fusion à la GST, les billes sont incubées pendant 15 min sur roue, à T°C ambiante, dans 500  $\mu L$  de tampon d'élution (HEPES-KOH 20 mM, pH 7,5 ; NaCl 300 mM ; DTT 1 mM ; Glutathion, réduit 10 mM). L'opération est répétée 2 fois. L'éluât est ensuite dialysé contre 2x 500 mL de tampon (HEPES-KOH 20 mM, pH 7,5 ; NaCl 300 mM ; DTT 1 mM) ou traité par gel-filtration, sur colonnes Superdex<sup>TM</sup>-Sépharose 75 ou 200 (GE Healthcare) selon la taille des protéines. En revanche, si la protéine doit être séparée de la GST, les billes sont incubées pendant 12 h sur roue, à 4°C, dans 500  $\mu L$  de tampon d'élution (HEPES-KOH 20 mM, pH 7,5 ; NaCl 300 mM ; DTT 1 mM) additionné de protéase GST-PréScission (produite au laboratoire) à raison de 1 unité pour 100  $\mu g$  de protéines (estimation sur gel). L'éluât obtenu peut alors être traité par gel-filtration, sur colonnes Superdex<sup>TM</sup>-Sépharose 75 ou 200 (GE Healthcare).

### 2.1.3.3. Purification de protéines fusionnées à une étiquette 6-His

La purification des protéines porteuses d'une étiquette 6xHis est réalisée par IMAC (*Immobilized Metal Affinity Chromatography*) sur la résine Talon-Sépharose (Clontech), chargée en ions cobalt. Pour l'équivalent de 1 L de culture, 300 µL de résine sont équilibrés dans le tampon de lyse, puis incubés avec l'extrait total de protéines, sur roue, pendant 2 h, à 4°C. La résine est ensuite lavée pendant 5 min sur roue, à 4°C, avec 15 mL de tampon (HEPES-KOH 20 mM, pH 7,5 ; NaCl 400 mM ; β-mercaptoéthanol 5 mM). Si la démarche expérimentale nécessite de conserver l'étiquette 6xHis, les billes sont incubées pendant 5 min sur roue, à 4°C, dans 500 µL de tampon d'éluion (HEPES-KOH 20 mM, pH 7,5 ; NaCl 300 mM ; β-mercaptoéthanol 5 mM ; Imidazole 500 mM). L'opération est répétée 2 fois. L'éluât est ensuite dialysé contre 2x 500mL de tampon (HEPES-KOH 20 mM, pH 7,5 ; NaCl 300 mM ; β-mercaptoéthanol 5 mM) ou traité par gel-filtration, sur colonnes Superdex<sup>TM</sup>-Sépharose 75 ou 200 (GE Healthcare), selon la taille des protéines ou des complexes purifiés. En revanche, si l'étiquette doit être éliminée, les billes sont incubées pendant 12 h sur roue, à 4°C, dans 500 µL de tampon d'éluion (HEPES-KOH 20 mM, pH 7,5 ; NaCl 300 mM ; β-mercaptoéthanol 5 mM) additionné de protéase GST-PréScission (produite au laboratoire) à raison de 1 unité pour 100 µg de protéines (estimation sur gel). L'éluât obtenu peut alors être traité par gel-filtration, sur colonnes Superdex<sup>TM</sup>-Sépharose 75 ou 200 (GE Healthcare).

### 2.1.4. Transcription/traduction *in vitro* de protéines radiomarquées

Certaines protéines, comme Nop58p, qui ne peuvent pas être obtenues sous forme recombinante, ont été produites sous forme radiomarquée, par transcription/traduction en extrait cellulaire S30 d'*E. coli*, à l'aide du kit *E. coli T7 S30 Extract System for Circular DNA* (Promega). Ces productions ont été réalisées, selon les instructions du fournisseur, en utilisant comme ADN matrice le plasmide pnCS recombinant et en incorporant de la méthionine marquée au <sup>35</sup>S comme acide aminé radiomarqué.

### 2.1.5. Transcription *in vitro* et marquage radioactif des ARN

#### 2.1.5.1. Transcription *in vitro* par l'ARN polymérase T7

La transcription d'ARN *in vitro* est réalisée à l'aide d'une ARN polymérase T7 produite au laboratoire. Dans un premier temps, une matrice d'ADN double-brin linéaire, portant un promoteur T7 en amont de la séquence codant l'ARN, est générée par PCR, à l'aide de l'ADN polymérase GoTaq (Promega). La réaction de transcription est réalisée pendant 12 h à 37°C, à partir de 7 µl de matrice ADN, dans du tampon de transcription (MgCl<sub>2</sub> 6,66 mM ; DTT 1,66 mM ; Spermidine 0,33 mM ; Triton X-100 0,0033% ; PEG 8000 26,66 mg/ml ; Tris/HCl 13,33 mM, pH 8,1) additionné de 22,5 unités d'ARN polymérase T7 et 4,16 mM de chaque NTP. A l'issue de cette réaction, l'ADN matrice est dégradé par ajout de 7,5 unités de DNase I et incubation pendant 1 h à 37°C. Les ARN sont ensuite dénaturés 2 min à 96°C en présence de bleu de formamide, puis purifiés sur gel de polyacrylamide dénaturant. Après visualisation des ARN par ombrage aux UV, les bandes de gel d'intérêt sont découpées et les ARN sont élués par incubation sous agitation, pendant 30 min à 37°C, dans 300 µL de tampon d'éluion (Tris/HCl 10 mM, pH 7,5 ; NaCl 300 mM ; EDTA 1 mM ; SDS 1%). Les ARN élués sont

ensuite isolés par précipitation alcoolique (sans ajout de sels), repris dans 20 µL d'eau ultra-pure, puis dosés par mesure de l'absorbance à 260 nm.

### 2.1.5.2. Marquage au $^{32}\text{P}$ des ARN en 5'

Le marquage des ARN au  $^{32}\text{P}$  au niveau de leur extrémité 5' nécessite une étape préalable de déphosphorylation. Une quantité de 1 µg d'ARN est donc traitée à la phosphatase alcaline CIAP (MBI Fermentas) selon les conditions recommandées par le fournisseur. Après extraction au phénol-chloroforme et précipitation alcoolique en présence de NaCl 300 mM, les ARN sont repris dans 10 µL d'eau ultra-pure. La réaction de marquage, par phosphorylation, est réalisée à l'aide de la T4-PNK (MBI-Fermentas), à partir de 2,5 µL d'ARN (soit environ 250 ng) et de 10 µCi de  $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ . Après incubation 1h à 37°C, les ARN sont ensuite dénaturés 2 min à 96°C en présence de bleu de formamide, puis purifiés sur gel de polyacrylamide dénaturant. Après visualisation des ARN par autoradiographie, les bandes de gel d'intérêt sont découpées et les ARN sont élués par incubation sous agitation, pendant 30 min à 37°C, dans 300 µL de tampon d'éluion (Tris/HCl 10 mM, pH 7,5 ; NaCl 300 mM ; EDTA 1 mM ; SDS 1%). Les ARN élués sont ensuite isolés par précipitation alcoolique (sans ajout de sels), repris dans 10 µL d'eau ultra-pure et dosés au compteur à scintillation (Beckman Coulter).

## 2.2. Techniques de manipulation des levures

### 2.2.1. Transformation de cellules de levures compétentes

La préparation de cellules de levure compétentes et leur transformation est réalisée à partir de 30 mL de cellules cultivées à 30°C, en phase exponentielle de croissance. Arrivé à une absorbance d'environ 0,5 à 0,8 unités  $A_{600\text{nm}}/\text{mL}$ , les cellules sont lavées dans 15 mL d'eau stérile, avant d'être reprises dans 500 µL de tampon (Acétate de lithium 0,1 M ; TE). Pour chaque transformation, un volume de 50 µL de suspension cellulaire est incubé 30 min à 30°C, en présence de 1 à 2 µg d'ADN, de 45 µg d'ADN de sperme de saumon dénaturé (15 min à 100°C) et de 300 µL de tampon (PEG 4000 40% ; Acétate de lithium 0,1 M ; TE). A l'issue de cette étape, les cellules sont incubées 10 min à 42°C, avant d'être lavées dans 1 mL d'eau stérile, puis étalées sur milieu sélectif solide et incubées à 30°C pendant plusieurs jours.

### 2.2.2. Modification génétique des souches de levure

Au cours de ce travail, certaines souches de levure ont été modifiées génétiquement, dans le but d'invalider certains gènes ou de modifier leur profil de résistance aux antibiotiques. Pour cela les levures ont été transformées avec 1 à 2 µg d'ADNdb linéaire, obtenus par PCR ou par restriction enzymatique, et permettant la modification d'un *locus* spécifique par recombinaison homologue. A l'issue de la transformation, les cellules sont incubées 24 h à 30°C, sur milieu YPD ou YPG solide, puis répliquées, à l'aide d'un velours, sur le milieu sélectif approprié. Si besoin, les clones positifs font l'objet d'un génotypage par extraction de l'ADNg et analyse ciblée par PCR.

Dans le cas de l'invalidation d'un gène, le fragment d'ADN utilisé pour la recombinaison est obtenu par PCR à partir du plasmide pAF6, porteur de la cassette de résistance Kan-MX6. Pour cela, des oligonucléotides amorces complémentaires du gène cible (en 5') et des extrémités la séquence kan-MK6 (en 3') sont utilisés. Après amplification, les fragments obtenus sont purifiés sur colonnes NucleoSpin® et dosés par mesure de l'absorbance à 260 nm.

Dans le cadre de la réalisation des cribles d'interaction génétique GIM, les souches de levure invalidées pour les gènes d'intérêts ont été modifiées en remplaçant la cassette de résistance Kan-MX2, par un gène de résistance à la nourseothricine. Pour cela, le plasmide pGID3 (aimablement fourni par C. Saveanu) a subi une double restriction enzymatique par les endonucléases *EcoRI* (MBI Fermentas) et *AscI* (Biolabs), afin de libérer deux fragments d'ADNdb linéaire, dont un portant le gène de résistance. Après purification sur colonnes NucleoSpin®, les produits de digestion sont dosés par mesure de l'absorbance à 260 nm.

### 2.2.3. Isolement de nouvelles souches de levure par sporulation et dissection de tétrades

Dans le but de générer des cellules de levures invalidées pour deux gènes d'intérêt, nous avons utilisé une approche de sélection par sporulation et dissection de tétrades. Pour cela, nous avons utilisé les souches de levure BY4741 (*Mat a*) et BY4742 (*Mat a*) dont un seul gène est délété. Ces souches sont croisées deux à deux sur milieu riche YPD, à une température de 30°C. Les cellules diploïdes sont ensuite sélectionnées sur milieu Met<sup>-</sup>Lys<sup>-</sup>, puis repiquées deux nouvelles fois sur milieu riche YPD. Les cellules sont alors placées dans des conditions de carence, qui conduisent une faible proportion de cellules (environ 5 à 20%) à entrer en méiose. Chez *S. cerevisiae*, ce processus se matérialise par un phénomène de sporulation, conduisant à la production de 4 spores haploïdes, maintenues sous la forme d'une tétrade, appelée asque. Pour cela, une colonie de cellule est incubée dans 2 mL de milieu de sporulation (Acétate de potassium 1% ; Acétate de zinc 0,005% ; Uracile 2 µg/mL ; Histidine 4 µg/mL ; Leucine 12 µg/mL), sur roue, pendant 5 jours à 25°C, suivis de 3 jours à 30°C. A l'issue de cette période, la dissection des tétrades impose de réaliser une digestion partielle de la paroi des asques, en utilisant la zymolyase, une enzyme qui hydrolyse les liaisons β-1,3-glucane des protéines glycosylées de la paroi. Afin de réaliser cette digestion ménagée 150 µL de cellules sont lavées à l'eau stérile, puis reprises dans un volume de 150 µL de Sorbitol 1 M, additionné de zymolyase à une concentration finale de 0,25 mg/mL. Le mélange est incubé à 33°C, pendant une durée de 15 min. La réaction est ensuite arrêtée par ajout lent de 250 µL d'eau stérile glacée. Enfin, la dissection des tétrades est réalisée à l'aide d'un micromanipulateur (Singer). Pour cela, 50 µL de cellules sont déposés sur le quart d'une boîte de milieu riche YPD et les ascospores provenant d'une même tétrade sont positionnées au niveau de coordonnées précises, indiquées par l'appareil. Après dissection de 10 à 15 tétrades, les boîtes sont incubées plusieurs jours à 30°C et le profil de croissance de chaque lignée cellulaire est analysé. Ces dernières font également l'objet d'un génotypage, par extraction de l'ADNg et analyse ciblée par PCR.

#### 2.2.4. Utilisation de systèmes génétiques permettant de réprimer l'expression de gènes essentiels à la viabilité cellulaires

Certaines protéines d'intérêt, comme Snu13p, Nhp2p ou Bcd1p, sont essentielles à la viabilité cellulaire. A défaut de pouvoir utiliser des cellules de levure invalidées pour les gènes codant ces protéines, nous avons donc recours à des souches permettant d'induire une répression transitoire de leur expression. Deux systèmes de régulation différents ont été utilisés au cours de mon travail de thèse : (i) le système GAL, dans lequel le gène d'intérêt est placé sous contrôle de l'élément régulateur UAS-Gal10, qui active la transcription en présence de galactose dans le milieu de culture (Watkins *et al.*, 2000), (ii) le système *Tet Off*, dans lequel le gène d'intérêt est placé sous contrôle de l'opérateur Tet, dont l'effet activateur nécessite la fixation de la protéine tTA, qui est réprimée en présence de Doxycycline (Peng *et al.*, 2003). La répression de l'expression de gènes placés sous contrôle de l'élément UAS Gal10 est réalisée en utilisant le glucose comme source carbonée. La répression par le système *Tet Off* est réalisée par ajout de Doxycycline à 50 µg/mL dans le milieu de culture. Dans ces deux cas, les cultures en milieu non permissif sontensemencées à une concentration cellulaire faible ( $A_{600nm}/mL < 0,05$ ), afin d'assurer la production d'un maximum de biomasse dans les conditions de répression. De plus, ces systèmes sont utilisés pour évaluer la fonctionnalité de fragments ou de variants protéiques, exprimés de façon ectopique à partir d'un vecteur d'expression approprié.

#### 2.2.5. Traitement de cellules en culture par les inhibiteurs de chaperonnes HSP90 et HSP70

L'inhibition des chaperonnes HSP90 (Hsp82p et Hsc82p chez la levure) dans des cellules en culture a été réalisée selon les conditions décrites par l'équipe de W.A. Houry (Zhao *et al.*, 2005). Pour cela, les différentes cultures sontensemencées à une concentration de 0,025 unités  $A_{600nm}/mL$ , en présence de 20 µg/mL de Geldanamycine (Sigma). Les cultures sont incubées à 30°C, sous agitation, jusqu'à l'obtention de valeurs d'absorbance supérieures à 4,5 unités  $A_{600nm}/mL$ . Afin de s'assurer de l'inhibition constante des protéines chaperonnes, les pré-cultures utilisées pour l'ensemencement sont également réalisées en présence d'une concentration équivalente en Geldanamycine. La Geldanamycine étant préparée dans du DMSO, cette expérience doit s'accompagner d'une culture témoin, réalisée en présence d'un volume de DMSO équivalent à celui ajouté dans les cultures traitées. Concernant l'inhibition des protéines HSP70, l'utilisation de la Pifithrine  $\gamma$  sur des levures en culture n'étant pas référencée, nous avons utilisé cette molécule dans les conditions décrites pour la Geldanamycine.

#### 2.2.6. Techniques de double-hybride et de triple-hybride

La technique de double-hybride (en abrégé Y2H) permet de révéler l'association de deux protéines dans des cellules de levure. Elle repose sur l'expression par la levure de deux protéines de fusion. La première protéine chimérique, codée par le plasmide pGBKT7 est composée du domaine de fixation à l'ADN de l'activateur transcriptionnel Gal4 (Gal4-BD) et de la protéine appât. La seconde, codée par le plasmide pACT2, est composée du domaine d'activation de la transcription de l'activateur Gal4 (Gal4-AD) et de la protéine proie. L'interaction entre les deux partenaires permet la reconstitution de l'activateur

transcriptionnel Gal4 et induit la transcription d'un gène rapporteur placé sous contrôle de l'opérateur *GALI*, qui constitue l'élément de réponse à l'activateur Gal4.

La technique de triple-hybride (en abrégé Y3H) repose sur le même principe et permet de mettre en évidence une association entre une séquence ARN et une protéine. Elle fait intervenir 3 partenaires : (i) une première protéine de fusion (exprimée par les souches R40 et L40 coat) composée de la protéine de l'enveloppe du phage MS2 fusionnée à la protéine bactérienne LexA, qui se fixe spécifiquement sur l'opérateur *lexA*, (ii) un ARN hybride (codé par le plasmide pIIIAMS2.2) composé de l'ARN "appât" et de l'ARN du phage MS2, qui interagit spécifiquement avec la protéine de l'enveloppe du phage MS2, (iii) une seconde protéine de fusion, codée par le plasmide pACT2 et composée du domaine Gal4-AD fusionné à la protéine proie. Comme pour le double-hybride, l'interaction couplée de ces trois partenaires conduit au recrutement du domaine Gal4-AD et permet l'expression du gène rapporteur.

Concernant la technique Y2H, la souche Y187 (*Mat a*) est transformée par le plasmide pGBKT7 (*TRP1*) qui code la protéine "appât" et la souche Y190 (*Mat a*) est transformée par le plasmide pACT2 (*LEU2*) qui code la protéine "proie". Dans la technique Y3H la souche R40 coat (*Mat a*) est transformée par le plasmide pIIIAMS2.2 (*URA3-ADE2*) qui code l'ARN "appât" et la souche L40 coat (*Mat a*) est transformée par le plasmide pACT2 (*LEU2*) qui code la protéine "proie". Les souches haploïdes transformées sont ensuite croisées sur milieu YPD solide, afin de générer des souches diploïdes contenant les deux plasmides d'intérêt. Ces dernières sont alors sélectionnées par repiquage sur un milieu sélectif (Y2H : Trp<sup>-</sup>-Leu<sup>-</sup> ; Y3H : Ura/Ade<sup>-</sup>-Leu<sup>-</sup>). Pour ces deux approches, nous avons tiré profit du gène rapporteur *HIS3*, dont le niveau d'expression dépend de la force d'interaction entre les deux partenaires d'intérêt. Ces interactions sont donc révélées sur un milieu sans histidine (Y2H : Trp<sup>-</sup>-Leu<sup>-</sup>-His<sup>-</sup> ; Y3H : Ura/Ade<sup>-</sup>-Leu<sup>-</sup>-His<sup>-</sup>), enrichi en 3-AT (3-aminotriazole) (Sigma), un inhibiteur spécifique de la voie de biosynthèse de l'Histidine, qui permet d'évaluer l'intensité de l'interaction étudiée. Les différentes phases d'incubation se déroulent à une température de 30°C.

## 2.3. Techniques d'extraction de matériel biologique

### 2.3.1. Préparation des extraits cellulaires

Au cours de ma thèse, plusieurs approches expérimentales ont nécessité la préparation d'extraits cellulaires. Dans la majorité des cas, ces derniers ont été préparés par action mécanique de billes de verre (Sigma). Pour une quantité de cellules équivalente à 30 unités  $A_{600nm}$ , la lyse des parois est réalisée dans 300  $\mu$ L de tampon (Tris-HCl 20 mM, pH 7,5 ; KCl 150 mM ; MgCl<sub>2</sub> 5 mM; TRITON-X100 0,05% ; Antiprotéases 1x (Roche)), en présence de 300  $\mu$ L de billes de verre. Après 3 fois 1 min d'agitation au vortex, interrompues par 2 min d'incubation dans la glace, le lysat est clarifié par deux centrifugations successives, pendant 5 min à 6000g. En cas de stockage, les extraits sont incubés quelques seconde dans de l'azote liquide, puis placés à -80°C.



### **2.3.2. Extraction des ARN totaux**

Pour l'extraction d'ARN totaux, une quantité de cellules équivalente à 10 unités d'absorbance à 600 nm sont reprises dans 200 µl de tampon d'extraction (Tris/HCl 50 mM, pH7,5 ; NaCl 100 mM ; EDTA 10 mM), additionnés de 300 µl de billes de verre (Sigma). Les échantillons sont agités à l'aide d'un vortex pendant 3 fois 1min, interrompues par 2 min d'incubation dans la glace. A l'issue de cette étape, 200 µl de tampon d'extraction, 50 µl de SDS 10% et 400 µl d'un mélange phénol/chloroforme/alcool isoamylique (25:24:1) sont ajoutés au lysat. L'échantillon est a nouveau agité à l'aide d'un vortex pendant 1 min, puis incubé 10 min à 65°C. Après centrifugation à 20000g, la phase aqueuse est récupérée et soumise à une seconde étape d'extraction dans 1 volume de mélange phénol/chloroforme/alcool isoamylique (25:24:1). Les ARN sont ensuite isolés par précipitation alcoolique en présence de 300 mM de NaCl, puis repris dans 50 µL d'eau ultra-pure et dosés par mesure de l'absorbance à 600 nm.

## **2.4. Techniques de détection à partir d'échantillons biologiques**

### **2.4.1. Technique de Western-blot**

La technique de Western-blot a été utilisée afin d'évaluer le taux de certaines protéines à partir d'extraits cellulaires (Préparation, cf. Matériel & Méthodes, 2.3.1). Les différents extraits sont dosés par la technique de Bradford, afin de normaliser la quantité de chaque échantillon à utiliser. Une quantité d'extrait, équivalent à 10 µL de l'échantillon le moins concentré, est ainsi fractionnée sur gel SDS-PAGE. Après migration, le transfert sur membrane de Nitrocellulose (Biorad) est réalisé dans un appareil Mini Trans-Blot (Biorad), pendant 1 h à 100 V, dans un tampon (Tris 3 g/L ; Glycine 14,4 g/l ; éthanol 20%). A l'issue du transfert, la membrane est bloquée par incubation sous agitation, pendant 1 h à température ambiante, dans un tampon de pré-hybridation (PBS/Tween-20 0,1% ; lait 5%). L'hybridation de l'anticorps primaire est réalisée sous agitation, pendant 1 h à température ambiante, dans le tampon d'hybridation (PBS/Tween-20 0,1%). La membrane est ensuite lavée 3 fois 15 min, puis incubée sous agitation, pendant 45 min avec l'anticorps secondaire approprié, couplé à la peroxydase (Invitrogen). Après deux lavages de 15 min dans le tampon d'hybridation, suivis de deux lavages de 5 min dans du tampon PBS, les protéines d'intérêt sont détectées par chemiluminescence à l'aide du kit ECL+ (GE Healthcare) et révélation par autoradiographie.

### **2.4.2. Technique de Northern-blot**

#### **2.4.2.1. Pour la détection des snoRNA**

Pour la détection des snoRNA, 5 µg d'ARN totaux sont dénaturés 2 min à 96°C dans du bleu de formamide, puis fractionnés par électrophorèse sur gel de polyacrylamide dénaturant 6% (acrylamide/bisacrylamide (38:2) ; urée 8 M ; TBE 0,5X). Après la migration, les ARN sont transférés sur une membrane de Nylon (ZetaProbe, Biorad), à l'aide du système de transfert Trans Blot SD SemiDry Transfer Cell (Biorad). Le transfert est réalisé dans du tampon TBE 0,5X, pendant 30 min. Les ARN sont ensuite fixés de façon covalente à la membrane par pontage photochimique aux UV, pendant 5 min. La membrane est pré-hybridée 1 h à 37°C sous agitation dans un tube à hybridation en présence de 15 ml de tampon d'hybridation (SSPE 6X ; Réactif de Denhardt 5X ; SDS 0,5%). A l'issue de cette étape, les sondes

radiomarquées au  $^{32}\text{P}$ , au niveau de leur extrémité 5', grâce à la T4-PNK (MBI Fermentas), sont ajoutés dans le tube à hybridation. L'hybridation se poursuit pendant 12 h à 37°C, puis la membrane est lavée plusieurs fois dans du tampon de lavage (SSC 6X) préchauffé à 42°C. La révélation des signaux radioactifs est réalisée à l'aide d'un appareil de détection Typhoon (GE Healthcare), après exposition sur un écran amplificateur.

#### 2.4.2.2. Pour la détection des ARNr

Pour la détection des ARNr et de leurs précurseurs, 5 µg d'ARN totaux sont dénaturés 1 h à 60°C, dans du tampon MOPS, additionné de 50% de DMSO et de 1 M de Glyoxal désionisé. Un volume de bleu de formamide est ajouté à l'issue de cette étape, et les ARN sont déposés sur gel d'agarose 1,2%, préparé dans du tampon MOPS (= l'agarose est fondu dans de l'eau et le MOPS est ajouté lorsque la température retombe à 55°C). La migration est réalisée à 30 V pendant au moins 12 h, dans du tampon MOPS. Le gel est ensuite lavé 10 min dans de l'eau ultra-pure, puis le transfert est réalisé sur membrane de Nylon Hybond+ (Amersham), à l'aide d'un système de pompe à vide, à une pression de 50 mbar. Le gel doit être imbibé de tampon pendant toute la durée du transfert. Il doit dans un premier temps être recouvert de NaOH à 75 mM pendant 20 min, afin d'hydrolyser partiellement les ARN et de faciliter leur transfert. Dans un second temps, le gel est recouvert pendant 2 périodes de 15 min de tampon (Tris/HCl 0,5 mM ; NaCl 1,5 M, puis pendant 10 min de tampon SSC 10X. Enfin, le transfert se poursuit pendant 2h30 à 3h avec un apport régulier de tampon SSC 10X. A l'issue du transfert, la membrane est traitée pendant 5 min aux UV, avant d'être pré-hybridée, puis hybridée avec les sondes d'intérêt, comme décrit dans la section précédente.

#### 2.4.3. Technique de RT-PCR

Au cours de ma thèse, la technique de RT-PCR a été employée afin de détecter les précurseurs du snoRNA U3, naturellement peu abondant et difficilement observable par Northern-blot. Cette approche a été optimisée afin d'assurer une spécificité de détection, ainsi qu'une sélectivité vis-à-vis des différents intermédiaires (cf. Résultats & Discussion, Chapitre 5). Dans un premier temps, les ARN totaux (1 µg) ou les ARN isolés par co-immunoprécipitation (la totalité de l'échantillon), sont traités dans un volume de 15 µL, par 1 unité DNase RQ1 (Promega) dans son tampon, pendant 30 min à 37°C. A l'issue de cette réaction, interrompue par ajout de 1,5 µL d'un tampon inhibiteur 10X (Promega), les ARN sont dénaturés pendant 10min à 65°C et placés dans la glace. 2 µL d'ARN traités à la DNase sont ensuite pré-incubés avec 1 µL d'oligonucléotide amorce de la réaction de transcription inverse à 20 µM. Dans le cas d'une amorce partiellement structurée en tige-boucle, celle-ci subit au préalable une étape de dénaturation/renaturation, par incubation successive 10 min à 65°C et 15 min à température ambiante. Dans un second temps, le volume réactionnel est élevé à 12,5 µL par ajout de 2 mM de chaque dNTP (MBI Fermentas) et 200 unités de transcriptase inverse M-MLV (Promega) dans son tampon. La réaction de transcription inverse est alors réalisée dans un thermocycleur par incubations successives 5 min à 16°C (Hybridation), 45 min à 42°C (Elongation) et 5 min à 85°C (Inactivation de l'enzyme). Un témoin négatif sans oligonucléotide amorce est réalisé afin de s'assurer de l'absence d'amorçage endogène de la réaction de transcription inverse. A l'issue de cette réaction, les ADNc obtenus sont analysés par PCR, en utilisant l'enzyme GoTaq (Promega), dans des conditions standard d'utilisation. La réaction se déroule dans un volume de 50 µL, à partir de 2 µL d'ADNc et en réalisant 30 à

35 cycles d'amplification. Les produits obtenus sont analysés par électrophorèse sur gel d'agarose 2,25%.

## 2.5. Techniques d'immunosélection

### 2.5.1. Approches standard d'immunoprécipitation

#### 2.5.1.1. Pour l'étude des snoRNP matures

L'analyse par Northern-blot des snoRNA associés *in vivo* aux protéines d'intérêt, a nécessité la mise en place d'une approche d'immunoprécipitation, basée sur l'utilisation de l'étiquette ProtA du TAP-tag. Chaque essai est réalisé à partir d'une quantité de cellules équivalente à 45 unités d'absorbance à 600 nm, prélevées en phase exponentielle de croissance. Les extraits cellulaires sont préparés comme décrit dans la section (2.3.1). Ils sont ensuite incubés, dans un volume total de 400 µL, avec 20 µL de billes IgG-Sépharose (Sigma), pré-équilibrées sur roue, pendant 1 h à 4°C, dans du tampon de lyse additionné de BSA acétylée à 20 µg/mL. La rétention des protéines étiquetées est alors réalisée sur roue à 4°C, pendant 2 h 30. Les billes sont ensuite lavées trois fois 5 min sur roue, à T°C ambiante, dans 1 mL de tampon IPP150 (Tris-HCl 10 mM, pH 8 ; NaCl 150 mM; 0.1% IGEPAL). A l'issue de cette étape, les billes sont reprises dans 300 µL tampon IPP150, additionnés de 15 µL de SDS 10% et de 150 µg de protéinase K (Roche). Les échantillons sont alors incubés à 37°C, pendant 30 min afin d'assurer la dégradation des anticorps IgG et des différentes protéines retenues sur les billes. Ils sont ensuite incubés pendant 10 min à 65°C, avant de subir deux extractions successives par traitement au phénol/chloroforme/alcool isoamylique (25:24:1). Enfin, les ARN sont isolés par précipitation alcoolique (ajuster NaCl 300 mM), en présence de 1 µL de glycogène, avant d'être repris dans 10 µL d'eau ultra-pure et stockés à -80°C. Pour chaque échantillon, 10% de l'extrait initial subit le traitement par la protéinase K, suivis de deux étapes d'extraction et de la précipitation alcoolique, afin de constituer l'input.

#### 2.5.1.2. Pour l'étude des pré-snoRNP

Pour l'analyse par RT-PCR des précurseurs du snoRNA U3 associés aux protéines d'intérêt, les extraits cellulaires sont réalisés dans le tampon (Tris-HCl 100 mM, pH 8 ; NaCl 100 mM ; glycérol 8% ; DTT 0,1 mM ; Antiprotéases 1x (Roche)) et les lavages sont réalisés dans le tampon de lavage (Tris-HCl 20 mM, pH 7,4 ; NaCl 100 mM; 0.1% IGEPAL ; DTT 0,1 mM). Avant l'étape de déprotéinisation à la protéinase K, les échantillons sont traités par 3 unités de DNase RQ1 (Promega) dans un volume de 40 µL, pendant 30 min à 37°C. Toutes les autres étapes se déroulent comme décrit dans la section précédente.

### 2.5.2. Purification de complexes pour analyse protéomique

L'analyse par spectrométrie de masse des protéines associées aux facteurs d'assemblage a nécessité d'adapter l'approche de purification par l'étiquette TAP (Rigaut *et al.*, 1999), afin de parer à leur faible abondance et à la nature transitoire des complexes d'assemblage. Pour cela seule l'étape de purification par l'étiquette ProtA a été réalisée et les complexes isolés ont été élués par coupure à la protéase TEV. Chaque essai est réalisé à partir d'une quantité de

cellules équivalente à 2500 unités d'absorbance à 600 nm, prélevées en phase exponentielle de croissance. Dans un premier temps, les cellules sont lysées dans 10 mL de tampon (Tris-HCl 100 mM, pH 8 ; NaCl 100 mM ; glycérol 8% ; DTT 0,1 mM ; Antiprotéases 1x (Roche)), à l'aide d'un désintégrateur de cellules (Constant Systems), suite à deux passages successifs à 20 kPsi (1379 bar). Les extraits sont ensuite clarifiés par centrifugation pendant 30 min à 20000g et 4°C. Ils sont ensuite additionnés de 75 µL de billes IgG-Sépharose (Sigma), pré-équilibrées sur roue, pendant 2 h 30 à 4°C, dans du tampon de lyse additionné de BSA acétylée à 20µg/mL. La rétention des protéines étiquetées est alors réalisée sur roue, à 4°C, pendant 24h. Ensuite, les billes sont lavées dans une colonne de filtration Mobicol (MoBiTec) par injection à la seringue de 10mL de tampon de lavage (Tris-HCl 20 mM, pH 7,4 ; NaCl 100 mM; 0.1% IGEPAL ; DTT 0,1 mM). A l'issue de cette étape, les billes sont reprises dans 100 µL de tampon (Tris-HCl 50 mM, pH 8 ; EDTA 0,5 mM ; DTT 1 mM), additionné de 75 unités de protéase TEV (produite au laboratoire). La protéolyse est réalisée sur roue, à 30°C, pendant 3 h dans les colonnes Mobicol, puis l'extrait est récupéré par élution à la seringue. Les 100 µL d'échantillons sont additionnés de bleu de dépôt, dénaturés à 96°C pendant 5 min, puis fractionnés par électrophorèse en gel SDS-PAGE 12,5%, Acryl-bisacryl (29:1). Après coloration au bleu coloïdale pendant 12 h à 4°C, et décoloration à l'eau ultra-pure, chaque piste est découpée de façon systématique tous les 2 mm. Les bandes, placées dans des plaques 96 puits, sont ensuite analysées par spectrométrie de masse MS/MS (J.M. Saliou, LSMBO, Strasbourg) (cf. Rothé *et al.*, Publication n°3).

### 2.5.3. Immunoprécipitation de la chromatine

Les expériences d'immunoprécipitation de la chromatine (ChIP) ont été réalisées à partir d'un protocole adapté de (Kuras & Struhl, 1999), en utilisant une quantité de cellules équivalente à 45 unités d'absorbance à 600 nm, prélevées en phase exponentielle de croissance.

#### Traitement des cellules au formaldéhyde :

Les cellules sont traitées par ajout de formaldéhyde à une concentration finale de 1%, afin d'induire un pontage covalente entre les protéines et l'ADN. Après une incubation de 15 min à T°C ambiante, l'effet du formaldéhyde est neutralisé par ajout de glycine à une concentration de 400 mM, pendant 5 min. Les cellules sont ensuite lavées deux fois dans du tampon (Tris/HCl 20 mM, pH 8) et une fois dans du tampon de lyse (HEPES/KOH 50 mM, pH 7,5 ; NaCl 150 mM ; EDTA 1 mM ; Triton 1% ; acide déoxycholique 0,1% ; SDS 0,05% ; Antiprotéases 1x (Roche)).

#### Extraction de la chromatine :

Les cellules sont reprises dans 500 µL de tampon de lyse, additionnés de 2 volumes de billes de verre (Sigma), puis elles sont agitées à l'aide d'un vortex pendant 10 séquences de 1min, interrompues par 2 min d'incubation dans la glace. Le microtube contenant l'échantillon est ensuite percé à sa base et le lysat est récupéré dans un second microtube par centrifugation 1 min à 3000g et 4°C. La chromatine est alors culottée par centrifugation à 20000g pendant 30 min, à 4°C, avant d'être reprise dans 1,5 mL de tampon lyse (HEPES/KOH 50 mM, pH 7,5 ; NaCl 150 mM ; EDTA 1 mM ; Antiprotéases 1x (Roche)).

#### Fragmentation de la chromatine :

La chromatine est fragmentée par traitement aux ultra-sons à l'aide d'un sonicateur Sonifer 250 (Branson). L'opération se déroule à 4°C, en réalisant 8 séquences de 10 sec de traitement (duty-cycle 50% ; Output 4), interrompues par des périodes de pause de 30 sec. Les

échantillons sont traités en absence de détergents, afin d'éviter la formation de mousse. En revanche, à l'issue de cette étape, 0,25 volume d'un tampon additif (HEPES/KOH 50 mM, pH 7,5 ; NaCl 150 mM ; EDTA 1 mM ; Triton 4% ; acide déoxycholique 0,4% ; SDS 0,2% ; Antiprotéases 1x (Roche)) est ajouté, afin de rétablir une composition équivalente à celle du tampon de lyse. Les extraits sont alors incubés sur roue, à 4°C pendant 30 min, afin de favoriser la solubilisation de la chromatine fragmentée. Puis, ils sont clarifiés par centrifugation pendant 10 min à 20000g et 4°C. Les surnageants de centrifugation sont distribués en 3 fractions de 400 µL (utilisées pour l'étape d'immunosélection) et 1 fraction de 200µL (utilisée pour constituer l'input). Les fractions peuvent éventuellement être incubées quelques secondes dans de l'azote liquide, puis stockées à -80°C.

#### Immunoprécipitation :

Chaque immunosélection est réalisée à partir de 400 µL de chromatine fragmentée. Ce volume est additionné de 20 µL de billes IgG-Sépharose (Sigma), pré-équilibrées sur roue, pendant 1 h à 4°C, dans du tampon de lyse additionné de BSA acétylée à 20 µg/mL. La rétention des protéines étiquetées est alors réalisée sur roue à 4°C, pendant 1 h 30. Les billes sont ensuite lavées par deux lavages successifs de 5 min, à T°C ambiante, dans 1 mL de tampon (HEPES/KOH 50 mM, pH 7,5 ; NaCl 500 mM ; EDTA 1 mM ; Triton 1% ; acide déoxycholique 0,1% ; SDS 0,05%), puis par un lavage de 5 min dans 1 mL de tampon (Tris-HCl 10 mM, pH8 ; LiCl 250 mM ; EDTA 1 mM ; IGEPAL 0,1% ; acide déoxycholique 0,5% ), enfin par un lavage final de 5 min dans 1 mL de tampon IPP150 (Tris-HCl 10 mM, pH 8 ; NaCl 150 mM ; IGEPAL 0,1%).

#### Isolement des fragments d'ADN co-immunoprécipités :

Après les lavages, les billes sont reprises dans 200 µL de tampon IPP150, additionné de SDS à une concentration de 0,5% et de 200 µg de protéinase K (Roche). Les échantillons sont incubés à 37°C pendant 1 h, avant d'être séparés des billes par centrifugation. La rupture des liaisons covalentes générées par le formaldéhyde est alors réalisée par incubation à 70°C pendant 12 h. A l'issue de cette étape, l'ADN est extrait par traitement avec un mélange phénol/chloroforme/alcool isoamylique (25:24:1). Puis, il est isolé par précipitation alcoolique en présence de 300 mM de NaCl, en présence de 1 µL de glycogène, avant d'être repris dans 20 µL d'eau ultra-pure et stocké, si nécessaire, à -80°C. Pour chaque échantillon, les 200 µL de chromatine fragmentée, correspondant aux inputs, sont traités de façon similaire, afin d'isoler la totalité de l'ADN fragmenté, présent avant l'immunosélection.

#### Analyse par PCR de l'ADN co-immunoprécipité :

L'ADN co-immunoprécipité est analysé par PCR à l'aide de couples d'oligonucléotides amorces, ciblant des *loci* spécifiques. Cette analyse est réalisée à l'aide de l'enzyme GoTaq (Promega), dans des conditions standard d'utilisation. Les réactions se déroulent dans un volume total de 25 µL, à partir de 1 µL d'échantillon dilué 5 fois pour les produits immunoprécipités et 500 fois pour les inputs. Après 20 cycles d'amplification, les produits obtenus sont analysés par électrophorèse sur gel d'agarose 1,5%.

## **2.6. Techniques d'étude de complexes macromoléculaires *in vitro***

### **2.6.1. Technique de retard sur gel**

Les interactions entre protéines recombinantes et ARN sont testées *in vitro* par la technique de retard sur gel. L'ARN radiomarqué au niveau de son extrémité 5', est au préalable dénaturé et

renaturé, par incubation successive 5 min à 65°C et 10 min à T°C ambiante. Pour chaque essai, environ 500 cpm d'ARN radiomarqué sont incubés sur glace, pendant 20 min, dans un volume de 5 µL de tampon D, comprenant 250 ng d'ARNt de levure et une concentration variable en protéine(s). A l'issue de cette incubation, un volume de bleu de dépôt natif est ajouté et les échantillons sont analysés par électrophorèse sur gel de polyacrylamide natif 6% (acrylamide/bisacrylamide 38:2), additionné de glycérol à une concentration de 0,5%. La migration est réalisée à 100 V pendant 2 h 30, puis le gel est séché et la radioactivité est révélée à l'aide d'un appareil de détection Typhoon (GE Healthcare), après exposition sur un écran amplificateur. Dans le cas d'une expérience de compétition, une quantité variable d'ARN compétiteur non radiomarqué est pré-incubé avec l'ARN radiomarqué, avant l'ajout des protéines recombinantes.

### 2.6.2. Technique d'empreintes enzymatiques

L'hydrolyse ménagée des ARN, a été réalisée avec 3 RNases différentes : (i) la RNase T1 qui hydrolyse spécifiquement les ARN en 3' des résidus guanines, préférentiellement dans les régions non appariées, (ii) la RNase T2 qui hydrolyse préférentiellement les régions en simple-brin, avec peu de spécificité vis-à-vis des nucléotides et qui, comme la RNase T1, libère des fragments avec des extrémités 3'-phosphate et 5'-hydroxyle, (iii) la RNase V1 qui hydrolyse préférentiellement les régions en double-brin et les régions structurées par empilement de bases, avec peu de spécificité vis-à-vis des nucléotides et qui libère des fragments avec les extrémités 3'-hydroxyle et 5'-phosphate.

Les complexes d'intérêt sont formés dans un volume de 10 µL, à partir de 5000 cpm d'ARN radiomarqué, dans les conditions décrites dans la section précédente. A l'issue de cette étape, le volume réactionnel est élevé à 11 µL, par ajout de 1 unité de RNase T1, 1 unité de RNase T2, ou  $2,5 \cdot 10^{-4}$  unités de RNase V1. Les échantillons sont alors incubés à T°C ambiante, pendant 6 min, puis la réaction d'hydrolyse est inhibée par ajout de 100 µL d'un mélange d'arrêt (H<sub>2</sub>O ultra-pure glacée ; ARNt de levure 200 ng/µL) et 100 µL d'un mélange phénol/chloroforme/alcool isoamylique (25:24:1). Après extraction, les ARN sont isolés par précipitation alcoolique en présence de 300 mM de NaCl, en présence de 1 µL de glycogène, avant d'être repris dans 10 µL d'eau ultra-pure et stockés, si nécessaire, à -80°C. En parallèle une échelle d'hydrolyse ménagée à la RNase T1 est réalisée en condition dénaturante afin de visualiser tous les résidus G du transcrit. Pour cela, 5000 cpm d'ARN radiomarqué et 2 µg d'ARNt de levure sont incubés 5 min à 65°C, dans 6 µL de tampon (Citrates de sodium 20 mM ; EDTA 1 mM ; Urée 7 M ; Xylène C. 0.05 % ; Bromophénol 0.05%). Puis, 2 µL de RNase T1 à 1 U/µL sont ajoutées et l'incubation à 65°C est prolongée de 10 min. De plus, une échelle d'hydrolyse alcaline ménagée est réalisée extemporanément afin de visualiser chaque position du transcrit. Pour cela, 5000 cpm d'ARN radiomarqué sont incubés pendant 2 min à 96°C en présence de 1 volume de tampon alcalin (Na<sub>2</sub>CO<sub>3</sub>/NaHCO<sub>3</sub> 100 mM, pH 9).

Enfin, environ 20 cpm d'ARN radioactifs, dénaturés pendant 2 min à 96°C, en présence de bleu de formamide, sont fractionnés par électrophorèse sur gel de polyacrylamide dénaturant 7%. La migration est réalisée sur un support de gel de séquençage, à une puissance de 100 W et une température limite de 50°C, pendant 2 h à 4 h. A l'issue de la migration, le gel est séché et la radioactivité est révélée à l'aide d'un appareil de détection Typhoon (GE Healthcare), après exposition sur un écran amplificateur.

### 2.6.3. Technique d'étude d'interactions protéiques par His pull-down

Cette approche a été utilisée pour tester l'interaction de protéines recombinantes porteuses d'une étiquette 6xHis, avec une protéine radiomarquée, transcrite et traduite en extrait cellulaire S30 d'*E. coli* (cf. section 2.1.4). La formation de complexes a été testée dans un volume de 20  $\mu$ L de tampon D, en présence de 4  $\mu$ L de protéine radiomarquée et de protéines recombinantes à une concentration de 2,5  $\mu$ M. Après 20 min d'incubation à 30°C, le volume total est élevé à 200  $\mu$ L par ajout de tampon (Tris-HCl 10 mM, pH 8 ; NaCl 150 mM; IGEPAL 0,1% ; Imidazole 5 mM) et 20  $\mu$ L de billes Cobalt-sepharose (GE Healthcare), pré-équilibrées dans du tampon IPP150 en présence de 20  $\mu$ g/mL de BSA acétylée. L'incubation est réalisée sur roue, à 4°C, pendant 1 h 30. A l'issue de cette étape, les billes sont lavées pendant 15 min, puis par deux lavages successifs de 5 min sur roue, à T°C ambiante, dans 1 mL de tampon. Enfin, les échantillons sont additionnés de bleu de dépôt, dénaturé à 96°C pendant 5 min, puis fractionnés par électrophorèse en gel SDS-PAGE 7,5%, Acryl-bisacryl. A l'issue de la migration, le gel est séché et la radioactivité est révélée à l'aide d'un appareil de détection Typhoon (GE Healthcare), après exposition sur un écran amplificateur.

### 2.6.4. Technique d'assemblage de RNP *in vitro* par His pull-down

Afin, d'étudier la formation de complexes de haut poids moléculaire entre un ARN radiomarqué et plusieurs protéines recombinantes, nous avons utilisé une approche d'assemblage *in vitro*, basée sur l'approche de pull-down. Après incubation des différents constituants, la formation d'une RNP peut en effet être vérifiée par rétention spécifique d'une protéine porteuse d'une étiquette 6xHis et détection de l'ARN radiomarqué. Pour cela, 10000 cpm d'ARN radiomarqués en 5' sont incubés dans un volume de 20  $\mu$ L de tampon D, en présence de 1  $\mu$ g d'ARNt de levure et de protéine(s) recombinante(s) à une concentration variable de 0,5 à 5  $\mu$ M. Après 30 min d'incubation à 30°C, le volume total est élevé à 200  $\mu$ L par ajout de tampon IPP150 (Tris-HCl 10 mM, pH 8 ; NaCl 150 mM; IGEPAL 0,1%), puis incubé sur roue, à T°C ambiante, en présence de 20  $\mu$ L de billes Nickel-sepharose (GE Healthcare), pré-équilibrées dans du tampon IPP150. A l'issue de cette étape, les billes sont lavées par 3 lavages successifs de 5 min sur roue, à T°C ambiante, dans 1 mL de tampon de lavage (Tris-HCl 10 mM, pH 8 ; NaCl 150 mM; IGEPAL 0,1% ; Imidazole 20 mM). Enfin, les billes sont reprises dans 300  $\mu$ L de tampon IPP150 et les ARN sont extraits par traitement avec un mélange phénol/chloroforme/alcool isoamylique (25:24:1). Après une étape de précipitation alcoolique en présence de 300 mM de NaCl, les ARN sont dénaturés pendant 2 min, à 96°C, dans du bleu de formamide, puis fractionnés sur gel de polyacrylamide dénaturant 7%. Une quantité d'ARN correspondant à 10% de la quantité de molécules utilisées par essai est déposée comme référence (Input). A l'issue de la migration, le gel est séché et la radioactivité est révélée à l'aide d'un appareil de détection Typhoon (GE Healthcare), après exposition sur un écran amplificateur.

### 2.6.5. Technique d'assemblage de RNP en extrait cellulaire

Afin de parer à l'insolubilité des protéines Nop56p et NOP58p lorsqu'elles sont produites sous forme recombinante chez *E. coli*, la reconstitution de particules à boîtes C/D a été réalisée en extrait cellulaire, en adaptant le protocole utilisé par l'équipe de R. Lührmann chez l'homme (Watkins 2002). Les différents essais ont été réalisés à partir de cellules de levure WT ou  $\Delta$ RSAl, cultivées jusqu'en phase exponentielle de croissance et exprimant de manière

ectopique les protéines flag-Nop1p, flag-Nop56p ou flag-Nop58p, à partir du vecteur p413 (*HIS3*). Chaque extrait est réalisé à partir d'une quantité de cellules équivalente à 30 unités d'absorbance à 600 nm, comme décrit dans la section (2.3.1). La concentration relative de chaque extrait en protéine porteuse de l'étiquette flag est estimée par Western-blot (cf. section 2.4.1). Ainsi, pour chaque série d'expérience, la quantité d'extrait utilisé correspond à l'équivalent de 15  $\mu$ L de l'extrait le moins concentré. Dans un premier temps, les extraits sont pré-incubés pendant 10 min à 30°C, dans un volume total de 20  $\mu$ L, en présence de tampon D, d'ATP à 1 mM et de 20  $\mu$ M d'un petit ARN porteur du motif B/C du snoRNA U3, qui en s'associant à la protéine Snu13p de l'extrait contribue à épuiser sa forme libre (Cléry 2007). En parallèle, 10000 cpm de snoRNA U14 radiomarqué et 60 nmol de protéine Snu13p recombinante sont co-incubés dans un volume de 5  $\mu$ L, selon les conditions décrites dans la section (2.6.1), afin de former un pré-complexe Snu13p-U14. Ce dernier, ainsi que 20 unités de RNAsine (Promega) sont ajoutés dans l'extrait. Après une incubation de 25 min à 30°C, 10% de chaque échantillon sont prélevés afin de constituer les inputs, tandis que les 90% restant sont utilisés pour immunosélectionner les protéines porteuses de l'étiquette flag. Pour cela, le volume de chaque échantillon est élevé à 300  $\mu$ L par ajout de tampon IPP150 (Tris-HCl 10 mM, pH 8 ; NaCl 150 mM; IGEPAL 0,1%) et de 20  $\mu$ L de billes anti-flag-Sépharose (Sigma), pré-équilibrées dans du tampon IPP150 en présence de 20  $\mu$ g/mL de BSA acétylée. L'incubation est réalisée sur roue, à 4°C, pendant 2 h 30. A l'issue de cette étape, les billes sont lavées par 3 lavages successifs de 5 min sur roue, à T°C ambiante, dans 1 mL de tampon de lavage (Tris-HCl 10 mM, pH 8 ; NaCl 150 mM; IGEPAL 0,1% ; Imidazole 20 mM). Enfin, les billes sont reprises dans 300  $\mu$ L de tampon IPP150 et les ARN sont extraits par traitement avec un mélange phénol/chloroforme/alcool isoamylique (25:24:1). Après une étape de précipitation alcoolique en présence de 300 mM de NaCl, les ARN sont dénaturés pendant 2 min, à 96°C, dans du bleu de formamide, puis fractionnés sur gel de polyacrylamide dénaturant 7%. A l'issue de la migration, le gel est séché et la radioactivité est révélée à l'aide d'un appareil de détection Typhoon (GE Healthcare), après exposition sur un écran amplificateur.



RESULTATS  
&  
DISCUSSION



## CHAPITRE 1

### Mise en évidence d'une machinerie protéique dédiée à la biogenèse des RNP de la famille L7Ae

#### 1. Présentation des travaux

Comme nous l'avons abordé au cours de l'introduction, plusieurs facteurs potentiellement impliqués dans la biogenèse des snoRNP à boîtes C/D ont été décrits dans la littérature et sont suspectés de pouvoir jouer un rôle dans l'assemblage des particules. Parmi ces facteurs, nous avons focalisé notre attention sur les protéines Rvb1p/TIP49, Rvb2p/TIP48 et Pih1p. En association avec la protéine Tah1p, ces dernières, ont en effet été identifiées au sein du complexe R2TP, vraisemblablement impliqué dans la régulation de l'activité de HSP90 chez la levure *S. cerevisiae* (Zhao *et al.*, 2005). La perspective de retrouver au sein d'un même complexe, plusieurs facteurs d'intérêt, nous a donc poussé à vérifier leurs liens avec les différents composants des snoRNP, ainsi que l'implication potentielle de la chaperonne HSP90. Ces travaux ont été menés en parallèle dans la levure et chez l'homme, pour qui la protéine SPAGH, homologue fonctionnelle de la protéine Tah1p, a été décrite pour la première fois. Les protéines Rsa1p et NUFIP, qui partagent le domaine commun PEP, ont également été intégrées à cette étude. Leurs liens avec la protéine Snu13p/15.5K et avec les autres facteurs d'intérêt ont été étudiés.

Ces travaux, réalisés dans le cadre de l'ANR blanche *RNPASSEMBLY BLAN06-2\_134682*, ont été menés en collaboration avec les équipes d'E. Bertrand (IGM, Montpellier) et d'A. Krol (IBMC, Strasbourg). Ils ont notamment abouti à une publication dans la revue *Journal of Cellular Biology* (Boulon *et al.*, 2008). Les résultats présentés dans cet article indiquent l'existence d'un large réseau d'interaction reliant physiquement les composants des snoRNP, les membres du complexe R2TP et le facteur Rsa1p/NUFIP, identifié lors de ces travaux. De façon intéressante, la protéine Rsa1p/NUFIP semble jouer un rôle de plateforme d'assemblage, puisque son interaction avec la protéine Snu13p/15.5K facilite le recrutement des autres protéines sur les pré-RNP ARN-Snu13p/15.5K. Les données obtenues chez l'homme indiquent que le complexe R2TP pourrait quant à lui favoriser l'intervention de la

chaperonne HSP90, dont l'inhibition par un traitement à la geldanamycine s'est avérée délétère pour l'accumulation de certains composants des snoRNP. Enfin, l'extension de cette étude aux snoRNP à boîtes H/ACA, à la snRNP U4, au complexe télomérase et aux mRNP SECIS, révèle que ce processus d'assemblage pourrait être commun à l'ensemble des RNP contenant des protéines de la famille L7Ae. Ce spectre d'action élargi, observé uniquement chez l'homme, repose entre autre sur la capacité de la protéine NUFIP à interagir à la fois avec les protéines 15.5K, NHP2 et SBP2.

## 2. Résultats

Les résultats concernant cette partie ont majoritairement été publiés dans l'article *The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery*. Quelques données supplémentaires, obtenues de façon ultérieure, seront présentées en complément.

### 2.1. Publication n°1

Ma contribution personnelle à cet article a concerné les expériences réalisées chez la levure. J'ai ainsi participé aux expériences de double-hybride, de triple-hybride et de triple-hybride *bridged*. J'ai réalisé la caractérisation des souches de levure  $\Delta RSA1$ ,  $\Delta PIH1$  et  $\Delta TAH1$  par la technique de Northern-blot. Enfin, j'ai également pris en charge la mise au point et la réalisation des expériences de retard sur gel et de co-immunoprécipitation, afin d'étudier l'association de la protéine Rsa1p avec les snoRNA à boîtes C/D.

---

---

Publication n°1

**The Hsp90 chaperone controls the biogenesis of  
L7Ae RNPs through conserved machinery**

Séverine Boulon, Nathalie Marmier-Gourrier, Bérengère, Pradet-Balade, Laurence Wurth,  
Céline Verheggen, Beáta E. Jády, Benjamin Rothé, Christina Pescia, Marie-Cécile Robert,  
Tamás Kiss, Barbara Bardoni, Alain Krol, Christiane Branlant, Christine Allmang,  
Edouard Bertrand, and Bruno Charpentier

**J Cell Biol. 2008 Feb 11;180(3):579-95**

---

---



# The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery

S  verine Boulon,<sup>1</sup> Nathalie Marmier-Gourrier,<sup>2</sup> B  reng  re Pradet-Balade,<sup>1</sup> Laurence Wurth,<sup>3</sup> C  line Verheggen,<sup>1</sup> Be  ta E. J  dy,<sup>4</sup> Benjamin Roth  ,<sup>2</sup> Christina Pescia,<sup>1</sup> Marie-C  cile Robert,<sup>1</sup> Tam  s Kiss,<sup>4</sup> Barbara Bardoni,<sup>5</sup> Alain Krol,<sup>3</sup> Christiane Branlant,<sup>2</sup> Christine Allmang,<sup>3</sup> Edouard Bertrand,<sup>1</sup> and Bruno Charpentier<sup>2</sup>

<sup>1</sup>Institute of Molecular Genetics of Montpellier, Centre National de la Recherche Scientifique, Unit   Mixte de Recherche 5535, Montpellier Cedex 5, France

<sup>2</sup>Maturation des ARN et Enzymologie Mol  culaire, Unit   Mixte de Recherche 7567, Centre National de la Recherche Scientifique, Universit   Henri Poincar  , Nancy Universit  , Facult   des Sciences et Techniques, 54506 Vandoeuvre-les-Nancy, France

<sup>3</sup>Architecture et R  activit   de l'ARN, Universit   Louis Pasteur de Strasbourg, Centre National de la Recherche Scientifique, Institut de Biologie Mol  culaire et Cellulaire, 67084 Strasbourg, France

<sup>4</sup>Laboratoire de Biologie Mol  culaire Eucaryote, Universit   Paul Sabatier, 31062 Toulouse, France

<sup>5</sup>Institute of Genetics and Molecular and Cellular Biology, 67404 Illkirch, France

**R**NA-binding proteins of the L7Ae family are at the heart of many essential ribonucleoproteins (RNPs), including box C/D and H/ACA small nucleolar RNPs, U4 small nuclear RNP, telomerase, and messenger RNPs coding for selenoproteins. In this study, we show that Nufip and its yeast homologue Rsa1 are key components of the machinery that assembles these RNPs. We observed that Rsa1 and Nufip bind several L7Ae proteins and tether them to other core proteins in the immature particles. Surprisingly, Rsa1 and Nufip also link assembling RNPs with

the AAA + adenosine triphosphatases hRvb1 and hRvb2 and with the Hsp90 chaperone through two conserved adaptors, Tah1/hSpagh and Pih1. Inhibition of Hsp90 in human cells prevents the accumulation of U3, U4, and telomerase RNAs and decreases the levels of newly synthesized hNop58, hNHP2, 15.5K, and SBP2. Thus, Hsp90 may control the folding of these proteins during the formation of new RNPs. This suggests that Hsp90 functions as a master regulator of cell proliferation by allowing simultaneous control of cell signaling and cell growth.

## Introduction

Noncoding RNPs are ancient devices that play fundamental roles in the cell. The L7Ae family of proteins comprises a set of related RNA-binding proteins that share a homologous RNA recognition domain (Koonin et al., 1994). Members of the L7Ae protein family occur in both Archaea and Eukarya. They have been found in numerous essential RNPs (Fig. 1), including the large ribosomal subunit, U4 spliceosomal small nuclear RNP (snRNP), box C/D and H/ACA small nucleolar RNPs (snoRNPs), small Cajal body RNPs (scaRNPs), telomerase, archaeal small RNP, and selenoprotein messenger RNPs (mRNPs). In humans, canonical members are ribosomal proteins L7A, L30, 15.5K,

hNHP2, and SBP2 (Koonin et al., 1994; Watkins et al., 2000; Allmang et al., 2002).

The snoRNPs play essential roles in ribosomal RNA maturation both during the cleavage steps and nucleotide modifications (for reviews see Kiss, 2002; Matera et al., 2007). The snoRNPs are divided into two major classes: the box C/D snoRNPs that catalyze ribose 2'-O-methylation and the box H/ACA snoRNPs that mediate pseudouridine formation. Each snoRNP contains a small nucleolar RNA (snoRNA) and a set of common snoRNP proteins. The box C/D snoRNPs contain 15.5K (Snu13 in yeast), Nop56, Nop58, and the methyltransferase fibrillarin (Nop1 in yeast). Box H/ACA snoRNPs contain hNHP2, Gar1, Nop10, and the pseudouridine synthase dyskerin (Cbf5 in yeast). Vertebrate cells also contain related RNPs like the scaRNPs that also share box C/D or H/ACA RNP proteins. Although most scaRNPs are involved in the posttranscriptional modification of spliceosomal small nuclear RNAs (snRNAs), the telomerase box H/ACA scaRNP functions in telomeric DNA synthesis. Selenoprotein synthesis requires cotranslational recoding of in-frame UGA codons. In eukaryotes, this process involves the assembly of RNA-protein complexes at specific stem

S. Boulon, N. Marmier-Gourrier, B. Pradet-Balade, and L. Wurth contributed equally to this paper.

Correspondence to Christine Allmang: c.allmang@ibmc.u-strasbg.fr; Edouard Bertrand: Edouard.Bertrand@igmm.cnrs.fr; or Bruno Charpentier: Bruno.Charpentier@maem.uhp-nancy.fr

Abbreviations used in this paper: IP, immunoprecipitation; mRNP, messenger RNP; qPCR, quantitative PCR; scaRNP, small Cajal body RNP; SECIS, selenocysteine insertion sequence; snoRNA, small nucleolar RNA; snoRNP, small nucleolar RNP; snRNA, small nuclear RNA; snRNP, small nuclear RNP; TPR, tetratricopeptide repeat; Y2H, yeast two hybrid; Y3H, yeast three hybrid.

The online version of this article contains supplemental material.

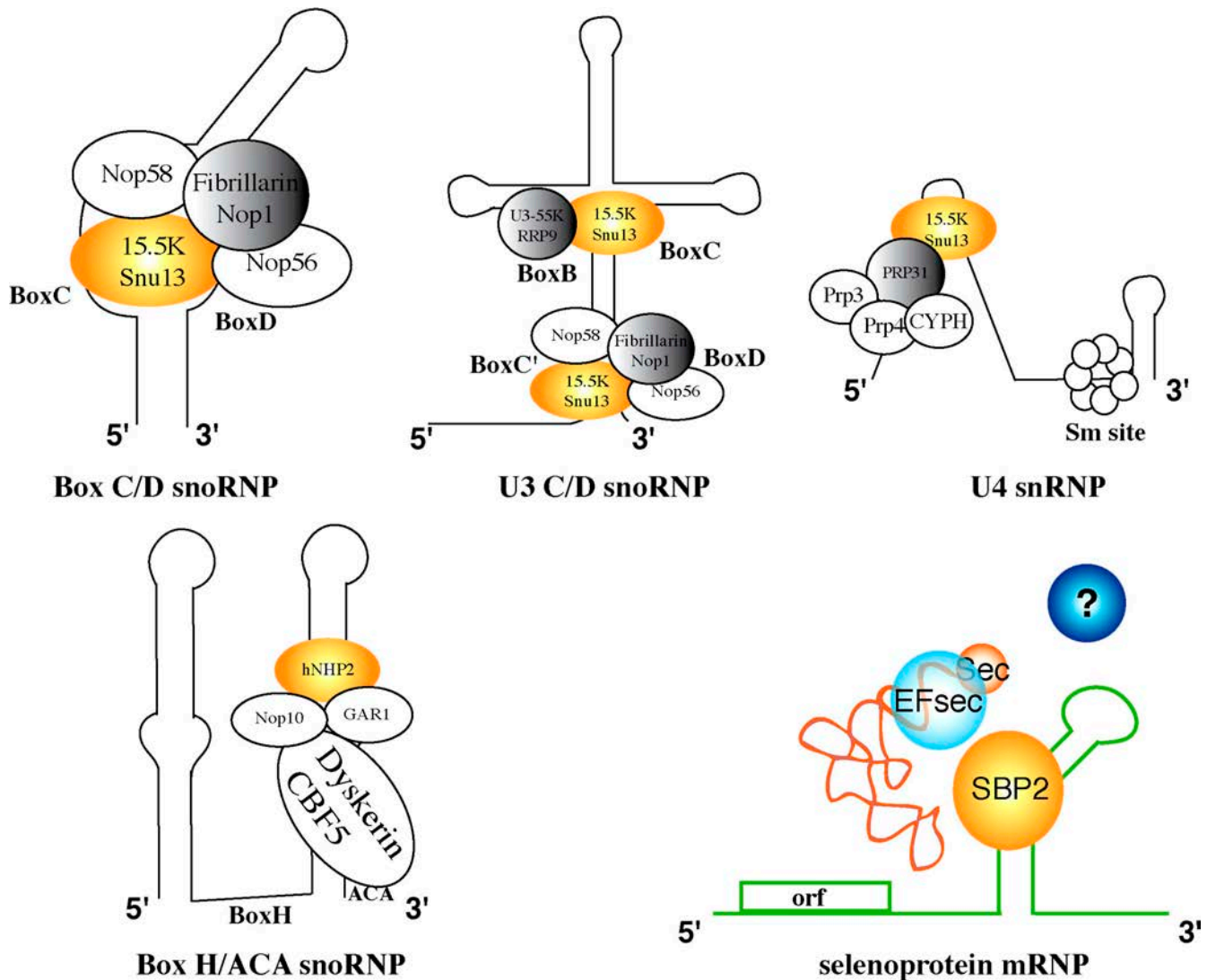


Figure 1. **Composition of L7Ae RNPs.** The various RNPs containing the L7Ae family members (yellow) studied here are depicted. The yeast names are indicated below the human names.

loops located in the 3' untranslated region of selenoprotein mRNAs, termed the selenocysteine insertion sequence (SECIS). SBP2 specifically binds the SECIS element and recruits the specialized elongation factor EFsec (Copeland et al., 2000; Fagegaltier et al., 2000).

Detailed structural information is available for several members of the L7Ae family (Vidovic et al., 2000; Li and Ye, 2006). They bind to a common RNA structure called the K-turn or K-loop motif, and binding exposes further interaction surfaces on both the RNA and the protein. Thus, upon RNA binding, L7Ae proteins recruit additional factors to construct complex RNPs. Remarkably, some L7Ae proteins participate in several RNPs. This is the case for Snu13/15.5K, which is at the heart of the U4 snRNP, box C/D snoRNPs, and a specialized B/C structure of the U3 snoRNP (Watkins et al., 2000). Although Snu13/15.5K has structurally similar binding sites in these RNAs, it recruits different sets of protein partners (or core proteins), such as Nop56, Nop58, and fibrillarin in box C/D snoRNPs (Watkins et al., 2002), PRP31 and the cyclophilin H-hPRP4-hPRP3

complex for the U4 snRNP (Nottrott et al., 2002), and RRP9 (U3-55K in the human) for the B/C motif of the U3 snoRNP (Granneman et al., 2002). Importantly, the binding of Snu13/15.5K is required for association of the other core proteins. For instance, hPRP31 binds neither 15.5K nor U4 snRNA alone, but it does bind the 15.5K-U4 complex, in which it directly contacts both the RNA and protein (Liu et al., 2007).

Although some RNPs can be reconstituted *in vitro*, in living cells they are formed through assembly and maturation pathways that are far more complex than initially anticipated (Yong et al., 2004; for review see Matera et al., 2007). The number of assembly factors often exceeds the number of proteins present in the mature particle. Assembly of most RNPs requires dynamic remodeling of the maturing RNP particle, and it is often accompanied by a complicated series of transient intermolecular interactions involving stably associated core components and transiently interacting processing factors. Although RNP assembly requires the proper folding of both the RNA and protein components, protein chaperones have not yet been implicated



in this process. Interestingly, some factors involved in yeast snoRNP biogenesis have been independently identified as cochaperones for Hsp90. Indeed, the R2TP complex, which is composed of Tah1, Pih1, and the two AAA + ATPases Rvb1 and Rvb2, is a cochaperone for Hsp90 (Zhao et al., 2005), and both Rvb2 and Pih1 are involved in box C/D snoRNP biogenesis (King et al., 2001; Gonzales et al., 2005), raising the possibility that Hsp90 and the R2TP complex are involved in the production of box C/D snoRNPs. Although no human equivalent of the R2TP complex has been identified, it has been shown that immature human U3 particles contain the homologues of yeast Rvb1 and Rvb2 (hRvb1 and hRvb2, also called Tip48 and Tip49; Newman et al., 2000).

Hsp90 has attracted a lot of interest because it is involved in cancer and in the control of nuclear receptors and protein kinases (Pearl and Prodromou, 2006; Caplan et al., 2007). Hsp90 is highly conserved from eubacteria to eukarya and performs essential cellular functions. Compared with other chaperones, Hsp90 has two unique features. First, it usually binds target proteins at a late stage of their folding in a near-native state. Second, it appears as a specialized chaperone that is mainly involved in the control of signal transduction. Like other chaperones, Hsp90 recognizes its substrate through cochaperones and cofactors.

In this study, we describe a conserved molecular machinery for the assembly of RNPs of the L7Ae family. This machinery exists in yeast and human cells and is composed of the R2TP proteins and an adaptor called Nufip (Rsa1 in yeast). Remarkably, we show that Nufip holds together the core components of the mature RNP and directly links RNP assembly with Hsp90.

## Results

### Nufip and its yeast homologue Rsa1 bind 15.5K and Snu13 through a short peptide motif

To characterize new factors involved in box C/D snoRNP biogenesis, we performed yeast two-hybrid (Y2H) and yeast three-hybrid (Y3H) screens. First, we screened a yeast genomic library using a minimal U3 B/C motif as bait. Out of  $10^4$  clones screened, we found two positive ones that coded for Rsa1, a nuclear protein previously found in a synthetic lethal screen with Dbp6 (Kressler et al., 1999). Interestingly, when Rsa1 was tested in Y3H against various RNAs, its binding correlated with that of Snu13 (Fig. S1, available at <http://www.jcb.org/cgi/content/full/jcb.200708110/DC1>), suggesting that it formed a complex with Snu13 bound to target RNAs (see the section Nufip and Rsa1 form a ternary complex...). Second, we screened a human liver library using 15.5K as bait. Upon screening of  $10^7$  clones, we detected 23 positives, three of which coded for Nufip, a nuclear RNA-binding protein previously shown to interact with nuclear isoforms of fragile X mental retardation protein (FMRP; Bardoni et al., 1999). No homology between Nufip and Rsa1 could be identified by BLAST search; however, a careful comparison of their sequence identified a conserved stretch of 32 amino acids (PEP; Fig. 2 A). Although no other homologies could be detected between the two proteins, BLAST searches performed with this oligopeptide sequence identified potential homologues

in fungi, plants, and animals (Fig. S2). To identify the domains of Nufip involved in the interaction with 15.5K, we generated several mutants: an N-terminal fragment truncated before its Zn finger (Nufip-Nter), a central fragment containing the Zn finger and the PEP motif (Nufip-Zn), and a C-terminal region (Nufip-Cter). Y2H tests revealed that the 15.5K interaction domain matched the central region containing the PEP motif (Fig. 2 B). Remarkably, deletion of PEP (Nufip- $\Delta$ PEP) abolished binding, whereas the PEP motif alone was sufficient for interaction with 15.5K. To test whether Rsa1 also interacts with Snu13, full-length and mutant forms of Rsa1 were tested in Y2H assays (Fig. 2 B). We found that Rsa1 indeed interacted with Snu13 and that the interaction site mapped to its PEP motif.

To verify that Y2H interactions reflected a direct physical association of Nufip and Rsa1 with 15.5K and Snu13, *in vitro* binding assays were performed (Fig. 2 C). We found that 15.5K translated *in vitro* in a bacterial S30 lysate interacted with immobilized GST-Nufip and that recombinant Snu13p produced in *Escherichia coli* bound to GST-Rsa1p.

To test whether Nufip associates with 15.5K in human cells, we performed co-immunoprecipitation (IP) assays. A construct expressing GFP-15.5K was transfected into HeLa cells, extracts were incubated with anti-GFP beads, and the selected proteins were analyzed by Western blots. Nufip was found in the GFP-15.5K pellet but not in the control (Fig. 2 D). Collectively, these results show that Nufip associates with 15.5K *in vivo* and that Nufip and Rsa1 are related proteins that physically interact with 15.5K and Snu13.

### Nufip binds hNHP2 and SBP2

The binding of Nufip to 15.5K prompted us to test whether it could also interact with other members of the L7Ae family. By using a Y2H system, we tested the interaction of Nufip with the box H/ACA snoRNP core protein hNHP2, with SBP2, which associates with mRNAs coding for selenoproteins, and also with the human ribosomal protein hL30, another member of this family. We found that Nufip interacted with hNHP2 and SBP2 but not with hL30 (Fig. 2 B). Although the PEP domain was required for these interactions, it was not sufficient to bind hNHP2, and it only weakly bound SBP2. *In vitro* binding assays confirmed that these interactions reflected direct physical associations (Fig. 2 C). Next, we transfected HeLa cells with a GFP-hNHP2 construct, and extracts were immunoprecipitated with an anti-GFP antibody. Western blot analysis revealed efficient co-IP of Nufip by GFP-hNHP2 (Fig. 2 D). Similarly, when nuclear HeLa cell extracts were passed through a column with immobilized anti-SBP2 peptide antibodies, we found that Nufip copurified with SBP2 (Fig. 2 D). These results show that Nufip associates not only with 15.5K but also with two other members of the L7Ae family, hNHP2 and SBP2.

### Nufip and Rsa1 form a ternary complex with 15.5K/Snu13 and target RNAs

Rsa1 was initially found in a Y3H screen with an RNA bait that bound Snu13, suggesting that it interacted with an Snu13-RNA complex. To verify this, we reconstituted the complex *in vitro* using gel-shift assays (Fig. 3 A, left). Snu13 alone could bind

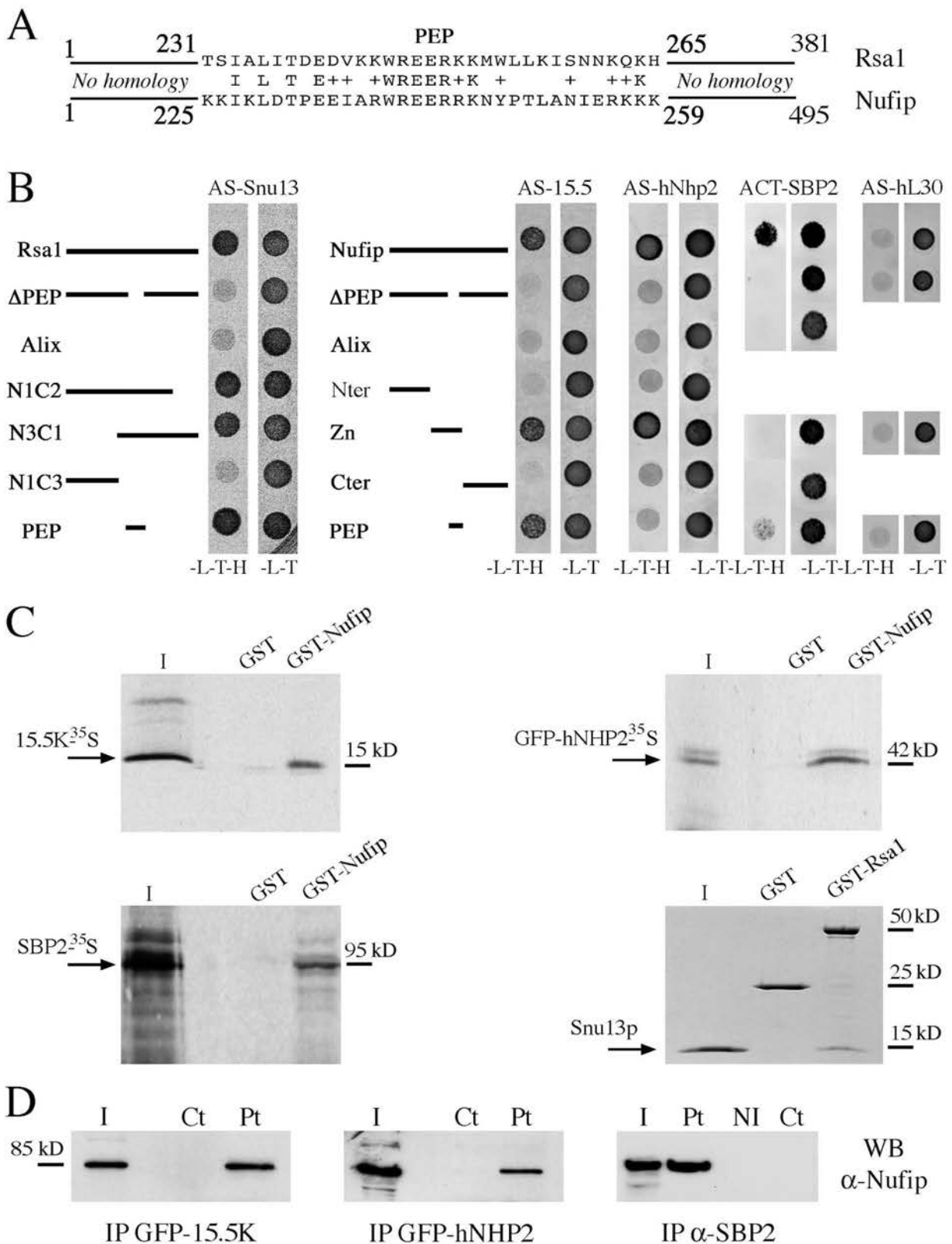


Figure 2. **Nufip and Rsa1 are related proteins that bind L7Ae family members through a short peptide motif.** (A) Alignment of Rsa1 with Nufip. The conserved motif, named PEP, is represented. (B) Y2H tests between the indicated proteins. Rsa1, Nufip, and their various domains are represented by lines. Growth on -Leu -His -Trp (-L -H -T) indicates an interaction. (C) GST pull-down experiments. 15.5K and SBP2 were translated in vitro in bacterial S30 lysate and assayed for binding to the indicated recombinant GST fusions. GFP-hNHP2 was translated in rabbit reticulocyte lysate, but similar results were obtained with bacterial lysate (not depicted). Snu13 was purified from *E. coli* and stained with Coomassie blue. (D) In vivo co-IP assays. GFP-15.5K and GFP-hNHP2 were expressed in HeLa cells and purified with anti-GFP antibodies. Endogenous SBP2 was immunopurified from nuclear HeLa cell extracts using anti-peptide antibodies. Bound proteins were analyzed by Western blots with anti-Nufip antibodies. I, input (10% of total); Pt, pellet; NI, preimmune control serum; Ct, control beads without antibodies. The sizes of the various products are indicated on the side of the gels.

U14 snoRNA, whereas two His-tagged fragments of Rsa1 (N3C1 and yPEP) did not. However, when Snu13 was added, complexes of higher molecular weight were obtained with both Rsa1 fragments. These complexes were supershifted by anti-His antibodies (Fig. 3 A), demonstrating that they contained N3C1 and yPEP. To check whether Nufip could also associate with Snu13–RNA complexes, we used Y3H assays (Fig. 3 A, right). As expected, we found that Nufip could specifically interact with the B/C motif in a PEP-dependent manner. These results demonstrate that Nufip and Rsa1 can form ternary complexes with Snu13 bound to RNA.

#### **Nufip associates with U4 snRNA, box C/D and H/ACA snoRNAs, a minimal U3 B/C RNA, and SBP2 mRNPs**

To test whether Nufip also associated with 15.5K RNPs in vivo, we performed a series of co-IP assays. First, we analyzed whether Nufip was associated with box C/D snoRNAs. HeLa cells were transfected with mouse U8, human U13, and rat U3 genes. Extracts were subjected to IP with anti-Nufip antibodies, and the pellets were analyzed by RNase protection with RNA probes specific for the 3'-terminal regions of the test RNAs. The U3, U8, and U13 snoRNAs are synthesized from their own promoters, and their precursors contain a short 3' extension that is trimmed during RNP assembly (Verheggen et al., 2002; Boulon et al., 2004). Remarkably, we observed that Nufip was associated with all precursors and mature forms of U3, U8, and U13 (Fig. 3 C). Identical results that were obtained with GFP-Nufip and anti-GFP antibodies excluded nonspecific binding of the antibody. To test whether Nufip also binds snoRNAs processed from introns, we transfected HeLa cells with a construct expressing an artificial C/D snoRNA inserted in the second intron of the  $\beta$ -globin gene (dBB). The anti-Nufip antibody efficiently precipitated this snoRNA (Fig. 3 C), indicating that a C/D motif is sufficient to direct Nufip binding.

U3 contains two binding sites for 15.5K (Watkins et al., 2000): the C'/D motif that is equivalent to the C/D motif in other snoRNAs and the B/C motif that recruits U3-55K. IP experiments showed that Nufip bound a U3 mutated in its C'/D motif but not a U3 mutant in which both the B/C and C'/D motifs had been inactivated (Fig. 3 C). 15.5K is also found in U4 snRNP, where it interacts with hPRP31 (Watkins et al., 2000). Again, co-IP experiments found endogenous Nufip and Nufip-GFP associated with both endogenous U4 and a transiently expressed tagged U4 RNA (Fig. 3 D).

The interaction of Nufip with hNHP2 suggested that it could also bind box H/ACA snoRNAs. RNAs immuno-precipitated with anti-Nufip antibodies from HeLa cell extracts were thus analyzed by RNase protection with probes specific for the U19 H/ACA snoRNA and for the U3 box C/D snoRNA as a control (Fig. 3 E). Nufip antibodies coprecipitated both U3 and U19 snoRNAs. Although this assay could not discriminate between mature and precursor forms of U19, the fact that Nufip is excluded from nucleoli in HeLa cells (unpublished data) is consistent with an association with forms of U19 not yet involved in rRNA maturation.

As Nufip was shown to interact with SBP2, we tested whether it could also bind selenoprotein mRNAs. Extracts were

prepared from 293FT cells expressing GFP-Nufip and RNAs coimmunoprecipitated with anti-GFP antibodies and were analyzed by RT-PCR. Anti-GFP antibodies coprecipitated endogenous selenoprotein glutathione peroxidase 4 and type 2 deiodinase mRNAs but not the  $\beta$ -actin mRNA that was tested as a negative control (Fig. 3 F). These results demonstrate that Nufip associates with box C/D and box H/ACA snoRNAs, a B/C RNA, U4, and SBP2 mRNPs. Nufip binds precursor forms of C/D snoRNAs, and binding is dependent on the presence of a binding site for 15.5K. Finally, Rsa1 also interacts with U3 precursors in yeast (Fig. 3 B).

#### **Nufip and Rsa1 are required for snoRNA biogenesis**

To further document its role in snoRNA biogenesis, Nufip was depleted from HeLa cells using two different siRNAs (Fig. 4 A). Northern blot analysis showed that depletion of Nufip reduced the level of the endogenous U3 snoRNA by 50%. As demonstrated by quantitative PCR (qPCR) analysis, depletion of Nufip had little effect on accumulation of the U4 snRNA. Although its depletion slightly reduced the levels of the U14 box C/D and U19 box H/ACA snoRNAs, it had a more dramatic effect on the accumulation of telomerase RNA that carries an H/ACA domain.

Next, we investigated the effect of removing Rsa1 in yeast. Primer extension analyses showed that the levels of the U14 and U3 box C/D snoRNAs were reduced in the  $\Delta$ Rsa1 strain by 66% and 42%, respectively (Fig. 4 B). As in mammalian cells, the level of U4 remained almost unaffected, and a slight decrease was observed in the levels of box H/ACA snoRNAs. To gain further insights into the role of Rsa1, we analyzed a truncated version of the U3 snRNA (U3del) that is incorporated into snoRNPs with low efficacy and, therefore, accumulates precursor RNPs that are devoid of Nop1, Nop56, and Nop58 and are stabilized by La and LSm proteins (Kufel et al., 2000). The lack of Rsa1 resulted in a greater accumulation of U3 precursors (three- to fourfold when normalized to mature U3del levels; Fig. 4 C), indicating a role for Rsa1 in U3 assembly and maturation.

#### **Nufip and Rsa1 tether 15.5K and Snu13 to the other core proteins of U4, box C/D, and box B/C RNPs**

Because Nufip and Rsa1 bound Snu13 and 15.5K and were involved in RNP assembly, we tested whether they could interact with other core proteins contained in the mature RNP. In Y2H assays (Table S1, available at <http://www.jcb.org/cgi/content/full/jcb.200708110/DC1>), we observed that Rsa1 and Nufip associated with yeast and human PRP31, respectively. In addition, Rsa1 interacted with Nop58, and Nufip associated with U3-55K and fibrillarin. Binding of Nufip to the endogenous Snu13 did not contribute to these interactions because the C-terminal domain of Nufip that does not bind Snu13 and 15.5K also interacted with hPRP31, U3-55K, and fibrillarin. To verify these results, we performed GST pull-down experiments with in vitro-translated <sup>35</sup>S-labeled proteins. In agreement with the Y2H results, we found that Nufip directly bound hPRP31, U3-55K,

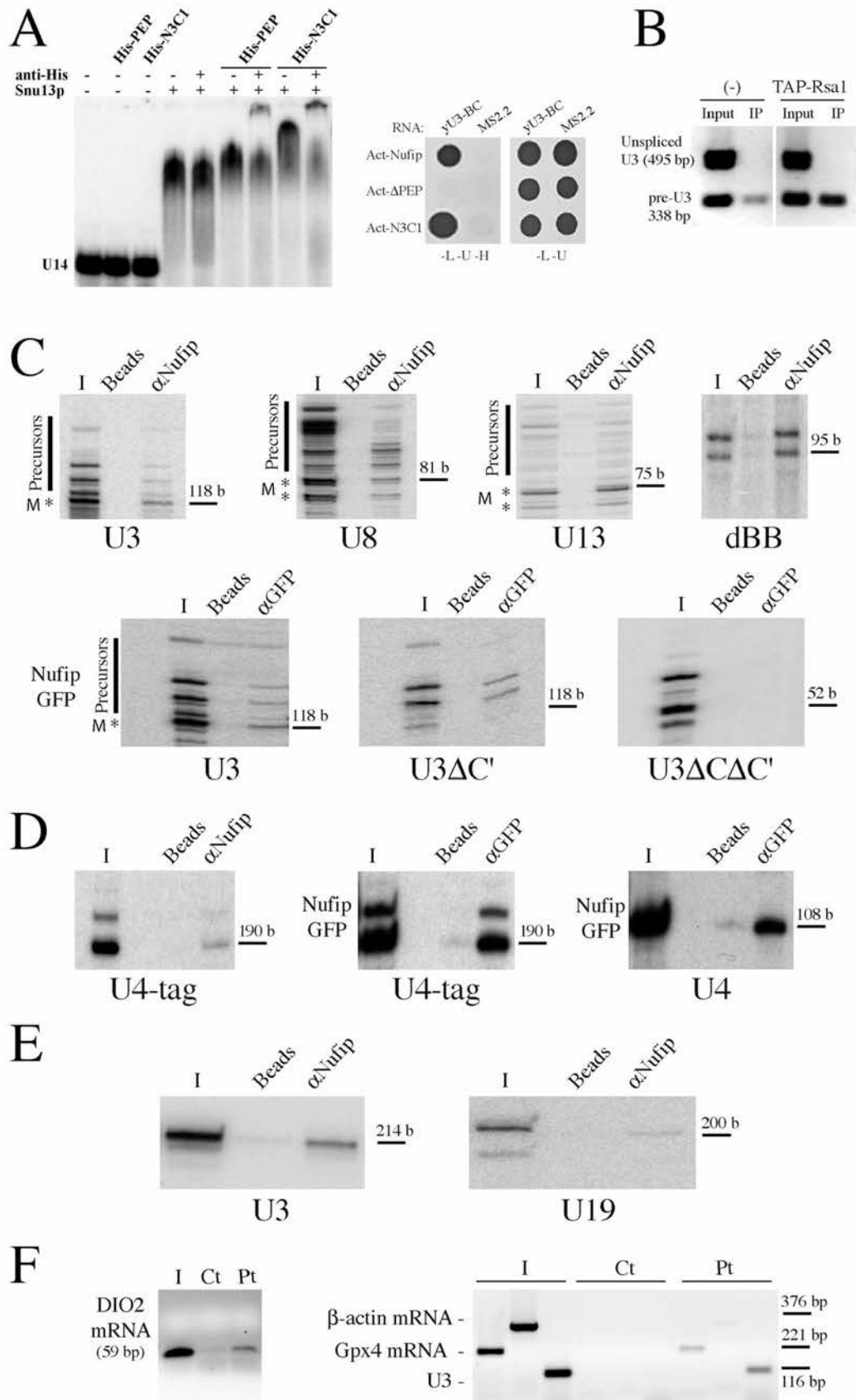


Figure 3. Nufip binds box C/D and H/ACA snoRNAs, a B/C-containing RNA, U4 snRNA, and mRNAs coding for selenoproteins. (A) In vitro interactions of Rsa1 and Nufip with Snu13-RNA complexes. (left) Gel-shift assays show that the Rsa1 N3C1 and PEP domains interact with Snu13 bound to RNA. Radiolabeled yeast U14 snoRNA was incubated with the indicated recombinant proteins and anti-His antibodies when indicated. (right) Y3H assays show that Nufip interacts with yU3-B/C RNA in a PEP-dependent manner. Plate -Leu -Ura (-L -U) shows the growth of the test strain. Growth on -Leu -Ura -His (-L -U -H)

and fibrillarin (Fig. 5, A and B). This suggested that Nufip can tether 15.5K to other core proteins. To test this more directly, we introduced it as a bridge in Y2H assays (Fig. 5 C). Although 15.5K fused to the Gal4 DNA-binding domain did not produce Y2H interactions with fibrillarin, hPrp31, or U3-55K fused to the Gal4-activating domain, introduction of Nufip into this strain allowed growth on selective media, indicating that the proteins formed a ternary complex. It was unlikely that endogenous yeast proteins contributed to the complex because little cross-reaction was observed between the yeast and human proteins in Y2H assays (unpublished data). In particular, ySnu13 did not interact with hPRP31 and fibrillarin. In addition, neither transformation of an empty vector nor expression of an Nufip- $\Delta$ PEP mutant produced an interaction. Thus, we conclude that Nufip can tether 15.5K with fibrillarin, hPrp31, and U3-55K.

Next, we tested whether Nufip could also link Snu13–RNA complexes to core RNP proteins. Using bridged Y3H assays, we found that Nufip could promote the interaction of hPRP31 and a fibrillarin C-terminal domain with yU3B/C RNA (Fig. 5 D). It should be noted that under in vitro conditions, B/C RNAs are not competent to form ternary complexes with 15.5K and either hPRP31 or fibrillarin (Schultz et al., 2006). This suggests that RNP complexes tethered by Nufip bind RNA primarily through Snu13/15.5K and that the other core proteins become associated by virtue of their binding to Nufip. Thus, we concluded that Nufip and Rsa1 could act as adaptors to bridge core RNP proteins to Snu13–RNA complexes.

#### Yeast R2TP proteins function in U3 assembly, and their human homologues, together with Hsp90, associate with immature U3 and U4 particles and SBP2 mRNPs

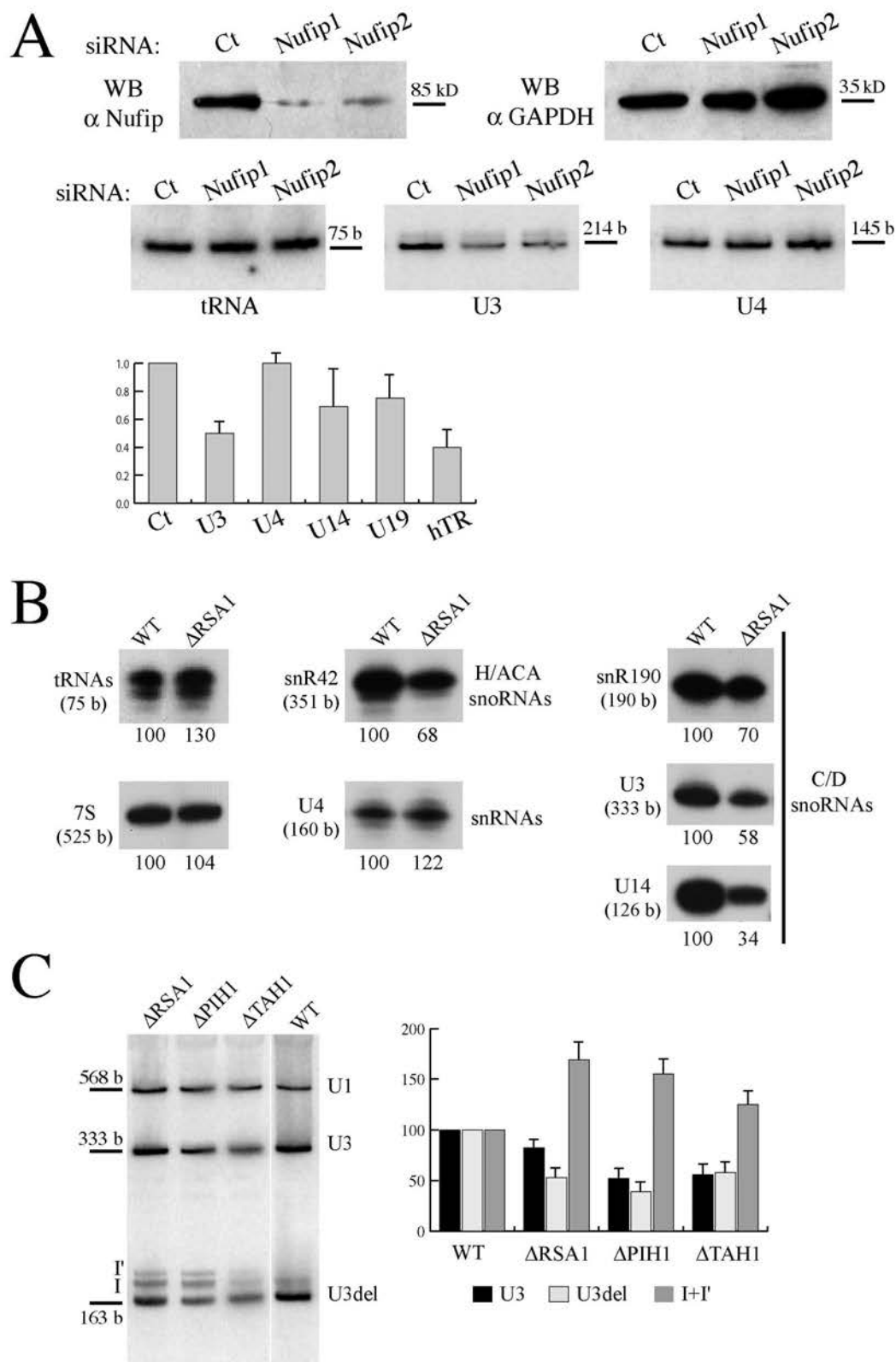
To define more precisely the function of Nufip and Rsa1, we characterized the composition of immature U3 and U4 particles in more detail. As described in the Introduction, yeast box C/D snoRNP accumulation and localization involve Rvb2 and Pih1 proteins that, together with Rvb1 and the yeast Hsp90 cochaperone Tah1, form the R2TP complex (Zhao et al., 2005). To confirm the role of Pih1 in the assembly of yeast U3 snoRNP and to demonstrate that Tah1 is also involved in this process, we analyzed the biogenesis of the truncated form of U3 (U3del) in strains lacking these proteins. As shown in Fig. 4 C, such strains had reduced levels of mature U3 but increased levels of U3 precursors, indicating a defect in the formation of U3 snoRNP.

The R2TP complex has not been identified in human cells, but a recent analysis of the human Hsp90 proteome showed that hRvb1, hRvb2, and the human homologue of Pih1 were associated with Hsp90 (Te et al., 2007). This raises the possibility that the R2TP complex may also occur and function as an Hsp90 cofactor in higher eukaryotes. The yeast Tah1 is a small protein mostly composed of two tetratricopeptide repeat (TPR) motifs. Given that the human genome encodes several TPR-containing proteins, a sequence comparison failed to identify the human functional homologue of Tah1. However, upon a closer examination of systematic interaction databases, we found that a *Drosophila melanogaster* TPR protein called Spaghetti had been reported to interact with both dHsp90 and dPih1 in Y2H screens (Giot et al., 2003). In addition, the human homologue of Spaghetti (FLJ21908, hereafter named hSpagh) was found in the Hsp90 proteome (Te et al., 2007), suggesting that it may represent the functional homologue of yeast Tah1. To test this possibility, we generated an antibody against hSpagh and performed IP analyses. Indeed, we found that hRvb1, hRvb2, and hPih1 were immunoprecipitated by this antibody (Fig. S3, available at <http://www.jcb.org/cgi/content/full/jcb.200708110/DC1>).

To determine whether these proteins associate with assembling U3 particles, we performed co-IP assays. Antibodies against Hsp90, hSpagh, and hPih1 were used to immunoprecipitate their respective proteins from extracts prepared from HeLa cells transiently expressing rat U3 snoRNA. The coprecipitated RNAs were analyzed by RNase protection (Fig. 6 A). Hsp90, hSpagh, and hPih1 bound precursor and mature forms of U3. Similar results were obtained with transiently expressed epitope-tagged proteins, confirming that they specifically interacted with U3 RNA (unpublished data). In similar experiments performed with U4 snRNA, we observed that antibodies against hPih1, hSpagh, Hsp90, hRvb1, and hRvb2 also coprecipitated the U4 snRNA (Fig. 6 B).

To test whether R2TP proteins can interact with other RNPs of the L7Ae family, we checked their association with SBP2. IP of Hsp90, GFP-Pih1, GFP-hRvb1, and GFP-hRvb2 coprecipitated SBP2 (Fig. 6 C). In addition, endogenous hRvb1 coimmunopurified with SBP2 from native HeLa nuclear extracts (Fig. 6 D). This suggests that these proteins are involved in the assembly of SECIS-containing RNPs. Collectively, these data provide a direct and remarkable link between the protein folding machinery and assembly of L7Ae RNPs. These observations also support the idea that the human homologues of yeast R2TP proteins function as Hsp90 cofactors during RNP assembly.

indicates a positive interaction. (B) In vivo association of Rsa1 with U3 precursors in yeast. Extracts from TAP-Rsa1 or wild-type (–) isogenic strains were purified on IgG beads and analyzed by RT-PCR with primers specific for U3 precursors. (C) In vivo interactions of Nufip with rat U3B.7 and other box C/D snoRNAs. HeLa cells were transfected with the indicated snoRNA gene either alone (top) or with an Nufip-GFP vector (bottom). Extracts were purified with anti-Nufip (top) or anti-GFP (bottom) antibodies or beads as a control, and bound RNAs were analyzed by RNase protection. U3 $\Delta$ C' and U3 $\Delta$ C $\Delta$ C' are mutated in the C' and in the C and C' boxes. dBB is an artificial intronic C/D snoRNA (see Results). I, input (10% of total); M\*, mature species. (D) In vivo association of Nufip with endogenous U4 snRNA (right) and a transfected, tagged U4 snRNA (left and middle). Legend as in C. (E) In vivo binding of Nufip with H/ACA snoRNAs. HeLa nuclear extracts were immunoprecipitated with anti-Nufip antibodies and analyzed by RNase protection with the indicated probes. Legend as in C. (F) Nufip associates with mRNAs coding for selenoproteins. Anti-GFP IP of extracts of 293FT cells transfected with SBP2 alone (lanes Ct) or together with Nufip-GFP (lane Pt). U3 and  $\beta$ -actin are positive and negative controls, and type 2 deiodinase and glutathione peroxidase 4 are two selenoproteins. Input, 10% of total.



**Figure 4. Nufip and Rsa1 are required for the production of snoRNAs.** (A) Nufip is required for the production of U3. HeLa cells were treated with siRNA duplexes against Nufip. Nufip depletion was verified by Western blotting (top), and the levels of the indicated RNAs were monitored by Northern blotting (bottom) or by qPCR (graph). The qPCR values represent ratios of the indicated RNA to GAPDH mRNA and are expressed as fractions of the levels obtained with the control siRNAs (Ct). (B) Primer extension was used to measure the levels of small RNAs in yeast cells lacking Rsa1 ( $\Delta$ Rsa1). Numbers represent the amount of RNA relative to wild type (W303). (C) Processing of a modified U3del gene was analyzed in strains deleted for Rsa1, Pih1, and Tah1 (BY4741 background). This U3 gene accumulates precursor species (I and I'), and this defect is more pronounced in mutant strains. (left) Northern blot. (right) Quantification of each species after normalization to levels in wild-type cells. Error bars represent SD.

### **Inhibition of Hsp90 reduces the levels of newly synthesized U3, U4, and telomerase RNA and leads to the loss of 15.5K, hNHP2, SBP2, and hNop58**

The aforementioned interactions suggested that Hsp90 could participate in RNP assembly; thus, we tested the effect of inhibiting Hsp90 on the production of L7Ae RNPs. The U3 and U4 RNPs have very long half-lives in human cells (Fury and Zieve, 1996), and they are essential RNPs, making the analysis of cells depleted from these RNPs problematic. In addition, inhibition of Hsp90 for long periods may result in indirect effects as a result of its pleiotropic functions. To specifically analyze newly synthesized RNAs, we thus resorted to a transient transfection approach. Human 293 cells were cotransfected with vectors encoding GFP as an internal control and either rat U3, tagged U4, or telomerase RNAs. 4 h after adding the DNA, cells were incubated with geldanamycin, a drug that inhibits the ATPase activity of Hsp90 and prevents folding of client proteins. After 16 h of treatment, the levels of transiently expressed RNAs were measured by RNase protection or qPCR. Remarkably, we found that the ectopically expressed RNAs under-accumulated in geldanamycin-treated cells (Fig. 7 A), indicating an important role for Hsp90 in the biogenesis of these RNPs. Interestingly, the transiently expressed U4 was more severely affected than the endogenous RNA, indicating that Hsp90 inhibition preferentially affected newly synthesized RNPs. This suggests that Hsp90 is involved in RNP assembly rather than in the maintenance of already formed RNPs.

Because Hsp90 is a protein chaperone, one likely possibility is that it acts by controlling the folding of core RNP proteins or assembly factors. Client proteins for Hsp90 often become unstable when its ATPase activity is inhibited by geldanamycin (Pearl and Prodromou, 2006; Caplan et al., 2007). To test which proteins could be Hsp90 substrates, we analyzed the effect of this drug on the accumulation of newly synthesized proteins using a similar transfection approach. Human 293 cells were cotransfected with vectors expressing GFP-tagged proteins and GFP alone as an internal control and were incubated with the drug for 4 h after adding the DNA, and protein levels were monitored after 16 h of expression. Remarkably, we found that some GFP-tagged proteins failed to accumulate when Hsp90 was inhibited (Fig. 7 B). This was the case for all L7Ae proteins (15.5K, hNHP2, and SBP2) and for hNop58. A mild effect was also seen for hNop56 and hPrp31, whereas U3-55K, fibrillarin, dyskerin, Nufip, hPih1, hRvb1, and hRvb2 appeared to be unaffected (unpublished data). Altogether, this indicates an essential role of Hsp90 in the biogenesis of C/D snoRNP, U4 snRNA, and telomerase RNPs. It also suggests that Hsp90 acts by controlling the folding of several core proteins during RNP assembly.

### **hSpagh links Hsp90 to hPih1 and Nufip**

To understand how Hsp90 is recruited to assembling L7Ae RNPs and is presented to its potential clients, we analyzed the interactions of the proteins involved by performing systematic Y2H interaction assays with the human and yeast proteins (Fig. 8 A and Table S1). Consistent with previous results (Zhao et al., 2005), we found that yeast Hsp90 bound Tah1 that, in turn, associated with Pih1. Rsa1 appeared to make extensive contacts with

the R2TP complex, as it interacted with Pih1, Rvb1, and Rvb2. For the human proteins, we defined a similar set of interactions. Hsp90 interacted with hSpagh, which bound hPih1. hSpagh was additionally connected to hRvb2 that bound hRvb1. Similar to Rsa1, Nufip was also tightly connected to the human homologues of the R2TP proteins because it interacted with hPih1, hRvb1, and hRvb2. Interactions between the human proteins were unlikely to be mediated by endogenous yeast factors because there was little cross-reaction between the yeast and human proteins in Y2H assays (unpublished data). To confirm that the Y2H interactions were direct, we performed GST pull-down experiments with proteins translated *in vitro* in a bacterial S30 lysate. The results confirmed the interactions of hSpagh with Hsp90, hPih1, and hRvb2 as well as the association of Nufip with hPih1, hRvb1, and hRvb2 (Fig. 8 A). We conclude that the human R2TP proteins form intermolecular interactions similar to those formed by their yeast counterparts and that hSpagh can bridge Hsp90 to hPih1 and Nufip.

### **hPih1 is associated with SBP2, hNop58, and hNop56**

Because yeast Nop58 also forms a Y2H interaction with Pih1, we tested whether human Pih1 also binds snoRNP proteins. Indeed, hNop58 and hNop56 translated *in vitro* in a rabbit reticulocyte lysate-bound hPih1 in GST pull-down assays (Fig. 8 B). However, this was not the case when these proteins were translated in a bacterial lysate (unpublished data), suggesting that the interaction is indirect or requires protein modification. Interestingly, in Y2H assays, hPih1 was also able to interact with SBP2 (Table S1). The direct physical association of SBP2 and hPih1 was confirmed by *in vitro* binding assay. SBP2 translated *in vitro* in a bacterial S30 lysate interacted with immobilized GST-hPih1 (Fig. 8 B). Thus, hPih1 can also interact with hNop58, hNop56, and SBP2.

### **R2TP proteins associate with core snoRNP proteins in large cytoplasmic complexes**

To test whether some assembly factors and core RNP proteins could associate together within large complexes, HeLa S100 extracts were separated onto linear 10–30% glycerol gradients and analyzed by Western blotting (Fig. 9 A and Fig. S4, available at <http://www.jcb.org/cgi/content/full/jcb.200708110/DC1>). As previously described, hRvb1 peaked in the middle of the gradient in complexes >670 kD (Makino et al., 1998), and a similar pattern was found for hRvb2. Nufip, hSpagh, hNop58, and fibrillarin peaked in slightly lighter fractions but were also present in heavier fractions where hRvb1 and hRvb2 accumulated. To test whether this cosedimentation reflected a physical association, these fractions were immunoprecipitated with anti-fibrillarin antibodies. Indeed, hRvb1 was found to be associated with fibrillarin in these large complexes (Fig. 9 B).

## **Discussion**

### **Assembly of L7Ae-type RNPs involves Nufip and a conserved set of proteins**

In this work, we have characterized a conserved machinery that is involved in the assembly of several L7Ae RNPs, including box C/D

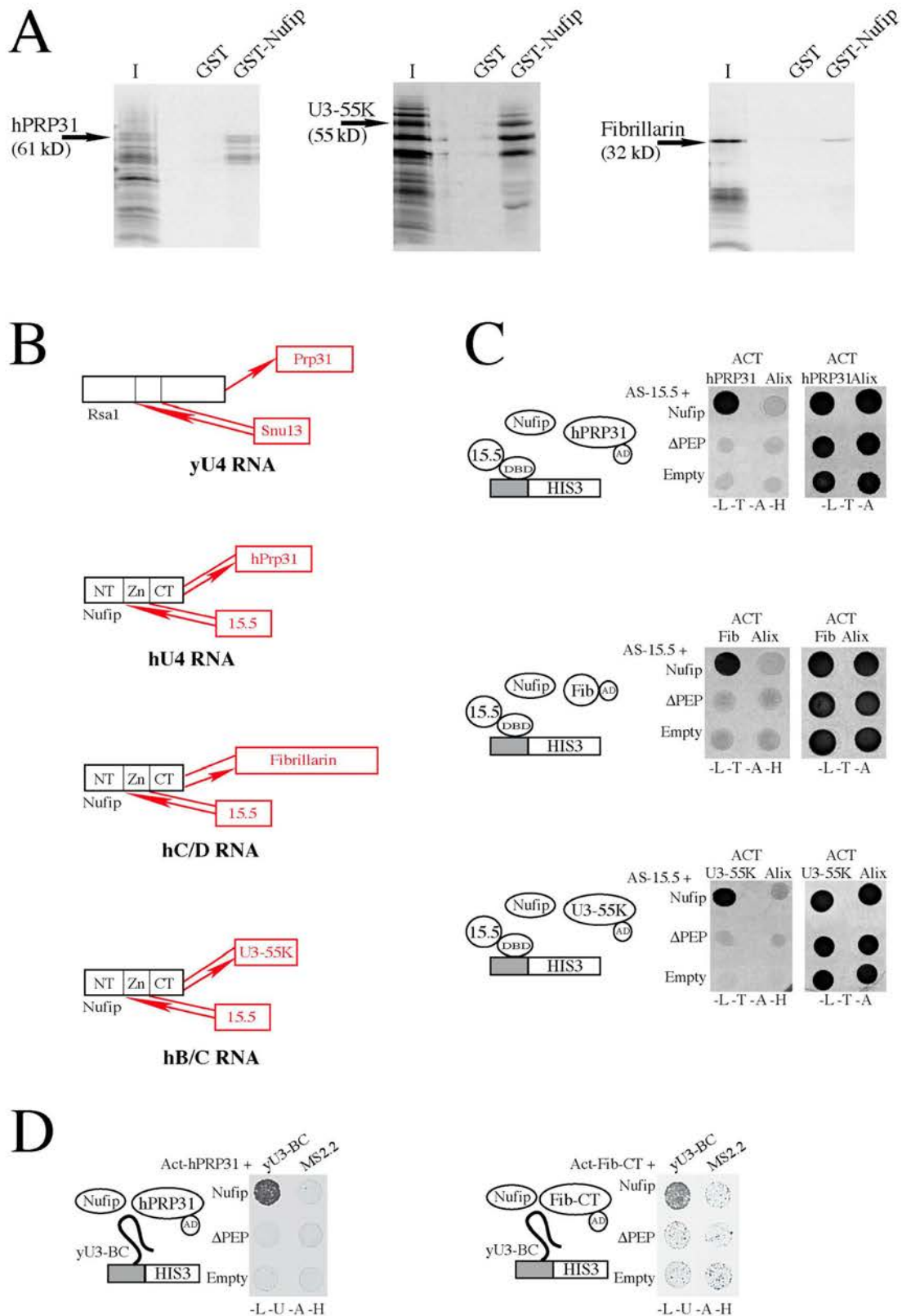
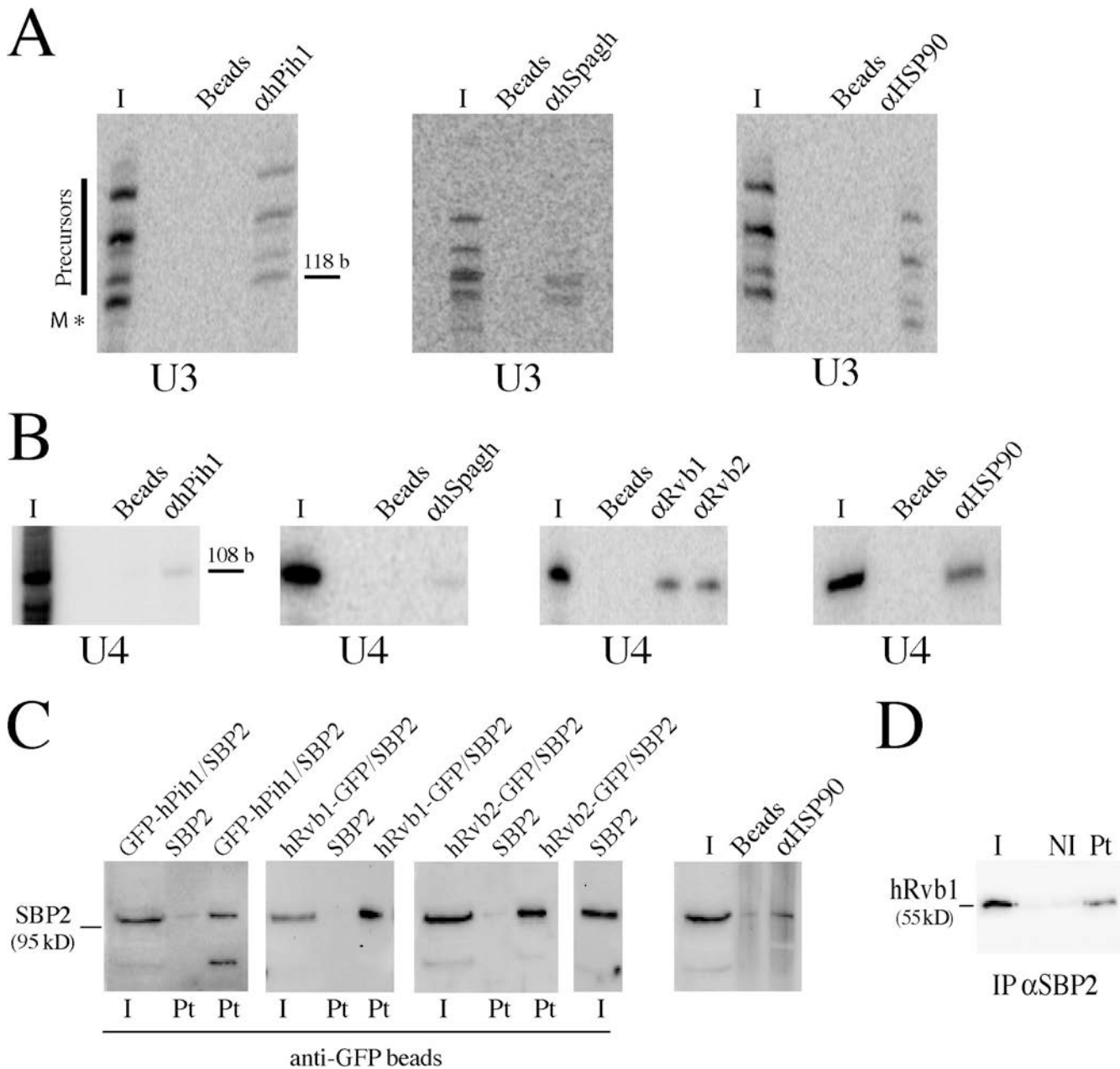


Figure 5. **Nufip tethers 15.5K to hPrp31, U3-55K, and fibrillarlin.** (A) Nufip interacts with hPrp31, U3-55K, and fibrillarlin in GST pull-down experiments. The indicated proteins were translated in vitro in bacterial S30 lysate and assayed for binding to recombinant GST-Nufip. I, input (10% of total). Arrows indicate the full-length protein. The other bands represent incomplete and read-through translation products that occur in the bacterial lysate. (B) Summary of the interactions obtained between Nufip and U4 or box C/D core proteins. Y2H interactions are indicated by arrows, and physical interactions observed in GST pull-downs are indicated by straight lines. (C) Nufip links 15.5K to fibrillarlin, hPrp31, and U3-55K in a bridged Y2H assay. Y2H strains containing the indicated pAS and pACT test plasmids were transformed with an ADE2 plasmid expressing Nufip, a mutant lacking the PEP domain, or no protein. Transformants were plated on triple or quadruple dropout media. Growth on  $-Leu -Trp -His -Ade$  ( $-L -T -H -A$ ) indicates the formation of a



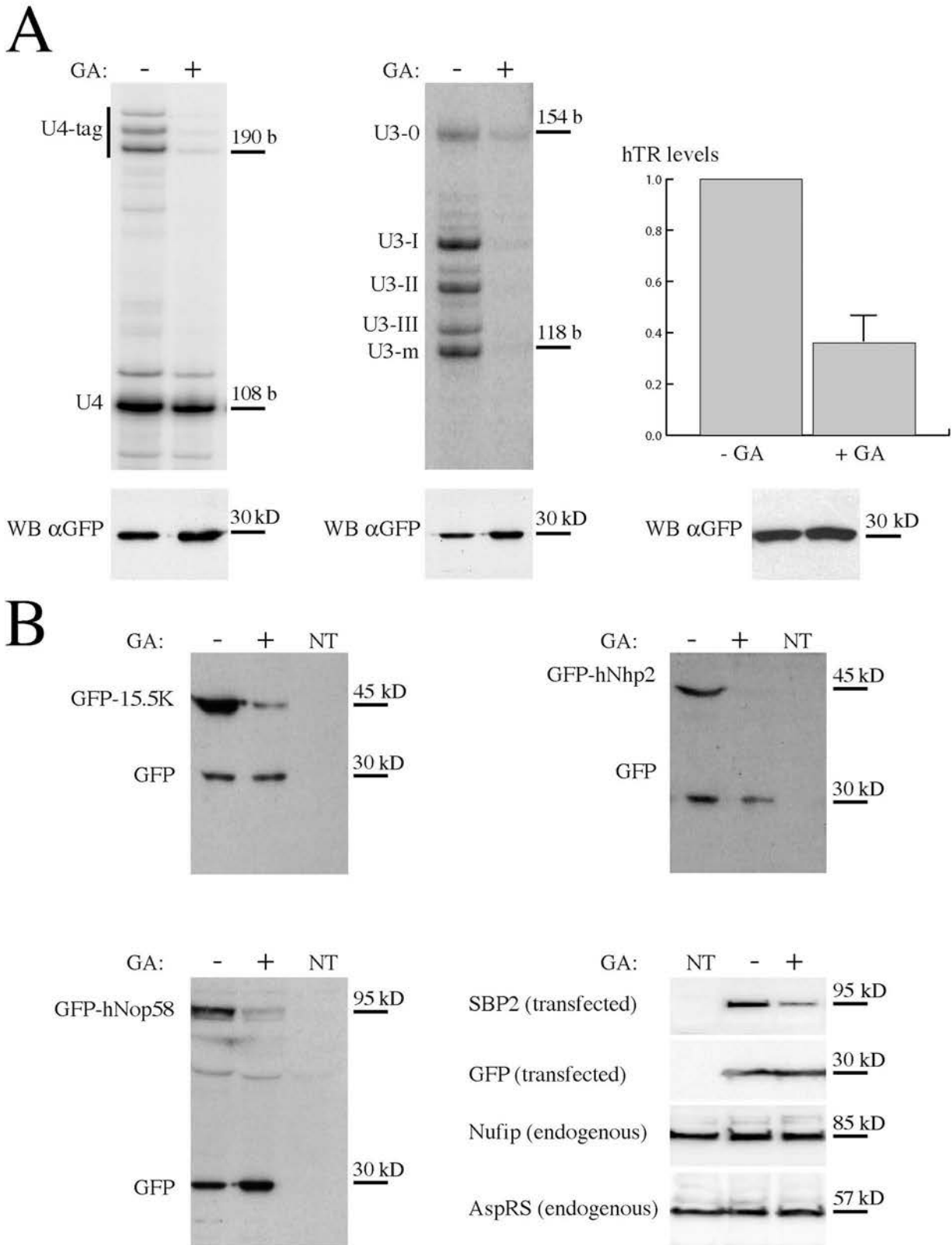


**Figure 6. Hsp90 and human homologues of R2TP proteins bind U3, U4, and SBP2.** (A) U3. Vectors expressing rU3B.7 were transfected in HeLa cells, extracts were immunoprecipitated with the indicated antibody, and RNAs bound to antibody-coated beads or beads alone as a control were analyzed by RNase protection. Precursors and mature forms of U3 are indicated by a bar and an asterisk, respectively. I, input (3% of total). (B) U4. Legend as in A except that binding was tested against the endogenous U4 RNA. I, input (10% of total). (C) SBP2. 293FT cells were transfected with vectors expressing SBP2 alone or with GFP-hPih1, hRvb1-GFP, or hRvb2-GFP, and extracts were immunoprecipitated with anti-GFP (left) or with anti-Hsp90 antibodies (right). Precipitates were analyzed with anti-SBP2 antibodies. I, input (5% of total); Pt, pellets. (D) hRvb1 is present in immunopurified SBP2 complexes. HeLa nuclear extracts were purified with anti-SBP2 antibodies (Pt) or nonimmune serum (NI) and analyzed by Western blotting with anti-hRvb1 antibodies. I, input (0.5% of total).

and H/ACA snoRNPs, the U4 snRNP, and, most strikingly, also the SECIS mRNPs that code for selenoproteins. In yeast, this machinery is composed of Rsa1 and four R2TP proteins: Tah1, Pih1, Rvb1, and Rvb2. In humans, it contains homologues of these proteins: Nufip, hSpagh, hPih1, hRvb1, and hRvb2. This conserved RNP assembly machine possesses several remarkable features: (1) it binds several

L7Ae family members (Snu13/15.5K, hNHP2, and SBP2) through Nufip and Rsa1, and this directs assembly factors onto nascent RNPs; (2) it tethers L7Ae proteins to other core proteins, thus likely increasing the efficiency and specificity of their assembly; and (3) it is associated with chaperones of the Hsp90 family, indicating that RNP assembly is closely linked to protein folding.

complex between the indicated proteins. (D) Nufip tethers hPRP31 and the C-terminal domain of fibrillarin to yU3B/C RNA. Y3H strains deleted for Rsa1 were transformed with the indicated plasmid and an ADE2 vector expressing Nufip, a mutant lacking the PEP domain, or no protein. Growth on quadruple dropout media indicates the formation of a complex between the indicated proteins and yU3B/C RNA.



**Figure 7. Inhibition of Hsp90 prevents the production of L7Ae RNPs and destabilizes hNop58, 15.5K, hNHP2, and SBP2.** (A) Geldanamycin prevents the accumulation of U4, U3, and telomerase RNAs. HeLa cells were cotransfected with a GFP expression vector and either a tagged U4 gene (U4 tag), a rat U3 gene, or a human telomerase gene and were incubated with 2  $\mu$ M geldanamycin for 16 h. The RNAs produced were then analyzed by RNase protection (left) or by qPCR (right). The different U3 precursors (U3-0 to U3-III) and the mature U3 (U3-m) are indicated. (bottom) Western blots against GFP as controls. The graph represents the levels of hTR RNA normalized to levels of GAPDH mRNA and expressed as the fraction of untreated cells. Values are the averages of three experiments. The error bar represents SD. (B) Geldanamycin inhibits the accumulation of GFP-15.5K, GFP-hNHP2, GFP-hNop58, and SBP2. 293 cells were cotransfected with vectors expressing the indicated proteins and GFP as a control and were incubated with 2  $\mu$ M geldanamycin for 16 h. Cell extracts were analyzed by Western blotting with anti-GFP antibodies (left) or with the indicated combination of antibodies. NT, nontransfected cells.

### **Hsp90 is required for RNP biogenesis and acts by controlling protein folding during RNP assembly**

Hsp90 plays an essential role in the formation of U3, U4, and telomerase RNP, as its inhibition with geldanamycin prevents their accumulation. Client proteins for Hsp90 are often degraded in cells treated by geldanamycin. We showed that this drug inhibits the accumulation of newly synthesized hNop58, 15.5K, hNHP2, and SBP2, thus suggesting that Hsp90 is required to chaperone their folding during RNP assembly. One possibility would be that these proteins adopt a particular conformation when present in the mature RNP complex but that this conformation is unstable or prone to aggregation when the proteins are unassembled and would thus require Hsp90. In this view, the role of chaperones and assembly factors would be to stabilize unassembled proteins, to bring them together on the nascent RNP, and to facilitate the transition to the mature complex. Thus, with the help of scaffolding factors, formation of an intermolecular complex would mimic intramolecular folding.

Although our data are consistent with a role for Hsp90 during RNP assembly, it is equally possible that it exerts its essential folding activity at an earlier step, before the core proteins bind RNA. In this case, the proper protein fold may be stabilized by assembly factors such as hPih1 and Nufip. It is also possible that some proteins like 15.5K bind RNA first and that assembly factors and chaperones would then join the nascent RNP. Finally, the effect of Hsp90 could be indirect, and binding of Hsp90 to U3 precursors may not be functionally relevant or may even be caused by reassociation of the chaperone during the extraction or IP procedure (Kittur et al., 2006).

We found that hPih1 associates with hNop58 in reticulocyte lysates, which are cytoplasmic extracts devoid of nucleolar RNAs. Similarly, Nufip is bound to 15.5K in cytoplasmic S100 extracts (unpublished data). Thus, we favor a model in which core RNP proteins would follow an assembly line in which newly translated proteins would first bind chaperone and co-chaperone complexes in the cytoplasm and then translocate into the nucleus, where they would assemble with nascent RNPs.

### **hSpagh, hPih1, and Nufip are probable cofactors for Hsp90**

Hsp90 recognizes many of its client proteins through adaptors, which frequently interact with the chaperone through their TPR domains, as in the case of Tah1 (Zhao et al., 2005) and possibly hSpagh as well. We observed that in human cells, RNP assembly factors remain stable when Hsp90 is inhibited, indicating that they are Hsp90 cofactors rather than targets. Many such cofactors are not essential but merely facilitate the recognition of client proteins. In agreement, although Tah1, Pih1, and Rsa1 are all involved in U3 assembly in yeast (Fig. 4 C), they are not essential genes. Interestingly, strains lacking Pih1 and Rsa1 display a thermosensitive phenotype, suggesting that more chaperone activity might be required at higher temperature and may render these proteins essential. Thus, rather than being absolutely required for snoRNP biogenesis, Hsp90 might simply ensure efficient RNP formation.

### **Assembling L7Ae RNPs: a remodeling event that involves AAA + ATPases and Hsp90?**

On the basis of the differential stability of immature and mature C/D snoRNP complexes toward salt, it has been proposed that pre-snoRNPs undergo a remodeling event to yield the mature particle (Watkins et al., 2004). Although Nufip is part of immature particles, it is absent from mature snoRNPs. Because Nufip is able to tether fibrillar and hNop58 to 15.5K either directly or through hPih1, this suggests that fibrillar and hNop58 are loaded on the assembling particle by virtue of their interaction with Nufip. Thus, the remodeling event that leads to the formation of the mature particle probably involves the transfer of hNop58 and fibrillar from hPih1 or Nufip to 15.5K. Similarly, in the case of U4 and B/C RNPs, it is tempting to speculate that hPRP31 and U3-55K first interact with Nufip bound to 15.5K–RNA complexes and that remodeling transfers hPrp31 and U3-55K to 15.5K.

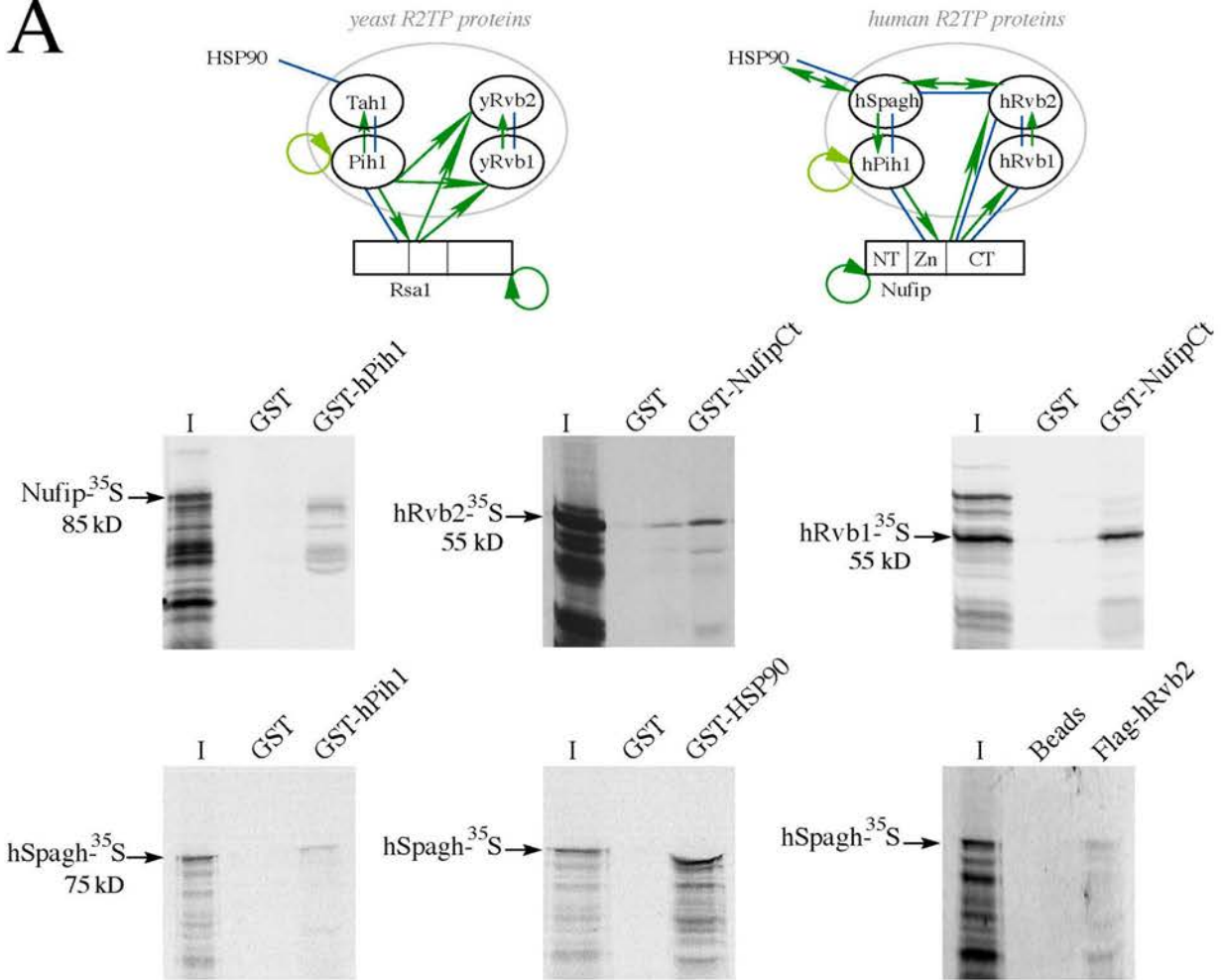
It is remarkable that the C-terminal domain of Nufip also binds the AAA + ATPases hRvb1 and hRvb2. This family of proteins is known to unfold client proteins, which can be presented by adaptors, and to drive the dissociation of protein complexes and protein aggregates (Hanson and Whiteheart, 2005). The C-terminal domain of Nufip also binds hPRP31, U3-55K, and fibrillar, and hRvb2 itself binds fibrillar (Table S1). Thus, one possibility would be that Nufip presents core RNP proteins to hRvb1 and hRvb2, which would partially unfold them to promote their transfer from Nufip to 15.5K–RNA complexes. Another contrasting possibility would be that hRvb2 brings fibrillar to the nascent RNP and that interaction of this complex with Nufip would trigger its transfer to 15.5K–RNA complexes. Hsp90 may also participate in the process by controlling protein folding during remodeling.

### **Hsp90, L7Ae RNPs, cell growth, and cancer**

Hsp90 is an essential and ubiquitous chaperone but with a rather specialized function in signal transduction (Pearl and Prodromou, 2006; Caplan et al., 2007). In this paper, we show that snoRNPs are targets of Hsp90 and that this is likely conserved from yeast to humans, which is indicative of an ancient function for Hsp90. Hsp90 is tightly linked to human cancer; its overexpression is a marker for cancer cells, and geldanamycin, a drug that specifically targets Hsp90, is a promising treatment for cancer (Pearl and Prodromou, 2006; Caplan et al., 2007). It was previously shown that Hsp90 binds and controls the activity of hTERT, the telomerase reverse transcription (Holt et al., 1999). Here, we show that Hsp90 is also required for accumulation of the telomerase RNA, a process that does not require hTERT. Thus, formation of an active telomerase is tightly controlled by Hsp90, as the chaperone is required at several steps during its biogenesis.

It has been proposed that Hsp90 behaves as a master regulator of cell proliferation by controlling many signal transduction pathways and that this underlies its role in cancers. Our results indicate that Hsp90 also controls cell growth through the synthesis of new ribosomes, replication through the production

**A**



**B**

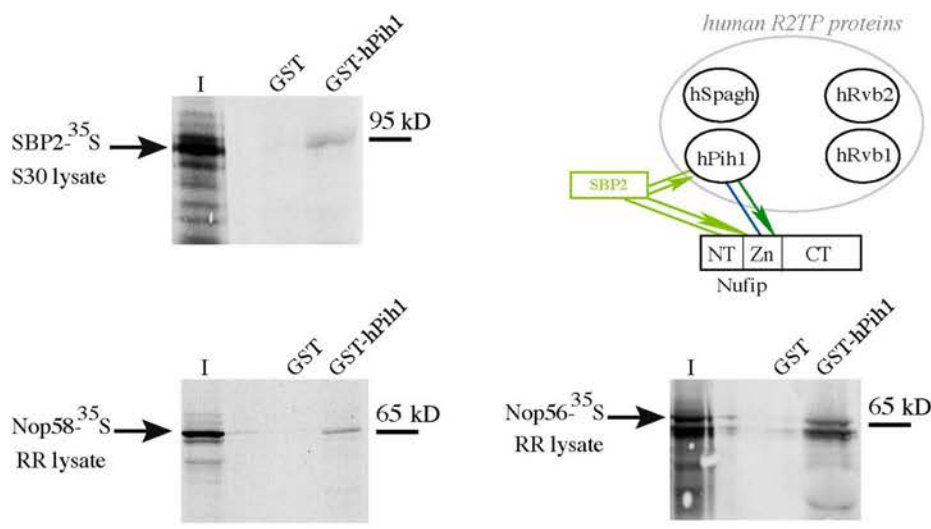


Figure 8. **hSpagh, hPih1, and Nufip mediate the interaction of Hsp90 with hNop58, 15.5K, hNHP2, and SBP2.** (A) Interactions between Rsa1, Nufip, and R2TP proteins. (top) Summary of interactions between yeast (left) and human (right) proteins. Y2H interactions are indicated by green arrows, and physical interactions observed in GST pull-downs are indicated by blue lines. (bottom) GST pull-down experiments. Legend as in Fig. 4. For the interaction between hSpagh and hRvb2, a FLAG-hRvb2 fusion was used. (B) SBP2, hNop56, and hNop58 interact with hPih1 in GST pull-down experiments. Legend as in A except that hNop56 and hNop58 were translated in rabbit reticulocyte lysates (RR). The schematic summarizes the interactions of SBP2 with Nufip and R2TP proteins (legend as in A).

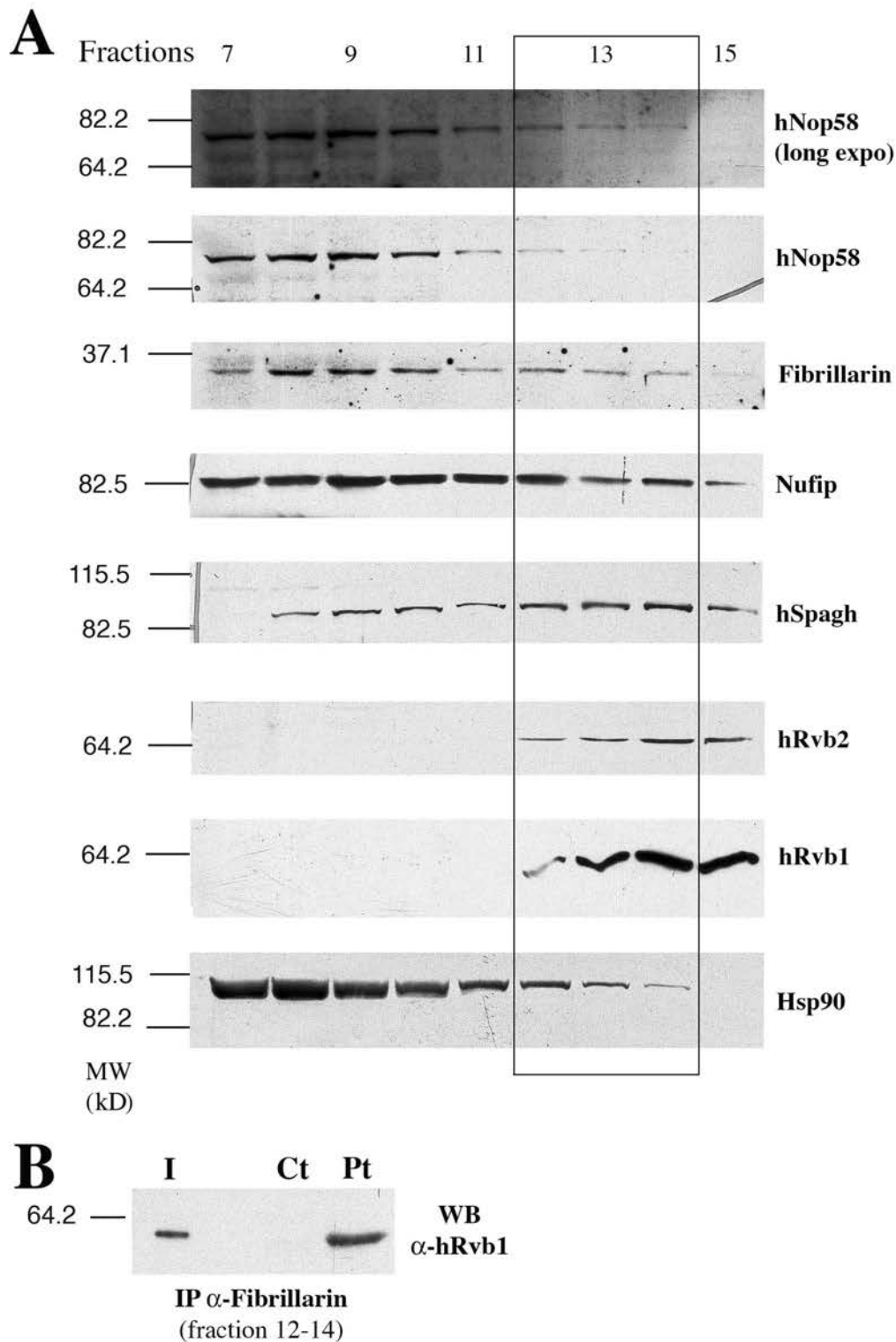


Figure 9. **Nufip and R2TP proteins are present within large cytoplasmic complexes.** (A) Sedimentation profile of Nufip, R2TP proteins, and core snoRNP proteins. S100 cytoplasmic extracts were fractionated on 10–30% glycerol gradients, and 23 fractions were collected. Fractions 7 (top)–16 (bottom) are shown. The boxed fractions contain all of the analyzed proteins. (B) Fibrillarin associates with hRvb1 within large cytoplasmic complexes. Fractions 12–14 were pooled, immunoprecipitated with fibrillarin antibodies, and analyzed by Western blots with anti-hRvb1 antibodies.

of telomerase RNA, and antioxidant defense through seleno-protein synthesis. Thus, Hsp90 may coordinate many of the events required for cell proliferation, explaining why its inhibition is so detrimental to tumor cells.

## Materials and methods

### Cells, plasmids, and antibodies

Plasmids were obtained by standard techniques and using Gateway technology (Invitrogen) with appropriate destination vectors (pACTII, pASII $\Delta$ ),

pDEST27, pET15b, pET32a, pGEX2T, and pSPOII-GFP). GFP-Nufip, yU3del, rat U3B.7, U8, and U13 were described previously (Bardoni et al., 1999; Boulon et al., 2004).

HeLa and 293 cells were cultivated at 37°C in DME containing 10% FCS. Transfections were performed with the calcium-phosphate procedure or with LipofectAMINE (Invitrogen) as recommended by the manufacturer. siRNAs against Nufip had the sequences 5'-GGAGCAGUAUUGACAA-CAA dT dT and 5'-GAGGAGAAACCACAACAUU dT dT. Cells were analyzed 48 h after transfection.

Antibodies against Nufip and fibrillarin were described previously (Bardoni et al., 1999; Verheggen et al., 2002). Anti-hRvb1 and anti-hRvb2 antibodies were purchased from ProteinTech Group. Anti-Hsp90 was monoclonal 3G3 (Affinity BioReagents). Polyclonal serum against hSpagh and hPih1 were produced by immunizing rabbit with peptides CTREENTKNRIKSYD-CONH2 and CIRSEGRPRIQELGDL-CONH2, respectively.

#### Electrophoresis mobility shift assays

Uniformly labeled yeast U14 snoRNA (yU14) was synthesized by T7 RNA polymerase in vitro transcription of a PCR fragment carrying the *Saccharomyces cerevisiae* U14 snoRNA sequence according to the conditions described previously (Marmier-Gourrier et al., 2003). For electrophoresis mobility shift assay, 200 nM of recombinant protein Snu13p, N3C1, or yPEP was mixed in various combinations at room temperature with ~5 fmol <sup>32</sup>P-radiolabeled yU14 RNA in 150 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, and 20 mM Hepes, pH 7.9. The RNA-protein complexes were resolved by native gel electrophoresis.

#### Yeast strains and measurements of RNA levels

Growth and handling of *S. cerevisiae* were by standard techniques. Total RNAs were extracted from exponentially growing yeast cultures. RNAs were analyzed by quantitative primer extension using 5 µg of total RNAs and 5 ng of appropriate radiolabeled oligonucleotide. Northern blot analyses were performed with oligonucleotides complementary to the RNA of interest. Sequences of oligonucleotides used as probes and in qPCR are listed in the section Oligonucleotides. qPCR was performed as described previously (Lutfalla and Uze, 2006).

#### Y2H, Y3H, and bridged assays

For Y2H assays, appropriate pACT-II and pAS2ΔΔ plasmids were introduced into haploid test strains (CG929 or Y190 and Y187), which were then crossed. Diploids were plated on double or triple selectable media (-Leu, -Trp, or -Leu, -Trp, and -His), and growth was assessed 3 d later. For Y2H tests involving pGBKT7-SBP2 or pGADT7-SBP2, plasmids were cotransformed into AH109 (Clontech Laboratories, Inc.). Diploids were plated on triple and quadruple selective media (-Leu -Trp -Ade or -Leu -Trp -Ade -His). For Y3H tests, L40-coat strains carrying appropriate pACTII plasmids were mated with R40-coat strains carrying the plasmid pIII<sub>MS2-2</sub>::yU3B/C or pIII<sub>MS2-2</sub> as a control. Diploids were selected on -Trp, -Leu, and -Ura and were tested in -Leu -Ura -His. For the bridged Y2H and Y3H assays, bridge proteins were expressed from the ADE2 multicopy p422 plasmid.

#### Recombinant proteins and GST pull-down assays

Radiolabeled proteins were synthesized in the presence of [<sup>35</sup>S]methionine, in *Escherichia coli* S30 lysate (Promega), or in rabbit reticulocyte lysate (TNT; Promega). Binding was performed with 5 µg GST-tagged protein in 20 mM Tris-HCl, pH 8, 40 mM KCl, 1 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 0.1% NP-40, and 10% glycerol. Washing was performed with the same buffer but with 100 mM KCl instead of 40 mM and 5 mM MgCl<sub>2</sub> instead of 1 mM. GST- and His-tagged proteins were produced by standard procedure.

#### Co-IP assays

HeLa cytoplasmic S100 extracts were purchased from 4CBiotech (Belgium). Total extracts of human cells were prepared in HNTG (Girard et al., 2006) except in the experiment of Fig. 3 E, where it was prepared in NET-2 (Boulon et al., 2004). IP was performed as previously described (Boulon et al., 2004). For RNA analyses, TRIzol (Invitrogen) was added to the beads, and RNAs were extracted and analyzed by RNase protection. For protein analyses, bound material was eluted in 1% SDS at RT, mixed with Laemmli buffer, and analyzed by Western blots. SBP2 co-IP was performed using µMACS Epitope Tag Protein Isolation kits (Miltenyi Biotec) according to the manufacturer's instructions.

For co-IP in yeast, total cellular extracts were prepared from exponentially growing cultures. Extracts were added to IgG-Sepharose beads (Sigma-Aldrich) and incubated at 4°C for 2 h in buffer IPP150 (NaCl 150 mM, 0.1% NP-40, and 10 mM Tris-HCl, pH 8). Beads were

washed four times in the same buffer, and RNAs were extracted and analyzed by RT-PCR.

#### Purification of endogenous SBP2 complexes

HeLa cell lysates (a gift from R. Lüthmann, Max Planck Institute of Biophysical Chemistry, Göttingen, Germany) were diluted in IPP250 (250 mM NaCl, 20 mM Hepes-NaOH, pH 7.9, 1.5 mM MgCl<sub>2</sub>, 0.5 mM DTT, and 0.05% NP-40) and incubated for 2 h at 4°C with protein A-Sepharose (GE Healthcare) charged with affinity-purified anti-SBP2 peptide antibodies. Beads were washed with 120 column volumes, and bound proteins were eluted using IPP250 supplemented with 5% glycerol and 0.6 mg/ml peptide.

#### Oligonucleotides

The oligonucleotides used in this study for qPCR are as follows: U3 (5'-TTCTCT-GAACGTGTAGAGCACCGA and 3'-GATCATCAATGGCTGACGGCAGTT), U4 (5'-GCTTTGCGCAGTGGCAGT and 3'-AGCAATAATCGCTCCCTCGG), U14 (5'-CCAACATTCGCAGTTCCACCAG and 3'-CTCACTCAGACATC-CAAGGAAGG), hTel RNA (5'-CTAACTGAGAAGGGCGTAGGC and 3'-TGCTCTAGAATGAACGGTGGGA), and U19 (5'-ATGTGGTGCCTGT-GATGGTGTAC and 3'-ACACTGCCCAAAGGTACTCAGTA). The oligonucleotides used in this study for primer extension are as follows: RTU3 (GGGTACAAGGTTAT), RTU14 (TCACTCAGACATCCTAG), RTsnR190 (CGAGGAAAGAAGAGACACCATATC), RTsnR42 (TCAAACAATAGGCT-CCTAAAGCATCACAA), RTU4 (TAAATTTCAACCGGAAACACCA-ATCTCGGACGAA), RTiRNA (TGGACGCAACCGGAATCGAACCG), and RT7S (CAGGACAAATTACGACGGAGGAA).

#### Online supplemental material

Fig. S1 shows Y3H interactions of Nufip with various B/C RNAs. Fig. S2 shows alignment of the RsaI PEP sequence with potential homologues in various organisms. Fig. S3 shows that hSpagh is associated with hRvb1, hRvb2, and hPih1. Fig. S4 shows cosedimentation of R2TP proteins and Nufip in glycerol gradients. Table S1 is a Y2H interaction map of yeast and human proteins. Online supplemental material is available at <http://www.jcb.org/cgi/content/full/jcb.200708110/DC1>.

We thank R. Lüthmann for the gift of HeLa extracts, M. Grigoriev for the recombinant Flag-hRvb2, and A. Schweigert for technical assistance. We also thank W. Houry for exchanging unpublished information, L. Trinkler-Mulcahy for critical reading of the manuscript, and A. Lamond for supporting experiments.

This study was supported by grants from la Ligue contre le Cancer (to E. Bertrand), Action Concertée Incitative (to A. Krol and C. Branlant), ToxNuc-E (to A. Krol), and the French research agency Agence Nationale de la Recherche.

Submitted: 15 August 2007

Accepted: 10 November 2007

## References

- Allmang, C., P. Carbon, and A. Krol. 2002. The SBP2 and 15.5 kD/Snu13p proteins share the same RNA binding domain: identification of SBP2 amino acids important to SECIS RNA binding. *RNA*. 8:1308-1318.
- Bardoni, B., A. Schenck, and J. Mandel. 1999. A novel RNA-binding nuclear protein that interacts with the fragile X mental retardation (FMR1) protein. *Hum. Mol. Genet.* 8:2557-2566.
- Boulon, S., C. Verheggen, B. Jady, C. Girard, C. Pescia, C. Paul, J. Ospina, T. Kiss, A. Matera, R. Bordonne, and E. Bertrand. 2004. PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. *Mol. Cell.* 16:777-787.
- Caplan, A.J., A. Mandal, and M. Theodoraki. 2007. Molecular chaperones and protein kinase quality control. *Trends Cell Biol.* 17:87-92.
- Copeland, P.R., J.E. Fletcher, B.A. Carlson, D.L. Hatfield, and D.M. Driscoll. 2000. A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J.* 19:306-314.
- Fagegaltier, D., N. Hubert, K. Yamada, T. Mizutani, P. Carbon, and A. Krol. 2000. Characterization of mSelB, a novel mammalian elongation factor for selenoprotein translation. *EMBO J.* 19:4796-4805.
- Fury, M.G., and G.W. Zieve. 1996. U6 snRNA maturation and stability. *Exp. Cell Res.* 228:160-163.
- Giot, L., J. Bader, C. Brouwer, A. Chaudhuri, B. Kuang, Y. Li, Y. Hao, C. Ooi, B. Godwin, E. Vitols, et al. 2003. A protein interaction map of *Drosophila melanogaster*. *Science*. 302:1727-1737.

- Girard, C., H. Neel, E. Bertrand, and R. Bordonne. 2006. Depletion of SMN by RNA interference in HeLa cells induces defects in Cajal body formation. *Nucleic Acids Res.* 34:2925–2932.
- Gonzales, F.A., N.I. Zanchin, J.S. Luz, and C.C. Oliveira. 2005. Characterization of *Saccharomyces cerevisiae* Nop17p, a novel Nop58p-interacting protein that is involved in Pre-rRNA processing. *J. Mol. Biol.* 346:437–455.
- Granneman, S., G. Pruijn, W. Horstman, W. van Venrooij, R. Luhrmann, and N. Watkins. 2002. The hU3-55K protein requires 15.5K binding to the box B/C motif as well as flanking RNA elements for its association with the U3 small nucleolar RNA in Vitro. *J. Biol. Chem.* 277:48490–48500.
- Hanson, P.L., and S.W. Whiteheart. 2005. AAA+ proteins: have engine, will work. *Nat. Rev. Mol. Cell Biol.* 6:519–529.
- Holt, S.E., D.L. Aisner, J. Baur, V.M. Tesmer, M. Dy, M. Ouellette, J. Trager, G. Morin, D.O. Toft, J.W. Shay, et al. 1999. Functional requirement of p23 and Hsp90 in telomerase complexes. *Genes Dev.* 13:817–826.
- King, T.H., W.A. Decatur, E. Bertrand, E.S. Maxwell, and M.J. Fournier. 2001. A well-connected and conserved nucleoplasmic helicase is required for production of box C/D and H/ACA snoRNAs and localization of snoRNP proteins. *Mol. Cell. Biol.* 21:7731–7746.
- Kiss, T. 2002. Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell.* 109:145–148.
- Kittur, N., X. Darzacq, S. Roy, R. Singer, and U. Meier. 2006. Dynamic association and localization of human H/ACA RNP proteins. *RNA.* 12:2057–2062.
- Koonin, E.V., P. Bork, and C. Sander. 1994. A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins and a ribosome modification enzyme? *Nucleic Acids Res.* 22:2166–2167.
- Kressler, D., M. Doere, M. Rojo, and P. Linder. 1999. Synthetic lethality with conditional dbp6 alleles identifies rsa1p, a nucleoplasmic protein involved in the assembly of 60S ribosomal subunits. *Mol. Cell. Biol.* 19:8633–8645.
- Kufel, J., C. Allmang, G. Chanfreau, E. Petfalski, D. Lafontaine, and D. Tollervey. 2000. Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. *Mol. Cell. Biol.* 20:5415–5424.
- Li, L., and K. Ye. 2006. Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature.* 443:302–307.
- Liu, S., P. Li, O. Dybkov, S. Nottrott, K. Hartmuth, R. Luhrmann, T. Carlomagno, and M. Wahl. 2007. Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP. *Science.* 316:115–120.
- Lutfalla, G., and G. Uze. 2006. Performing quantitative RT-PCR experiments. *Methods Enzymol.* 410:386–400.
- Makino, Y., T. Mimori, C. Koike, M. Kanemaki, Y. Kurokawa, S. Inoue, T. Kishimoto, and T. Tamura. 1998. TIP49, homologous to the bacterial DNA helicase RuvB, acts as an autoantigen in human. *Biochem. Biophys. Res. Commun.* 245:819–823.
- Marmier-Gourrier, N., A. Cléry, V. Senty-Ségault, B. Charpentier, F. Schlotter, F. Leclerc, R. Fournier, and C. Branlant. 2003. A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA. *RNA.* 9:821–838.
- Matera, A.G., R. Terns, and M. Terns. 2007. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.* 8:209–220.
- Newman, D.R., J.F. Kuhn, G.M. Shanab, and E.S. Maxwell. 2000. Box C/D snoRNA-associated proteins: two pairs of evolutionarily ancient proteins and possible links to replication and transcription. *RNA.* 6:861–879.
- Nottrott, S., H. Urlaub, and R. Luhrmann. 2002. Hierarchical, clustered protein interactions with U4/U6 snRNA: a biochemical role for U4/U6 proteins. *EMBO J.* 21:5527–5538.
- Pearl, L.H., and C. Prodromou. 2006. Structure and mechanism of the Hsp90 molecular chaperone machinery. *Annu. Rev. Biochem.* 75:271–294.
- Schultz, A., S. Nottrott, N. Watkins, and R. Luhrmann. 2006. Protein-protein and protein-RNA contacts both contribute to the 15.5K-mediated assembly of the U4/U6 snRNP and the box C/D snoRNPs. *Mol. Cell. Biol.* 26:5146–5154.
- Te, J., L. Jia, J. Rogers, A. Miller, and S. Hartson. 2007. Novel subunits of the mammalian Hsp90 signal transduction chaperone. *J. Proteome Res.* 6:1963–1973.
- Verheggen, C., D. Lafontaine, D. Samarsky, J. Mouaikel, J. Blanchard, R. Bordonne, and E. Bertrand. 2002. Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. *EMBO J.* 21:2736–2745.
- Vidovic, I., S. Nottrott, K. Hartmuth, R. Luhrmann, and R. Ficner. 2000. Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol. Cell.* 6:1331–1342.
- Watkins, N.J., V. Segault, B. Charpentier, S. Nottrott, P. Fabrizio, A. Bachi, M. Wilm, M. Rosbash, C. Branlant, and R. Luhrmann. 2000. A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP. *Cell.* 103:457–466.
- Watkins, N.J., A. Dickmanns, and R. Luhrmann. 2002. Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Mol. Cell. Biol.* 22:8342–8352.
- Watkins, N.J., I. Lemm, D. Ingelfinger, C. Schneider, M. Hossbach, H. Urlaub, and R. Luhrmann. 2004. Assembly and maturation of the U3 snoRNP in the nucleoplasm in a large dynamic multiprotein complex. *Mol. Cell.* 16:789–98.
- Yong, J., L. Wan, and G. Dreyfuss. 2004. Why do cells need an assembly machine for RNA-protein complexes? *Trends Cell Biol.* 14:226–232.
- Zhao, R., M. Davey, Y. Hsu, P. Kaplanek, A. Tong, A. Parsons, N. Krogan, G. Cagney, D. Mai, J. Greenblatt, et al. 2005. Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the hsp90 chaperone. *Cell.* 120:715–727.





## 2.2. Données complémentaires

Dans l'article *The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery*, les approches de double et triple-hybride *bridged* nous ont permis de mettre en évidence le rôle de la protéine humaine NUFIP en tant que médiateur des associations entre constituants de plusieurs RNP (Boulon *et al.*, 2008). En effet, l'interaction par double-hybride de la protéine 15.5K avec les protéines PRP31, Fibrillarine et U3-55K, n'est effective que lorsque NUFIP est co-exprimée, de manière ectopique, dans des cellules de levure invalidées pour le gène *RSA1*. De la même façon, l'association par triple-hybride de la Fibrillarine et de PRP31 avec un ARN porteur d'un motif en *K-turn*, repose sur la présence de NUFIP. Enfin, l'effet médiateur du facteur d'assemblage nécessite la présence du domaine PEP, ce qui suggère que son interaction avec la protéine 15.5K, dans le cas du double-hybride, ou avec la protéine Snu13p endogène, dans le cas du triple-hybride, est essentielle pour assurer cette fonction. Comme discuté dans l'article, ces résultats suggèrent donc que NUFIP pourrait jouer un rôle de plateforme d'assemblage. Par le biais de sa capacité à interagir avec la protéine 15.5K, elle pourrait en effet faciliter le recrutement et l'arrimage des différentes protéines constitutives sur les RNP en cours d'assemblage.

Les techniques de double et triple-hybride *bridged* chez la levure constituent des approches intéressantes dans le cadre de l'étude de systèmes hétérologues. En revanche, pour des raisons d'interférences avec les protéines endogènes, ces approches sont difficilement transposables à l'étude des protéines de levure, comme c'est le cas du facteur Rsa1p. Afin de compléter les données présentées dans l'article et dans le but de confirmer ce rôle de protéine plateforme, nous avons donc mis au point une technique de reconstitution de la snoRNP U14 en extrait cellulaire. Cette approche, réalisée dans des extraits de levures WT ou invalidées pour le gène *RSA1*, nous a ainsi permis d'évaluer l'influence de ce facteur sur l'assemblage de la snoRNP. Avant de réaliser ces expériences, nous avons néanmoins tenu à vérifier la capacité de Rsa1p à interagir directement avec les différentes protéines constitutives des snoRNP à boîtes C/D.

### 2.2.1 Rsa1p interagit avec Nop58p et Prp31p

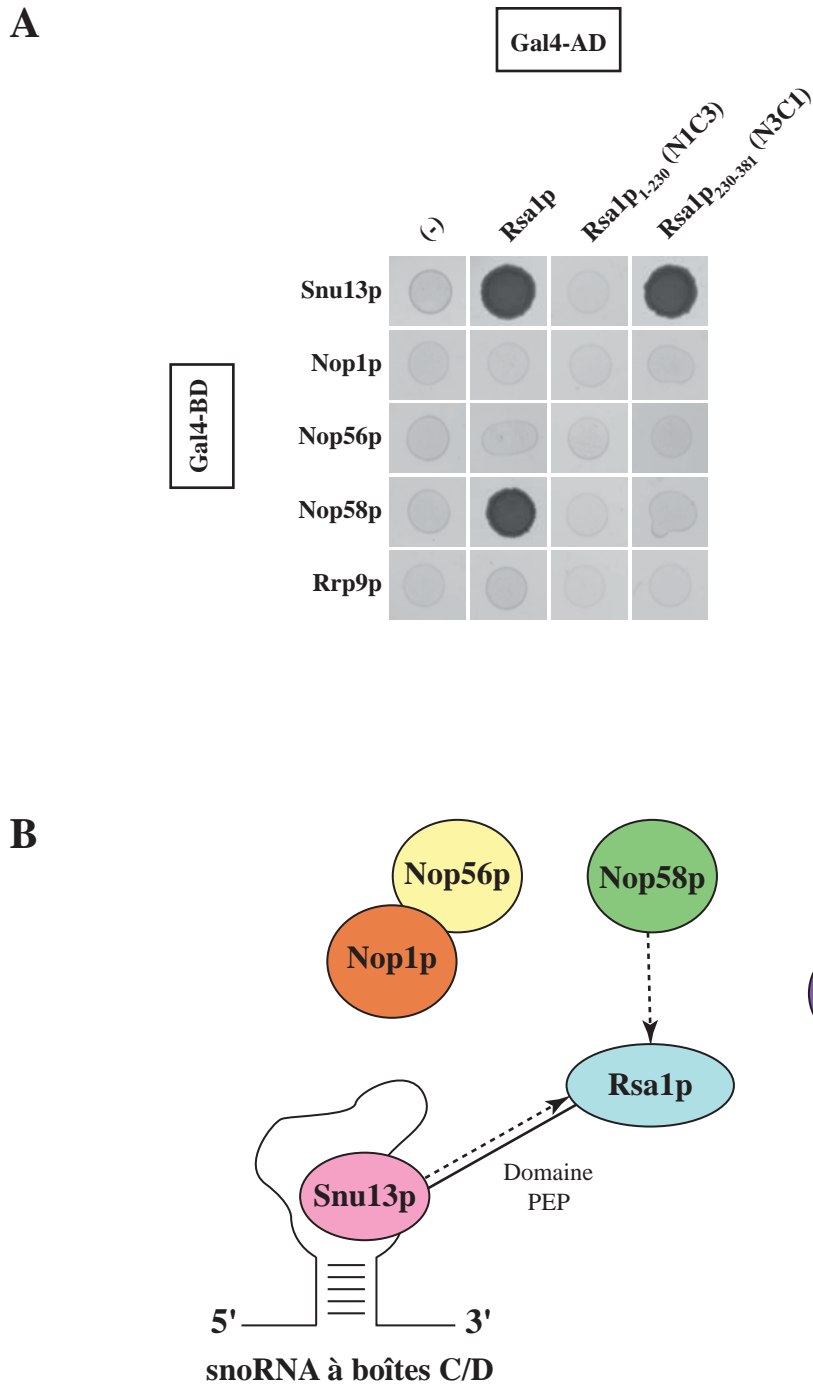
En utilisant les approches de GST pull-down et de retard sur gel, nous avons décrit l'association physique des protéines Rsa1p et Snu13p (Boulon *et al.*, 2008). Cette interaction,

conservée chez l'homme entre le facteur NUFIP et la protéine 15.5K, implique le domaine d'homologie PEP et peut s'établir indépendamment de l'ARN ou au sein d'une RNP. Néanmoins, parmi les protéines constitutives des snoRNP à boîtes C/D, seule la protéine Snu13p peut être aisément étudiée *in vitro*. En effet, les protéines Nop1p, Nop56p et Nop58p, qui sont instables lorsqu'elles sont exprimées chez *E. coli*, n'ont pas pu être purifiées sous la forme recombinante. L'interaction de Rsa1p avec ces protéines a donc été testée exclusivement par l'approche de double-hybride (Boulon *et al.*, 2008). Comme présenté dans la **Figure 54A**, cette démarche nous a permis de mettre en évidence une interaction entre Rsa1p et Nop58p. Cette interaction est cependant faible, puisqu'elle n'est visible qu'en absence de 3-AT, un inhibiteur de la voie de biosynthèse de l'histidine, qui permet d'évaluer la force de l'interaction conduisant à l'activation du gène rapporteur *HIS3* (cf. Matériel & Méthodes). Contrairement à la forte stabilité de l'association observée entre Rsa1p et Snu13p, ce résultat suggère donc que l'interaction entre Rsa1p et Nop58p pourrait être de nature instable ou indirecte. D'autre part, seule la protéine Rsa1p entière présente la capacité d'interagir avec Nop58p, ce qui suggère que des déterminants importants pour cette interaction pourraient se situer dans la région N-terminale de Rsa1p.

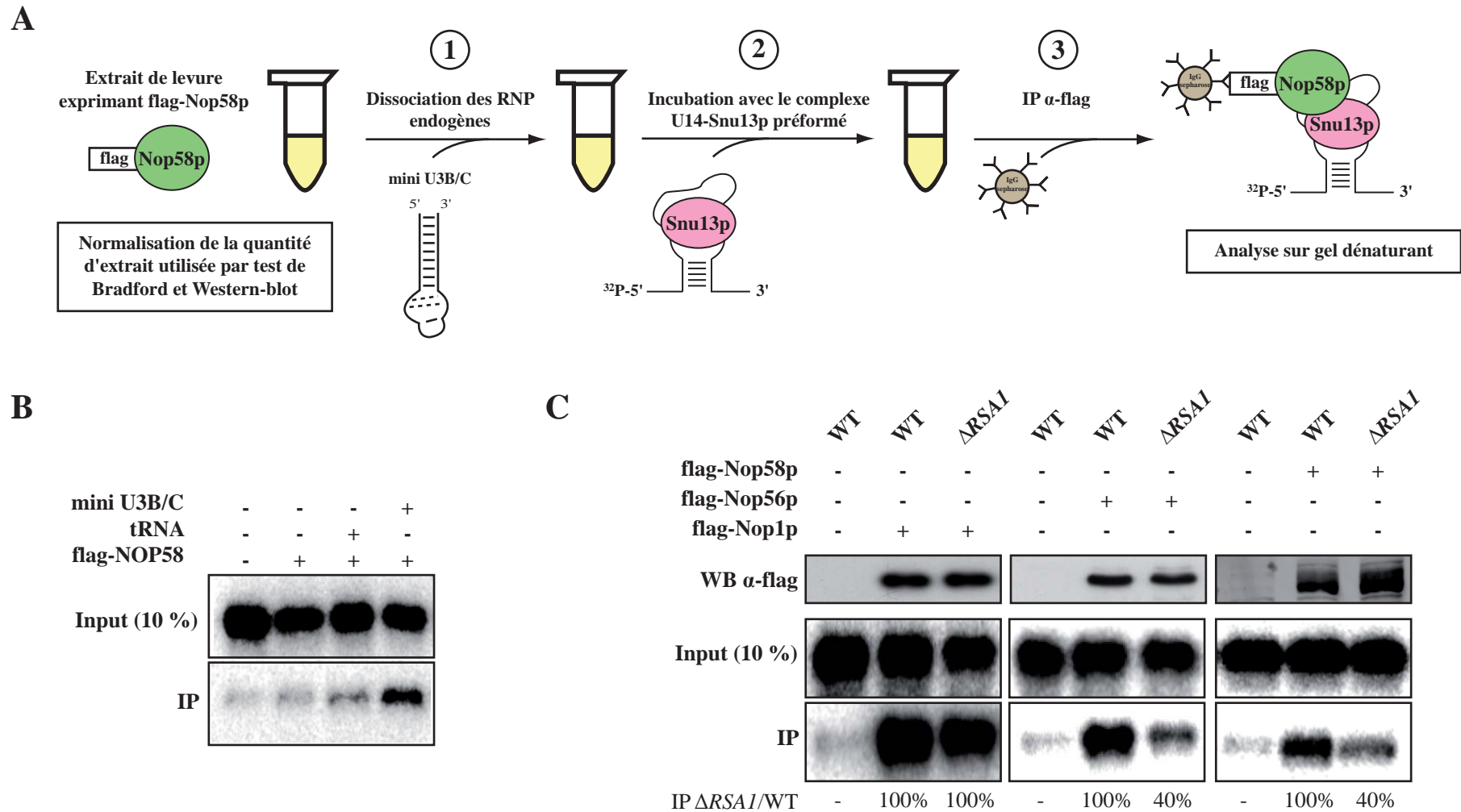
Outre cette association, l'équipe d'E. Bertrand a également mis en évidence une interaction de Rsa1p avec la protéine Prp31p de la snRNP U4 (Boulon *et al.*, 2008). Etant donné la forte homologie existant entre Prp31p et Nop58p (**Fig. 30**), il est probable que leur association avec la protéine Rsa1p repose sur des déterminants communs. Ceci est d'autant plus envisageable, que l'interaction entre Prp31p et Rsa1p requiert également l'intégralité de la séquence du facteur d'assemblage. De ce fait, et compte tenu de la construction des snoRNP à boîtes C/D et de la snRNP U4 sur une base commune (cf. Introduction, Chapitre 2), la protéine Rsa1p pourrait être également impliquée dans l'assemblage de la UsnRNP spliceosomale.

### 2.2.2 Rsa1p facilite l'assemblage de la snoRNP U14

En dépit de l'impossibilité d'étudier la formation des snoRNP à partir de protéines recombinantes, nous avons donc entrepris la mise au point d'une approche d'assemblage de la snoRNP U14 en extrait cellulaire (**Fig. 55A**) (cf. Matériel & Méthodes). Pour cela nous avons réalisé des extraits de cellules de levures exprimant de façon ectopique les protéines Nop1p, Nop56p ou Nop58p porteuses de l'étiquette flag au niveau de leur extrémité N-terminale. Ces



**Figure 54 : Bilan des interactions établies entre Rsa1p et les constituants protéiques des snoRNP à boîtes C/D.** (A) Tests d'interaction par double-hybride entre Rsa1p, Rsa1p<sub>230-381</sub> (alias N1C3), Rsa1p<sub>230-381</sub> (alias N3C1) et les protéines Snu13p, Nop1p, Nop56p, Nop58p des snoRNP à boîtes C/D et de la protéine Rrp9p, spécifique de la snoRNP U3. Rsa1p et ses domaines sont exprimés à partir du vecteur pACT2 (Gal4-AD). Les protéines constitutives des snoRNP sont exprimées à partir du vecteur pGBKT7 (Gal4-BD). Les tests d'interaction ont été réalisés par croisement des souches CG1945 (pACT2) et Y187 (pGBKT7), et étalement sur milieu sélectif sans histidine et sans 3-aminotriazol (cf. détails Matériel & Méthodes). (B) Bilan des interactions établies entre Rsa1p et les protéines des snoRNP à boîtes C/D. Les interactions détectées par double-hybride sont représentées par des flèches en pointillées (orientée en direction de la protéine proie Gal4-AD). Les interactions physiques sont représentées en traits pleins.



**Figure 55 : Rsa1p est requise pour l'assemblage optimal des protéines Nop56p et Nop58p sur le complexe U14-Snu13p.** (A) Approche utilisée pour la reconstitution de la snoRNP U14 en extrait de levure (cf. détails Matériel & Méthodes). (B) Mise au point de l'approche expérimentale. La pré-incubation de l'extrait cellulaire avec l'ARN compétiteur mini U3B/C est nécessaire pour dissocier les RNP endogènes et libérer la protéine étiquetée. (C) Reconstitution de la snoRNP U14 à partir des protéines flag-Nop1p, flag-Nop56p et flag-Nop58p, dans un extrait WT ou ΔRSAI. La quantité d'extrait utilisée pour chaque essai a été normalisée par Western-blot anti-flag.

extraits ont été incubés à 30°C en présence d'un complexe préformé entre le snoRNA U14 radiomarqué et la protéine Snu13p recombinante. Les différentes protéines étiquetées ont ensuite été isolées sur des billes de sépharose greffées à des anticorps anti-flag et le matériel purifié a été analysé sur gel dénaturant. Comme décrit initialement par l'équipe de R. Lührmann, la mise au point d'une telle approche nécessite cependant de dissocier préalablement les RNP endogènes afin de libérer les protéines d'intérêt (Watkins *et al.*, 2002). Avant l'ajout du complexe U14-Snu13p, nous avons donc pré-incubé de façon systématique les extraits à 30°C en présence d'un petit ARN porteur du motif B/C du snoRNA U3 (*i.e.* mini U3B/C) (Clery *et al.*, 2007b). Ce dernier, incorporé en large excès, agit en effet comme compétiteur pour la fixation de la protéine Snu13p endogène, il permet ainsi de déstabiliser les RNP endogènes et de libérer les protéines flag-Nop1p, flag-Nop56p et flag-Nop58p. L'importance de cette étape est démontrée dans la **Figure 55B**. En effet, l'absence d'ARN compétiteur ou l'addition d'ARN non spécifique, comme un mélange d'ARNt, s'avère incompatible avec la rétention du snoRNA U14 par la protéine flag-Nop58p. Inversement, l'ajout de l'ARN U3BC favorise une rétention importante du snoRNA par les différentes protéines étiquetées, suggérant ainsi la formation de RNP *de novo* dans l'extrait (**Fig. 55B et C**).

Dans le but de tester l'importance de la protéine Rsa1p pour l'assemblage de la snoRNP U14, nous avons donc utilisé ce système sur la base d'extraits préparés à partir de levures WT ou invalidées pour le gène *RSA1* (**Fig. 55C**). Afin d'apporter la même quantité de protéines étiquetées dans chaque essai, les volumes d'extrait utilisés ont été normalisés par des tests de Bradford et par des analyses par Western-blot réalisées à l'aide d'anticorps anti-flag. La comparaison des résultats obtenus à partir des extraits WT et  $\Delta$ *RSA1* indique que la rétention optimale du snoRNA U14 par les protéines flag-Nop56p et flag-Nop58p nécessite la présence de Rsa1p. En effet, seulement 40% de l'ARN radiomarqué est co-immunoprécipité par ces deux protéines dans les extraits invalidés pour le gène *RSA1*. En revanche, le recrutement de la protéine Nop1p sur l'ARN ne semble pas affecté par l'absence de Rsa1p, puisque 100% des particules U14 sont formées dans cette condition. Ces résultats confirment donc le rôle de Rsa1p dans l'assemblage des snoRNP à boîtes C/D. Au même titre que NUFIP chez l'homme, Rsa1p semble jouer un rôle de protéine plateforme en facilitant le recrutement des protéines Nop56p et Nop58p sur les pré-complexes formés par le snoRNA et la protéine Snu13p.

Afin de caractériser plus finement cette fonction, nous avons essayé de compléter le défaut d'assemblage observé dans les extraits  $\Delta RSAI$ , par ajout du fragment Rsa1p<sub>230-381</sub> (alias N3C1) recombinant. Cependant, ni sa pré-incubation avec le complexe U14-Snu13p, ni son addition dans l'extrait, ne permettent de rétablir l'assemblage correct de la particule (données non montrées). Ce résultat suggère donc qu'en dépit de sa capacité à interagir avec la protéine Snu13p, le fragment C-terminal de la protéine n'est pas suffisant pour assurer cette fonction. Etant donné les résultats obtenus par double-hybride, il est envisageable que l'intégralité de la séquence protéique, requise pour l'interaction avec la protéine Nop58p, soit également nécessaire pour promouvoir son assemblage sur la particule. Cette hypothèse n'a pas pu être testée, puisque la protéine Rsa1p entière présente une forte tendance à l'agrégation, qui empêche sa purification sous forme recombinante. Cependant, une autre possibilité est que l'absence de Rsa1p dans les cellules de levure conduise à des effets délétères, non réversibles par la simple addition de la protéine recombinante dans l'extrait. Cette hypothèse est d'autant plus envisageable que le fragment Rsa1p<sub>230-381</sub>, suffisant pour compléter le défaut de croissance d'une souche invalidée pour le gène *RSAI*, constitue le domaine minimal requis pour assurer la fonction de la protéine *in vivo*.

Enfin, même si l'approche de double-hybride n'a révélé aucune interaction entre ces deux protéines, Nop56p est également dépendante de Rsa1p pour son assemblage sur l'ARN. Compte tenu de la forte homologie existant entre Nop56p, Nop58p et Prp31p, il est en effet probable que Rsa1p puisse exercer une influence similaire sur ces trois protéines. Selon toute vraisemblance, Rsa1p présente une affinité intrinsèquement limitée pour ces différentes protéines. En revanche, dans le contexte d'une pré-RNP, cette faible affinité, additionnée aux différents contacts établis entre les autres composants, pourrait constituer un apport thermodynamiquement non négligeable pour la formation d'un complexe. Ainsi, en interagissant avec la pré-particule ARN-Snu13p, Rsa1p pourrait faciliter le recrutement des autres protéines constitutives en augmentant de façon substantielle les surfaces de contact dédiées à leur arrimage et à leur incorporation au sein de la RNP.

### 3. Conclusion & Perspectives

#### 3.1. Bilan

Ces travaux révèlent donc que l'assemblage des RNP de la famille L7Ae, repose sur l'existence d'une machinerie protéique spécialisée, conservée entre la levure et l'homme (**Tableau 5**). Cette dernière inclue des facteurs généraux, comme la chaperonne HSP90 et les hélicases Rvb1p/TIP49 et Rvb2p/TIP48, qui interviennent dans de nombreux processus cellulaires. Elle fait également intervenir des facteurs vraisemblablement plus spécifiques, comme les protéines Rsa1p/NUFIP, Pih1p/PIH1 et Tah1p/SPAGH. Les bases moléculaires de son fonctionnement n'ont pas encore été mises à jour, mais il semblerait que cette machinerie se décompose en deux modules fonctionnels distinct : (i) le facteur Rsa1p/NUFIP, qui reconnaît spécifiquement les protéines de la famille L7Ae et facilite le recrutement des autres protéines sur les pré-RNP, (ii) le complexe R2TP qui, en association avec HSP90, mettrait à profit les activités hélicases et chaperonne afin d'assurer le bon déroulement du processus d'assemblage. De plus, il semblerait que ces deux modules puissent agir de façon commune, puisqu'une interaction conservée entre Rsa1p/NUFIP et Pih1p/PIH1 a été détectée *in vitro*.

En parallèle de nos travaux, l'équipe de N.J. Watkins a également entrepris la caractérisation des facteurs requis pour l'assemblage des snoRNP à boîtes C/D humaines (McKeegan *et al.*, 2007). Cette étude a également permis de mettre en évidence un rôle de plateforme d'assemblage pour la protéine NUFIP. En effet, comme nous l'avons démontré pour la protéine Rsa1p, NUFIP s'est avérée essentielle pour permettre le recrutement des protéines NOP56 et NOP58 sur le snoRNA U14 en présence de la protéine 15.5K *in vitro* (McKeegan *et al.*, 2007). De plus, cette équipe a également caractérisé plusieurs interactions entre les facteurs d'assemblage et les protéines constitutives des snoRNP (McKeegan *et al.*, 2007) (**Tableau 5**). La protéine recombinante GST-NUFIP interagit notamment avec les protéines 15.5K, Fibrillarine, NOP56 et NOP58, tandis que les hélicases TIP49 et TIP48 interagissent avec la Fibrillarine et la protéine 15.5K. De façon intéressante, le spectre d'interaction des deux hélicases pourrait même être régulé par son cycle d'hydrolyse d'ATP, puisque l'interaction de la protéine 15.5K avec TIP48 ou TIP49 n'est observée qu'en présence d'ATP.

D'autre part, une étude, réalisée par l'équipe de W.A. Houry et publiée conjointement à nos travaux, a également révélé l'implication du complexe R2TP et de HSP90 dans la biogenèse des snoRNP à boîtes C/D chez la levure *S. cerevisiae* (Zhao *et al.*, 2008). La caractérisation

Facteurs d'assemblage		Interactions détectées		Effets KO/Inhibition sur la stabilité des constituants des RNP					
		Protéines constitutives des RNP	Facteurs d'assemblage	C/D	H/ACA	U4	TERC	mRNA <sup>sel</sup>	Protéines affectées
<b>Rsa1p</b> <b>NUFIP</b>	Levure	Snu13p, Nop58p, Prp31p	Pih1p, Rvb1p, Rvb2p	↓↓	(↓)	→			non testé
	Homme	15.5K, NHP2, SBP2, Fibrillarine, NOP56*, NOP58*, PRP31, U3-55K	PIH1, TIP49, TIP48	↓↓	(↓)	→	↓↓	non testé	non testé
<b>Pih1p</b> <b>PIH1</b>	Levure	Nop58p, Prp31p, Rrp9p	Tah1p, Rvb1p, Rvb2p, Rsa1p	↓↓**	→**	non testé			aucun effet**
	Homme	SBP2, Fibrillarine*	SPAGH, TIP49*, NUFIP	↓↓*	non testé	non testé	non testé	non testé	non testé
<b>Tah1p</b> <b>SPAGH</b>	Levure		Pih1p	(↓)**	(↓)**	non testé			aucun effet**
	Homme		PIH1, TIP48, HSP90	non testé	non testé	non testé	non testé	non testé	non testé
<b>Rvb1p</b> <b>TIP49</b>	Levure	Prp31p	Rsa1p, Pih1p, Rvb2p	↓↓**	↓↓**	non testé			aucun effet**
	Homme	15.5K*, Fibrillarine*, NOP56*, NOP58*	TIP48, PIH1*	↓↓*	non testé	non testé	non testé	non testé	non testé
<b>Rvb2p</b> <b>TIP48</b>	Levure	Prp31p	Rsa1p, Pih1p, Rvb1p	↓↓**	↓↓**	non testé			aucun effet**
	Homme	15.5K*, Fibrillarine*, NOP56*, NOP58*	TIP49, SPAGH	↓↓*	non testé	non testé	non testé	non testé	non testé
<b>Hsp82p</b> <b>HSP90</b>	Levure	non testé	Tah1p	(↓)**	(↓)**	non testé			aucun effet**
	Homme		SPAGH	↓↓	non testé	(↓)	↓↓	non testé	15.5K, SBP2, NHP2, NOP58

**Tableau 5 : Tableau récapitulatif des liens établis entre les facteurs d'assemblage et les composants des RNP de la famille L7Ae chez la levure *S. cerevisiae* et chez l'homme.** Les noms de protéines mentionnés en bleu indiquent que l'interaction correspondante a été vérifiée *in vitro* par une approche de pull-down. Les noms en italique indiquent que l'interaction correspondante requiert la présence d'ATP. Les noms qui apparaissent en vert correspondent à des interactions observées uniquement par double-hybride. Seules les interactions double-hybride observées à partir des protéines de pleine taille sont mentionnées. Les effets "KO" ("Knock-Out") correspondent à la caractérisation de souches de levures invalidées pour le gène d'intérêt, ou de cellules humaines traitées par des siRNA spécifiques du facteur étudié. L'effet "inhibition" correspond à la caractérisation de cellules dans lesquelles la chaperonne HSP90 est inhibée par la geldanamycine. Le symbole ↓↓ indique un effet majeur sur le taux d'ARN, le symbole (↓) indique un effet modéré et/ou nécessitant des expériences complémentaires. Le symbole → indique l'absence d'effet négatif sur le taux d'ARN. Les cases grisées correspondent aux RNP de la famille L7Ae non conservées chez la levure. \*Résultats obtenus par l'équipe de N.J. Watkins (Watkins *et al.*, 2004; McKeegan *et al.*, 2007 et 2009). \*\*Résultats publiés dans Zhao *et al.*, 2008. Les autres données sont relatives à la publication Boulon *et al.*, 2008.



de souches de levures mutantes a notamment permis de mettre en évidence que la protéine Pih1p est requise pour préserver un taux optimal de snoRNA à boîtes C/D, tandis que les hélicases Rvb1p et Rvb2p sont requises pour assurer à la fois la stabilité des snoRNA à boîtes C/D et H/ACA (**Tableau 5**). Le rôle joué par la protéine Tah1p et par la chaperonne HSP90 (*i.e.* Hsc82p et Hsp82p) s'est en revanche avéré plus discret. En comparaison avec le phénotype  $\Delta PIH1$ , des analyses sur puces à ADN ont en effet montré que leur absence conduit à une diminution faible et relativement hétérogène du taux de snoRNA à boîtes C/D. De plus, ces effets sont uniquement observés au cours de la phase stationnaire de croissance, qui correspond à une période de stress, au cours de laquelle la protéine Hsp82p est surexprimée. Durant cette même phase, l'absence de Pih1p conduit également à une augmentation de l'association entre Nop56p et Hsp82p, au détriment de l'association entre Nop56p et Snu13p (Zhao *et al.*, 2008). Ce résultat suggère donc un rôle de maintenance de la chaperonne, qui pourrait contribuer à préserver l'intégrité des différents constituants des snoRNP, en cas de défaut d'assemblage et/ou de conditions défavorables. Cependant, la nature précise de cette fonction demeure incertaine, puisque l'absence de la chaperonne, comme celle des différents constituants du complexe R2TP, n'a aucun effet sur le taux des protéines Snu13p, Nop1p, Nop56p et Nop58p (Zhao *et al.*, 2008). Nous reviendrons plus en détail sur ces aspects dans le quatrième chapitre des résultats.

### **3.2. Quel mode d'action pour la machinerie d'assemblage des RNP ?**

L'ensemble de ces données apporte donc des premiers éléments de réflexions quant à l'établissement d'un modèle pour l'assemblage des RNP de la famille L7Ae et plus particulièrement des snoRNP à boîtes C/D. Le rôle de plateforme joué par le facteur Rsa1p/NUFIP semble clairement établi, en revanche, la fonction précise du système R2TP-HSP90 et des activités qui lui sont associées, sont encore incertaines.

Le co-fractionnement, dans des extraits cytoplasmiques de cellules humaines, de facteurs d'assemblage et de protéines constitutives des snoRNP à boîtes C/D, suggère que le processus d'assemblage pourrait débuter par la formation de pré-complexes au niveau du cytoplasme (Boulon *et al.*, 2008). Certains facteurs pourraient en effet assurer la prise en charge des protéines nouvellement synthétisées et participer à leur import vers le noyau. Chez la levure, il a notamment été montré que l'absence des protéines Rvb2p et Pih1p conduit à un défaut de

la localisation nucléolaire des protéines des snoRNP à boîtes C/D et à leur accumulation diffuse au niveau du nucléoplasme et du cytoplasme (Gonzales *et al.*, 2005; King *et al.*, 2001). Une hypothèse privilégiée est donc que le complexe R2TP assure la prise en charge de certaines protéines, depuis leur site de synthèse jusqu'à leur incorporation au sein des RNP. La phase d'assemblage à proprement parler, qui implique le facteur Rsa1p/NUFIP, semble également nécessiter les activités hélicases des protéines Rvb1p/TIP49 et Rvb2p/TIP48 (McKeegan *et al.*, 2007; Watkins *et al.*, 2004). Les changements conformationnels, induits par le biais de l'hydrolyse d'ATP, pourraient notamment permettre de dissocier certaines associations entre composants des RNP et facteurs d'assemblage, afin de favoriser la mise en place de nouvelles interactions prévalentes au sein des particules matures (Boulon *et al.*, 2008; McKeegan *et al.*, 2007; Watkins *et al.*, 2004). Ces événements de remodelage pourraient ainsi favoriser la transition des pré-RNP vers un état stable, indépendant de la machinerie d'assemblage et efficient pour l'adressage nucléolaire.

Concernant l'action de la chaperonne, deux hypothèses principales sont envisagées. Sa capacité à co-immunosélectionner les intermédiaires de maturation du snoRNA U3 humain suggère que HSP90 pourrait exercer son activité tout au long du processus de biogenèse, y compris au sein du noyau (Boulon *et al.*, 2008). Néanmoins, étant donné sa forte abondance dans la cellule, il est également possible que la co-immunosélection des pré-snoRNA résulte d'une association artificielle se déroulant lors de la préparation de l'extrait cellulaire (Boulon *et al.*, 2008; Kittur *et al.*, 2006). Une autre possibilité est donc que HSP90 intervienne uniquement au cours de la phase cytoplasmique du processus de biogenèse. En association avec le complexe R2TP, la chaperonne pourrait notamment assurer une prise en charge précoce de certaines protéines, afin de maintenir leur stabilité et de promouvoir la formation de pré-complexes d'assemblages (Boulon *et al.*, 2008).

### **3.3. Questions liées à la mise en évidence de la machinerie d'assemblage**

Dans ce premier chapitre, nous avons abordé la mise à jour des principaux facteurs impliqués dans l'assemblage des RNP de la famille L7Ae. Ces données ont notamment révélé que les différents acteurs de ce processus pourraient agir de façon coordonnée, au sein d'une même « machinerie ». Les bases moléculaires de son fonctionnement sont cependant encore peu connues et de nombreux points restaient à éclaircir. Plusieurs questions majeures, ainsi que les

perspectives qui en découlent, sont énoncées ci-dessous. La plupart d'entre elles ont été abordées au cours de ma thèse et feront l'objet des prochains chapitres.

- **Quel est le spectre d'action de la machinerie d'assemblage ?**

Les résultats obtenus chez l'homme attestent de l'intervention des facteurs d'assemblage dans la biogenèse de plusieurs RNP de la famille L7Ae : les snoRNP à boîtes C/D et H/ACA, la snRNP U4, le complexe télomérase et les mRNP SECIS (Boulon *et al.*, 2008). Ce spectre d'action semble guidé par la capacité de la protéine NUFIP à interagir à la fois avec les protéines 15.5K, NHP2 et SBP2. Chez la levure en revanche, seul leur rôle dans la biogenèse des snoRNP à boîtes C/D a clairement été démontré (**Tableau 5**). A l'exception des hélicases Rvb1p et Rvb2p, les autres facteurs ne semblent pas exercer une réelle influence sur l'accumulation des snoRNA à boîtes H/ACA. De plus, à l'exception d'une interaction détectée entre Rsa1p et Prp31p, il existe peu d'informations quant à leur rôle potentiel dans l'assemblage de la snRNP U4. Il serait donc intéressant d'établir si la machinerie d'assemblage de levure joue un rôle restreint à l'assemblage des snoRNP à boîtes C/D ou si elle exerce, comme chez l'homme, une fonction étendue à d'autres RNP de la famille L7Ae. Afin de répondre à cette question, il sera nécessaire de tester de façon plus systématique l'importance des différents facteurs d'assemblage pour l'accumulation des snoRNA à boîtes H/ACA et du snRNA U4. Il sera également important de vérifier si Rsa1p possède la capacité d'interagir avec la protéine Nhp2p. Enfin, nous aborderons également la question plus spécifique du mode d'assemblage de la protéine Rrp9p sur le motif B/C du snoRNA U3 de levure.

- **Quels éléments gouvernent la reconnaissance spécifique des protéines de la famille L7Ae par le facteur Rsa1p/NUFIP ?**

Dans l'article *The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery*, nous avons mis en évidence la capacité du facteur Rsa1p/NUFIP à interagir avec plusieurs protéines de la famille L7Ae (*i.e.* Snu13p chez la levure, 15.5K, NHP2, SBP2 chez l'homme) (Boulon *et al.*, 2008). Cette propriété unique lui confère un rôle de facteur spécialisé, capable de cibler les pré-RNP contenant une protéine de la famille L7Ae, afin de faciliter leur assemblage. Des données préliminaires indiquent que le domaine PEP, conservé entre Rsa1p et NUFIP, constitue un module dédié à la reconnaissance des protéines de la

famille L7Ae (Boulon *et al.*, 2008). En effet, ce domaine est nécessaire et suffisant pour permettre l'interaction conservée entre Rsa1p/NUFIP et Snu13p/15.5K. Chez l'homme, le domaine PEP s'est révélé non suffisant, mais néanmoins nécessaire pour permettre l'interaction de NUFIP avec les protéines NHP2 et SBP2. Enfin, à titre de comparaison, aucune interaction n'a été détectée entre NUFIP et la protéine L30. Ces différentes données suggèrent donc que l'intervention spécifique du facteur Rsa1p/NUFIP pourrait reposer sur la présence de déterminants conservés en surface de certaines protéines de la famille L7Ae. Dans le but de comprendre le mécanisme d'action ciblé de la machinerie d'assemblage, il sera donc nécessaire de caractériser le mode de reconnaissance des protéines de la famille L7Ae par le domaine PEP.

- **Quels liens existent-t-il entre Rsa1p/NUFIP et le complexe R2TP ?**

Le facteur Rsa1p/NUFIP et le complexe R2TP constituent vraisemblablement les deux principaux modules de la machinerie d'assemblage, cependant il existe encore peu de données fonctionnelles attestant de leur coopération. Les tests d'interactions réalisés chez la levure et chez l'homme ont toutefois révélé des liens potentiels entre ces deux entités (Boulon *et al.*, 2008) (**Tableau 5**). Une interaction directe entre Rsa1p/NUFIP et Pih1p/PIH1, a notamment été détectée chez la levure et chez l'homme. Une perspective intéressante sera donc d'étudier plus en détail le lien entre ces deux protéines et de déterminer si Rsa1p/NUFIP et les membres du complexe R2TP interagissent fonctionnellement au cours des événements d'assemblage des RNP.

- **Quelle est la nature des complexes d'assemblage des snoRNP à boîtes C/D ?**

L'étude de la biogenèse des UsnRNP et des snoRNP à boîtes H/ACA a révélé l'existence de différents sous-complexes intermédiaires, impliqués de façon séquentielle dans la formation des RNP (cf. Introduction, Chapitre 3). Concernant les snoRNP à boîtes C/D, plusieurs travaux ont également décrit l'existence de complexes macromoléculaires, constitués de différents facteurs associés aux composants des RNP matures (Boulon *et al.*, 2008; Watkins *et al.*, 2004) (cf. Introduction, Chapitre 3, 4.3.2). Une perspective majeure de ce travail sera donc de réaliser une caractérisation fine de ces différents complexes. L'utilisation couplée des techniques d'immunosélection et d'analyse systématique par spectrométrie de masse pourra notamment fournir une image plus exhaustive de leur composition protéique. De plus,

l'analyse, par des approches de Northern-blot ou de RT-PCR, des ARN co-immunoprécipités pourra nous renseigner sur la nature des ARN précurseurs présents dans ces complexes.

- **La machinerie d'assemblage intervient-elle de façon co-transcriptionnelle ?**

Comme nous l'avons abordé dans l'introduction, l'assemblage des snoRNP à boîtes H/ACA débute de façon co-transcriptionnelle, tandis que l'assemblage des UsnRNP est un processus exclusivement post-transcriptionnel. Concernant les snoRNA à boîtes C/D intronique, le recrutement des protéines de la snoRNP semble obéir à un processus co-transcriptionnel (cf. Introduction, Chapitre 3, 4.5). En revanche, peu de données sont actuellement disponibles quant à l'assemblage des particules contenant un snoRNA à boîte C/D cistronique. L'utilisation de l'approche d'immunoprécipitation de la chromatine (ChIP), à partir des protéines constitutives des snoRNP et des différents facteurs d'assemblage, constituent donc une perspective de travail majeure. D'autre part, il sera important de caractériser les étapes se déroulant de façon post-transcriptionnelle en étudiant l'association des protéines d'intérêt avec les différents intermédiaires de maturation des snoRNA. Chez l'homme, les facteurs NUFIP, TIP49, TIP48, PIH1, SPAGH et HSP90 ont été détectés en association avec les formes intermédiaires et matures de certains snoRNA à boîtes C/D et du snRNA U4 (Boulon *et al.*, 2008; Watkins *et al.*, 2004). Chez la levure en revanche, seule la protéine Rsa1p, qui s'associe avec un précurseur du snoRNA U3, a fait l'objet d'une telle étude (Boulon *et al.*, 2008).

- **Quel est la nature précise du rôle joué par les activités hélicases et chaperonnes ?**

Afin de comprendre le mécanisme d'action de la machinerie d'assemblage, il sera également nécessaire d'étudier comment les activités hélicases et chaperonnes sont mises à profit au cours du processus de biogenèse. Etant donné le caractère essentiel et ubiquitaire de ces protéines, l'étude découlée des processus qui leur sont associés constitue une tâche difficile. L'implication des protéines Rvb1p/TIP49 et Rvb2p/TIP48 dans des événements de remodelage des RNP, pourrait notamment être analysée grâce à l'utilisation de formes variantes, délétères pour l'activité ATPase. La résolution de la structure cristallographique de la protéine humaine TIP49 offre en effet la possibilité d'utiliser différents mutants protéiques et de tester leur impact sur le processus d'assemblage (Matias *et al.*, 2006). Concernant la chaperonne HSP90, la situation apparaît plus favorable, puisque la protéine Tah1p/SPAGH

pourrait être un cofacteur spécifique des processus d'assemblage des RNP (Boulon *et al.*, 2008; Zhao *et al.*, 2008). L'utilisation d'une souche de levure invalidée pour le gène *TAH1* ou la caractérisation de mutants d'interactions avec la chaperonne, pourraient notamment constituer des outils intéressants dans l'optique d'étudier de façon spécifique le rôle de HSP90 dans la biogenèse des RNP.

- **L'intervention de cette machinerie est-elle essentielle ?**

Enfin, parmi les facteurs d'assemblage caractérisés chez la levure, plusieurs acteurs majeurs, comme les protéines Rsa1p et Pih1p, se sont révélés non essentiels à la viabilité cellulaire. Leur absence conduit en effet à un simple phénotype de croissance ralentie, exacerbé à 37°C (Gonzales *et al.*, 2005; Kressler *et al.*, 1999). Il semble donc que ces facteurs ne soient pas indispensables pour la biogenèse des particules, dont certaines comme les snoRNP U3 et U14, sont requises pour assurer la viabilité cellulaire. Cette observation pose donc la question du caractère essentiel de la machinerie d'assemblage ? Il est possible que la richesse du réseau d'interactions mis à jour permette de compenser l'absence de certaines de ses sous-unités. Une perspective intéressante sera donc de caractériser des souches de levures invalidées pour plusieurs gènes d'intérêt afin de vérifier l'importance de la machinerie d'assemblage pour la biogenèse des RNP et pour la viabilité cellulaire.

## CHAPITRE 2

### Caractérisation structurale et fonctionnelle de l'interface d'interaction entre Rsa1p/NUFIP et Snu13p/15.5K

#### 1. Présentation des travaux

Parmi les différents systèmes d'assemblage caractérisés à ce jour (cf. Introduction, Chapitre 3), le facteur Rsa1p/NUFIP possède la caractéristique singulière de reconnaître plusieurs protéines appartenant à une même famille et exerçant des rôles distincts au sein de plusieurs catégories de RNP. Chez l'homme, NUFIP interagit en effet avec les protéines 15.5K, NHP2 et SBP2, tandis que la protéine Rsa1p de levure interagit avec la protéine Snu13p (Boulon *et al.*, 2008). Les résultats présentés dans le chapitre précédent indiquent que cette propriété implique un domaine protéique dédié à la reconnaissance des protéines de la famille L7Ae. Ce dernier, baptisé domaine PEP, est en effet essentiel pour l'ensemble des interactions caractérisées à ce jour entre Rsa1p/NUFIP et les protéines de cette famille (Boulon *et al.*, 2008). De plus, ce domaine constitue également l'un des rares fragments d'homologie entre Rsa1p et NUFIP qui, en dépit d'une grande homologie fonctionnelle, présentent de grandes divergences de séquence (**Fig. 53**).

Afin de mieux comprendre le mécanisme d'action ciblé de Rsa1p/NUFIP et de la machinerie d'assemblage qui leur est associée, nous avons décidé de caractériser en détail l'implication du domaine PEP dans la reconnaissance des protéines de la famille L7Ae. La conservation de ce domaine au sein du *phylum* eucaryote (**Fig. 53**), ainsi que la grande homologie structurale existant entre les protéines de la famille L7Ae (cf. Introduction, Chapitre 2, 1.3), suggère que ce mode de reconnaissance pourrait reposer sur des éléments de structure relativement bien conservés au cours de l'évolution. En collaboration avec l'équipe d'E. Bertrand (IGM, Montpellier) et avec B. Maigret (INRIA, Nancy), nous avons donc entrepris leur identification grâce à la mise en place d'une stratégie de mutagenèse, couplée à diverses approches de biochimie et à l'emploi des techniques de spectrométrie RMN et de modélisation moléculaire. Cette démarche, centrée sur l'interaction entre Rsa1p/NUFIP et Snu13p/15.5K, nous a permis d'établir un modèle tridimensionnel de l'interface d'interaction entre les deux partenaires de

levure. Dans ce modèle, des acides aminés exposés en surface de l'hélice  $\alpha 3$  de Snu13p et conservés au sein de la famille L7Ae, forment un réseau d'interactions électrostatiques avec des résidus du domaine  $\gamma$ PEP. De façon intéressante, la mutation ponctuelle des résidus apportant une contribution majoritaire à ce réseau, abolit l'interaction des deux protéines et conduit à une diminution spécifique du taux de snoRNA à boîtes C/ D, associée à un défaut de croissance des cellules de levure. Des travaux complémentaires, visant à élargir cette étude aux autres protéines de la famille L7Ae, ont été réalisés. Ils nous ont notamment permis de mettre en évidence que le spectre de reconnaissance de la protéine Rsa1p de levure se restreint vraisemblablement à la protéine Snu13p.

## 2. Résultats

L'ensemble de ce travail est décrit dans un article qui sera soumis pour publication dans la revue *Nucleic Acid Research* et dans lequel je suis co-premier auteur. Des données supplémentaires, concernant en particulier la protéine Nhp2p, seront également présentées.

### 2.1. Publication n°2



---

---

Publication n°2

**Formation of electrostatic and hydrophobic interactions between alpha helices of the Rsa1p assembly factor and the Snu13p snoRNP protein is required for efficient snoRNP assembly**

Benjamin Rothé\*, Régis Back\*, Xavier Manival, Marie-Cécile Robert, Marc Quinternet, Magali Blaud, Christophe Romier, Michael Leichter, Bernard Maigret, Bruno Charpentier, Edouard Bertrand and Christiane Branlant

**En préparation pour *Nucleic Acid Research***

\* Contribution équivalente

---

---



# Formation of electrostatic and hydrophobic interactions between alpha helices of the Rsa1p assembly factor and the Snu13p snoRNP protein is required for efficient snoRNP assembly

Benjamin Rothé<sup>1</sup>, Régis Back<sup>1</sup>, Xavier Manival<sup>1\*</sup>, Marie-Cécile Robert<sup>2</sup>, Marc Quinternet<sup>1</sup>, Magali Blaud<sup>1</sup>, Christophe Romier<sup>3</sup>, Michael Leichter<sup>4</sup>, Bernard Maigret<sup>5</sup>, Bruno Charpentier<sup>1\*</sup>, Edouard Bertrand<sup>2\*</sup>, and Christiane Branlant<sup>1</sup>

<sup>1</sup>AREMS (ARN, RNP, structure-fonction-maturation, Enzymologie Moléculaire et Structurale), Université de Lorraine, Unité Mixte de Recherche 7214 CNRS-UHP Nancy 1, Faculté de Médecine, Bâtiment Biopôle, BP 184, Vandoeuvre-lès-Nancy Cédex, France.

<sup>2</sup>IGMM (Institute of Molecular Genetics of Montpellier), Centre National de la Recherche Scientifique, Unité Mixte de Recherche 5535, Montpellier Cédex 5, France.

<sup>3</sup>IGBMC (Institut de Génétique et Biologie Moléculaire et Cellulaire), Département de Biologie et Génomique Structurales, Université de Strasbourg, CNRS, INSERM, 1 Rue Laurent Fries, BP 10142, 67404 Illkirch Cédex, France.

<sup>4</sup>Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, Institut de Biologie Moléculaire et Cellulaire, 15 rue Descartes, 67084 Strasbourg Cédex, France.

<sup>5</sup>LORIA, Groupe ORPAILLEUR, Nancy Université, Campus scientifique, BP 70239, 54506 Vandoeuvre-lès-Nancy Cédex, France.

\*To whom correspondence should be addressed. Tel: +33 3 83 68 55 08 ; Fax: +33 3 83 68 43 07; Email: [bruno.charpentier@maem.uhp-nancy.fr](mailto:bruno.charpentier@maem.uhp-nancy.fr). Correspondence may also be addressed to Xavier Manival. Tel: +33 3 83 68 55 65. Fax: +33 3 83 68 43 07; E-mail: [xavier.manival@maem.uhp-nancy.fr](mailto:xavier.manival@maem.uhp-nancy.fr). Correspondence may also be addressed to Edouard Bertrand. Tel: +33 4 67 61 36 47. Fax: +33 4 67 04 02 31; E-mail: [edouard.bertrand@igmm.cnrs.fr](mailto:edouard.bertrand@igmm.cnrs.fr)

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

Present address : Magali Blaud, Laboratoire de Cristallographie et RMN Biologiques, Université Paris Descartes, CNRS-UMR8015, Paris Cédex, France. Michael Leichter, Molecular Cell Physiology, Faculty for Biology and Institute for Genome Research and System Biology, Bielefeld University, Universitaetsstrasse 25, D-33615 Bielefeld, Germany.

## ABSTRACT

Proper assembly of non-coding RNAs into RNPs is essential to their function. Protein NUFIP and its yeast ortholog Rsa1p participate to the assembly of C/D box small nucleolar ribonucleoprotein particles (snoRNPs) and U4 small nuclear ribonucleoprotein particle (snRNP). These RNPs share a common L7Ae-like protein (15.5K in human, Snu13p in yeast). By interaction with 15.5K/Snu13p, NUFIP/Rsa1p is proposed to scaffold immature particles and to recruit HSP90 through a specific co-chaperone, the R2TP complex. However, the precise mode of interaction of 15.5K/Snu13p with NUFIP/Rsa1p is not known. Here, we used various experimental approaches to identify key residues in this interaction and by computer modeling we build a structural model of the Snu13p-Rsa1p interaction. In this model, amino acids R<sub>249</sub> and R<sub>246</sub> from Rsa1p and E<sub>72</sub> and D<sub>73</sub> from Snu13p form three ion pairs stabilized by the hydrophobic residues W<sub>245</sub> and L<sub>69</sub>. By mutation of interacting residues in yeast, we determined which ones contribute to cell growth. We also showed the importance of the 15.5K EDK motif for sub-cellular localization and based on the U4-15.5K-PRP31<sup>78-333</sup> structure, we predict that NUFIP binds 15.5K opposite to the RNA and at the border of the PRP31 interaction, suggesting an external scaffold role during RNP assembly. Altogether, this provides molecular bases to understand the NUFIP/Rsa1p functions in RNP assembly.

## INTRODUCTION

Ribonucleoprotein particles (RNPs) containing proteins of the L7Ae family (L7Ae-like protein) are essential for many cellular processes. Indeed, this class of RNPs includes: i) small nucleolar RNPs (box C/D and box H/ACA snoRNPs), which are involved in ribosomal RNA (rRNA) maturation and post-transcriptional modification; ii) the

spliceosomal U4 small nuclear RNP (U4 snRNA), which is essential for pre-mRNA splicing; iii) small Cajal body-specific RNPs (box C/D and H/ACA scaRNPs), which modifies post-transcriptionally spliceosomal UsnRNPs; iv) the vertebrate telomerase holoenzyme that is required for telomere synthesis; and v) specific messenger RNPs formed in the 3' UTR regions of selenoprotein mRNAs (SECIS mRNP), which are required for selenocysteine incorporation. In their active forms,

these RNPs consist of an RNA molecule associated with a specific set of proteins, including an L7Ae-like protein: 15.5K in human (Snu13p in yeast) for U4 snRNP, box C/D snoRNPs and scaRNPs (1,2); protein NHP2 (Nhp2p in yeast) for box H/ACA snoRNPs, scaRNPs, and the telomerase RNP (3,4); and protein SBP2 for selenocysteine mRNPs (5,6).

Recent studies revealed an unexpected complexity of RNP biogenesis pathways (7,8). RNAs are produced in a precursor form that is subsequently processed and assembled with proteins, and this often occurs in specific sub-cellular compartments and requires the aid of dedicated assembly factors. For instance, in the case of UsnRNPs, the pre-UsnRNA is exported to the cytoplasm for assembly with the Sm proteins and 3' end and 5' end processing (9). Then, the pre-UsnRNPs are imported to nucleus, in Cajal bodies to complete RNA maturation and protein assembly, and finally delivered to sites of splicing. Association of Sm proteins with UsnRNAs is mediated by the survival of motor neurons (SMN) complex (10,11). This complex binds both the RNA and the various Sm proteins, holds these proteins together from their outside surface to form an open ring, and present them to the RNA. The SMN complex thus works as an external scaffold to promote RNP formation, and also ensures the specificity of RNP assembly.

SnoRNPs provide a different paradigm for RNP assembly. For instance, studies on H/ACA snoRNPs have shown that the assembly factor NAF1/Naf1p functions by mimicking and replacing a core protein in immature particles, rather than as an external scaffold (12-14). RNPs containing 15.5K/Snu13p are able to form three different RNPs depending on the sequence that surrounds the 15.5K/Snu13p binding site. They thus provide an excellent model to understand how the specificity of assembly is achieved. 15.5K/Snu13p proteins specifically recognize K-turn motifs that include two helices, I and II, which are separated by a 3-nt long bulge loop containing a protruding U residue at position 3. This protruding residue and two non-canonical sheared G:A base pairs in helix II are essential for the interaction with 15.5K/Snu13p (15,16). Binding of 15.5K/Snu13p allows further recruitment of core RNP proteins: PRP31 on U4 snRNA, U3-55K (Rrp9p in yeast) on the U3 B/C motif, and NOP56, NOP58, and Fibrillarin (Nop1p in yeast) on C/D snoRNAs (17-19). The nucleotide sequence of helix II determines which protein is recruited after 15.5K/Snu13p binding (17-19), and the recent determination of the crystal structure of the PRP31 fragment encompassing amino acids 78

to 346 (PRP31<sub>78-333</sub>) in complex with 15.5K and a fragment of U4 snRNA (20), and the crystal structures of archaeal box C/D RNPs has provided some clues for the specificity of assembly of this pair of proteins (21-23).

Recent data show that *in vivo* assembly of RNPs containing L7Ae-like proteins also involves assembly factors, and we showed previously that the yeast protein Rsa1p and its human counterpart NUFIP play an important role in this process (24-26). These proteins directly bind 15.5K/Snu13p, and we proposed two functions for these interactions: (i) to hold together core proteins in the immature particles, and (ii) to recruit the HSP90 chaperone through interactions with its R2TP co-chaperone complex. This complex is composed of four proteins (24,26,27): hSpagh (also named RPAP3, and in yeast Tah1p), which binds directly to HSP90; PIH1/NOP17 (Pih1p or Nop17p in yeast), which binds directly to NUFIP/Rsa1p; and two AAA+ ATPases TIP48 (also known as Reptin or p50 in vertebrates, and Rvb2p in yeast) and TIP49 (also known as Pontin or p55 in vertebrates, and Rv1bp in yeast). Inhibition of HSP90 induced the disappearance of newly-synthesized U4, box C/D and telomerase RNAs, as well as of 15.5K, NOP58 and NHP2 proteins, suggesting that these proteins are clients of the chaperone (24).

NUFIP and Rsa1p proteins directly bind 15.5K and Snu13p. We have no information on how this interaction takes place at the atomic level, although this knowledge is a prerequisite to understand the mechanisms of action of the NUFIP/Rsa1p assembly factor. In this work, we used yeast genetics, biochemical assays, NMR spectroscopy, and molecular modeling experiments to characterize the interactions between NUFIP/Rsa1p and 15.5K/Snu13p at the atomic level and to define its function. By substitution in yeast cells of the interacting amino acids, we demonstrated their importance for cell growth and the stability of box C/D snoRNAs. Altogether, we bring new important insight on the mechanism of action of an RNP assembly factor involved in both snoRNP and U4 snRNP assembly.

## MATERIAL AND METHODS

**Recombinant plasmids used in this study.** Plasmids were manipulated by standard techniques. Mutations were generated using the QuickChange mutagenesis system according to the manufacturer's instructions (Stratagene<sup>®</sup>) and appropriate primer pairs. The bacterial expression vector pGEX-6P-1::SNU13 was already described

(15). To express the C-terminal 152 amino acids fragment of Rsa1p in *E. coli*, a *BglII-XhoI* restriction fragment was released by digestion of plasmid pACTII::Rsa1p<sub>230-381</sub> (previously designated as pACTII::N3C1) (24) and we cloned it into plasmid pET30c (Novagen<sup>®</sup>). A 111 bp fragment containing the Rsa1p coding sequence of the amino acids 230 to 266 was obtained by PCR amplification using two primers that generated a *NheI* site at one and an *XhoI* site at the other end. This fragment was inserted between the *NheI* and *XhoI* sites of plasmid pET28a (Novagen<sup>®</sup>) to obtain the *E. coli* expression vector pET28a::Rsa1p<sub>230-266</sub>.

For protein co-expression in *E. coli*, DNA fragments encoding the full length Rsa1p, the Rsa1p<sub>230-381</sub> fragment or protein Pih1p were amplified by standard PCR procedures, using primers that generated *NdeI* and *BamHI* restriction sites, and they were inserted in the pnCS vector. On the other hand, the Snu13p and Tah1p ORFs were inserted in the pnEA-3CH vector in order to produce His-Tagged proteins (28-30).

For expression of WT or mutated Snu13p or Rsa1p protein in yeast, the *S. cerevisiae* *SNU13* and *RSAl* ORFs were PCR amplified using primers that generated a *BamHI* site at the ORF 5' end and an *XhoI* site at its 3' end, and were cloned between the *BamHI* and *SalI* sites of a centromeric version of the pG1 (*TRP1*) expression vector (31,32). The resulting plasmids were designated as pG1::SNU13 and pG1::RSAl.

The recombinant plasmids used for Y2H assays were already described (24). Briefly, the ORFs were cloned by the Gateway system (Invitrogen<sup>™</sup>) in plasmid pACTII downstream from the *GAL4* transactivation domain coding sequence, or in plasmid pAS2 downstream from the *GAL4* DNA binding domain coding sequence.

The plasmids coding for mRFP-LacI-NUFIP and GFP-15.5K were generated with the Gateway technology with destination vectors previously described (33).

**Production of recombinant proteins.** The recombinant pGEX-6P-1::SNU13 plasmid was used to produce Snu13p fused to the glutathione S-transferase (GST) (15) in the *E. coli* CodonPlus strain (Novagen<sup>®</sup>). The recombinant GST-Snu13p protein was purified from cellular extract under native conditions, using Glutathione-Sepharose 4B as recommended by the manufacturer (GE Healthcare<sup>®</sup>). The GST-Snu13p proteins were recovered after overnight cleavage on the beads using the PreScission protease (GE Healthcare<sup>®</sup>). The cleaved recombinant Snu13 proteins had two

additional amino acid residues at their N-terminus (G and P) compared to the native protein. The recombinant pET30c::Rsa1p<sub>230-381</sub> and pET28a::Rsa1p<sub>230-266</sub> expression vectors were used to produce the Rsa1p fragments encoding amino acids 230 to 381, and 230 to 266, respectively, with an His6 N-terminal tag in the *E. coli* CodonPlus strain (Novagen<sup>®</sup>). Note that the peptide produced by plasmid pET28a::Rsa1p<sub>230-266</sub> contains at its N-terminus the His6-tag sequence and a 15 amino acid sequence resulting from the cloning into the pET28a plasmid. The recombinant His6Rsa1p<sub>230-266</sub> and His6Rsa1p<sub>230-381</sub> proteins were purified under native and denaturing/refolding conditions, respectively, using a 5 mL His-Trap<sup>™</sup> HP column (GE Healthcare<sup>®</sup>). Proteins were eluted with 0.5 M imidazole at pH 8. When necessary, the His<sub>6</sub>-tag was removed by cleavage with the Thrombin protease. All recombinant proteins, either after Thrombin cleavages, or imidazole elution, were dialyzed against the suitable buffer.

**Test of protein-protein interactions by their co-expression in *E. coli*.** Co-expression tests in *E. coli* were carried out using a standard procedure described previously (29). The two recombinant plasmids pET15b and pACYC18b-11b expressing, respectively, His6Snu13p and full-length Rsa1p or its Rsa1p<sub>230-381</sub> fragment were used simultaneously to transform pRare2 *E. coli* host cells (Novagen). Growth was performed in 500 mL of 2X concentrated LB medium at 37°C to an absorbance at 600 nm of 0.8. Then, the temperature was switched to 20°C. At this point, co-expression of the proteins was induced by addition of isopropyl-β-D-thiogalactopyranoside (Euromedex) at a 1 mM final concentration. Cells were further grown overnight at 20°C. They were collected by low speed centrifugation, resuspended in buffer A (20 mM HEPES-NaOH, pH 7.5; 400 mM NaCl; 5 mM β-mercaptoethanol), and lysed by sonication. The soluble fraction recovered by high speed centrifugation was mixed with 100 μl of Talon resin (Clontech). After 1 h of incubation, the supernatant was removed and the resin washed extensively with buffer A. Proteins were eluted from the column using buffer A containing 500 mM imidazole. Eluted proteins were analyzed by SDS-PAGE.

**Electrophoresis Mobility Shift Assays (EMSA).** Yeast U14 snoRNA was synthesized by *in vitro* transcription with T7 RNA polymerase of a PCR fragment carrying the *S. cerevisiae* snoRNA sequences in the conditions previously described

(15). After transcription, RNAs were 5'-end labelled with T4 polynucleotide kinase in the presence of [ $\gamma$ - $^{32}$ P]ATP and purified by gel electrophoresis. For EMSA, WT or variant Snu13p were mixed with His6Rsa1p<sub>230-266</sub> or His6Rsa1p<sub>230-381</sub> protein at a final concentration of 0.5  $\mu$ M, and incubated for 20 min at 4°C with 5 fmol of  $^{32}$ P-radiolabeled U14 RNA in buffer D (20 mM HEPES-KOH, pH 7.9; 150 mM KCl; 1.5 mM MgCl<sub>2</sub>; 0.2 mM EDTA; 10% glycerol). The RNA-protein complexes formed were resolved by native gel electrophoresis in conditions described previously (34).

**Circular dichroism spectroscopy.** Circular dichroism (CD) spectroscopy analyses were carried out using a Jobin-Yvon Mark VI circular dichrograph. Quartz spare split-compartment cuvettes with 0.44 cm path length per compartment were used for binding measurements. Use of these cuvettes allows accurate assessment of spectral differences between proteins in a bound and a free state or specific modifications in the RNA spectrum caused by its interaction with proteins. The relevant protein solution was placed in one compartment of the cuvette and the target protein solution in the other compartment. CD spectra were recorded before and after mixing the cuvette contents and blanks were run for each interaction and subtracted from the raw data. At least two spectra were collected before and after mixing, an average spectrum was established in order to increase the signal-to-noise ratio. Spectra were recorded from 200 to 260 nm to test protein-protein interaction with a scan speed of 0.2 nm.s<sup>-1</sup>. Proteins were adjusted at a 4  $\mu$ M final concentration in 10 mM Tris-HCl buffer, pH 7.8; 150 mM NaCl. The results are presented as normalized  $\Delta\epsilon$  values. For protein-protein interactions we used amino acid mean residue mass of 110 Da. Taking into account the residue concentration and the optical path-length of the cuvette, the precision of measurements was of  $\delta(\Delta\epsilon) = \pm 0,03 \text{ dm}^3 \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ .

#### **NMR Spectroscopy and structure calculation.**

For solubility reasons, a synthetic peptide of Rsa1p (Rsa1p<sub>238-252</sub>) was used for NMR analyses (GeneCust®). The NMR sample of Rsa1p<sub>238-252</sub> contained 3 mM of peptide in 10 mM Phosphate buffer, pH 6.4; 95% H<sub>2</sub>O; 5% D<sub>2</sub>O. NMR data were acquired at 20°C on a Bruker Avance III spectrometer equipped with a Z axis gradients TCI cryoprobe. For proton resonance assignment, TOCSY spectra using 2 mixing times (30 and 60 ms) and NOESY spectra using 3 mixing times (75,

150 and 300 ms) were recorded. Assignment of heteronuclei was achieved using  $^1\text{H}$ - $^{15}\text{N}$  HSQC,  $^1\text{H}$ - $^{13}\text{C}$  HSQC and  $^1\text{H}$ - $^{13}\text{C}$  HSQC-TOCSY (with a mixing time of 60 ms) experiments in natural abundance. All spectra were processed with Topspin 3.0 (Bruker BioSpin) and analyzed with CARA (35).

Distance restraints were derived from NOESY spectra and torsion angle restraints for the  $\phi$  angle were derived from measurements of the  $^3\text{J H}^{\text{N}}\text{-H}^{\alpha}$  coupling constant on 1D proton spectra. 200 structures of Rsa1p<sub>238-252</sub> were calculated with CYANA 2.1 (36). The 50 best structures were used for explicit water-refinement in CNS 1.2 (37). Finally, the 10 structures with the lowest energies were selected as the most representative structures of Rsa1p<sub>238-252</sub>.

**Molecular modeling.** The NMR structure of the Rsa1p<sub>238-252</sub> was docked on the established Snu13p 3D structure (PDB code 2ALE). This model was placed in a water box of 80Å<sup>3</sup> containing 16257 water molecules and one chlorine ion for ensuring the electrostatic neutrality of the system. All Arg, Lys, Asp, and Glu side chains were considered ionic. The resulting system, consisting in a total number of 51,121 atoms, was energy refined. The MD program NAMD (38) was employed in combination with the CHARMM27 force field (39) in order to describe the whole molecular system. Coulomb forces were evaluated with the particle-mesh Ewald method. Chemical bonds between hydrogen and heavy atoms were constrained to their equilibrium value by means of the SHAKE algorithm. The refinement procedure was based on several energy minimization steps carried out by the conjugate gradient method: first, the backbone of the protein complex was fixed such that only the protein side chains and water molecule movements were variable; next, the whole system was relaxed. Finally, a 5-ns MD simulation was carried out in the isobaric-isothermal ensemble, maintaining the pressure and the temperature at 1.0 atm and 300.0 K, respectively, by means of Langevin dynamics and the Langevin piston approach. The equations of motion were integrated with a 2-fs time step, using the r-RESPA algorithm to update short- and long-range contributions at different frequencies.

**Yeast two-hybrid assays.** For yeast two-hybrid assays (Y2H), appropriate pACTII and pAS2 plasmids were used to transform haploid cells (strain CG929 or Y190, and strain Y187, respectively), which were then crossed. Diploids were selected on media suitable for double selection

(Leu<sup>-</sup>, Trp<sup>-</sup>) and then plated on media suitable for triple selection (Leu<sup>-</sup>, Trp<sup>-</sup>, His<sup>-</sup>). 3-Amino-1,2,4-triazol (3AT) was used to evaluate the strength of the interactions. Growth was assessed after three days of incubation at 30°C.

**Tests of the functionality of variant Snu13p and Rsa1p proteins in yeast.** Standard *S. cerevisiae* growth and handling techniques were employed. The *S. cerevisiae* strain YPH499-GAL::*SNU13* (*Mata trp1-Δ63; his3-Δ200; ura3-52; lys2-801; ade2-101; leu2-Δ1; HIS3::GAL1::SNU13*) (2) was used to test whether Snu13p variants can complement depletion of the essential endogenous *SNU13* gene. To this end, strain YPH499-GAL::*SNU13* was transformed with recombinant pG1::*SNU13* plasmids encoding WT or mutated Snu13p. Cells were grown overnight at 30°C in a medium containing galactose as the carbon source (YPG). They were collected by centrifugation, washed, and diluted in YPD medium containing glucose as the carbon source. Cells were grown at 30°C and constantly maintained in exponential phase by dilution with pre-warmed YPD. The doubling time was monitored by measuring the absorbance at 600 nm (OD<sub>600</sub>).

The *S. cerevisiae* Y190-Δ*RSAl* (*YPL193W::kanMX2*) strain was obtained by homologous recombination. A DNA fragment carrying the kanamycin cassette flanked by sequences amplified from the *RSAl* gene was used to transform the haploid Y190 strain. Insertion of the kanamycin cassette at the *RSAl* locus was then verified by PCR amplifications. The *RSAl* knock-out strain obtained was transformed with plasmid pG1::*RSAl* encoding WT or variant Rsa1p, and grown in YPD medium.

**Evaluation of snoRNA levels in yeast.** Total RNAs were extracted from exponentially growing yeast cultures. Northern-blot analyses were performed in standard conditions using 5 μg of total RNAs and specific 5'-end <sup>32</sup>P-radiolabeled antisense probes : OG-U1 (5'-GACCAAGGAGTTTGCATCAATGA-3'); OG-snr63 (5'-TTATGTTGGCCACTCATCAC-3'); OG-snr190 (5'-CGAGGAAAGAAGAGACACCATTATC-3'); OG-snr45 (5'-CCTCAGATCGCTCCGAGA-3'); U14 (5'-TCACTCAGACATCCTAG-3'); OG-snr13 (5'-AGCTTGAGTTTTTCCACACC-3'); OG-U18 (5'-TGTGATAGTCAACCACATCTC-3').

**His-pull down assays of in vitro assembled RNPs.** 100 fmol of 5'-end <sup>32</sup>P-radiolabeled U14 RNA prepared according to (15) were incubated

with various sets of recombinant proteins during 30 min in buffer D at 30°C, 1 μM of protein Snu13p and of the His<sub>6</sub>Tah1p-Pih1p complex were used, while the Rsa1p<sub>230-381</sub> fragment was used at a 0.25, 0.5 μM, 1 μM, or 2 μM concentration. The RNP complexes formed were incubated with Nickel-sepharose beads (GE Healthcare©) for 15 minutes at room temperature in 200 μL of binding buffer (10 mM Tris-HCl, pH 8.0; 150 mM NaCl; 0.1% IGEPAL). The beads were washed three times for 10 min with 1 mL of binding buffer. RNAs were extracted by phenol-chloroform treatments, ethanol precipitated, and fractionated on a 7% denaturing polyacrylamide gel.

**Subcellular localization assay of WT and mutated 15.5K proteins in HeLa cells.** The co-recruitment assay was done as previously described (33).

## RESULTS

Although NUFIP and Rsa1p are expected to have similar functions in snoRNP assembly, sequence alignment of these two proteins revealed a low level of sequence similarity except for a 37 amino acids stretch, termed the PEP domain (Figure 1A). According to our previous data, this common peptide is sufficient for 15.5K and Snu13p binding (24). We therefore decided to identify the 15.5K and Snu13p amino acid sequences which interact with PEP. According to the high degree of similarity of 15.5K and Snu13p (Figure 1B), we expected that they develop similar interactions with the human (NUFIP<sub>223-260</sub>) and yeast (Rsa1p<sub>230-266</sub>) PEP domains, respectively.

### 15.5K-NUFIP association depends upon the 15.5K helix α3

To screen for the NUFIP interacting domain, surface residues in 15.5K were mutagenized into alanines. A series of ten 15.5K mutants was generated, and the resulting proteins were tested for association with NUFIP and NUFIP<sub>223-260</sub> by yeast two-hybrid (Y2H) assays. Although, the interactions with NUFIP<sub>223-260</sub> were generally weaker than with the full-length protein, the mutations in protein 15.5K induced similar variations of affinity for the entire NUFIP protein and the NUFIP<sub>223-260</sub> domain (Figure 1C). Nine of the tested variants still interacted with NUFIP and NUFIP<sub>223-260</sub>, while one had lost these interactions: variant D (E<sub>74</sub>D<sub>75</sub>K<sub>76</sub>/AAA), which affects the C-terminal part of helix α3 of 15.5K (Figure 1B). Next, we refined the identification of the interaction

site by generating a new set of substitutions. We first tested each individual position of the EDK sequence, and the E<sub>74</sub>/A mutation turned-out to be sufficient to abolish the Y2H interaction with NUFIP and NUFIP<sub>223-260</sub>, while the D<sub>75</sub>/A and K<sub>76</sub>/A mutations had no adverse effect (Figure 1C). Interestingly, E<sub>74</sub> is conserved in Snu13p, NHP2/Nhp2p and SBP2, but not in other proteins containing a K-turn recognition domain, like protein L30 (Figure 1B). We then tested residues in the other part of the 15.5K helix  $\alpha$ 3. Residues E<sub>64</sub>, I<sub>65</sub>, H<sub>68</sub>, and L<sub>71</sub> were of particular interest, since in the U4 snRNP, E<sub>64</sub>, I<sub>65</sub>, and H<sub>68</sub> were found to make contacts with the PRP31 Nop domain (20), and L<sub>71</sub> is spatially close to E<sub>74</sub> but not conserved in NHP2, where it is replaced by a valine (Figure 1B). Variants E<sub>64</sub>/A, I<sub>65</sub>/A, H<sub>68</sub>/A, and L<sub>71</sub>/A still interacted with NUFIP, although the interaction was weaker for the L<sub>71</sub>/A variant (Figure 1C and data not shown). When tested with NUFIP<sub>223-260</sub>, the E<sub>64</sub>/A and I<sub>65</sub>/A 15.5K variants interacted with this small motif, while H<sub>68</sub>/A and L<sub>71</sub>/A had lost the interaction (Figure 1C), reflecting the generally weaker interactions obtained with NUFIP<sub>223-260</sub> compared to NUFIP. Altogether, these data indicate that residue E<sub>74</sub> of helix  $\alpha$ 3 of 15.5K is essential to bind NUFIP, and that residue L<sub>71</sub> modulates this interaction.

#### **The 15.5K E<sub>74</sub>D<sub>75</sub>K<sub>76</sub>/AAA mutant has an altered intra-cellular localization and does not bind NUFIP *in vivo***

Having shown the functional importance of the Snu13p E<sub>72</sub>D<sub>73</sub>K<sub>74</sub> motif in yeast, we then developed a spatial recruitment assay in HeLa cells to test whether the E<sub>74</sub>D<sub>75</sub>K<sub>76</sub> motif in helix  $\alpha$ 3 of 15.5K is also involved in NUFIP binding *in vivo* (Figure 2). The approach developed circumvents artifacts often occurring during co-immunoprecipitation assays, in particular, the loss of weak interaction or on the contrary the re-association of proteins in the test tube. Our assay was based on the LacI/LacO system (12,40). A HeLa cell line containing a tandem array of LacO sites integrated in its genome was transfected with a mRFP1-LacI-NUFIP expression vector. In the U2OS parental cell line devoid of LacO sites, the mRFP1-LacI-NUFIP had a diffuse localization in the nucleoplasm, similar to wild-type NUFIP protein. In contrast, in the cell line containing the LacO sites, it accumulated in a bright spot in the nucleoplasm, which corresponded to the LacO repeats. When GFP-15.5K was co-expressed with mRFP1-LacI-NUFIP in U2OS cells, it localized in the nucleoli, Cajal bodies and nucleoplasm as

previously described, and no particular co-localization was observed with mRFP1-LacI-NUFIP. In contrast, in the cell line containing the LacO array, GFP-15.5K co-localized with the bright spots formed by the binding of mRFP1-LacI-NUFIP to the LacO sites. This indicated that, as expected, NUFIP interacted with 15.5K *in vivo*. We then tested the 15.5K E<sub>74</sub>D<sub>75</sub>K<sub>76</sub>/AAA variant in the same assay. In parental U2OS cells, the localization of the mutant protein differed from its wild-type counterpart: it accumulated to lower levels in nucleoli and it had a more diffuse localization throughout the nucleoplasm and the cytoplasm, consistent with a defect in its incorporation into U4 and box C/D RNPs. Moreover, in cells containing mRFP1-LacI-NUFIP bound to the LacO array, GFP-15.5K E<sub>74</sub>D<sub>75</sub>K<sub>76</sub>/AAA was unable to accumulate at the LacO sites, indicating that it was defective for NUFIP binding *in vivo*. Therefore, the data demonstrated the importance of the extremity of helix  $\alpha$ 3 of 15.5K for *in cellulo* binding of protein 15.5K to NUFIP. Then, for further characterization of the 15.5K/Snu13p-NUFIP/Rsa1p interaction we turned to yeast because of the possibilities offered by yeast genetics.

#### **The Snu13p E<sub>72</sub>D<sub>73</sub>K<sub>74</sub> sequence is required for direct interaction of Snu13p with Rsa1p**

We hypothesized that the E<sub>72</sub>D<sub>73</sub>K<sub>74</sub> motif in the Snu13p helix  $\alpha$ 3, which corresponds to the E<sub>74</sub>D<sub>75</sub>K<sub>76</sub> sequence in 15.5K, may be important for Rsa1p binding to Snu13p (Figures 1 and 3A) and tested this hypothesis by yeast two-hybrid (Y2H) assays. Snu13p interacted with Rsa1p and its functional C-terminal half (Rsa1p<sub>230-381</sub>) (24), while variants Snu13p with the entire E<sub>72</sub>D<sub>73</sub>K<sub>74</sub> sequence or only E<sub>72</sub> converted into alanine did not (Figure 3B). As interactions of yeast proteins detected by Y2H assays may be either direct or mediated by an endogenous yeast protein, we performed complementary experiments, to test whether the EDK motif in Snu13p and the Rsa1p<sub>230-266</sub> domain of Rsa1p establish direct contacts.

First, we co-expressed His-tagged WT or mutated Snu13p together with Rsa1p full-length or Rsa1p<sub>230-381</sub> in *E. coli* and investigated whether the two proteins could be co-purified by Immobilized Metal Ion Affinity Chromatography (IMAC), (Figure 3C). Fractionation of the IMAC purified proteins by SDS-PAGE revealed the formation of a His<sub>6</sub>-Snu13p-Rsa1p complex with an apparent 1:1 molar stoichiometry (Figure 3C lane 2), which was a strong indication for a direct interaction of the two proteins. Snu13p also co-purified with the C-



terminal domain of Rsa1p that contains the Rsa1p<sub>230-266</sub> domain (Rsa1p<sub>230-381</sub> fragment), but not with the Rsa1p N-terminal domain (data not shown). No co-purification of Rsa1p was detected with the His-tagged E<sub>72</sub>D<sub>73</sub>K<sub>74</sub>/AAA Snu13p variant (Figure 3C, lane 4). These data strongly suggested that the EDK motif is needed for formation of a direct interaction between Snu13p and Rsa1p. Accordingly, when GST-Snu13p or its E<sub>72</sub>D<sub>73</sub>K<sub>74</sub>/AAA variant were immobilized on glutathion-sepharose beads and incubated with the recombinant His<sub>6</sub>Rsa1p<sub>230-266</sub>, GST-Snu13p retained significant amount of His<sub>6</sub>Rsa1p<sub>230-266</sub> (Supplementary data Figure S1, lane 6), while GST alone, or the E<sub>72</sub>D<sub>73</sub>K<sub>74</sub>/AAA Snu13p variant did not (Supplementary data Figure S1, lane 8). We observed similar effects of the Snu13p mutations when the Snu13p-Rsa1p<sub>230-266</sub> or Snu13p-Rsa1p<sub>230-381</sub> interactions were studied in the presence of the RNA. As previously observed (24), the Rsa1p<sub>230-381</sub> and Rsa1p<sub>230-266</sub> super-shifted the complex formed by Snu13p and an *in vitro* transcribed U14 C/D box snoRNA (Figure 3D, lanes 3 and 4). However, they did not shift the stable complexes formed by association of the Snu13p E<sub>72</sub>D<sub>73</sub>K<sub>74</sub>/AAA or E<sub>72</sub>/A variants with U14 snoRNA (Figure 3D, lanes 7-8, 11-12). We concluded that the Snu13p E<sub>72</sub>D<sub>73</sub>K<sub>74</sub> sequence and especially residue E<sub>72</sub> in this sequence is required for formation of a direct interaction between Snu13p and Rsa1p.

### **The EDK motif of Snu13p interacts directly with the Rsa1p<sub>230-266</sub> motif of Rsa1p**

Then we used circular dichroism (CD) experiments to determine whether the E<sub>72</sub>D<sub>73</sub>K<sub>74</sub> motif directly contacts the Rsa1p<sub>230-266</sub> domain, or whether it has an indirect effect, for instance by maintaining Snu13p in a proper conformation. Far UV CD spectroscopy provides information about protein secondary structure. Therefore, it can reveal conformational changes occurring upon generation of mutations in a protein or upon its interaction with a partner. We first recorded the CD spectra of Snu13p WT and Snu13p E<sub>72</sub>D<sub>73</sub>K<sub>74</sub>/AAA from 205 to 260 nm. They were very similar and both showed negative peaks at 208 and 222 nm, indicating the presence of  $\alpha$ -helices as the major secondary structure element (data not shown). Therefore the three substitutions in Snu13p E<sub>72</sub>D<sub>73</sub>K<sub>74</sub>/AAA did not have a major effect on the protein structure.

Direct protein-protein interaction can be studied by CD spectroscopy through the use of a cuvette with two compartments. We used this approach to test for the effect of mutation in the

EDK motif on the Snu13p-Rsa1p<sub>230-266</sub> interaction. One compartment was filled with wild-type Snu13p and the other one with Rsa1p<sub>230-266</sub> using a protein to peptide molar ratio of 1 and we compared the CD spectra recorded before and after mixing the material contained in the two compartments of the cuvette. We detected a significant variation of the signal intensity at about 222 nm in the spectrum collected after mixing Snu13p with the Rsa1p<sub>230-266</sub>, indicating the formation of a direct interaction and a modification of the structure of one or both partners upon complex formation (Figure 4A). In contrast, when the Snu13p E<sub>72</sub>D<sub>73</sub>K<sub>74</sub>/AAA variant was used in the experiment, the CD spectra recorded before and after mixing the two proteins were identical (Figure 4A).

Therefore, altogether, the data of these CD experiments strongly reinforced the idea that the E<sub>72</sub>D<sub>73</sub>K<sub>74</sub> motif at the surface of Snu13p interacts directly with the Rsa1p<sub>230-266</sub> motif of Rsa1p.

### **Rsa1p residues W<sub>245</sub> and R<sub>249</sub> are required for Snu13p binding**

Then, we tried to identify residues in the Rsa1p<sub>230-266</sub> fragment which are important for interaction with Snu13p. Alignment of the human NUFIP<sub>223-260</sub> and yeast Rsa1p<sub>230-266</sub> sequences revealed the presence of a strictly conserved five amino acids stretch (WREER, positions 245 to 249 in Rsa1p) (Supplementary data Figure S2). To test whether this conserved amino acid sequence was involved in the interaction with Snu13p, we performed individual substitutions of some of these residues by alanine in both the Rsa1p<sub>230-266</sub> and Rsa1p<sub>230-381</sub> proteins (Figure 5A) and investigated the effects of these mutations on Rsa1p<sub>230-266</sub> and Rsa1p<sub>230-381</sub> binding to Snu13p by using different approaches. First, we used the specific modification of the CD signal observed upon Rsa1p<sub>230-266</sub> binding to Snu13p as a reporter signal of complex formation (Figure 4A). While similar decreases in the CD signal at 222 nm were recorded for complex formation when using Rsa1p<sub>230-266</sub> WT or Rsa1p<sub>230-266</sub> E<sub>247</sub>A, no significant modification was detected when using the Rsa1p<sub>230-266</sub> W<sub>245</sub>A, R<sub>246</sub>A, and R<sub>249</sub>A variants (Figure 4B). These data strongly suggested the involvement of residues W<sub>245</sub>, R<sub>246</sub>, and R<sub>249</sub> in the Snu13p- Rsa1p<sub>230-266</sub> interaction. Accordingly, when we tested the ability of Rsa1p<sub>230-266</sub> variants to form trimeric complexes with Snu13p and the U14 RNA by gel-shift assays, only Rsa1p<sub>230-266</sub> E<sub>247</sub>A could super-shift the Snu13p-U14 RNA complex (Figure 5B). Interestingly, when we introduced the amino acid substitutions in the WREER motif of Rsa1p<sub>230-381</sub>

and tested the ability of the variant proteins to interact with Snu13p by Y2H assays, only the W<sub>245</sub>A or R<sub>249</sub>A substitutions impaired the Y2H interaction (Figure 5C). The R<sub>246</sub>A substitution had no visible effects. Similarly, when we co-expressed the WT or variant Rsa1p or Rsa1p<sub>230-381</sub> proteins with His<sub>6</sub>Snu13p in *E. coli* and co-purified His<sub>6</sub>Snu13p with its associated protein by IMAC, the R<sub>249</sub>A substitution in both Rsa1p and Rsa1p<sub>230-381</sub> was found to abolish the interaction with Snu13p, whereas variant R<sub>246</sub>A retained some binding activity (Figure 5D).

The lower effect of the R<sub>246</sub>A substitution in our Y2H and co-expression assays compared to super-shift assays may be explained by the fact that the former assays were made with Rsa1p<sub>230-381</sub> while the latter assays were performed with Rsa1p<sub>230-266</sub>. The Snu13p- Rsa1p<sub>230-266</sub> interaction may be less stable than the Snu13p-Rsa1p<sub>230-381</sub> interaction (Figure 1) and therefore more strongly destabilized by the R<sub>246</sub>A amino acid substitution. Altogether, the data identified two residues in Rsa1p (W<sub>245</sub> and R<sub>249</sub>) that play a key role in the interaction with Snu13p.

### **Rsa1p<sub>238-252</sub> folds into an alpha helix**

Having obtained strong data in favour of a direct contact between Snu13p and the Rsa1p<sub>230-266</sub> domain and having identified residues in these motifs that are essential for the interaction, we decided to build a 3D structure model of this interaction. First, we attempted to establish the 3D structure of the entire Rsa1p<sub>230-266</sub> fragment but we encountered strong solubility problems at the concentration used for NMR studies. Therefore, we determined the 3D structure of a peptide of the Rsa1p<sub>230-266</sub> domain, T<sub>238</sub>DEDVKKWREERKKM<sub>252</sub> (named Rsa1p<sub>238-252</sub>) which contained the three residues W<sub>245</sub>, R<sub>246</sub> and R<sub>249</sub> important for interaction with Snu13p and which solubility was suitable to obtain well-resolved <sup>1</sup>H-NMR spectra. Thus, excluding carbonyl, carboxyl and guanidium groups as well as tertiary amine of lysine side chains, only the NMR frequencies of the backbone amide groups of T<sub>238</sub> and D<sub>239</sub> and of C<sup>δ</sup> and C<sup>ε</sup> of K<sub>244</sub> remained undetermined, meaning that the rest of the <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N nuclei were unambiguously assigned. A schematic view of the NMR distances restraints and structural statistics for the calculation are displayed in figure S3 and table S1 (supplementary data). As shown by the superimposition of the 10 most representative structures (Figure 6A) and the low values of RMSD, Rsa1p<sub>238-252</sub> is folded into an  $\alpha$ -helix. Interestingly, the side chains of residues W<sub>245</sub>,

R<sub>246</sub> and R<sub>249</sub> are positioned on the same face of the helix, in a configuration favourable for the establishment of interaction with Snu13p amino acids (Figure 6B). Conversely, the side chain of the residue E<sub>247</sub>, which is not required for Snu13p binding (Figure 4B), is positioned on the opposite side of the helix (Figure 6B).

### **Modeling the 3D structure of the Rsa1p<sub>238-252</sub>-Snu13p interaction**

By using the established crystal structure of Snu13p (PDB accession number: 2ALE, (41)), the solution structure of Rsa1p<sub>238-252</sub> and our knowledge on the amino-acids important for the Snu13p-Rsa1p interaction, we build a structural model of the interaction. In this docking experiment, according to our mutagenesis data, we pre-positioned the peptide at the extremity of the Snu13p helix  $\alpha$ 3 containing the EDK motif. The docking procedure included the minimization of electrostatic and Van der Waals energies of interaction. Then, in order to improve the 3D model obtained, it was placed it in water and a 5 ns Molecular Dynamic (MD) simulation was carried out as described in Material and Methods to obtain the 3D structural model presented in figure 7.

In this model, the Rsa1p<sub>238-252</sub>  $\alpha$ -helix and the Snu13p helix  $\alpha$ 3 interact in a parallel orientation, so that a coiled-coil structure is formed. Our model predicts Van der Waals contacts between the aromatic side chain of residue W<sub>245</sub> in Rsa1p<sub>238-252</sub> and the hydrophobic side chain residue L<sub>69</sub> in Snu13p. The aromatic ring of residue W<sub>245</sub> stacks on the side chain of residue R<sub>249</sub> from Rsa1p<sub>238-252</sub>, which orients its guanidinium groups in order that they form two electrostatic bonds with the carboxylate groups of residues E<sub>72</sub> and D<sub>73</sub> from Snu13p. Finally, the side chain of residue E<sub>72</sub> in Snu13p forms a salt bridge with one guanidinium group of R<sub>246</sub> in Rsa1p<sub>238-252</sub>. This ion-pair network in which residues R<sub>249</sub> and E<sub>72</sub> play a central role likely explains why these two residues are required for establishment of the Snu13p-Rsa1p interaction. These ionic interactions might be protected from water by the Van der Waals interaction between residues W<sub>245</sub> in Rsa1p<sub>238-252</sub> and L<sub>69</sub> in Snu13p, explaining the importance of the residue W<sub>245</sub>.

As the model predicted a stabilization effect of residue L<sub>69</sub> in Snu13p, we tested the effects of its substitution by alanine on the Snu13p-Rsa1p<sub>230-266</sub> interaction by gel shift-assay. As illustrated in figure 8, the L<sub>69</sub>A mutation abolished the supershift of Rsa1p<sub>230-266</sub> with U14 RNA, which demonstrates experimentally the stabilization effect of residue L<sub>69</sub>.

### **The interaction between Rsa1p and Snu13p is important for cell growth**

Deletion of *RSAl* gene in *S. cerevisiae* leads to a growth defect, especially at high temperature (42). Hence, we decided to test whether abolition of the Snu13p-Rsa1p interaction is responsible for this growth phenotype. To this end, we used an engineered yeast strain containing a single *SNU13* ORF placed under the control of the conditional *GAL1* promoter. These cells were transformed with a centromeric pG1 plasmid allowing constitutive expression of the WT or mutant Snu13p protein. Cell growth was monitored in liquid media (Figure 9A). In the presence of glucose in the medium (absence of expression of the genomic *SNU13* gene), a doubling time of 3.26 h was observed for the ectopic expression of WT Snu13p from the pG1 plasmid. This doubling time was increased to 3.55 h and 3.56 h, respectively, when the E<sub>72</sub>D<sub>73</sub>K<sub>74</sub>/AAA and E<sub>72</sub>A Snu13p variants were expressed. These data strongly supported the idea of a functional importance of the identified Snu13p-Rsa1p interaction. To complete the demonstration, WT and variant Rsa1p were expressed from a centromeric plasmid in an *RSAl* KO strain, and growth rates were measured (Figure 9B). The KO strain had a doubling time ~44% slower compared to WT cells. Ectopic expression of Rsa1p restored the wild-type growth rates. In contrast, only a partial restoration of cell growth was obtained for the Rsa1p R<sub>249</sub>A and W<sub>245</sub>A variants that did not bind to Snu13p (doubling time ~25% slower than that of WT cell). The growth reduction observed for this Rsa1p variant was similar to those obtained for the E<sub>72</sub>D<sub>73</sub>K<sub>74</sub>/AAA and E<sub>72</sub>A Snu13p variants. Interestingly, the R<sub>246</sub>A substitution in Rsa1p which did not abolish binding to Snu13p as well as the L<sub>69</sub>A mutation in Snu13p, restored growth to wild-type levels.

### **The interaction between Rsa1p and Snu13p is important for the stability of C/D box snoRNAs**

Then, we tested whether the amounts of the C/D box snoRNAs could be affected in a strain impaired for the Snu13p-Rsa1p interaction. By using northern blot analysis, we measured the amounts of several snoRNAs in a WT strain and the *RSAl* KO strain complemented or not by the ectopic expression of Rsa1p WT or R<sub>249</sub>A (Figure 9C). The results confirmed that the *RSAl* gene disruption is associated to a drastic decrease of the amounts of the C/D box snoRNAs snR63, snR190, snR45, snR13 and U18. Ectopic expression of Rsa1p WT restored normal snoRNAs amounts, while the expression of the variant R<sub>249</sub>A Rsa1p did not. Only

the level of U14 snoRNA seems unaffected by the mutation. This result may be a consequence of the essential feature of this RNA which is involved in the nucleolytic maturation step of the 18S rRNA and which could be more stable than others snoRNAs. Furthermore, this observation could explain the fact that Rsa1p mutation does not completely reproduce the growth defect observed with the *RSAl* KO strain. Altogether, these data support the functional importance of the identified Snu13p-Rsa1p interaction for C/D box snoRNP biogenesis and normal cell growth in yeast.

### **Rsa1p recruits the Pih1-Tah1p dimer on a snoRNA through its interaction with Snu13p**

Previously, we showed that Rsa1p<sub>230-381</sub> interacts with Pih1p, an integral component of the R2TP complex, which was found to be required for snoRNP assembly (24). We tested *in vitro* whether recruitment of Pih1p on the RNA-protein complex analyzed in Figure 3D, i.e. U14-Snu13p-Rsa1p, was dependent on the interaction between Rsa1p and Snu13p. As the Pih1 protein presents a strong tendency to aggregate in bacteria, it was co-expressed with the His-tagged Tah1 protein, allowing purification of a stable His<sub>6</sub>Tah1p-Pih1p complex (43). Association of His<sub>6</sub>Tah1p-Pih1p with the RNA-protein complex containing Rsa1p was tested by His-tag pull-down assays (Figure 10). After formation of the RNA-protein complex, the His<sub>6</sub>Tah1p-Pih1p heterodimer was added, and complexes were purified by Immobilized Metal Ion Affinity. Retention of radiolabeled RNAs was analysed by autoradiography after electrophoresis on a denaturing gel. As shown in Figure 10A, a very little amount of RNA was co-purified in the absence or the presence of Snu13p, suggesting a weak capacity of the His<sub>6</sub>Tah1p-Pih1p to form a stable complex with U14 snoRNA or the Snu13p-U14 complex. In contrast, in the presence of the Rsa1p<sub>230-381</sub> C-terminal domain of Rsa1p, the retention of the snoRNA was substantially improved, indicating that Rsa1p can mediate the recruitment of the Tah1p-Pih1p complex on the Snu13p-U14 complex. To evaluate the functional importance of the Rsa1p-Snu13p interaction for recruitment of the His<sub>6</sub>Tah1p-Pih1p complex, we repeated the experiment with the E72/A Snu13p variant. As shown in panel B, the absence of the Snu13p-Rsa1p<sub>230-381</sub> interaction abolished the U14 retention on the beads coated with the His<sub>6</sub>Tah1p-Pih1p complex. Thus, through its capacity to interact both with Snu13p and with Pih1p, Rsa1p can facilitate the recruitment of the R2TP complex on pre-snoRNPs.

## DISCUSSION

Several proteins have been proposed to participate in snoRNP biogenesis. However, the mechanism of action of only a few of them has been clearly established. NUFIP and its yeast homolog Rsa1p were identified as key factors involved in the biogenesis of RNPs containing L7Ae-like proteins. In their absence, a marked decrease of the C/D box snoRNPs level is observed (24). NUFIP and Rsa1p were proposed to interact with proteins 15.5K and Snu13p, respectively, to recruit HSP90 and its R2TP co-chaperone on nascent RNPs (24), and to act as scaffolding proteins (24,25). In this study, by NMR and 3D structure modeling, and site-directed mutagenesis, we define at the atomic level how Snu13p interacts with Rsa1p, and demonstrate the functional importance of the implicated amino acids. The established 3D structural model brings insights on how NUFIP and Rsa1p can discriminate L7Ae members within the L30 protein super-family, and supports the idea that NUFIP/Rsa1p function as external scaffolds during RNP assembly.

### **Both electrostatic and Van der Waals interactions are required for stable 15.5K-NUFIP and Snu13p-Rsa1p interactions**

Here, by producing a large series of Snu13p and 15.5K protein variants that were mutated in helix  $\alpha_3$ , we identified the amino acids in this helix which are important for interaction with proteins Rsa1p and NUFIP. A large substitution in the human 15.5K helix  $\alpha_3$  had also been previously produced (the entire EDKN motif was substituted for a RKMT motif). This mutation had been found to abolish the binding of the recombinant protein PRP31 to the U4-15.5K complex as well as of snoRNP proteins and TIP48, TIP49 proteins on snoRNA-15.5K complexes (44). However, the defect of activity of this protein may be due to its incorrect folding. Indeed, in contrast to the WT 15.5K protein this variant protein was unable to interact with a recombinant PRP31 protein (44). This is an intriguing observation because according to the 3D structure established for a U4-15.5K-PRP31<sub>78-333</sub> complex, the 15.5K region involved in the interaction with protein PRP31 does not contain the EDKN motif which is mutated in the variant protein (20) (Figure 11). As the interaction of PRP31<sub>78-333</sub> with the U4-15.5K complex is highly stable, the N-terminal domain absent in the recombinant PRP31<sub>78-333</sub> protein used for X-ray crystallography is not expected to have a major contribution to the authentic 15.5K-PRP31

interaction. In addition, the amino acid sequence of this N-terminal region is not conserved in yeast Prp31p. Therefore, in the absence of clear information on the structure of the previous 15.5K variant mutated in the EDKN motif, it is difficult to take into consideration the data obtained with this variant protein in the present discussion. While the EDKN motif was substituted by a RKMT motif in the variant previously produced, here, in both 15.5K and Snu13p, the EDK motif was substituted by three alanines. Such a substitution was not expected to disturb the protein 3D structure, first, because according to the 15.5K and Snu13p 3D structures (16,41), the EDK residues are not interacting with other residues in the protein and secondly, because the side chain of alanine residues cannot establish stable interactions with other residues in the protein. Therefore, the integrity of helix  $\alpha_3$  and the overall 3D structure of 15.5K and Snu13p were expected to be conserved after mutation, and we indeed verified by circular dichroism that the generated mutations had no dramatic effect on the Snu13p protein structure. Furthermore, we refined the study by production of variant proteins containing punctual amino acid substitutions and all data obtained, as well by *in vitro* as by *in cellulo* assays, or by computer modeling are in good agreement one with the other and they support our model. Indeed, through 3D structure modeling, site-directed mutagenesis, and by performing various *in vitro* and *in cellulo* experiments, we provide compelling evidences for a functional direct interaction of the E<sub>72</sub> residue in Snu13p with residue R<sub>249</sub> in Rsa1p and of the corresponding interaction in human cells (E<sub>74</sub> residue in 15.5K with residue R<sub>243</sub> in NUFIP). We also define the respective roles of the surrounding residues.

Altogether, our experimental data on both the human and yeast systems, demonstrate that in the EDK motif of proteins 15.5K and Snu13p, only the E residue plays a crucial role for interaction with NUFIP or Rsa1p. Individual mutations of the D and K residues had no detectable effects on binding to NUFIP. According to our 3D structure model the major importance of residues E<sub>72</sub> in Snu13p and R<sub>249</sub> in Rsa1p may be explained by the fact that both of them establish ionic interactions with two amino acids in the partner protein, which is in contrast to residues D<sub>73</sub> in Snu13p and R<sub>246</sub> in Rsa1p that only establish one ionic interaction with amino acids in the partner protein (Figure 7). According to our model, residue W<sub>245</sub> favors the formation of the electrostatic interaction between residues E<sub>72</sub> in Snu13p and R<sub>249</sub> in Rsa1p by

stacking onto residue R<sub>249</sub> and interacting by Van der Waals contact with the side chain of residue L<sub>69</sub> in Snu13p. The fact that residue W<sub>245</sub> in Rsa1p, which is important for cell growth, stabilizes the ionic interaction formed by the carboxylate group of E<sub>72</sub> and the guanidium group of R<sub>249</sub> again strengthens the importance of this electrostatic interaction.

Interestingly, in our model, residue L<sub>69</sub> in Snu13p also develops hydrophobic interaction with residue R<sub>246</sub> in Rsa1p. The absence of effect on cell growth of the L<sub>69</sub>A mutation in Snu13p suggests that this hydrophobic interaction is not essential to the RNP assembly mechanism in the growth conditions that we used. However, several of our observations strongly suggest that the presence of residues L<sub>69</sub> in Snu13p and its counterpart L<sub>71</sub> in 15.5K participates to the stability of the interaction: i) the presence of these residues increases the strength of the interactions established with Rsa1p<sub>230-381</sub> and NUFIP, respectively, and ii) these residues are respectively required to get stable Snu13p-Rsa1p<sub>230-266</sub> and 15.5K-NUFIP<sub>223-260</sub> binding (Figure 1C and Figure 8). Therefore, we propose that the Snu13p-Rsa1p and 15.5K-NUFIP interactions consist in ionic interactions caged by hydrophobic interactions.

### **NUFIP and Rsa1p discriminate L7Ae members through helix $\alpha$ 3**

Our data reveal that both the Snu13p-Rsa1p and 15.5K-NUFIP interactions depend upon residues located at one extremity of helix  $\alpha$ 3. Previous data also implicated the 15.5K E<sub>124</sub>R<sub>125</sub> motif in NUFIP recruitment, because the E<sub>124</sub>R<sub>125</sub>/SS mutation was found to abolish the recruitment of NUFIP from a cell extract by a GST-15.5K mutant retained on glutathione-sepharose beads (44,45). We show here that the E<sub>124</sub>R<sub>125</sub>/SS mutant has the capability to interact efficiently with NUFIP and the small NUFIP<sub>223-260</sub> domain in yeast two-hybrid assays (Mut I in Figure 1C). Furthermore, these residues are not conserved in Snu13p, which also interacts with Rsa1p through the Rsa1p<sub>230-266</sub> domain. Thus, the effect previously observed for these two amino acid substitutions was probably not linked to the disruption of the direct interaction of 15.5K with NUFIP. It may be due to an alteration of the structure of the recombinant protein produced *in vitro* or to an indirect effect in nuclear extract. Our data strongly suggest that residue E<sub>74</sub> in the 15.5K helix  $\alpha$ 3 is the key player in NUFIP recruitment during C/D box snoRNPs assembly. We previously showed that NUFIP interacts with NHP2 and SBP2 (24). It is interesting to note that among the known

members of the L30 family, which includes the L7Ae family (46), residue E<sub>74</sub> is only conserved in proteins 15.5K, NHP2 and SBP2 (Figure 1B). Furthermore, data of our co-expression and yeast two-hybrid assays failed to detect an interaction between NUFIP and other members of the L30 family that contain no counterparts of residue E<sub>74</sub> ((24) and data not shown). This suggests that the presence of E<sub>74</sub> is a good predictor of the interaction with NUFIP or Rsa1p.

Interestingly, amino acid sequence alignments show the replacement of the 15.5K L<sub>71</sub> residue by a valine (V<sub>96</sub>) in human NHP2. As discussed above, residue L<sub>69</sub> in Snu13p and its counterpart L<sub>71</sub> in 15.5K are expected to stabilize the interactions with Rsa1p and NUFIP by formation of Van der Waals interactions, suggesting that the presence of residue V<sub>96</sub> in NHP2 may explain the difference of affinity of NUFIP<sub>223-260</sub> for 15.5K and NHP2. Indeed, NHP2 does not interact with NUFIP<sub>223-260</sub> in yeast two-hybrid assays, while the V<sub>96</sub>L substitution restores the interaction (Figure 1C). Conversely, the mutation L<sub>71</sub>V in 15.5K decreases the interaction with NUFIP<sub>223-260</sub>. Altogether, these data support the idea of an important role of the Van der Waals interactions established by residue L<sub>71</sub> in the discrimination of Snu13p and 15.5K from other L7Ae-like proteins. Accordingly, residue L<sub>71</sub> is also not conserved in SBP2, and SBP2 does not bind NUFIP<sub>223-260</sub>. This suggests the involvement of other amino acids of SBP2 and NHP2 in binding to NUFIP (Figure 1). In agreement with this hypothesis, the isolated NUFIP<sub>223-260</sub> fails to interact with both proteins.

Altogether, this suggests that the ancestral role of NUFIP was likely restricted to the biogenesis of the U4 snRNP and C/D snoRNPs, and that its role in H/ACA snoRNPs and SBP2 mRNPs has been acquired later.

### **A model of immature U4 RNP**

In the 3D structure established by NMR, Rsa1p<sub>238-252</sub> adopts a  $\alpha$ -helical structure, and in the model that we built, interacts through a parallel coiled-coil interaction involving helix  $\alpha$ 3 of Snu13p. In this model, in agreement with our experimental data showing the absence of interaction of Rsa1p with the RNA, NUFIP and Rsa1p bind the 15.5K and Snu13p proteins at the side opposite to their respective RNA-binding domains. This is consistent with the possibility to form a ternary complex containing the RNA, 15.5K/Snu13p and NUFIP/Rsa1p. To get information on the possibility to form such a ternary complex, we took advantage of the available 3D structure of a fragment of U4

snRNP, which contains the 5'-terminal stem-loop of U4 snRNA and proteins 15.5K and PRP31<sub>78-333</sub> (PDB accession number: 2OZB). Based on our 3D model of the Snu13p-Rsa1p<sub>238-252</sub> interaction, we built a model of the 15.5K-NUFIP<sub>223-260</sub> interaction. Then, we incorporated this modeled interaction in the crystal structure established for the U4-15.5K-PRP31<sub>78-333</sub> complex (Figure 11). Interestingly, according to this model, the NUFIP<sub>223-260</sub> domain is predicted to interact with 15.5K without disturbing the 15.5K-PRP31<sub>78-333</sub> interaction. Indeed, although adjacent, the NUFIP<sub>223-260</sub> interacting site on 15.5K is not overlapping the PRP31<sub>78-333</sub> interacting site. Accordingly, we found that individual mutations of the 15.5K residues (E<sub>64</sub>, I<sub>65</sub>, and H<sub>68</sub>) involved in PRP31 binding do not affect the interaction with NUFIP (Figure 1C). Only residue H<sub>68</sub> was found to have an effect when the NUFIP<sub>223-260</sub> domain alone was used, which might be due to an indirect modification of helix  $\alpha$ 3 (see model). Obviously, NUFIP<sub>223-260</sub> only represents a very small piece of NUFIP, and the N-terminal part of PRP31 is missing in the 3D structure. However, the fact that the NUFIP and PRP31 binding sites on 15.5K are not overlapping sites opens the possibility that the two proteins might be accommodated together on 15.5K bound to U4 snRNA. Therefore, the proposed interaction between PRP31 and the C-terminal domain of NUFIP (24) might take place within a tri-protein complex (15.5K, PRP31 and NUFIP) bound to the RNA, which would be in accordance with the proposed role of NUFIP and Rsa1p in scaffolding the nascent RNPs (24,25). Such a scaffolding mechanism is reminiscent of the case of the SMN complex, which provides an external scaffold for the nascent Sm ring (10,11). It should also be noted that the HSP90 co-chaperone R2TP is recruited on the nascent snoRNPs through an interaction between PIH1 and NUFIP, and that the interaction site in NUFIP lies in the 50 amino acids upstream of NUFIP<sub>223-260</sub> (24). According to the model in Figure 11, HSP90 and its co-chaperone are expected to be located in a strategic position at proximity of both 15.5K and PRP31.

#### **Mutations in the EDK sequences of Snu13p or NUFIP have marked negative effects *in cellulo***

For the first time, we show the *in cellulo* effects of mutations in proteins 15.5K can abolish its interaction with NUFIP. The 15.5K mutant protein unable to bind NUFIP shows a substantially altered intracellular localization, with a stronger cytoplasmic signal that suggests a partial defect in incorporation into U4 and C/D RNPs. In yeast, deletion of the *RSAl* gene was known to induce a

marked temperature sensitive phenotype (42). However, Rsa1p was proposed to have several functions in yeast, in particular, it may be involved in biogenesis of the large ribosomal subunit (42). Here, based on the detailed mapping of the Rsa1p-Snu13p interaction, we could create point mutations in Rsa1p and Snu13p interface and test the *in vivo* importance of the individual electrostatic or Van der Waals interactions that we identified.

We observed that cells expressing Snu13p proteins impaired in their capacity to bind Rsa1p have a doubling time ~10% longer than their wild-type counterpart. Conversely, cells expressing Rsa1p mutants unable to bind Snu13p also grow ~25% less rapidly than the wild-type and exhibit a specific decrease of C/D snoRNAs level. Furthermore, we showed that binding of the Pih1p-Tah1p dimer on pre-snoRNP is dependent on Rsa1p and mutation of residue E<sub>72</sub> in Snu13p impairs this binding. Interestingly, mutation of residue E<sub>72</sub> in Snu13p or R<sub>249</sub> in Rsa1p found to form the essential electrostatic interactions are sufficient to get the growth phenotype. This is remarkable and indicates that this single protein-protein interaction and more precisely the electrostatic interaction of residues E<sub>72</sub> and R<sub>249</sub> and the hydrophobic contacts they established with the surrounding amino acids likely play a significant role in cell growth through the facilitation of the R2TP complex recruitment and the C/D RNPs assembly. The stronger phenotype of the complete deletion of the Rsa1p gene may be due both to the other interactions Rsa1p makes with the R2TP complex and core RNP proteins, which may partially compensate for the loss of Rsa1p binding to Snu13p, and to the additional functions of Rsa1p in yeast.

#### **SUPPLEMENTARY DATA**

Supplementary Table S1, Supplementary Figures S1, S2 and S3.

#### **ACCESSION NUMBERS**

The resonance assignments have been deposited in the BMRB with accession codes 17556. The corresponding coordinates and restraints for structure calculations have been deposited in the PDB with accession code 2LBE.

#### **ACKNOWLEDGEMENT**

C. Bobo is acknowledged for her technical assistance. A. Krol, C. Allmang-Cura are thanked for sharing data on protein SBP2.

## FUNDING

This work was supported by the Agence Nationale de la Recherche [ANR-06-BLAN-0208], the ARC, the french Ministère de l'Enseignement Supérieur et de la Recherche [doctoral fellowships to B.R. and R.B.]; the Centre National de la Recherche Scientifique ; the Pôle de Recherche Scientifique et Technologique "Bioingénierie" of Région lorraine. The authors are grateful to the Service Commun de Biophysicochimie des Interactions Moléculaires (SCBIM) of Nancy University for access to the CD spectrometer.

## REFERENCES

1. Nottrott, S., Hartmuth, K., Fabrizio, P., Urlaub, H., Vidovic, I., Ficner, R. and Luhrmann, R. (1999) Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5' stem-loop of U4 snRNA. *EMBO J*, **18**, 6119-6133.
2. Watkins, N.J., Segault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C. and Luhrmann, R. (2000) A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP. *Cell*, **103**, 457-466.
3. Henras, A., Henry, Y., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Gelugne, J.P. and Caizergues-Ferrer, M. (1998) Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *EMBO J*, **17**, 7078-7090.
4. Wang, C. and Meier, U.T. (2004) Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *EMBO J*, **23**, 1857-1867.
5. Allmang, C., Carbon, P. and Krol, A. (2002) The SBP2 and 15.5 kD/Snu13p proteins share the same RNA binding domain: identification of SBP2 amino acids important to SECIS RNA binding. *RNA*, **8**, 1308-1318.
6. Copeland, P.R., Fletcher, J.E., Carlson, B.A., Hatfield, D.L. and Driscoll, D.M. (2000) A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J*, **19**, 306-314.
7. Matera, A.G., Terns, R.M. and Terns, M.P. (2007) Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol*, **8**, 209-220.
8. Yong, J., Wan, L. and Dreyfuss, G. (2004) Why do cells need an assembly machine for RNA-protein complexes? *Trends Cell Biol*, **14**, 226-232.
9. Kolb, S.J., Battle, D.J. and Dreyfuss, G. (2007) Molecular functions of the SMN complex. *J Child Neurol*, **22**, 990-994.
10. Chari, A., Paknia, E. and Fischer, U. (2009) The role of RNP biogenesis in spinal muscular atrophy. *Curr Opin Cell Biol*, **21**, 387-393.
11. Pellizzoni, L. (2007) Chaperoning ribonucleoprotein biogenesis in health and disease. *EMBO Rep*, **8**, 340-345.
12. Darzacq, X., Kittur, N., Roy, S., Shav-Tal, Y., Singer, R.H. and Meier, U.T. (2006) Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *J Cell Biol*, **173**, 207-218.
13. Dez, C., Noaillac-Depeyre, J., Caizergues-Ferrer, M. and Henry, Y. (2002) Naf1p, an essential nucleoplasmic factor specifically required for accumulation of box H/ACA small nucleolar RNPs. *Mol Cell Biol*, **22**, 7053-7065.
14. Fatica, A., Dlakic, M. and Tollervey, D. (2002) Naf1p is a box H/ACA snoRNP assembly factor. *RNA*, **8**, 1502-1514.
15. Marmier-Gourrier, N., Clery, A., Senty-Segault, V., Charpentier, B., Schlotter, F., Leclerc, F., Fournier, R. and Branlant, C. (2003) A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA. *RNA*, **9**, 821-838.
16. Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R. and Ficner, R. (2000) Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol Cell*, **6**, 1331-1342.
17. Clery, A., Senty-Segault, V., Leclerc, F., Raue, H.A. and Branlant, C. (2007) Analysis of sequence and structural features that identify the B/C motif of U3 small nucleolar RNA as the recognition site for the Snu13p-Rrp9p protein pair. *Mol Cell Biol*, **27**, 1191-1206.
18. Schultz, A., Nottrott, S., Hartmuth, K. and Luhrmann, R. (2006) RNA structural requirements for the association of the spliceosomal hPrp31 protein with the U4 and U4atac small nuclear ribonucleoproteins. *J Biol Chem*, **281**, 28278-28286.
19. Watkins, N.J., Dickmanns, A. and Luhrmann, R. (2002) Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Mol Cell Biol*, **22**, 8342-8352.
20. Liu, S., Li, P., Dybkov, O., Nottrott, S., Hartmuth, K., Luhrmann, R., Carlomagno, T. and Wahl, M.C. (2007) Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP. *Science*, **316**, 115-120.
21. Lin, J., Lai, S., Jia, R., Xu, A., Zhang, L., Lu, J. and Ye, K. (2011) Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. *Nature*, **469**, 559-563.
22. Xue, S., Wang, R., Yang, F., Terns, R.M., Terns, M.P., Zhang, X., Maxwell, E.S. and Li, H. (2010) Structural basis for substrate placement by an archaeal box C/D ribonucleoprotein particle. *Mol Cell*, **39**, 939-949.
23. Ye, K., Jia, R., Lin, J., Ju, M., Peng, J., Xu, A. and Zhang, L. (2009) Structural organization of box C/D RNA-guided RNA methyltransferase. *Proc Natl Acad Sci U S A*, **106**, 13808-13813.
24. Boulon, S., Marmier-Gourrier, N., Pradet-Balade, B., Wurth, L., Verheggen, C., Jady, B.E., Rothe, B., Pescia, C., Robert, M.C., Kiss, T. et al. (2008) The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J Cell Biol*, **180**, 579-595.
25. McKeegan, K.S., Debieux, C.M., Boulon, S., Bertrand, E. and Watkins, N.J. (2007) A dynamic scaffold of pre-snoRNP factors facilitates human box C/D snoRNP assembly. *Mol Cell Biol*, **27**, 6782-6793.

26. Zhao, R. and Houry, W.A. (2005) Hsp90: a chaperone for protein folding and gene regulation. *Biochem Cell Biol*, **83**, 703-710.
27. Zhao, R., Kakihara, Y., Gribun, A., Huen, J., Yang, G., Khanna, M., Costanzo, M., Brost, R.L., Boone, C., Hughes, T.R. *et al.* (2008) Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation. *J Cell Biol*, **180**, 563-578.
28. Diebold, M.L., Fribourg, S., Koch, M., Metzger, T. and Romier, C. (2011) Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. *J Struct Biol*, **175**, 178-188.
29. Fribourg, S., Romier, C., Werten, S., Gangloff, Y.G., Poterszman, A. and Moras, D. (2001) Dissecting the interaction network of multiprotein complexes by pairwise coexpression of subunits in *E. coli*. *J Mol Biol*, **306**, 363-373.
30. Romier, C., Ben Jelloul, M., Albeck, S., Buchwald, G., Busso, D., Celie, P.H., Christodoulou, E., De Marco, V., van Gerwen, S., Knipscheer, P. *et al.* (2006) Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. *Acta Crystallogr D Biol Crystallogr*, **62**, 1232-1242.
31. Bitter, G.A. and Egan, K.M. (1984) Expression of heterologous genes in *Saccharomyces cerevisiae* from vectors utilizing the glyceraldehyde-3-phosphate dehydrogenase gene promoter. *Gene*, **32**, 263-274.
32. Stutz, F. and Rosbash, M. (1994) A functional interaction between Rev and yeast pre-mRNA is related to splicing complex formation. *EMBO J*, **13**, 4096-4104.
33. Boulon, S., Pradet-Balade, B., Verheggen, C., Molle, D., Boireau, S., Georgieva, M., Azzag, K., Robert, M.C., Ahmad, Y., Neel, H. *et al.* (2010) HSP90 and Its R2TP/Prefoldin-like Cochaperone Are Involved in the Cytoplasmic Assembly of RNA Polymerase II. *Mol Cell*, **39**, 912-924.
34. Charpentier, B., Muller, S. and Branlant, C. (2005) Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res*, **33**, 3133-3144.
35. Keller, R.L.J. (2004) The Computer Aided Resonance Assignment Tutorial. *First edition ed. CANTINA Verlag, Zürich*.
36. Guntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol*, **278**, 353-378.
37. Brunger, A.T. (2007) Version 1.2 of the Crystallography and NMR system. *Nat Protoc*, **2**, 2728-2733.
38. Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L. and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem*, **26**, 1781-1802.
39. MacKerell, A.D., Jr., Banavali, N. and Foloppe, N. (2000) Development and current status of the CHARMM force field for nucleic acids. *Biopolymers*, **56**, 257-265.
40. Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R.H. *et al.* (2004) From silencing to gene expression: real-time analysis in single cells. *Cell*, **116**, 683-698.
41. Oruganti, S., Zhang, Y. and Li, H. (2005) Structural comparison of yeast snoRNP and spliceosomal protein Snu13p with its homologs. *Biochem Biophys Res Commun*, **333**, 550-554.
42. Kressler, D., Doere, M., Rojo, M. and Linder, P. (1999) Synthetic lethality with conditional dbp6 alleles identifies rsa1p, a nucleoplasmic protein involved in the assembly of 60S ribosomal subunits. *Mol Cell Biol*, **19**, 8633-8645.
43. Eckert, K., Saliou, J.M., Monlezun, L., Vigouroux, A., Atmane, N., Caillat, C., Quevillon-Cheruel, S., Madiona, K., Nicaise, M., Lazereg, S. *et al.* The Pih1-Tah1 cochaperone complex inhibits Hsp90 molecular chaperone ATPase activity. *J Biol Chem*, **285**, 31304-31312.
44. Schultz, A., Nottrott, S., Watkins, N.J. and Luhrmann, R. (2006) Protein-protein and protein-RNA contacts both contribute to the 15.5K-mediated assembly of the U4/U6 snRNP and the box C/D snoRNPs. *Mol Cell Biol*, **26**, 5146-5154.
45. McKeegan, K.S., Debieux, C.M. and Watkins, N.J. (2009) Evidence that the AAA+ proteins TIP48 and TIP49 bridge interactions between 15.5K and the related NOP56 and NOP58 proteins during box C/D snoRNP biogenesis. *Mol Cell Biol*, **29**, 4971-4981.
46. Koonin, E.V., Bork, P. and Sander, C. (1994) A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins and a ribosome modification enzyme? *Nucleic Acids Res*, **22**, 2166-2167.
47. Schrodinger, LLC. (2010).
48. Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J Mol Graph*, **14**, 33-38, 27-38.

## TABLE AND FIGURES LEGENDS

**Figure 1.** Residues in 15.5K helix  $\alpha 3$  are involved in the 15.5K-NUFIP<sub>223-260</sub> interaction. **(A)** Schematic representation of the human protein NUFIP and the yeast protein Rsa1p. The homologous domain between the two proteins (named hPEP in human and yPEP in yeast) is colored in gray. The various fragments of NUFIP and Rsa1p used in this study are represented. **(B)** The RNA binding domain shared by members of the L7Ae family. Multiple amino acid sequence alignments were constructed using CLUSTALW. Secondary structure elements of human protein 15.5K are labeled on top of the sequences. Numbering correspond to residues of 15.5K. Identical residues are highlighted in light grey and similar residues are in dark grey. **(C)** Yeast two hybrid interactions between 15.5K and NUFIP, NUFIP<sub>223-260</sub> or Alix as a negative control. The following residues of 15.5K were changed into alanines except when indicated. MutA: H<sub>17</sub>L<sub>18</sub>T<sub>19</sub>K<sub>20</sub>K<sub>21</sub>; MutB: K<sub>33</sub>Q<sub>34</sub>L<sub>35</sub>R<sub>36</sub>; MutC: E<sub>61</sub>P<sub>62</sub>E<sub>64</sub>; MutD: E<sub>74</sub>D<sub>75</sub>K<sub>76</sub>; MutE: R<sub>84</sub>S<sub>85</sub>K<sub>86</sub>;



MutF: K<sub>107</sub>E<sub>108</sub>K<sub>113</sub>A<sub>114</sub>; MutG: Q<sub>27</sub>Q<sub>28</sub>; MutH: N<sub>31</sub>Y<sub>32</sub>K<sub>33</sub>; MutI: E124S R125S; MutJ: N<sub>40</sub>T<sub>43</sub>N<sub>47</sub>.

**Figure 2.** 15.5K E<sub>74</sub>A does not bind NUFIP *in vivo* and is not efficiently incorporated into RNPs. U2OS and U2OS LacO cells were transfected with mRFP-Laci-NUFIP and either GFP-15.5K (upper two panels), or its EDK->AAA variant (bottom two panels). WT 15.5K accumulates in nucleoli, in Cajal bodies and is also diffusely present in the nucleoplasm (top left panel). Furthermore, when mRFP1-Laci-NUFIP accumulates at the LacO sites, WT 15.5K is also enriched there (see insets in the top panel). In contrast, while the mutant 15.5K is also present in nucleoli, Cajal bodies and in the nucleoplasm, it is also present in large amount in the cytoplasm, indicating a defect in RNP assembly (bottom left panel). In addition, this mutant fails to be recruited to the LacO sites by mRFP-Laci-NUFIP (bottom panels).

**Figure 3.** Functional analysis of the Snu13p-Rsa1p interaction. (A) Comparison of the 15.5K and Snu13p proteins. 3D structure representations of the complex formed between human protein 15.5K and the 5' stem-loop of U4 snRNA (PDB accession number: 1E7K) (16), and of the *S. cerevisiae* Snu13p protein (PDB accession number: 2ALE) (41) were obtained by using the PyMOL software (47). Residues of the COOH terminus of helix  $\alpha$ 3 are indicated by black letters. On the U4-15.5K structure, residues in pink interact with PRP31 in the U4-15.5K-PRP31<sub>78-333</sub> complex (20). Alanine substitutions performed in Snu13p are displayed in the insert. (B) Yeast two-hybrid assays between Snu13p and its variants used as the bait and the preys Rsa1p or its minimal functional fragment Rsa1p<sub>230-381</sub>. A control was performed with an empty vector (-). Growth on medium lacking histidine but containing 3-Amino-1,2,4-triazol (3AT), a competitive inhibitor of the product of the *HIS3* reporter gene, indicates an interaction. (C) Co-expression and co-purification assays. The wild type or mutated His-tagged protein Snu13p was co-expressed in *E. coli* with Rsa1p. His-pull down assays from the crude cellular extracts were performed by IMAC. The soluble fraction (Input, In) and the purified proteins (Eluate, E) were fractionated by SDS-PAGE and the gel stained with Coomassie blue. The ladder (L) is indicated on the left. (D) Gel-shift assays. Radiolabeled U14 snoRNA was incubated with the wild type Snu13p or one of its  $\alpha$ 3 helix variants and the indicated His-tagged Rsa1p<sub>230-381</sub> or Rsa1p<sub>230-266</sub>. GST is used as a negative control.

**Figure 4.** Study of protein-protein interaction by circular dichroism. A two-compartment cuvette was used and spectra were recorded as described in Materials and Methods. (A) and (B), Snu13p-Rsa1p<sub>230-266</sub> interaction. For each interaction experiments, the CD spectra recorded before mixing the samples are in dark blue and the spectra in light blue (WT Snu13p) or in dark green were obtained after mixing the two compartments.

**Figure 5.** Functional analysis of residues of peptide Rsa1p<sub>230-266</sub>. (A) The amino acid sequence of Rsa1p<sub>230-266</sub>, which is conserved between NUFIP and Rsa1p, was mutated. Alanine substitutions in the various Rsa1p variants are underlined. (B) Analysis of the substitutions by EMSA in the same conditions as in Figure 3D. (C) Yeast two-hybrid assays using Snu13p used as the bait and WT or variant Rsa1p<sub>230-381</sub> fragments of Rsa1p as the preys. Analyses were performed as in Figure 2B. (D) Co-expression and co-purification assays. Same experiments as in Figure 3C but with His<sub>6</sub>Snu13p co-expressed in *E. coli* with one of the Rsa1p or Rsa1p<sub>230-381</sub> variants.

**Figure 6.** NMR structure of Rsa1p<sub>238-252</sub>. (A) Backbone superimposition of the ten best structures of Rsa1p<sub>230-266</sub>. (B) Side chains representation of the lowest-energy Rsa1p<sub>230-266</sub> structure. The helicoidal conformation is highlighted using a cartoon representation. The hydrogen atoms are not represented. The figure has been done with PyMOL software (47).

**Figure 7.** 3D structure model of the Snu13p and Rsa1p interaction interface. (A) and (B), Structure viewed in two opposite directions rotated of 180°. Modeling was performed as explained in Materials and Methods. The Rsa1p<sub>230-266</sub> peptide and the Snu13p  $\alpha$ 3 helix are displayed by the orange and green ribbons, respectively. The lateral chains at the interaction interface are in a stick representation and are labeled. Electrostatic interactions are indicated by black lines. This representation of the interaction was obtained with the VMD 1.8.6 software (48).

**Figure 8:** Role of residue L<sub>69</sub> in the Snu13p-Rsa1p interaction. Formation of U14 RNA-Snu13p-Rsa1p<sub>230-266</sub> complexes were analyzed by gel-shift assays as in Figure 3D.

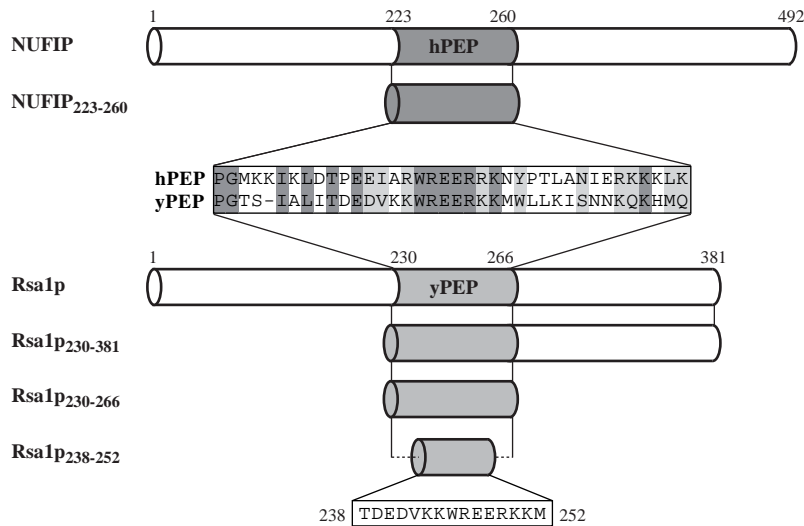
**Figure 9:** Effect of substitutions within the Snu13p helix  $\alpha$ 3 or the Rsa1p<sub>230-266</sub> domain on cell growth. (A) Yeast strain YPH499-GAL::*SNU13* was transformed with various recombinant pG1::*SNU13*

vectors and grown in medium containing glucose. Effect of the ectopic expression of Snu13p and its variant forms on cell growth was monitored by measurement of OD at 600 nm and by maintaining cell in exponential growth by successive dilutions. **(B)** Effects of ectopic expression of WT and variant Rsa1p from recombinant pG1::*RSAl* plasmid on growth of a *RSAl* KO strain. *C*, *C/D* snoRNAs levels in cells with ectopic expression of WT and R249/A variant Rsa1p from recombinant pG1::*RSAl* plasmid.

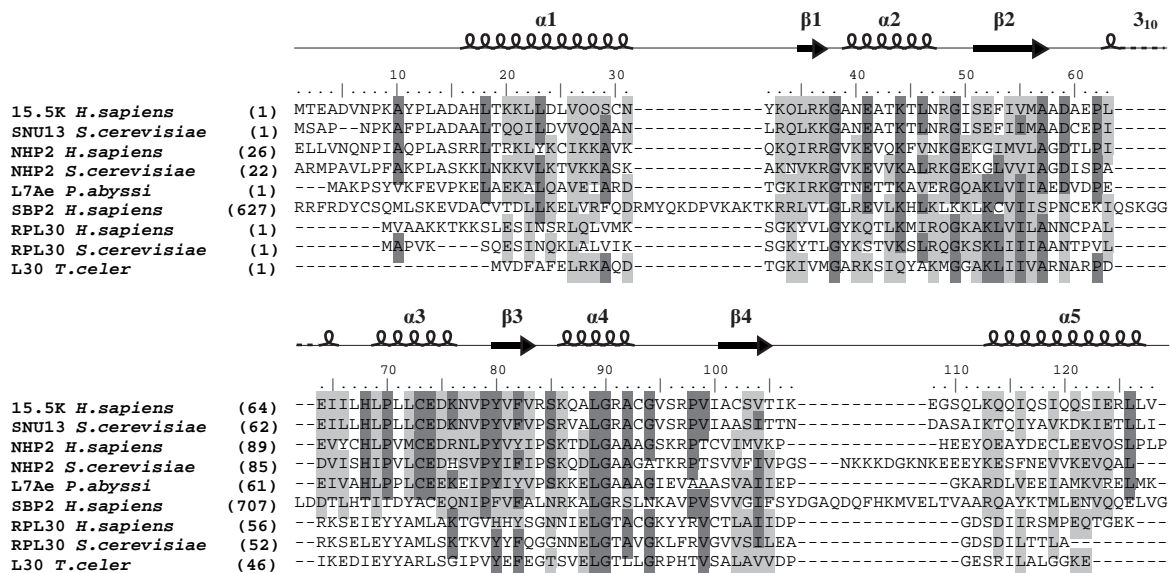
**Figure 10:** *In vitro* interaction of Pih1p with the pre-RNP U14 snoRNA-Snu13p-Rsa1p<sub>230-381</sub>. **(A)** His-pull down assay with the His<sub>6</sub>Tah1p-Pih1p protein complex. The radiolabeled U14 snoRNA was incubated with the recombinant Snu13 protein and the His<sub>6</sub>Tah1p-Pih1p dimer. The same experiment was realized in presence of an increasing concentration of the Rsa1p<sub>230-381</sub> fragment of Rsa1p. The mixture was incubated with nickel-sepharose beads to allow binding of His<sub>6</sub>Tah1p-Pih1p, and retention of the RNA was analyzed by denaturing gel electrophoresis. The input corresponds to 10% of the total quantity of radiolabeled RNA that was used in the binding assays (on the left). **(B)** His-pull down performed in the same conditions with recombinant Snu13p WT and Snu13p E72/A.

**Figure 11:** Tentative model of the possible interaction between the U4-SL RNA (represented in green), 15.5K (displayed in purple) and PRP31<sub>78-333</sub> (displayed in orange) (crystal structure) and NUFIP<sub>223-260</sub> (displayed in gray). The 15.5K-NUFIP<sub>223-260</sub> complex was positioned by superimposing the helix  $\alpha$ 3 of Snu13p-Rsa1p<sub>230-266</sub> model with the same helix of 15.5K in the crystal structure of the U4-15.5K-PRP31<sub>78-333</sub> complex (PDB accession number: 2OZB). Then, the representation of Snu13p-Rsa1p<sub>230-266</sub> structure was hidden. This figure was obtained by using the PyMOL software (47). The side chains of residues in contact are displayed in bold and stick representation. Lines and dashed lines show respectively 15.5K-NUFIP<sub>223-260</sub> and 15.5K-PRP31<sub>78-333</sub>.

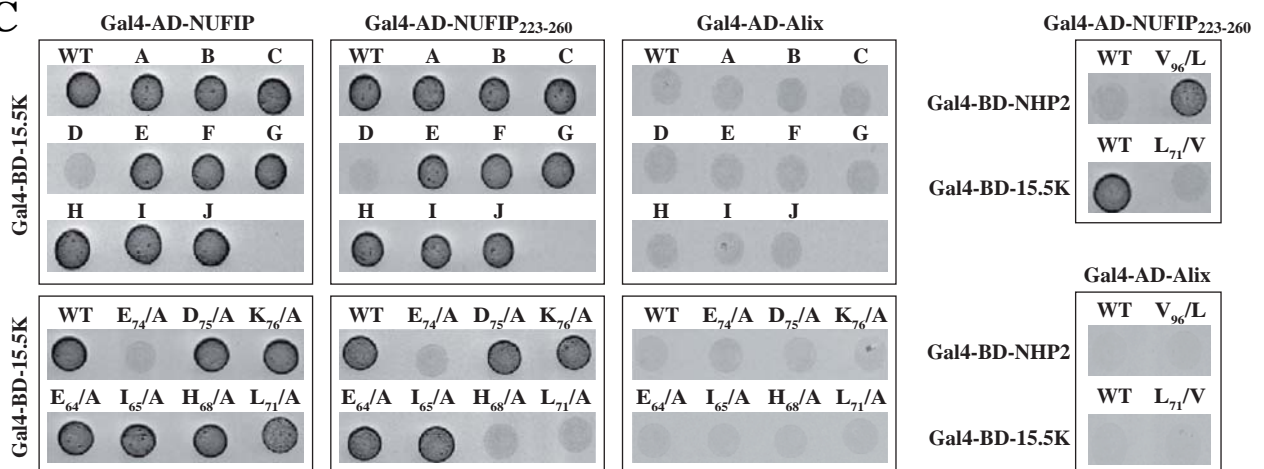
**A**



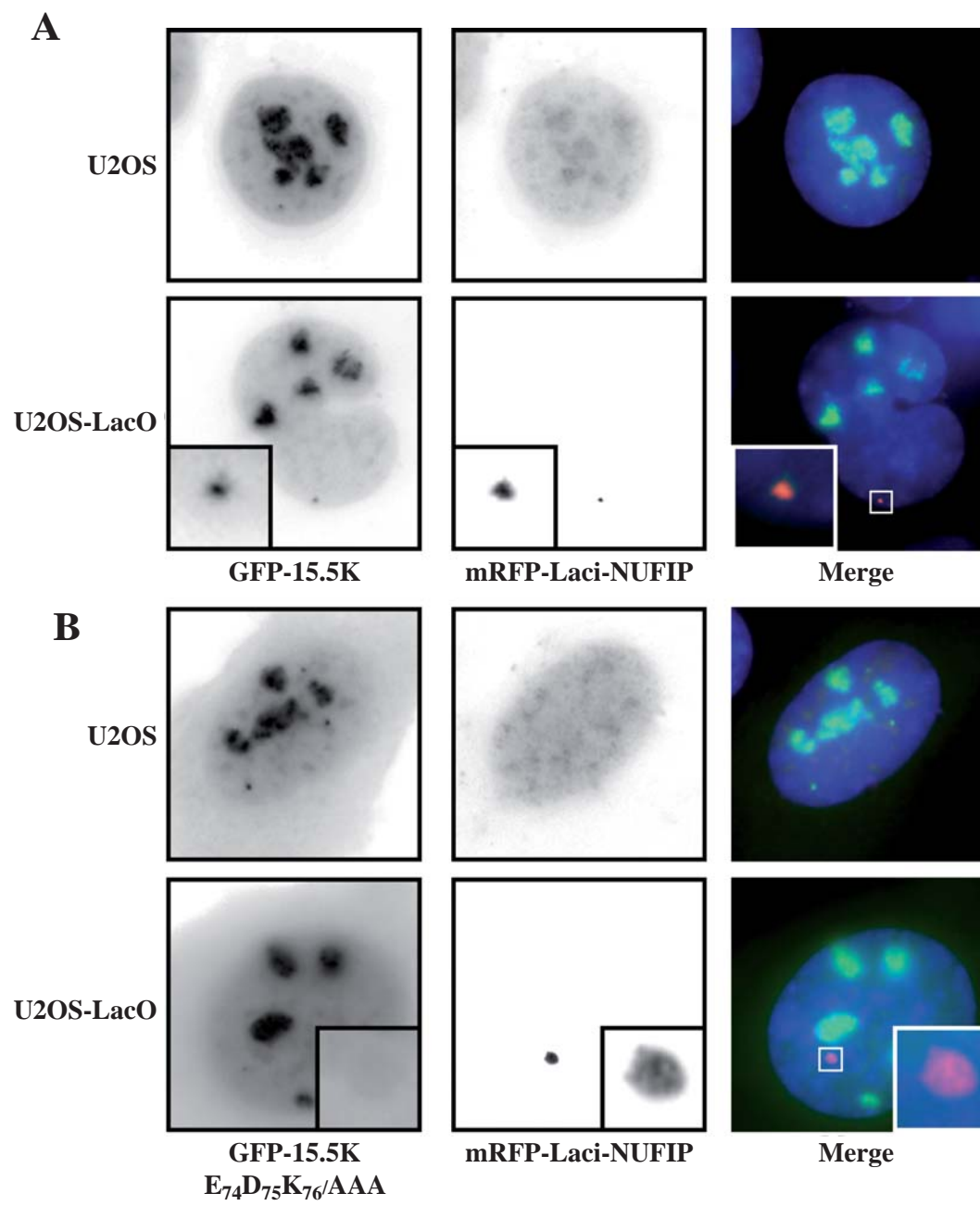
**B**



**C**

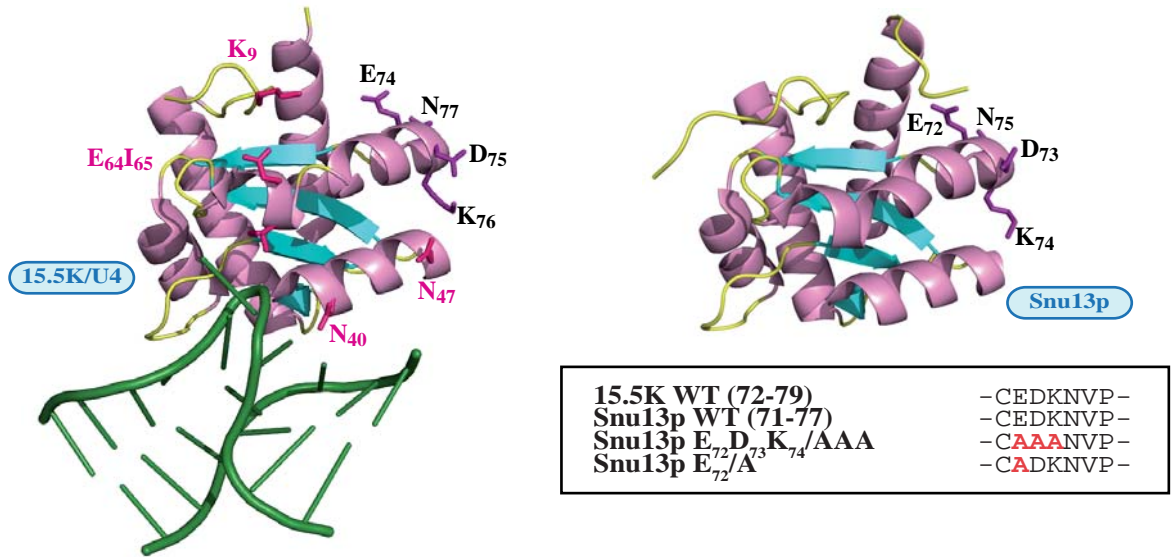


**Figure 1.**

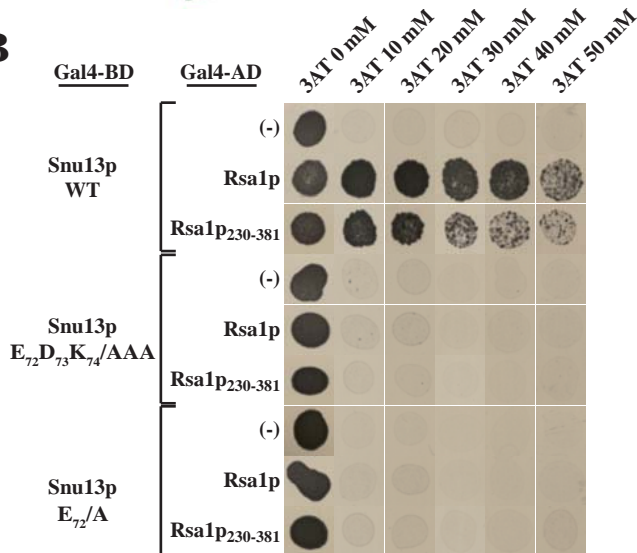


**Figure 2.**

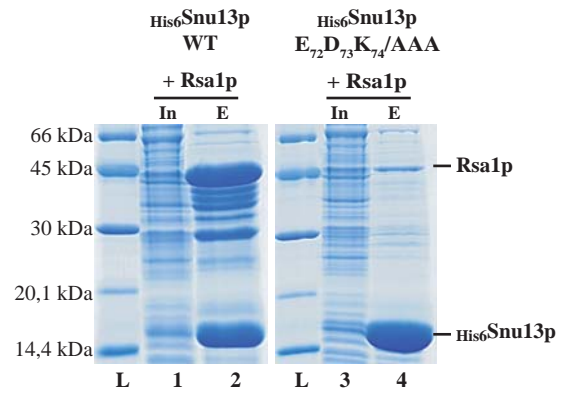
**A**



**B**



**C**



**D**

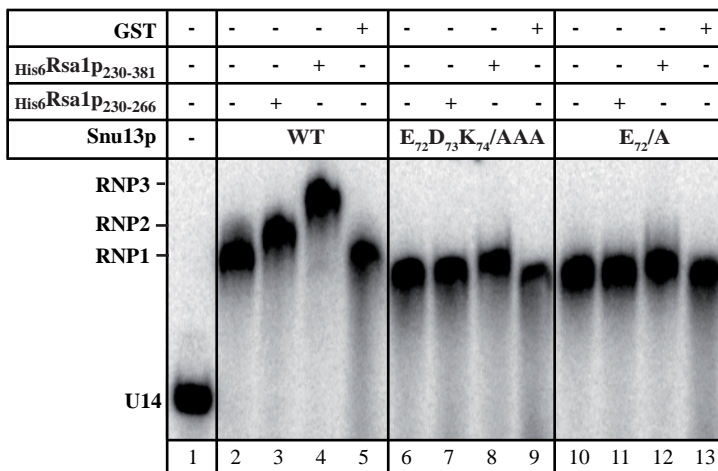
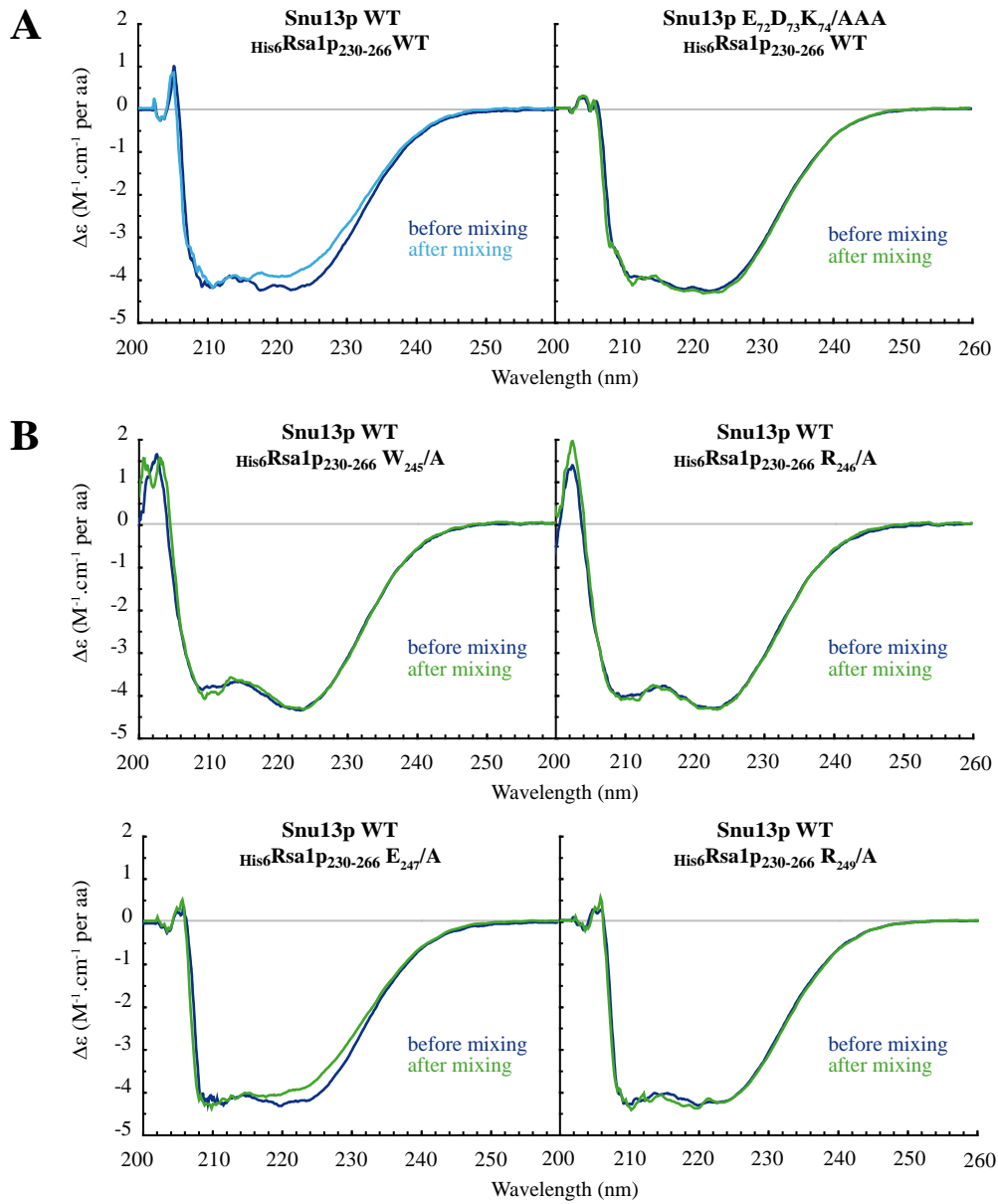


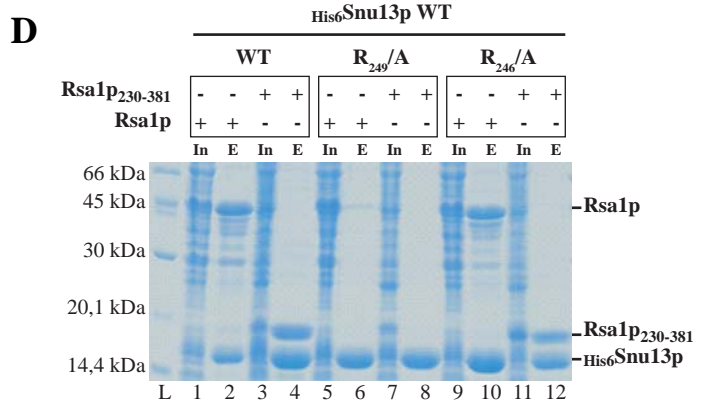
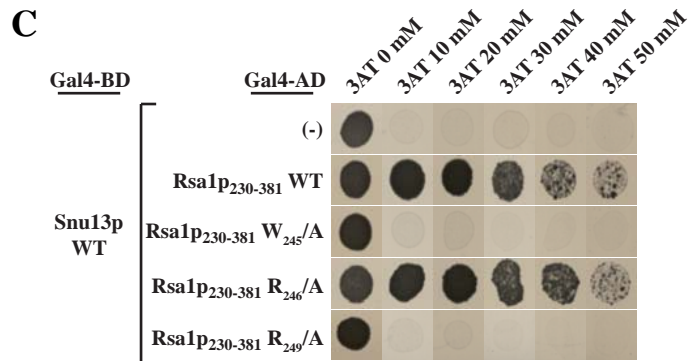
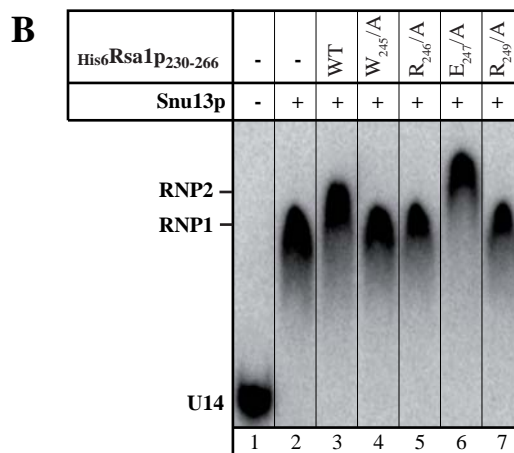
Figure 3.



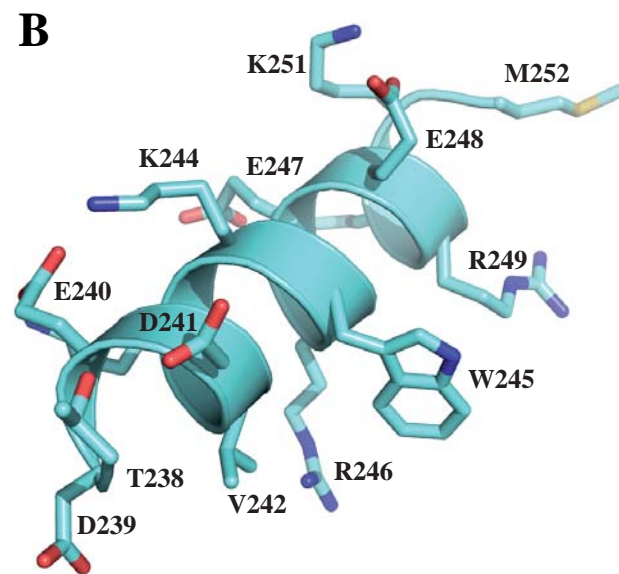
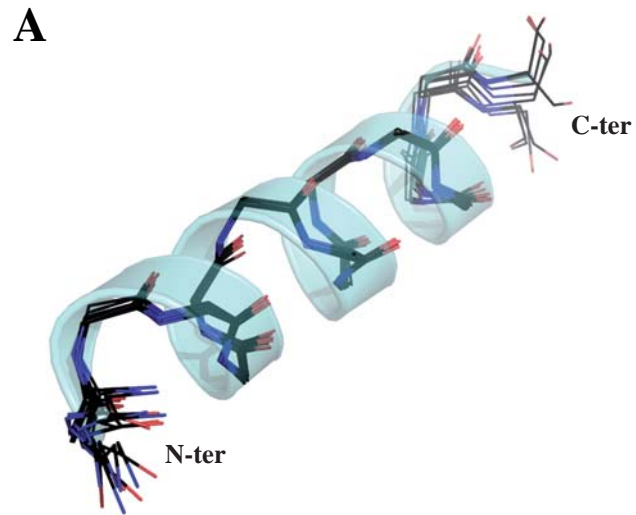
**Figure 4.**

**A**

NUFIP	238	RWREERRK	245
Rsa1p WT	244	KWREERKK	251
Rsa1p W <sub>245</sub> /A	244	K <b>A</b> REERKK	251
Rsa1p R <sub>246</sub> /A	244	KW <b>A</b> EERKK	251
Rsa1p E <sub>247</sub> /A	244	KWR <b>A</b> EERKK	251
Rsa1p R <sub>249</sub> /A	244	KWREER <b>A</b> KK	251

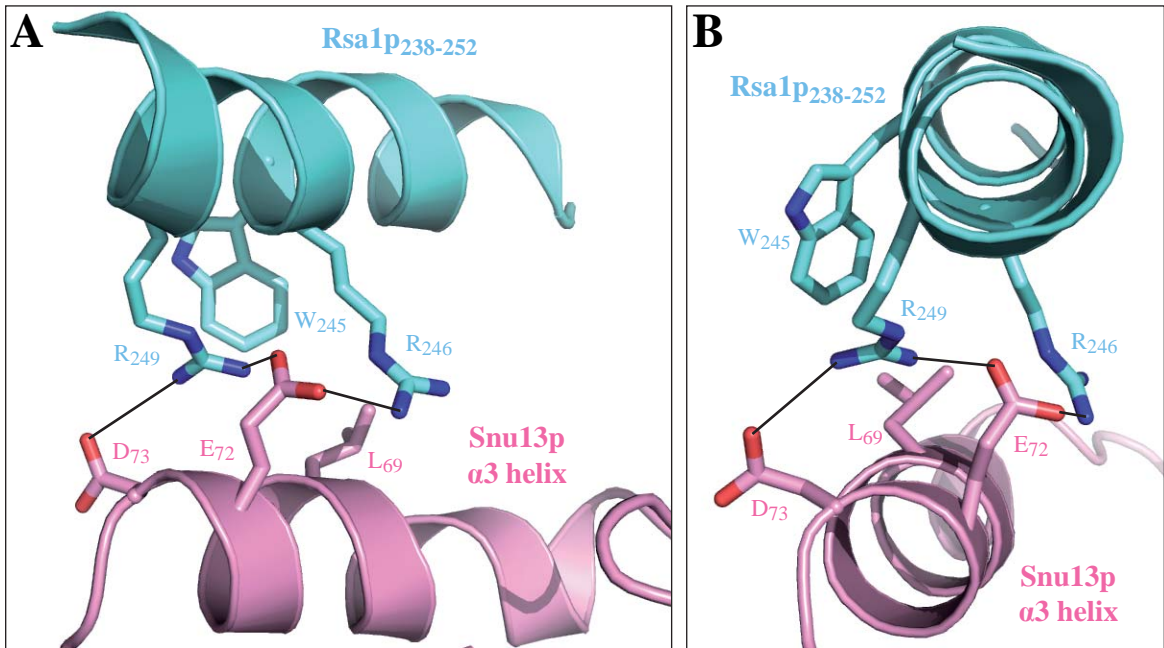


**Figure 5.**

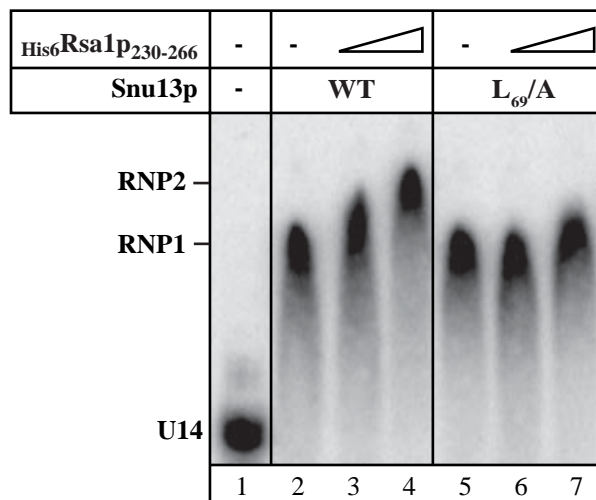


**Figure 6.**





**Figure 7.**



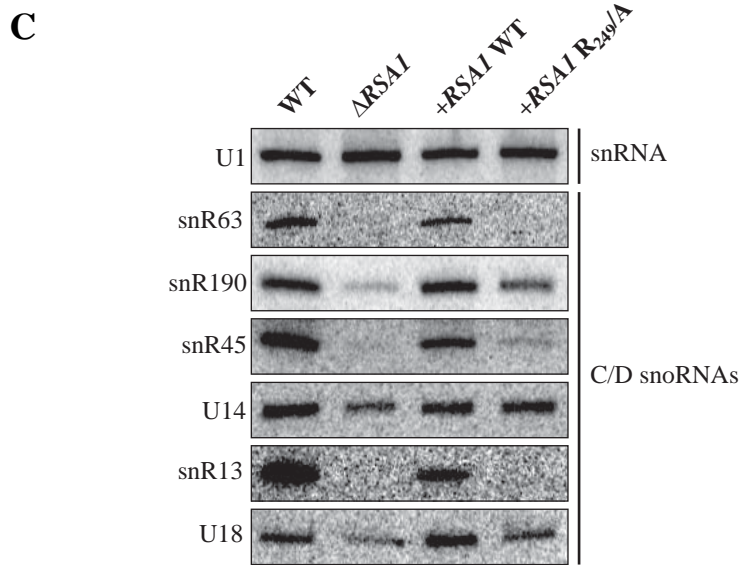
**Figure 8.**

**A**

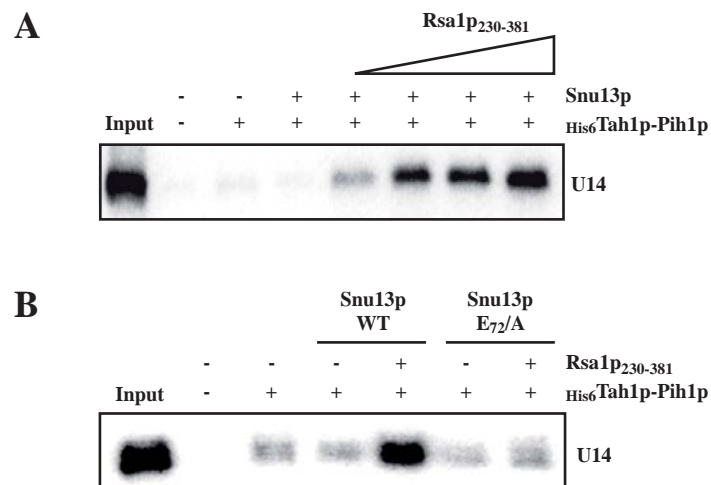
Strain <i>Gal::SNU13</i>	G (h)	Growth defect (% WT)
pGI:: <i>SNU13</i> -WT	3.26 ±0.04	
pGI:: <i>SNU13</i> -E72D73K74/AAA	3.56 ±0.05	-9.2
pGI:: <i>SNU13</i> -E72/A	3.58 ±0.09	-9.8
pGI:: <i>SNU13</i> -L69/A	3.22 ±0.09	+1.2

**B**

Strain $\Delta$ <i>RSAI</i>	G (h)	Growth defect (% WT)
pGI::(-)	3.74 ±0.18	-44.2
pGI:: <i>RSAI</i> - WT	2.6 ±0.13	
pGI:: <i>RSAI</i> - W245/A	3.06 ±0.10	-17.8
pGI:: <i>RSAI</i> - R246/A	2.55 ±0.10	+1.7
pGI:: <i>RSAI</i> - R249/A	3.24 ±0.03	-24.8



**Figure 9.**



**Figure 10.**

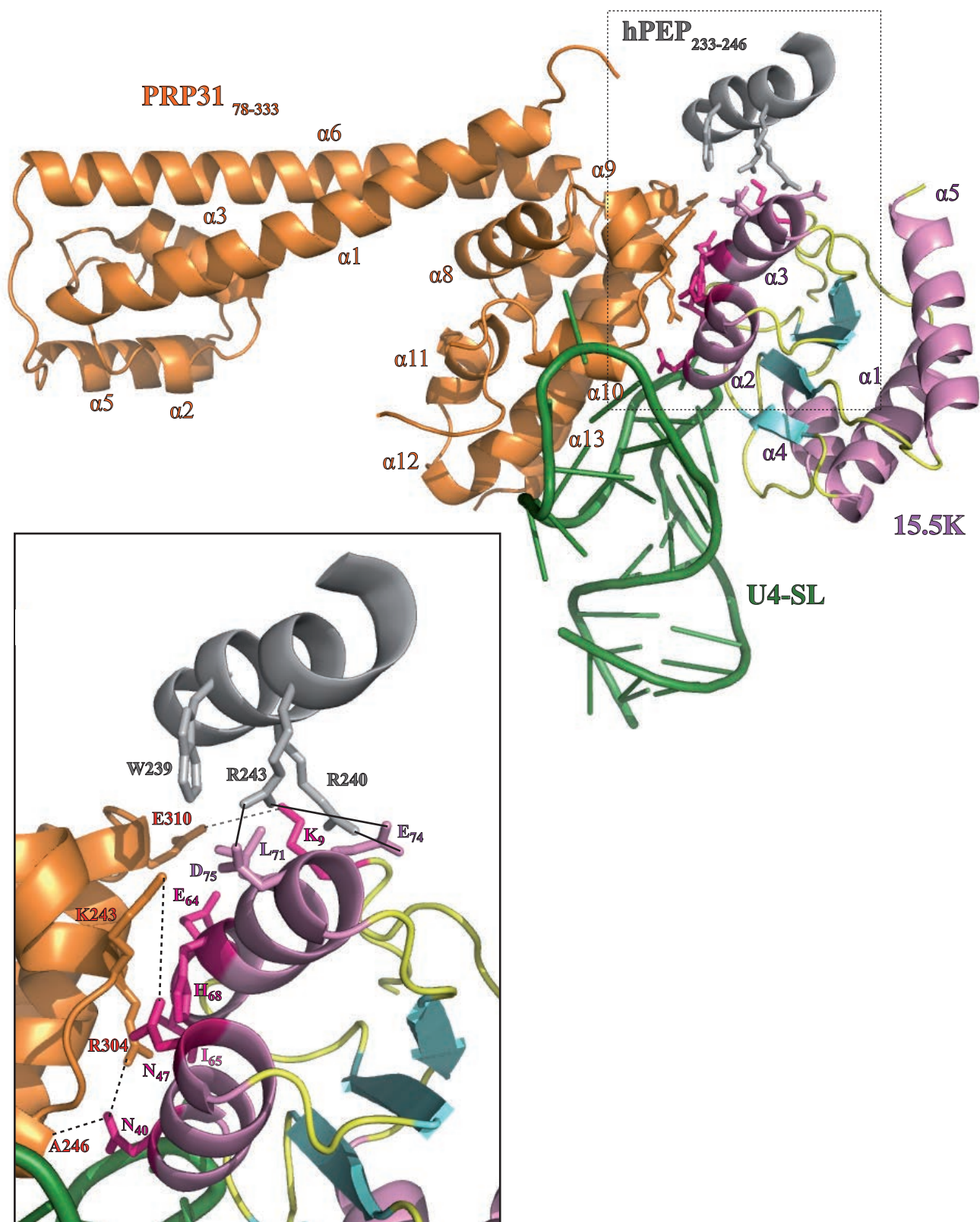


Figure 11.

## SUPPLEMENTARY DATA

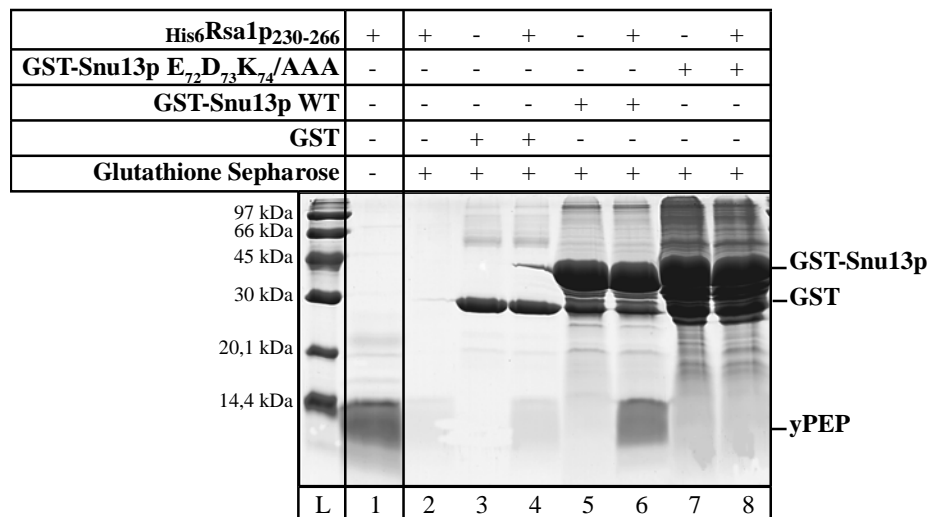
### SUPPLEMENTARY METHODS

#### GST pull-down assays

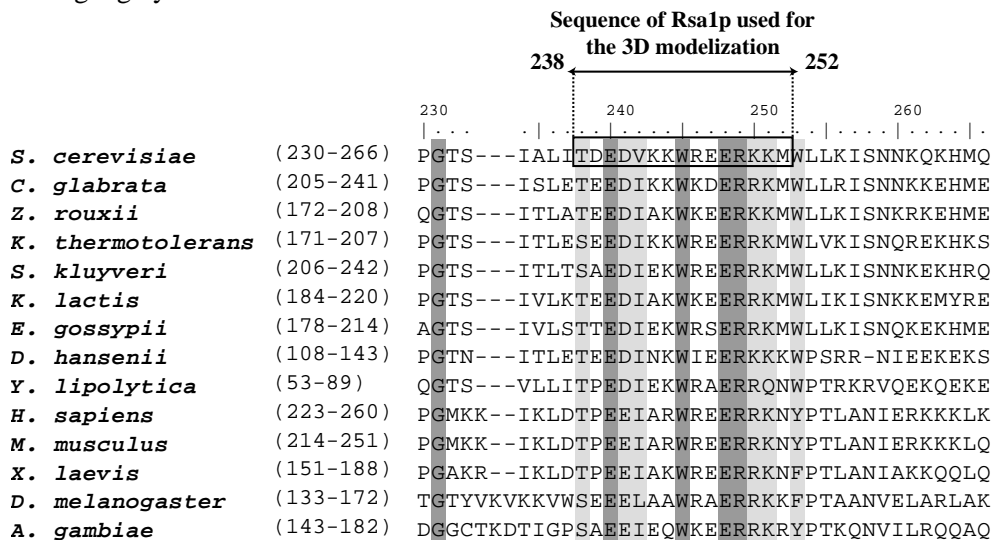
Recombinant proteins GST-Snu13p, wild-type or variant E<sub>72</sub>D<sub>73</sub>K<sub>74</sub>/AAA, (0.35 nmol) were loaded onto 15 µl of beads of Glutathione-Sepharose 4B (GE Healthcare®) in 100 µl of binding buffer (10 mM phosphate pH 7.0, 150 mM NaCl). Recombinant His<sub>6</sub>Rsa1p<sub>230-266</sub> (0.7 nmol) was treated in the same way. The protein partners were incubated for 30 min at room temperature. Beads loaded with proteins were washed three times with the binding buffer. Proteins were eluted by boiling in Laemmli loading buffer. Protein fractions were resolved by SDS-PAGE on 15% polyacrylamide gels and visualized by Coomassie staining.

### SUPPLEMENTARY FIGURES (S1-S3)

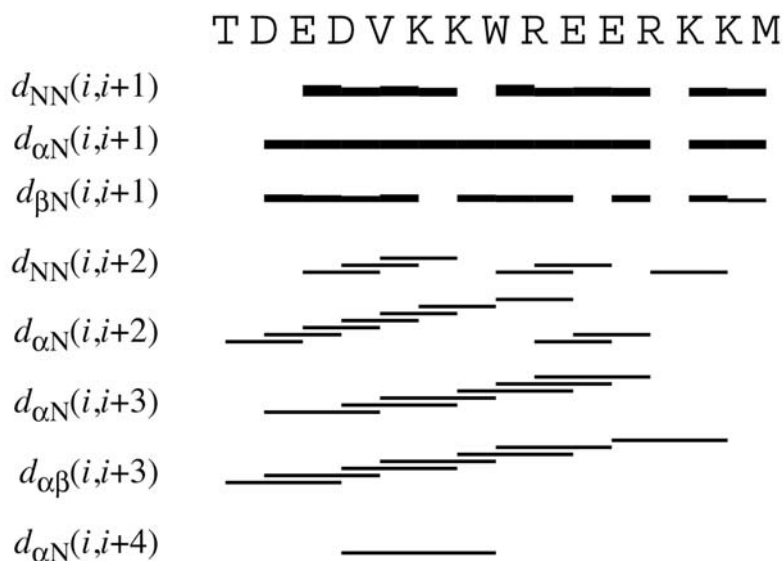
**Figure S1.** GST-pull down experiments. A His-tagged version of the conserved peptide yPEP within the yeast Ras1p protein was over-expressed and purified in *E. coli*. It was assayed for binding to the indicated recombinant GST fusions. The gel was stained with Coomassie blue.



**Figure S2.** The PEP sequence. Multiple amino acid sequence alignments were constructed using CLUSTALW. Numbering corresponds to residues of the *S. cerevisiae* Rsa1p protein. Identical residues are highlighted in light grey.



**Figure S3.** NOE connectivities for amide proton amide (N), alpha ( $\alpha$ ) and beta ( $\beta$ ) protons of Rsa1p<sub>238-252</sub>. The NOE intensities are represented by different line thicknesses, stronger lines meaning stronger intensities.



#### SUPPLEMENTARY TABLE (S1)

**Table S1.** NMR-Derived Geometrical Restraints and Structural Statistics of Rsa1p<sub>238-252</sub>.

Restrains		Number
Distance restraints		
	total of distance restraints	169
	intra	43
	short	65
	medium	61
dihedral angles		
	$\phi$	12
Violation		Occurrence
	distance violation > 0.2 Å	0
	dihedral violation > 5 degrees	0
Ramachandran statistics (PROCHECK-NMR)		
	residues in the most favorable region	96.2 %
	residues in additionally allowed regions	3.8 %
	residues in generously allowed regions	0 %
	residues in disallowed regions	0 %
Atoms (residues E <sub>240</sub> to K <sub>250</sub> )		rmsd (Å)
	backbone	0.23 ± 0.07
	heavy	1.31 ± 0.42

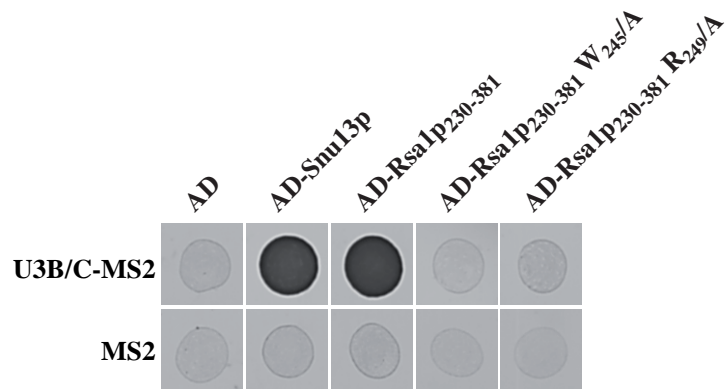
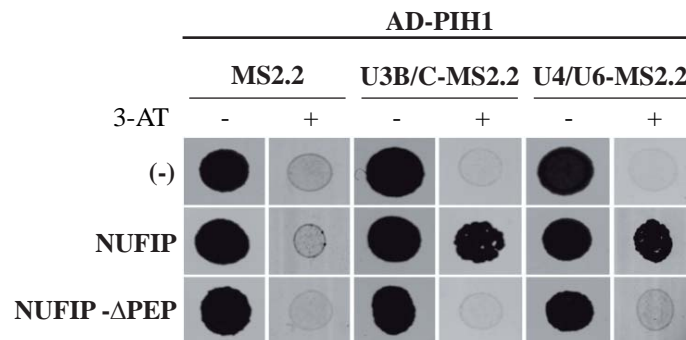
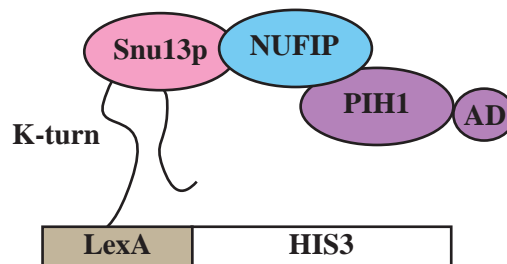


## 2.2. Données complémentaires

Dans l'article *Formation of electrostatic and hydrophobic interactions between alpha helices of the Rsa1p assembly factor and the Snu13p snoRNP protein is required for efficient snoRNP assembly*, nous avons mis en évidence l'importance de l'interaction physique établie entre Rsa1p et Snu13p pour la production normale des snoRNA à boîtes C/D et pour la croissance cellulaire. Des expériences réalisées *in vitro* nous ont également permis de montrer que la perte de cette interaction abolit le recrutement du complexe Pih1p-Tah1p sur les pré-snoRNP. Cette donnée révèle donc qu'en plus de son rôle de plateforme pour l'assemblage de Nop56p et Nop58p, Rsa1p pourrait faciliter le recrutement des autres facteurs d'assemblage sur les pré-RNP. Afin de confirmer ces données *in cellulo*, nous avons réalisé des expériences supplémentaires par l'approche de triple-hybride. De plus, étant donné la forte conservation de séquence de l'hélice  $\alpha 3$  au sein de la famille de protéines L7Ae, nous avons vérifié la capacité de Rsa1p à interagir avec la protéine Nhp2p.

### 2.2.1. *In cellulo*, les résidus W245 et R249 de Rsa1p sont requis pour son recrutement sur les snoRNA

L'impact des mutations du domaine yPEP sur l'interaction avec Snu13p a été vérifié par les approches de double-hybride, de pull-down et de dichroïsme circulaire (Rothé, Back *et al.*, Publication n°2). La technique de retard sur gel nous a également permis de confirmer qu'un défaut d'interaction entre les deux protéines abolit totalement le recrutement de Rsa1p sur les complexes snoRNA-Snu13p (Rothé, Back *et al.*, Publication n°2). Dans le but de confirmer cette donnée *in cellulo*, nous avons mis à profit l'approche de triple-hybride initialement utilisée pour l'identification de la protéine Rsa1p (Boulon *et al.*, 2008). Nous avons ainsi testé l'interaction du fragment Rsa1p<sub>230-381</sub> WT ou muté, avec le motif B/C du snoRNA U3 (**Fig. 56A**). En comparaison avec la protéine WT, la substitution en alanine des résidus W245 et R249 de Rsa1p abolit totalement la croissance des cellules sur milieu sélectif et révèle un défaut de recrutement du facteur d'assemblage sur l'ARN. Ce résultat confirme donc que l'interaction observée par triple-hybride entre Rsa1p et l'ARN U3B/C s'effectue grâce à l'effet médiateur de la protéine Snu13p endogène. De plus, il illustre l'importance cruciale de ces résidus pour l'incorporation de Rsa1p au sein d'une pré-snoRNP. En effet, même dans un

**A****B****C**

**Figure 56 : *In cellulo*, le domaine PEP de Rsa1p/NUFIP constitue une interface clé entre la machinerie d'assemblage et les pré-RNP. (A)** Expériences de triple-hybride réalisées entre les protéines Rsa1p et Rsa1p<sub>230-381</sub> (alias N3C1), fusionnées au domaine Gal4-AD (pACT2) et l'ARN U3B/C-MS2 (pIIIAMS2). L'intensité des interactions a été évaluée en présence d'une concentration de 10 mM en 3-AT. **(B)** Expériences de triple-hybride *bridged* réalisées entre la protéine humaine PIH1 (pACT2) et les ARN U3B/C et U4/U6 (pIIIAMS2), dans une souche de levure  $\Delta$  *RSAl* exprimant les protéines NUFIP WT et  $\Delta$ PEP (p422). L'intensité des interactions a été évaluée en présence d'une concentration de 0,5 mM en 3-AT. **(C)** Schéma illustrant le phénomène de *bridge* observé par triple-hybride.



contexte cellulaire, leur simple mutation produit un effet totalement délétère, non compensé par la présence d'autres partenaires potentiels.

### **2.2.2. Chez l'homme, le recrutement de PIH1 sur les snoRNA est également dépendant de l'interaction entre NUFIP et 15.5K**

Afin d'évaluer l'impact d'une perte d'interaction entre Snu13p et Rsa1p sur le mécanisme d'assemblage des snoRNP, nous avons réalisé des expériences d'assemblage de pré-RNP *in vitro* à partir de protéines recombinantes WT ou mutées (Rothé, Back *et al.*, Publication n°2). Cette démarche nous a permis de montrer que le recrutement du complexe Pih1p-Tah1p sur les snoRNA est dépendant du domaine C-terminal de Rsa1p et de son interaction avec la protéine Snu13p. Dans le but de confirmer cette donnée *in cellulo* et afin d'élargir notre questionnement au système humain, nous avons testé le mode de recrutement de la protéine humaine PIH1 sur les ARN, par l'approche de triple-hybride *bridged* (Fig. 56B). Pour cela, nous avons utilisé deux ARN appâts : l'ARN U3B/C et un ARN de fusion mimant le duplex U4/U6. Comme nous l'avons déjà constaté pour les protéines Fibrillarine et PRP31, l'interaction de PIH1 avec l'ARN U3B/C est dépendante de la protéine NUFIP. En effet, PIH1 n'interagit avec l'ARN appât, que lorsque le facteur d'assemblage est co-exprimé dans les cellules. De plus, cette interaction est perdue lorsque NUFIP est tronquée de son domaine PEP, suggérant que ce phénomène repose sur la formation d'un complexe entre l'ARN U3B/C, la protéine Snu13p endogène et les facteurs NUFIP et PIH1 (Fig. 56C). De plus, des résultats identiques ont été obtenus à partir de l'ARN U4/U6, suggérant que le recrutement des facteurs d'assemblage sur les ARN dépend principalement des spécificités d'interaction de la protéine Snu13p/15.5K.

L'ensemble de ces données suggère donc que l'action ciblée de la machinerie d'assemblage dans la biogenèse des snoRNP à boîtes C/D et de la snRNP U4, pourrait reposer sur un couplage physique et conservé entre les protéines Snu13p/15.5K, Rsa1p/NUFIP et Pih1p/PIH1. Une perspective intéressante serait donc d'étudier plus en détail l'interaction entre Rsa1p/NUFIP et Pih1p/PIH1. Cette dernière pourrait en effet constituer une interface clé entre les différents composants de la machinerie d'assemblage et la caractérisation de variants pourrait constituer un outil intéressant dans le but d'étudier les processus d'assemblage *in vivo*.

### 2.2.3. Chez la levure, Rsa1p n'interagit pas avec la protéine Nhp2p

Un point important de notre travail était également de vérifier l'implication de Rsa1p dans la biogenèse des snoRNP à boîtes H/ACA. Chez l'homme, NUFIP interagit en effet avec la protéine NHP2 et se trouve associée *in vivo* à des snoRNA à boîtes H/ACA (Boulon *et al.*, 2008). Chez la levure en revanche, cette possibilité n'avait pas encore été testée.

Dans l'article *Formation of electrostatic and hydrophobic interactions between alpha helices of the Rsa1p assembly factor and the Snu13p snoRNP protein is required for efficient snoRNP assembly*, nous avons montré que l'interaction entre Rsa1p et Snu13p repose sur certains acides aminés conservés au niveau de l'hélice  $\alpha 3$  de Snu13p. En effet, les résidus E72 et D73 sont impliqués dans un réseau d'interactions électrostatiques avec deux arginines du domaine yPEP, tandis que le résidu L69 contribue à créer un environnement hydrophobe stabilisateur. De façon intéressante, la séquence de l'hélice  $\alpha 3$  est relativement bien conservée dans Nhp2p, les résidus E et D sont strictement conservés et la leucine est remplacée par une valine (**Fig. 57A**). Cependant, en dépit de la bonne conservation de ces résidus clés, aucune interaction entre Rsa1p et Nhp2p n'a été détectée par double-hybride (**Fig. 57B**). Afin de parer à un éventuel problème de faux-négatif lié à l'expression de Nhp2p sous forme de protéine de fusion, nous avons décidé de tester l'importance de l'hélice  $\alpha 3$  de Nhp2p sur la croissance cellulaire. Nous avons donc réalisé au niveau de Nhp2p, la mutation E<sub>95</sub>D<sub>96</sub>H<sub>97</sub>/AAA, équivalente à la mutation E<sub>72</sub>D<sub>73</sub>H<sub>74</sub>/AAA de Snu13p (**Fig. 57A**). La fonctionnalité de ce variant a été testée *in vivo* dans une souche *Gal::NHP2*, cultivée à 30°C en milieu non permissif (**Fig. 57C**). Alors que la déplétion de Nhp2p conduit à un phénotype létal, l'expression ectopique de la protéine Nhp2p WT à partir du plasmide pG1 permet de restaurer la croissance des cellules (G=3,21 h). Par comparaison, l'expression du variant E<sub>72</sub>D<sub>73</sub>H<sub>74</sub>/AAA est associée à un niveau de croissance normal (G=3,23 h), indiquant que les déterminants présumés d'interaction avec Rsa1p ne sont pas requis pour la fonctionnalité de la protéine Nhp2p. En complément du résultat obtenu par double-hybride, cette observation suggère donc fortement que la protéine Rsa1p n'interagit pas avec Nhp2p chez la levure.

Afin de comprendre les bases moléculaires de cette sélectivité, nous avons poussé nos investigations en nous intéressant aux résidus non conservés entre Snu13p et Nhp2p. Nous avons ainsi opté pour une approche plus systématique en inversant les hélices  $\alpha 3$  de Snu13p et Nhp2p (PVLCEDHS  $\Leftrightarrow$  PLLCEDKN) (**Fig. 57A**). De façon intéressante, l'hélice  $\alpha 3$  de

**A**

Hélice  $\alpha 3$

<b>Snu13p wt (65-75)</b>	-LHLPLLCEDKN-
<b>Nhp2p wt (88-98)</b>	-SHIPVLCEDHS-
<b>Nhp2p E<sub>95</sub>D<sub>96</sub>H<sub>97</sub>/AAA</b>	-SHIPVLC <b>AAAS</b> -
<b>Nhp2p <math>\alpha 3</math>-Snu13p</b>	-SHIP <b>LL</b> CEDKN-

<b>Nhp2p wt (88-98)</b>	-SHIPVLCEDHS-
<b>Snu13p wt (65-75)</b>	-LHLPLLCEDKN-
<b>Snu13p E<sub>72</sub>/A</b>	-LHLPLLC <b>ADKN</b> -
<b>Snu13p <math>\alpha 3</math>-Nhp2p</b>	-LHLP <b>VLC</b> ED <b>HS</b> -

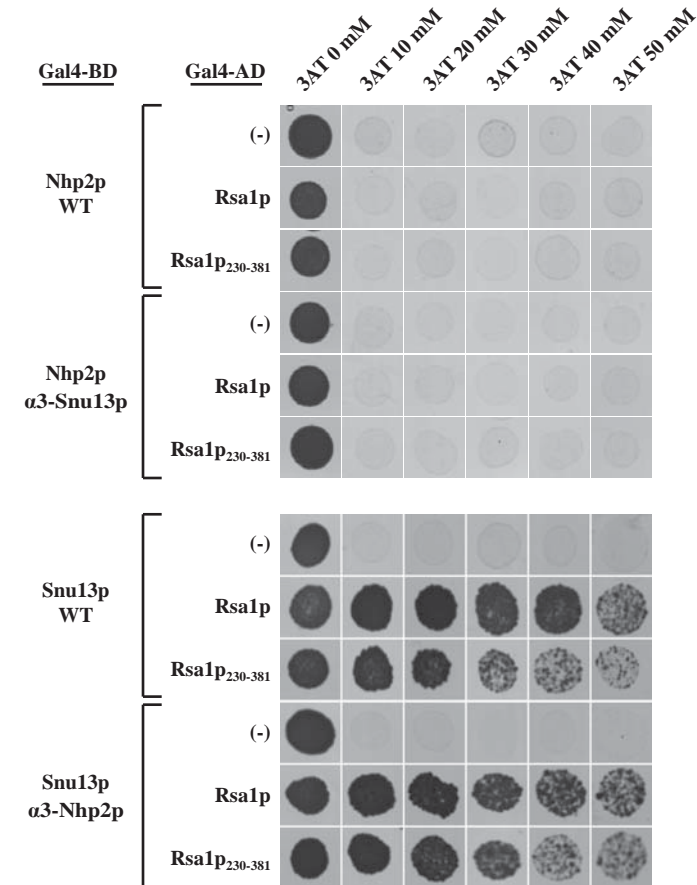
**C**

Souche GAL:: <i>NHP2</i>	G (heures)	Défaut de croissance (%WT)
pG1:: <i>NHP2</i> -WT	3,21	
pG1:: <i>NHP2</i> -E <sub>95</sub> D <sub>96</sub> H <sub>97</sub> /AAA	3,23	+0,62
pG1:: <i>NHP2</i> - $\alpha 3$ -Snu13p	3	-6,5

Souche GAL:: <i>SNU13</i>	G (heures)	Défaut de croissance (%WT)
pG1:: <i>SNU13</i> -WT	3,26 $\pm$ 0,04	
pG1:: <i>SNU13</i> -E <sub>72</sub> /A	3,58 $\pm$ 0,09	+9,8
pG1:: <i>SNU13</i> - $\alpha 3$ -Nhp2p	3,27	+0,3

**B**



**Figure 57 : Chez la levure *Saccharomyces cerevisiae*, Rsa1p n'interagit pas avec la protéine Nhp2p.** (A) Mutations réalisées au niveau de l'hélice  $\alpha 3$  de Snu13p et Nhp2p (annotées en rouge). Les mutations Nhp2p- $\alpha 3$ Snu13p et Snu13p- $\alpha 3$ Nhp2p correspondent à une inversion de l'hélice  $\alpha 3$  entre les deux protéines. (B) Expériences de double-hybride réalisées entre les protéines Nhp2p et Snu13p WT et chimériques, fusionnées au domaine Gal4-BD (pGBKT7), et les protéines Rsa1p et Rsa1p<sub>230-381</sub>, fusionnées au domaine Gal4-AD (pACT2). L'intensité des interactions a été évaluée par ajout de 3-AT en concentration croissante. (C) Temps de génération des souches de levure GAL::*NHP2* et GAL::*SNU13* complémentées respectivement par les plasmides pG1::*NHP2* WT/mut et pG1::*SNU13* WT/mut et cultivées à 30°C en milieu YPD. Les valeurs ne présentant pas d'écarts-types correspondent à des expériences n'ayant pas fait l'objet d'un duplicat.

Snu13p n'est pas suffisante pour conférer à Nhp2p la propriété d'interagir avec Rsa1p, et inversement, la capacité de Snu13p à interagir avec Rsa1p est strictement conservée en présence de l'hélice  $\alpha 3$  de Nhp2p (**Fig. 57B**). D'autre part, aucun effet notoire sur la croissance n'a été observé suite à l'expression de ces protéines chimères dans les cellules de levure (**Fig. 57C**). Ces résultats révèlent donc deux points importants. Tout d'abord, ils confirment la solidité du modèle d'interaction établi entre Snu13p et yPEP (Rothé, Back *et al.*, Publication n°2), puisqu'en comparaison avec le variant E<sub>72</sub>/A, la mutation de résidus non impliqués dans notre modèle d'interaction, ne produit aucun effet délétère. D'autre part, ils indiquent que l'absence d'interaction entre Rsa1p et Nhp2p n'est pas liée à la simple séquence de l'hélice  $\alpha 3$ , mais que d'autres paramètres d'influence doivent probablement exister.

D'autre part, en collaboration avec l'équipe d'E. Bertrand, les tests d'interaction par double-hybride ont été étendus à la protéine Pih1p et aux autres constituants des snoRNP à boîtes H/ACA. Les résultats obtenus n'ont révélé aucune interaction entre les facteurs d'assemblage Rsa1p et Pih1p, et les protéines Nhp2p, Gar1p, Nop10p et Cbf5p (Données non montrées). L'ensemble de ces données confirment donc que la machinerie d'assemblage des snoRNP à boîtes C/D, n'intervient vraisemblablement pas chez la levure, dans le processus de biogenèse des snoRNP à boîtes H/ACA.

### 3. Conclusion & Perspectives

#### 3.1. Bilan

Les résultats présentés dans cette partie ont permis d'identifier une interface d'interaction conservée entre les protéines Rsa1p/NUFIP et Snu13p/15.5K. Celle-ci implique les résidus WR--R du domaine PEP et le motif L--ED situé dans la partie C-terminale de l'hélice  $\alpha 3$  de la protéine Snu13p/15.5K. Ces acides aminés sont en effet impliqués dans la formation d'un réseau stable d'interactions électrostatiques, au sein duquel les deux arginines du domaine PEP interagissent avec les deux résidus acides de la protéine Snu13p/15.5K, tout en bénéficiant de l'environnement hydrophobe créé par les résidus tryptophane et leucine. De façon remarquable, la simple mutation des résidus E72 de Snu13p et R249 de Rsa1p, engagés chacun dans deux interactions électrostatiques, suffit à abolir l'interaction. *In vivo*, cette perte

d'interaction se matérialise par un défaut de croissance d'environ 10 à 20% et par une diminution spécifique du taux de certains snoRNA à boîtes C/D.

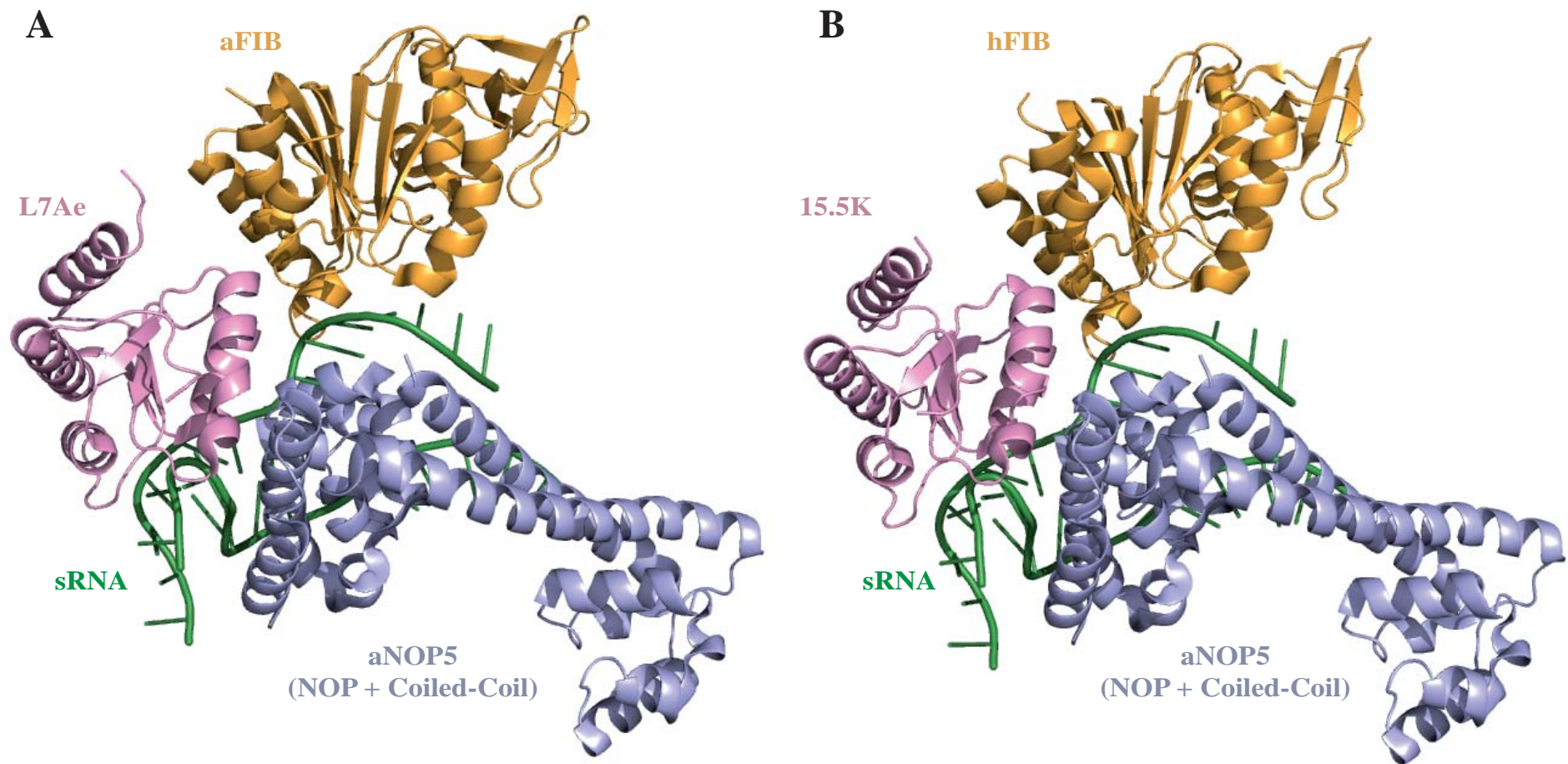
Afin de mieux comprendre les conséquences de ces mutations sur le processus d'assemblage, nous avons étudié la formation de pré-RNP *in vitro* et *in cellulo*. Les expériences réalisées nous ont notamment permis de montrer qu'une perte d'association entre les protéines Rsa1p/NUFIP et Snu13p/15.5K inhibe le recrutement du facteur Pih1p/PIH1 sur l'ARN. Cette donnée suggère donc qu'en plus de son rôle de plateforme pour l'assemblage des protéines Nop56p/NOP56 et Nop58p/NOP58, Rsa1p/NUFIP pourrait également faciliter le recrutement du complexe R2TP. Cette propriété repose vraisemblablement sur sa capacité à interagir à la fois avec Snu13p/15.5K et avec Pih1p/PIH1. Ainsi, le couplage physique entre ces trois protéines pourrait constituer un axe majeur du système d'assemblage, en permettant le ciblage spécifique des pré-particules ARN-Snu13p/15.5K. Ce résultat constitue donc une avancée importante, puisqu'il met en évidence pour la première fois un lien fonctionnel entre les deux sous-unités présumées de la machinerie d'assemblage : Rsa1p/NUFIP et le complexe R2TP.

Outre le recrutement de la protéine Pih1p/PIH1, il est fortement envisageable que la perte d'interaction entre Rsa1p/NUFIP et Snu13p/15.5K conduise également à un défaut d'assemblage des protéines Nop56p/NOP56 et Nop58p/NOP58. Le système d'assemblage en extrait de levure, mis au point sur la snoRNP U14 et présenté dans le chapitre précédent, pourra notamment être utilisé pour tester l'impact des mutations générées sur les protéines Snu13p et Rsa1p. A ce titre, des travaux similaires ont déjà abordé cette question au sujet des protéines humaines 15.5K et NUFIP (McKeegan *et al.*, 2009; Schultz *et al.*, 2006b). A l'origine de ce travail, la caractérisation de résidus conservés en surface de la protéine 15.5K avait permis de mettre en évidence l'importance du motif E<sub>74</sub>D<sub>75</sub>K<sub>76</sub>N<sub>77</sub> pour l'assemblage des snoRNP à boîtes C/D et de la snRNP U4 en extrait cellulaire (Schultz *et al.*, 2006b). Une explication moléculaire à cette observation a ensuite été apportée par l'équipe de N.J. Watkins, puisque des expériences de pull-down ont révélé que la mutation de ces résidus affecte spécifiquement l'interaction de la protéine 15.5K avec le facteur d'assemblage NUFIP (McKeegan *et al.*, 2009).

### 3.2. Eléments de réflexion sur la structure des snoRNP à boîtes C/D

Les résultats obtenus par l'équipe de R. Lührmann, à partir du variant E<sub>74</sub>D<sub>75</sub>K<sub>76</sub>N<sub>77</sub>/RKMT de la protéine 15.5K, pourraient résumer la dualité du rôle joué par le facteur Rsa1p/NUFIP. En effet, cette mutation inhibe le recrutement des protéines NOP56, NOP58, Fibrillarine, TIP48 et TIP49 sur la pré-snoRNP U14, ainsi que le recrutement des protéines PRP31, PRP3, PRP4 et CYPH sur la pré-snRNP U4 (Schultz *et al.*, 2006b). En analysant ces résultats *a posteriori*, il est donc tentant de penser que la mutation du motif EDKN, abolisse l'intervention de NUFIP, avec pour double conséquence un défaut d'assemblage des RNP et un défaut de recrutement des autres facteurs d'assemblages (*i.e.* : le complexe R2TP). Cependant, une telle interprétation soulève quelques points d'incertitudes. En effet, malgré les données de triple-hybride *bridged*, le rôle de Rsa1p/NUFIP pour l'incorporation de la Fibrillarine au sein des snoRNP n'a jamais été clairement démontré. Chez l'homme, seul le cas des protéines NOP56 et NOP58 a été rapporté (McKeegan *et al.*, 2007), tandis que chez la levure, nos expériences ont montré que l'assemblage de la protéine Nop1p sur le snoRNA U14 est totalement indépendant de Rsa1p (cf. Résultats, Chapitre 1, 2.2.2). D'autre part, l'impact de la mutation E<sub>74</sub>D<sub>75</sub>K<sub>76</sub>N<sub>77</sub>/RKMT sur l'assemblage de la protéine PRP31 n'a pas été observé en extrait cellulaire, mais à partir de protéines recombinantes (Schultz *et al.*, 2006b). Dans un tel contexte, il est donc évident que l'effet de la mutation n'est pas seulement lié à l'intervention de NUFIP et que le motif EDKN pourrait remplir plusieurs fonctions.

Afin d'apporter des éléments de réponse, nous avons donc cherché à savoir si la mutation des résidus E<sub>74</sub>D<sub>75</sub>K<sub>76</sub>N<sub>77</sub> de la protéine 15.5K pouvait avoir d'autres conséquences qu'un simple défaut d'interaction avec NUFIP. Une possibilité était notamment que les résidus touchés puissent être impliqués dans des contacts au sein de la RNP mature et que leur mutation influe également sur l'incorporation de la Fibrillarine. Dans le but de vérifier cette hypothèse, nous avons mis à profit les récentes avancées structurales réalisées dans le cadre de l'étude des sRNP à boîtes C/D d'archées (cf. Introduction, Chapitre 2, 3.1.1.3). Sur la base de la structure tridimensionnelle de la sRNP de *Sulfolobus solfataricus* (Lin *et al.*, 2011) et en utilisant les structures cristallographiques disponibles pour les protéines humaines 15.5K (Vidovic *et al.*, 2000) et Fibrillarine (Wu ; non publié), nous avons construit un modèle hybride de RNP (**Fig. 58**). Pour ce faire nous avons conservé une seule des deux particules cristallisée et nous avons remplacé, à l'aide du logiciel PyMol, les protéines L7Ae et Fibrillarine par leurs homologues humaines. Ce remplacement moléculaire a été réalisé sur la base de l'homologie de structure



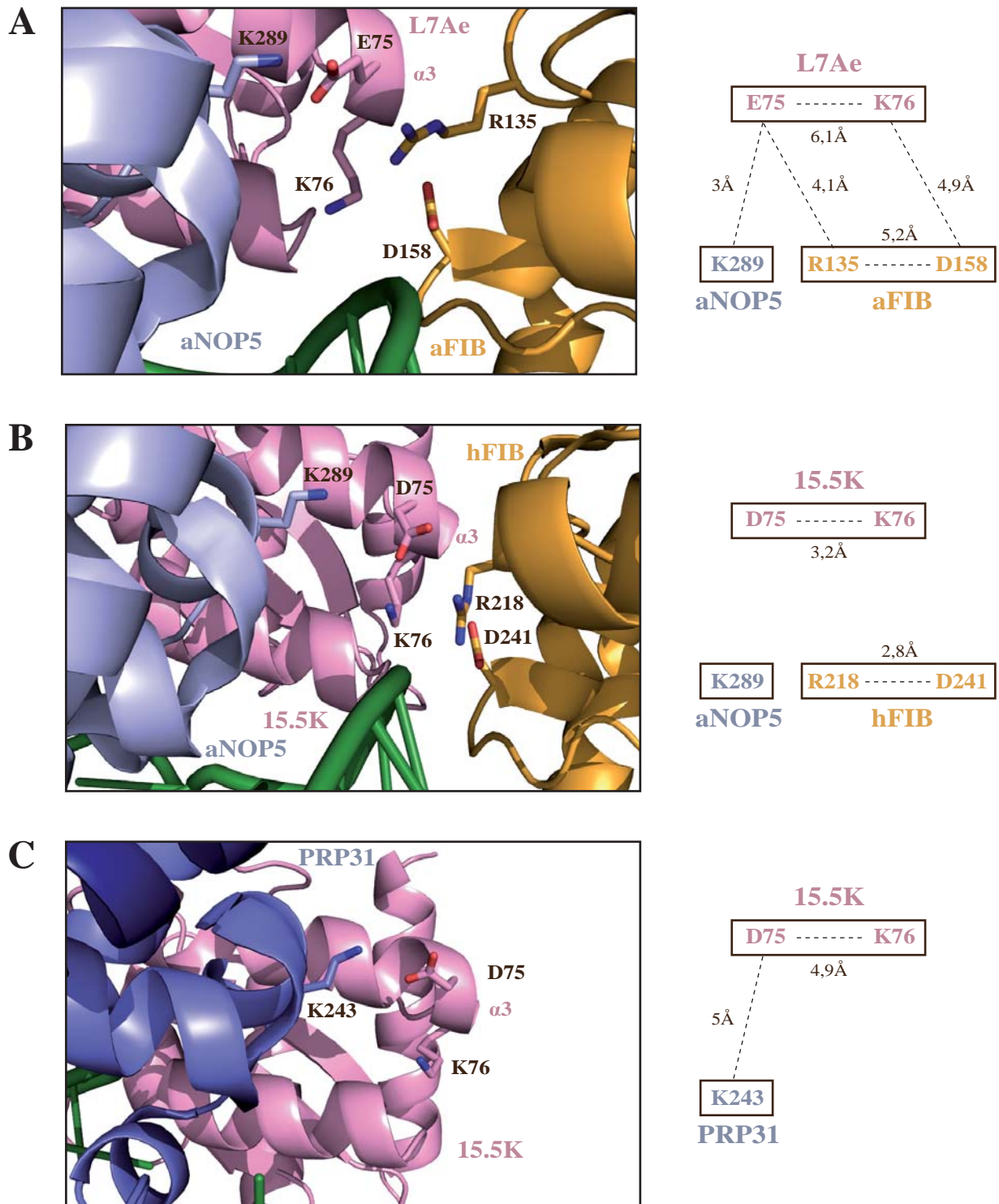
**Figure 58 : Modèle structural hybride entre une sRNP à boîtes C/D d'archées et les protéines humaines 15.5K et Fibrillarine.** (A) Représentation partielle de la structure tridimensionnelle de la sRNP à boîtes C/D de *Sulfolobus solfataricus* (Lin *et al.*, 2011; PDB: 3PLA). Un seul couple de boîtes C/D est représenté. Seuls les domaines NOP et Coiled-coil de NOP5, conservés dans les protéines NOP56 et NOP58 eucaryotes, sont figurés. (B) Remplacement moléculaire des protéines L7Ae et aFIB par leurs homologues humains 15.5K et hFIB, au sein de la sRNP à boîtes C/D. Le remplacement a été réalisé à l'aide du logiciel PyMol, en utilisant les structures cristallographiques des protéines humaines 15.5K (Vidovic *et al.*, 2000; PDB : 1E7K) et Fibrillarine (Wu *et al.*, non publié; PDB : 2IPX). Les nomenclatures aFIB et hFIB sont employées pour distinguer la protéine d'archées (aFIB) de son homologue humaine (hFIB).

entre les protéines, sans étape d'affinement par minimisation des énergies. D'autre part, seule la région (135-372) de la protéine NOP5, fortement conservée dans les protéines NOP56, NOP58 et PRP31 (**Fig. 30**), a été conservé dans le modèle.

Une fois établi, ce modèle hybride nous a permis d'analyser l'environnement potentiel du motif EDKN au sein de la RNP, tout en établissant des comparaisons avec les structures de la sRNP native (Lin *et al.*, 2011) et du complexe U4-15.5K-PRP31 humain (Liu *et al.*, 2007b) (**Fig. 59**). Dans la protéine L7Ae de *S. solfataricus*, le motif EDKN est remplacé par une séquence DEKK (**Fig. 60A**). L'analyse détaillée de la structure de la sRNP d'archées révèle que les deux résidus centraux de cette séquence, E75 et K76, sont localisés dans un environnement électrostatiquement riche, situé à l'interface avec les protéines NOP5 et Fibrillarine (**Fig. 59A**). En effet, les chaînes latérales des résidus chargés R135 et D158 de la Fibrillarine, et K289 de NOP5, sont orientées en direction des acides aminés E75 et K76 de L7Ae. Afin d'évaluer les interactions potentielles entre ces résidus, nous avons relevé les distances entre les groupements chargés les plus proches. La distance de 3Å mesurée entre l'acide glutamique 75 de L7Ae et la lysine 289 de NOP5 indique l'existence d'une interaction électrostatique entre ces deux résidus. Parmi les autres combinaisons possibles, seule la distance entre l'acide glutamique 75 de L7Ae et l'arginine 135 de la Fibrillarine, proche du seuil de 4Å, semble compatible avec la formation d'une interaction stable. Ces observations indiquent donc, qu'au sein de la sRNP d'archées, l'extrémité terminale de l'hélice  $\alpha 3$  de L7Ae, est prise en sandwich entre les protéines NOP5 et Fibrillarine, avec lesquelles elle pourrait établir plusieurs contacts de nature électrostatique.

De façon tout à fait intéressante, la même analyse, réalisée à partir de la RNP hybride archée/homme a révélé une configuration relativement proche (**Fig. 59B**). Les deux résidus centraux, D75 et K76, du motif EDKN de la protéine 15.5K sont également retrouvés à proximité du résidu K289 de NOP5. De plus, les résidus R135 et D158 de la Fibrillarine archée, sont strictement conservés dans la protéine humaine (*i.e.* : R218 et D241) (**Fig. 60B**). Au sein du modèle hybride, ces deux acides aminés, sont exposés au niveau de l'interface d'interaction avec la protéine 15.5K et pourraient également être impliqués dans des contacts de nature électrostatique. L'absence d'étape de minimisation des énergies lors de la construction du modèle, nous empêche cependant de réaliser une évaluation précise des distances inter-résidus et l'interface présentée dans la figure 60B doit être considérée à titre indicatif.





**Figure 59 : Localisation des résidus C-terminaux de l'hélice  $\alpha 3$  des protéines L7Ae et 15.5K au sein des RNP. (A)** Structure tridimensionnelle de l'interface d'interaction entre les protéines L7Ae, NOP5 et aFIB au sein la sRNP à boîtes C/D de *Sulfolobus solfataricus* (Lin *et al.*, 2011; PDB: 3PLA). **(B)** Structure tridimensionnelle de l'interface d'interaction entre les protéines 15.5K, NOP5 et hFIB au sein la RNP hybride construite par remplacement moléculaire et présentée dans la figure précédente. **(C)** Structure tridimensionnelle de l'interface d'interaction entre les protéines 15.5K et PRP31 au sein la snRNP U4 humaine (Liu *et al.*, 2007; PDB: 2OZB). Les distances inter-résidus issues de données expérimentales sont indiquées en Ångström (Å) sur la droite des panneaux.



Enfin, une vue similaire a été réalisée à partir de la structure du complexe humain U4-15.5K-PRP31 (Liu *et al.*, 2007b) (**Fig. 59C**). Cette particule ne contient aucun équivalent de la protéine Fibrillarine, en revanche la protéine PRP31 présente une forte homologie avec les protéines NOP56, NOP58 et NOP5 au niveau de ses domaines NOP et *Coiled-coil* (**Fig. 30**). De façon intéressante, le résidu K289 de NOP5 est conservé dans la protéine PRP31 humaine (*i.e.* : K243) et la fonction carbonyle de sa chaîne principale est engagée dans une liaison hydrogène avec la chaîne latérale du résidu N47 de l'hélice  $\alpha 2$  de la protéine 15.5K (Liu *et al.*, 2007b) (**Fig. 32**). Bien que séparés par une distance de 5Å, les résidus K243 de PRP31 et D75 de la protéine 15.5K sont orientés en direction l'un de l'autre et pourraient apporter une contribution électrostatique à l'interaction entre les deux protéines, une fois la snRNP U4 ou la di-snRNP U4/U6 complètement assemblées.

La comparaison de ces différents modèles soulève donc des éléments de réflexion nouveaux puisque l'extrémité C-terminale de l'hélice  $\alpha 3$  de la protéine eucaryote Snu13p/15.5K pourrait être dédiée à la fois à l'interaction avec le facteur Rsa1p/NUFIP et à la formation de contacts intermoléculaires au sein des RNP matures. Afin de s'assurer du bien fondé de cette hypothèse, nous avons cherché à savoir si les résidus mis à jour dans les différentes protéines, sont conservés au sein du *phylum* eucaryote. Pour chacun de ces trois partenaires, nous avons donc réalisé un alignement de séquence avec diverses protéines homologues eucaryotes (**Fig. 60**). Ainsi, nous avons constaté que les résidus D75 et K76 de la protéine 15.5K, les résidus R218 et D241 de la Fibrillarine, ainsi que la lysine du domaine NOP sont strictement conservés dans toutes les séquences homologues testées. Une seule exception a été constatée, elle concerne la protéine PRP31 de *S. cerevisiae*, dans laquelle le résidu lysine est remplacé par une arginine, un acide aminé présentant néanmoins des propriétés électrostatiques similaires. Ces observations constituent donc un argument supplémentaire quant à l'intérêt fonctionnel de ces résidus, qui pourraient avoir hérité d'un rôle structural important pour la stabilité des RNP à boîtes C/D et de la snRNP U4 eucaryote.

Ces informations nous permettent donc de poser un regard nouveau sur les résultats obtenus par l'équipe de R. Lührmann à partir du variant E<sub>74</sub>D<sub>75</sub>K<sub>76</sub>N<sub>77</sub>/RKMT. Etant donné la possible implication des résidus D75 et K76 dans la formation de contacts intermoléculaires avec les protéines à domaine NOP et avec la Fibrillarine, il est envisageable que leur remplacement par des acides aminés présentent des propriétés radicalement différentes conduise à des effets délétères majeurs pour la formation des RNP. Ainsi, cette mutation qui affecte le recrutement

de la protéine NUFIP, pourrait également déstabiliser la particule mature et conduire, dans les conditions expérimentales utilisées, à une diminution drastique du taux d'ARN co-immunosélectionné par les différents partenaires protéiques.

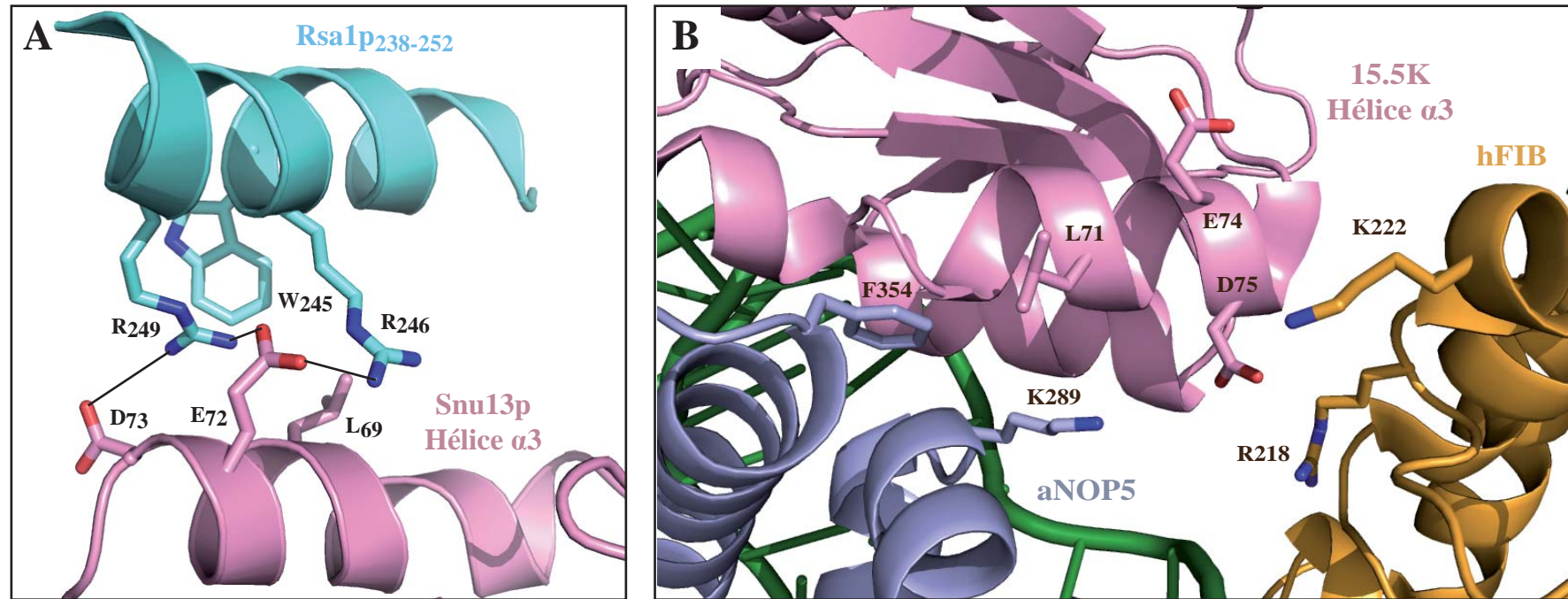
### 3.3. Rsa1p/NUFIP et stratégie de mimétisme moléculaire ?

Dans l'article *Formation of electrostatic and hydrophobic interactions between alpha helices of the Rsa1p assembly factor and the Snu13p snoRNP protein is required for efficient snoRNP assembly*, nous avons montré que le recrutement du facteur d'assemblage Rsa1p/NUFIP repose sur un motif conservé L--ED, situé au niveau de l'extrémité C-terminale de l'hélice  $\alpha 3$  de certaines protéines de la famille L7Ae. De façon intéressante, l'exploitation des données structurales acquises chez les archées suggère fortement que le résidu D de ce motif, dont le caractère acide (D ou E) est strictement conservé dans les protéines L7Ae d'archées, est également mis à profit au sein des RNP matures. En effet, ce dernier pourrait établir des contacts de nature électrostatique avec des résidus lysine et arginine, respectivement conservés au sein des protéines à domaine NOP et de la Fibrillarine.

De façon intéressante, l'extrémité C-terminale de l'hélice  $\alpha 3$  de la protéine Snu13p/15.5K semble donc constituer une zone de chevauchement entre le site de fixation du facteur Rsa1p/NUFIP et celui dédié à l'interaction avec les autres protéines constitutives de la RNP mature. Une double interrogation qui se pose désormais est de savoir comment ce site est occupé au cours de la biogenèse des RNP et quel est l'intérêt fonctionnel d'un tel recouvrement ?

Afin d'affiner notre analyse, nous avons entrepris de comparer l'environnement proximal du motif L--ED entre le modèle de la RNP hybride et celui établi pour l'interface d'interaction entre Rsa1p et Snu13p (**Fig. 61**). Cette démarche nous a permis de réaliser un constat marquant, puisqu'il existe une réelle analogie contextuelle entre ces modèles d'interaction. Tous deux semblent en effet, reposer sur une double composante électrostatique et hydrophobe :

- (i) Au sein de l'interface entre Rsa1p et Snu13p les résidus E72 et D73 sont impliqués dans un réseau d'interaction électrostatique avec deux arginines du domaine PEP



**Figure 61 : Comparaison de l'environnement du motif L--ED de Snu13p/15.5K au sein des interfaces d'interaction Rsa1p-Snu13p et 15.5K-FIB-NOP. (A) Modèle tridimensionnel de l'interface d'interaction entre les protéines Snu13p et Rsa1p de la levure *S. cerevisiae* (Rothé *et al.*, Publication n°2). Les traits en pointillés représentent les interactions électrostatiques. (B) Modèle tridimensionnel de l'interface d'interaction entre les protéines 15.5K, NOP5 et hFIB au sein la RNP hybride construite par remplacement moléculaire et présentée dans la figure 59.**

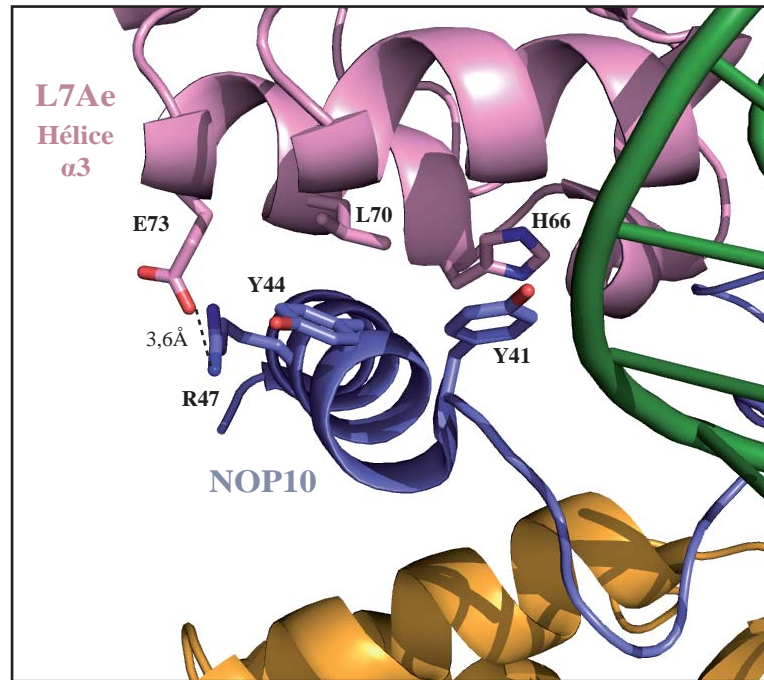
(**Fig. 61A**), tandis qu'au sein de la RNP hybride, les résidus E et D équivalents sont encadrés par les résidus basiques K289 de NOP5 et R218 de la Fibrillarine (**Fig. 61B**). De plus, le résidu K222 de la Fibrillarine humaine est également retrouvé dans le même espace. Cette lysine, absente chez l'archée *S. solfataricus*, mais strictement conservée dans les protéines Fibrillarine eucaryotes (**Fig. 60**), constitue un candidat intéressant, qui pourrait apporter une contribution supplémentaire pour la mise en place d'un réseau d'interactions électrostatiques entre ces protéines.

- (ii) Au sein de l'interface entre Rsa1p et Snu13p, nous avons également montré l'importance de la leucine du motif L<sub>69</sub>--E<sub>72</sub>D<sub>73</sub>. En établissant des interactions de Van der Waals avec un résidu tryptophane du domaine PEP, cette dernière pourrait en effet créer un environnement hydrophobe, contribuant à isoler les résidus chargés des molécules d'eau environnantes (**Fig. 61A**). De façon similaire, au sein de la RNP hybride le résidu L71 de la protéine 15.5K est situé à proximité de deux résidus de la protéine NOP5 présentant des propriétés hydrophobes (**Fig. 61B**) : la chaîne latérale aliphatique du résidu lysine K289 et la phénylalanine F355. De façon intéressante, même si cette dernière n'est pas strictement conservée chez les eucaryotes, il est important de noter que dans la majorité des cas elle est remplacée par un autre résidus hydrophobe tel une leucine ou une tyrosine (**Fig. 60**).

Cette comparaison révèle donc de grandes similitudes entre les deux modèles d'interactions et soulève une perspective de premier plan, puisqu'à l'image du facteur Naf1p/NAF1, le fonctionnement de Rsa1p/NUFIP pourrait également faire appel à la notion de mimétisme moléculaire. D'après nos observations, il est en effet envisageable que le mode d'interaction du domaine PEP avec l'hélice  $\alpha 3$  de Snu13p/15.5K mime tout ou partiellement la situation de cette même hélice au sein de la RNP mature.

### 3.4. Analogie avec le système H/ACA

A titre de comparaison, nous avons également utilisé les structures tridimensionnelles obtenues chez l'archée, afin de vérifier les contacts établis par les résidus de l'hélice  $\alpha 3$  de la protéine L7Ae au sein des sRNP à boîtes H/ACA (**Fig. 62**). Chez *Pyrococcus furiosus*, le



**Figure 62 : Contacts intermoléculaires impliquant l'extrémité C-terminale de l'hélice  $\alpha_3$  de la protéine L7Ae au sein des sRNP à boîtes H/ACA.** Cette représentation a été réalisée à l'aide du logiciel PyMol, en utilisant la structure cristallographique de la sRNP à boîtes H/ACA de l'archée *Pyrococcus furiosus* (d'après Duan *et al.*, 2009; PDB : 3HAY). L'agrandissement correspond à la zone de contact entre l'hélice  $\alpha_3$  de L7Ae et l'hélice unique de NOP10.

résidu E73, équivalent aux résidus D75 de la protéine humaine 15.5K, établit une interaction électrostatique avec le résidu R47 de NOP10, tandis que les résidus H66 et L70 de L7Ae font respectivement l'objet d'un empilement hydrophobe avec les résidus Y41 et Y44 de NOP10 (Duan *et al.*, 2009). L'hélice  $\alpha 3$  de L7Ae contribue donc également à l'architecture des particules à boîtes H/ACA par le biais de contacts électrostatiques et hydrophobes établis au sein de la sRNP. Les déterminants impliqués dans ces interactions ne sont cependant pas strictement identiques, puisque l'interaction avec NOP10 semble impliquer le motif H---L--E.

Au laboratoire, J.B. Fourman et M. Blaud ont étudié l'importance des contacts établis entre L7Ae et NOP10 au sein de la sRNP à boîte H/ACA de *Pyrococcus abyssi* (Fourmann, Blaud *et al.*, Résultats non publiés). Ces travaux ont démontré que la mutation des résidus HLE de L7Ae ou des deux tyrosines de NOP10 n'inhibe pas l'assemblage des sRNP *in vitro*. En revanche, ces mutations conduisent à une sévère diminution de l'activité de pseudouridylation des particules. L'utilisation d'approches d'empreinte et de spectrophotométrie a permis de mettre en évidence que les contacts entre L7Ae et NOP10 exercent une influence capitale sur la structuration de l'ARN guide et sur le placement de l'ARN substrat dans la poche de pseudouridylation.

Dans le cas des sRNP à boîtes H/ACA, il semblerait donc que l'extrémité C-terminale de l'hélice  $\alpha 3$  de L7Ae soit impliquée dans des ajustements structuraux fins conduisant à l'acquisition d'une activité optimale. De façon intéressante, la localisation de cette hélice au sein des particules à boîtes C/D suggère qu'un tel rôle est également envisageable dans le cas de la 2'-O-méthylation. En effet, elle pourrait participer à l'ancrage des protéines NOP et Fibrillarine à l'apex du motif en *K-turn* afin d'assurer leur positionnement optimal. Ainsi, l'occupation transitoire de ce site par le facteur Rsa1p/NUFIP pourrait contribuer à maintenir la particule dans une conformation inactive jusqu'à la fin du processus d'assemblage. L'interface hélice  $\alpha 3$ -NOP-Fibrillarine pourrait alors fonctionner comme un « bouton de pression » permettant la transition finale des RNP vers une conformation active.

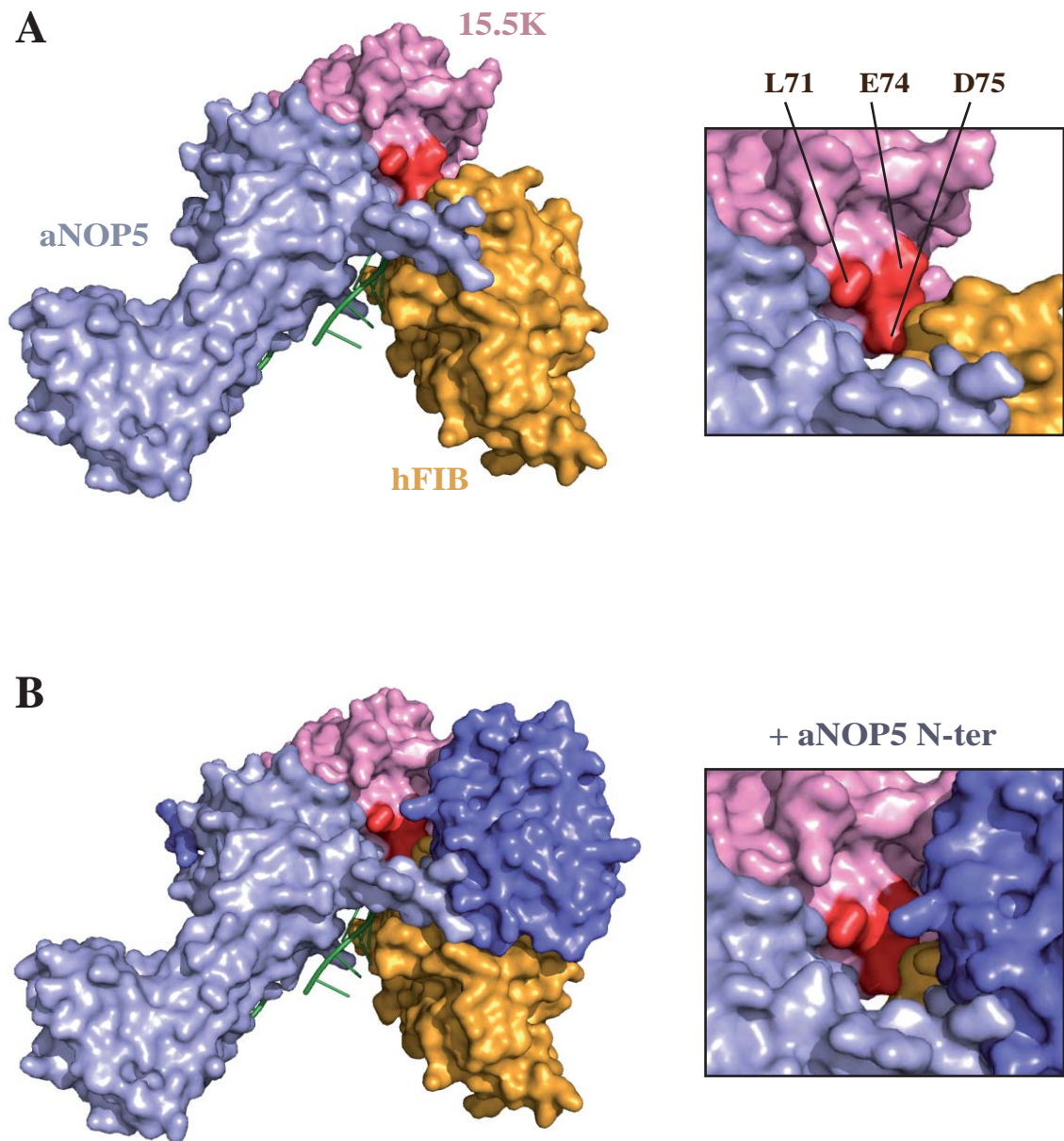
### 3.5. Proposition d'un modèle pour l'assemblage des snoRNP à boîtes C/D

Une implication majeure de cette observation serait de pouvoir reconstituer la chronologie des événements d'assemblage des snoRNP à boîtes C/D. En effet, il serait intéressant de savoir si



l'interaction de Rsa1p/NUFIP avec la protéine Snu13p/15.5K est exclusive ou si elle peut coexister avec celle des autres protéines. A ce stade, plusieurs cas de figure sont en effet envisageables : (i) soit Rsa1p/NUFIP et les deux protéines constitutives (Nop1p/Fibrillarine et Nop56p/NOP56 ou Nop58p/NOP58) peuvent interagir simultanément avec la protéine Snu13p/15.5K, (ii) soit la présence de Rsa1p/NUFIP empêche la fixation correcte des deux protéines constitutives sur Snu13p/15.5K, (iii) soit les protéines Rsa1p/NUFIP et Nop56p/NOP56 ou Nop58p/NOP58 peuvent co-interagir avec Snu13p/15.5K au détriment de la Fibrillarine, (iv) enfin, dans la situation inverse c'est la protéine à domaine NOP qui pourrait être exclue par la présence du facteur d'assemblage.

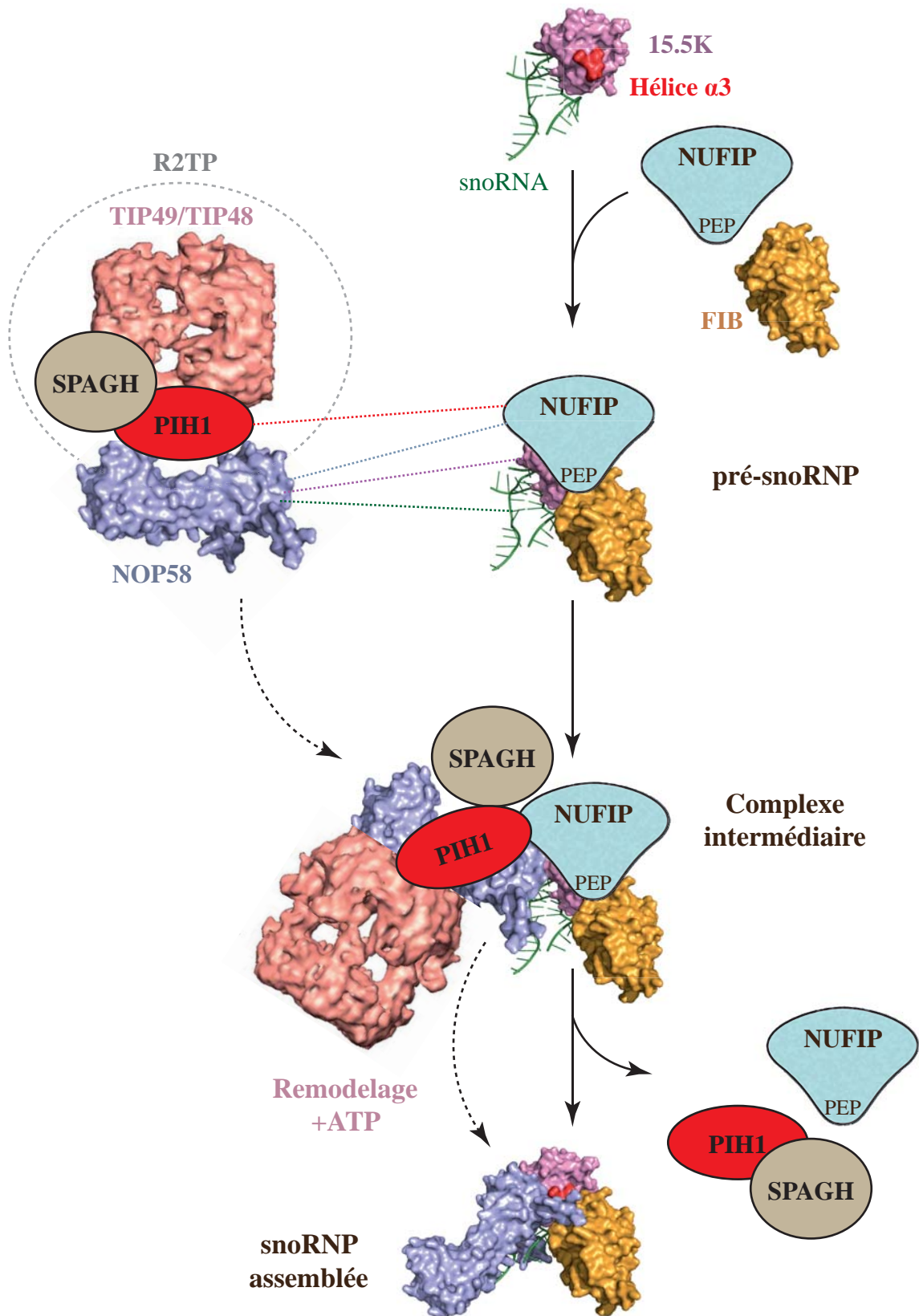
Afin d'apporter des premiers éléments de réponses, nous avons utilisé le modèle de RNP hybride archée/homme, en faisant apparaître les surfaces protéiques exposées au solvant (**Fig. 63**). Cette représentation permet en effet d'estimer l'accessibilité du motif L--ED pour une éventuelle interaction avec le facteur Rsa1p/NUFIP. Dans le contexte d'une RNP mature, comprenant à la fois la Fibrillarine et une protéine à domaine NOP, les résidus L, E et D sont situés au niveau d'une poche délimitée par les interfaces d'interaction entre les trois protéines constitutives et présentent une accessibilité relativement restreinte (**Fig. 63A**). De plus, un autre paramètre d'influence est la présence dans les protéines eucaryotes Nop56p/NOP56 et Nop58p/NOP58 d'un large domaine N-terminal d'environ 150 résidus. Bien que la séquence de cette région diverge de façon drastique dans la protéine NOP5, nous avons tenu à faire apparaître la partie N-terminale de la protéine d'archée, afin d'estimer la portion de l'espace potentiellement occupée par un tel domaine (**Fig. 63B**). Ainsi, comme dans la structure de la sRNP native (**Fig. 37**), ce domaine vient coiffer les protéines Fibrillarine et 15.5K. Il occupe donc un large espace situé dans le prolongement du domaine NOP et pourrait recouvrir tout ou partiellement la poche contenant le motif L--ED. En dépit de leur caractère spéculatif, ces observations suggèrent donc qu'au sein de la RNP mature, le site d'ancrage du facteur Rsa1p/NUFIP, pourrait être rendu inopérant, et ce pour deux raisons majeures : d'une part l'engagement des résidus L, E et D dans plusieurs interactions avec les protéines Fibrillarine et NOP56/58, d'autre part, un encombrement stérique important dans la zone d'accès au motif L--ED. De plus, l'inaccessibilité apparente du motif L--ED au sein de la RNP mature semble cohérente avec le fait que le facteur d'assemblage soit systématiquement exclu des particules matures.



**Figure 63 : Accessibilité des résidus L71, E74 et D75 de la protéine 15.5K au sein du modèle de RNP hybride sRNA-15.5K-hFIB-aNOP5. (A) Surface accessible au solvant du modèle de RNP hybride sRNA-15.5K-hFIB-aNOP5 (NOP + Coiled-coil). Les résidus L71, E74 et D75 de la protéine 15.5K sont représentés en rouge. (B) Représentation similaire incluant (en bleu foncé) le domaine N-terminal (3-124) de la protéine aNOP5. Ces représentations ont été réalisées à l'aide du logiciel PyMol.**

Parmi les autres possibilités énoncées, l'éventualité d'un conflit entre Rsa1p/NUFIP et la Fibrillarine est peu envisageable. Nos expériences d'assemblage en extrait de levure ont en effet démontré que le recrutement de la protéine Nop1p est totalement indépendant du facteur d'assemblage, l'absence de Rsa1p n'influençant ni positivement, ni négativement l'association de l'enzyme avec le snoRNA. De plus, il existe dans la littérature de nombreuses données indiquant que la protéine Nop1p/Fibrillarine est incorporée de façon précoce au sein des snoRNP (cf. Introduction, Chapitre 3, 3.5). Enfin, il a été montré chez la levure que le recrutement de la protéine Nop56p sur le snoRNA requiert la présence de la protéine Nop1p, suggérant que l'incorporation de l'enzyme au sein de la RNP pourrait être concomitante ou même préalable à celle de Nop56p (Lafontaine *et al.*, 2000). Concernant les protéines Nop56p/NOP56 et Nop58p/NOP58, la situation semble en revanche relativement différente puisque, chez la levure comme chez l'homme, leur assemblage sur le snoRNA est dépendant du facteur Rsa1p/NUFIP ((McKeegan *et al.*, 2007) ; Résultats & Discussion, Chapitre 1). D'autre part, la comparaison des interfaces d'interaction potentielles yPEP-Snu13p et NOP-15.5K au sein des deux modèles révèle beaucoup de similarités (**Fig. 61**). En effet, à la différence de la Fibrillarine, le domaine NOP pourrait interagir avec Snu13p/15.5K par l'intermédiaire de contacts électrostatiques, mais également par le biais de contacts hydrophobes, comme c'est le cas du domaine PEP. Ceci suggère donc, qu'en dépit de leur dépendance vis-à-vis de Rsa1p/NUFIP pour leur incorporation au sein des snoRNP, les protéines Nop56p/NOP56 et Nop58p/NOP58 pourraient être en conflit avec le facteur d'assemblage pour l'occupation de l'hélice  $\alpha 3$  de la protéine Snu13p/15.5K.

La synthèse de ces différents éléments de réflexions nous a donc conduit à proposer un modèle de travail pour l'assemblage des snoRNP à boîtes C/D dépendant de Rsa1p/NUFIP (**Fig. 64**). Ce modèle privilégie un recrutement indépendant du facteur d'assemblage et de la protéine Nop1p/Fibrillarine sur la pré-snoRNP, Rsa1p/NUFIP occupant l'hélice  $\alpha 3$  de la protéine Snu13p/15.5K indépendamment de la présence de l'enzyme. Le recrutement des protéines Nop56p/NOP56 ou Nop58p/NOP58 pourrait alors être influencé par plusieurs paramètres : (i) des contacts avec le facteur Rsa1p/NUFIP (cf., Résultats & Discussion, Chapitre 1, 2.2), (ii) des contacts du domaine NOP avec l'hélice II du motif en *K-turn* (**Fig. 32 et 38**), (iii) des contacts du domaine NOP avec l'hélice  $\alpha 2$  de la protéine Snu13p/15.5K (**Fig. 32 et 38**), (iv) l'interaction entre Rsa1p/NUFIP et la protéine Pih1p/PIH1 du complexe R2TP (Rothé, Back *et al.*, Publication n°2).



**Figure 64 : Proposition de modèle pour l'assemblage des snoRNP à boîtes CD.**  
 (cf. détail dans le texte : Résultats & Discussion, Chapitre 2, 3.4)

A ce stade du processus, la pré-snoRNP contiendrait donc l'ensemble de ses constituants, mais présenterait une conformation intermédiaire de faible stabilité, influencée par la présence des facteurs d'assemblage. De façon intéressante, de tels complexes contenant des précurseurs du snoRNA U3 ont été isolés par l'équipe de R. Lührmann (Watkins *et al.*, 2004). Comme nous l'avons abordé dans l'introduction, ces particules contenaient en effet les protéines constitutives de la snoRNP, ainsi que plusieurs facteurs impliqués dans la biogenèse des RNP, dont les facteurs TIP49 et TIP48 (cf. Introduction, Chapitre 3, 3.4.2). Etant donné la différence notoire de stabilité entre ces pré-snoRNP et les snoRNP matures isolées du nucléole, il avait été proposé que les deux hélicases de la famille AAA+ catalysent des événements de remodelage conduisant à la formation de particules stables (Watkins *et al.*, 2004). Compte-tenu de notre proposition pour l'architecture des pré-snoRNP, un tel processus pourrait notamment permettre de déplacer le facteur Rsa1p/NUFIP de son site d'interaction avec la protéine Snu13p/15.5K, afin de favoriser le positionnement correct des protéines NOP. La formation de l'interface d'interaction tripartite entre les protéines Snu13p/15.5K, Nop1p/Fibrillarine et Nop56p/NOP56 ou Nop58p/NOP58 pourrait alors constituer l'une des étapes finales du processus d'assemblage, conférant à la snoRNP un gain de stabilité et l'acquisition d'une conformation active.

Récemment, des travaux réalisés chez l'homme par l'équipe de N.J. Watkins ont apporté un éclairage nouveau sur ces événements de remodelage (McKeegan *et al.*, 2009). En utilisant des protéines recombinantes, ils ont en effet mis en évidence la capacité des hélicases TIP49 et TIP48 à ponter l'interaction des protéines NOP56 et NOP58 avec la protéine 15.5K *in vitro*. En revanche, les auteurs ont montré que l'ajout de la protéine NUFIP interfère négativement avec ce processus, en abolissant totalement ce phénomène. Cet effet de compétition, qui repose selon toute vraisemblance, sur la forte affinité de NUFIP pour la protéine 15.5K, révèle donc un antagonisme intéressant entre NUFIP et les deux hélicases. De plus, des tests d'interaction ont également révélé la capacité de TIP49 et TIP48 à interagir avec les protéines constitutives 15.5K, Fibrillarine, NOP56 et NOP58, ainsi qu'avec les facteurs NUFIP et PIH1 (McKeegan *et al.*, 2007; McKeegan *et al.*, 2009) (**Tableau 5**). De façon remarquable, ce large spectre d'interaction semble même régulé par le cycle d'hydrolyse d'ATP des hélicases, puisque leur association avec PIH1 est constatée seulement en absence d'ATP, tandis que leur interaction avec la protéine 15.5K nécessite la présence du nucléotide (McKeegan *et al.*, 2007; McKeegan *et al.*, 2009). Ces résultats semblent donc confirmer le rôle moteur joué par les deux hélicases dans les étapes d'assemblage des snoRNP à boîtes C/D. En effet, sous

l'effet de la fixation ou de l'hydrolyse d'ATP, ces dernières pourraient se dissocier du complexe R2TP et engager une interaction avec la protéine Snu13p/15.5K, afin de favoriser le positionnement correct des protéines NOP et de provoquer le départ des facteurs d'assemblage Rsa1p/NUFIP et Pih1p/PIH1.

Ce modèle, basé sur nos résultats, sur les données de la littérature et sur les nouvelles données de modélisation informatique, constitue donc un axe de réflexion majeur dans le cadre de nos perspectives de travail. Sa validation, nécessitera notamment de reproduire les approches *in silico* en intégrant des étapes d'affinements par minimisation des énergies. Sur la base d'un modèle informatique plus solide, la définition de nouveaux variants protéiques constituera ainsi un outil de choix pour de futures études *in vitro* et *in vivo*.

Concernant la protéine Snu13p/15.5K, il semble que le résidu E du motif EDKN constitue le pivot de l'interaction avec le facteur d'assemblage Rsa1p/NUFIP, tandis que les résidus D et K pourraient d'avantage contribuer aux contacts avec les autres protéines constitutives de la particule (**Fig. 59 et 61**). A ce titre, un résultat préliminaire, obtenu par l'équipe d'E. Bertrand a permis de montrer que les mutations individuelles D<sub>75</sub>/A et K<sub>76</sub>/A de la protéine humaine 15.5K n'ont pas d'effet sur l'interaction avec NUFIP par double-hybride (Rothé, Back *et al.*, Publication n°2). Cette donnée ouvre une perspective intéressante, puisqu'en mutant sélectivement certains résidus du motif EDKN, nous pourrions éventuellement dissocier les effets dus à une perte d'interaction avec Rsa1p, de ceux dus à une perte de contact avec les protéines NOP et Fibrillarine. De la même façon, la mutation des déterminants d'interaction potentiels, relevés au sein du domaine NOP et de la Fibrillarine, fournira deux angles d'approche supplémentaires afin d'étudier séparément l'assemblage de ces protéines sur la snoRNP.

Le système de reconstitution en extrait de la snoRNP U14 présenté dans le chapitre précédent, pourra notamment être utilisé afin d'évaluer, dans un contexte WT ou  $\Delta RSA1$ , l'impact des différentes mutations. Une fois leur intérêt établi, les variants des protéines Snu13p, Nop1p, Nop56p et Nop58p pourront être caractérisés *in vivo* en utilisant les souches de levures à expression conditionnelle GAL::*SNU13*, GAL::*NOPI*, GAL::*NOP56* et GAL::*NOP58*. L'effet cumulé de ces mutations et des variants déjà caractérisés pour la protéine Rsa1p, pourra également être évalué dans ces différentes souches après invalidation du gène *RSAL*. L'utilisation de ces systèmes génétiques pourra nous permettre d'étudier les conséquences des

différentes mutations sur la croissance cellulaire, sur l'accumulation des snoRNA, mais également sur la fonctionnalité des particules, en étudiant spécifiquement la maturation nucléolytique des ARNr et leur taux de méthylation. De plus, la possibilité de bloquer certaines étapes du processus d'assemblage pourrait offrir l'opportunité d'accumuler et de purifier des pré-snoRNP. Des approches d'immunosélection (cf. Résultats & Discussion, Chapitre 3), couplées à des techniques de séparation de complexes sur gradients de glycérol, pourront être utilisées à cet effet. La caractérisation de leur contenu protéique pourra être réalisée de façon ciblée par Western-blot ou de façon plus exhaustive par spectrométrie de masse, tandis que les nouvelles approches de séquençage à haut-débit pourront être utilisées afin d'identifier de façon systématique les composants ARN de ces pré-snoRNP.

Enfin, l'importance des hélicases dans ces phénomènes pourra également être étudiée. Dans un premier temps, l'approche de reconstitution de la snoRNP U14 en extrait cellulaire devra être reproduite en présence de différentes concentrations en ADP, en ATP ou en analogue non hydrolysable de l'ATP. Cette démarche nous permettra en effet d'estimer l'influence potentielle des activités ATPase sur la stabilité des particules immunosélectionnées. Par la suite, des variants affectant la fonctionnalité des protéines Rvb1p et Rvb2p, devront être utilisés afin de caractériser leur rôle précis dans ces processus. L'expression de tels variants dans des souches *GAL::RVB1* et *GAL::RVB2* nous permettra d'évaluer l'importance respective des deux hélicases pour l'assemblage en extrait de la particule U14 et pour le déroulement du processus de biogenèse des snoRNP *in vivo*. En utilisant des approches de purification et d'analyse par Western-blot, il sera également intéressant de vérifier les conséquences d'un dysfonctionnement des hélicases sur la composition des pré-snoRNP et en particulier sur la rétention du facteur Rsa1p.





## CHAPITRE 3

### Caractérisation fonctionnelle de la machinerie d'assemblage et mise en évidence du rôle joué par la protéine Hit1p

#### 1. Présentation des travaux

Dans le chapitre précédent, nous avons abordé la caractérisation physique et fonctionnelle de l'interaction entre Rsa1p/NUFIP et Snu13p/15.5K. Ce travail, nous a notamment conduit à poser les bases d'un nouveau modèle pour l'assemblage des snoRNP à boîtes C/D. La protéine Rsa1p/NUFIP pourrait en effet contribuer à la nucléation de la pré-RNP, en favorisant le recrutement des protéines à domaine NOP et du facteur Pih1p/PIH1 sur les pré-complexes snoRNA-Snu13p/15.5K. De plus, l'interaction de Rsa1p/NUFIP avec l'hélice  $\alpha 3$  de la protéine Snu13p/15.5K pourrait maintenir la particule dans une conformation intermédiaire, préfigurant les étapes finales de remodelage et la formation des RNP matures. Le facteur Rsa1p/NUFIP et le complexe R2TP constituent donc à ce jour, les deux principales entités impliquées dans le processus d'assemblage. Leur action coordonnée semble notamment reposer sur l'interaction entre Rsa1p/NUFIP et Pih1p/PIH1, qui pourrait constituer une interface clé pour le fonctionnement de la machinerie d'assemblage.

Afin d'analyser plus en détail la relation entre ces deux entités, nous avons entrepris la copurification et l'analyse exhaustive des protéines associées aux facteurs Rsa1p et Pih1p de levure. De façon surprenante, nous avons constaté que Rsa1p et le complexe R2TP ne font pas l'objet d'une association stable *in vivo*. En revanche, cette démarche nous a permis d'identifier la protéine Hit1p comme nouveau partenaire du facteur Rsa1p. Cette protéine, conservée chez les eucaryotes, interagit physiquement avec le facteur d'assemblage et s'est avérée requise pour maintenir son taux dans les cellules. Au cours de cette étude, nous avons également analysé l'association des facteurs d'assemblage avec les précurseurs du snoRNA U3. Ces travaux nous ont permis de mettre en évidence le recrutement de Rsa1p, Hit1p et des membres du complexe R2TP sur les mêmes pré-snoRNA U3, suggérant leur présence transitoire au sein des mêmes pré-snoRNP. D'autre part, la caractérisation phénotypique des souches de levures invalidées pour les différents gènes d'intérêt a révélé l'existence de liens fonctionnels forts entre les deux sous-unités de la machinerie d'assemblage, puisque le gène

*PIHI* s'est révélé synthétique létal avec les gènes *RSA1* et *HIT1*. Enfin, les résultats obtenus suggèrent une perspective nouvelle, puisqu'à la différence de Pih1p, les rôles de Rsa1p et Hit1p pourraient être dédiés spécifiquement à une sous-population de snoRNA à boîtes C/D.

## **2. Résultats**

L'ensemble de ce travail a abouti à la rédaction d'un article en préparation pour la revue *Molecular & Cellular Biology* et dans lequel je suis premier auteur. Des résultats d'analyse phénotypique, d'immunoprécipitation de la chromatine, de pull-down, de double-hybride et de cribles d'interactions génétiques sont également présentés en complément de l'article.

### **2.1. Publication n°3**

---

---

Publication n°3

**Functional link between proteins Rsa1 and Hit1  
during box C/D snoRNPs biogenesis in yeast**

Benjamin Rothé, Jean-Michel Saliou, Régis Back, Xavier Manival, Alain Van Dorssealer,  
Sarah Sanglier-Cianferani, Christiane Branlant and Bruno Charpentier

**En préparation pour *Molecular and Cellular Biology***

---

---



# Functional link between proteins Rsa1 and Hit1 during box C/D snoRNPs biogenesis in yeast

Benjamin Rothe<sup>1</sup>, Jean-Michel Saliou<sup>2</sup>, Régis Back<sup>1</sup>, Xavier Manival<sup>1</sup>, Alain Van Dorssealer<sup>2</sup>, Sarah Sanglier-Cianferani<sup>2</sup>, Christiane Branlant<sup>1</sup> and Bruno Charpentier<sup>1</sup>

AREMS<sup>1</sup> (ARN, RNP, structure-fonction-maturation, Enzymologie Moléculaire et Structurale), Nancy Université, Faculté de Médecine, 9, avenue de la forêt de Haye, BP 184, 54500 Vandoeuvre-Lès-Nancy, Cédex, France. LMBSO<sup>2</sup> (Laboratoire de Spectrométrie de Masse BioOrganique, Strasbourg, Université, IPHC, 25 rue Becquerel, 67087 Strasbourg, France.

<sup>§</sup>Address correspondence to: Bruno Charpentier, +33 383 68 55 08, Fax: +33 383 68 55 09, E-mail [bruno.charpentier@maem.uhp-nancy.fr](mailto:bruno.charpentier@maem.uhp-nancy.fr)

## ABSTRACT

**Biogenesis of box C/D small nucleolar ribonucleoprotein particles (snoRNPs) relies on conserved *trans*-acting factors, which are proposed to facilitate the assembly of the core proteins Snu13p, Nop58p, Nop56p, and Nop1p/Fibrillarin on box C/D small nucleolar RNAs (snoRNAs). In the yeast assembly machinery, protein Rsa1 acts as a platform for direct interaction with the RNA-binding core protein Snu13, and with protein Pih1, which forms the R2TP complex by association with the Hsp82-adaptor Tah1p and the two helicases Rvb1p, Rvb2p. Here, we identified by a proteomic-based approach protein Hit1 as a novel Rsa1p-interacting partner, and we show it is required to maintain steady-state levels of Rsa1p and box C/D snoRNAs. Moreover, we found presence of Hit1p with the other known assembly factors on precursors of U3 box C/D snoRNA, and observed that synthetic lethality arises upon combination of null mutants for *RSA1* and *PIH1* genes, or *HIT1* and *PIH1* genes, but not for *RSA1* and *HIT1* genes. Moreover, core protein Nop58 is shown here to directly interact *in vitro* with purified Rsa1p-Hit1p heterodimer or Snu13p-Rsa1p-Hit1p heterotrimer. Together, our observations conduct to classify Hit1p as a new C/D snoRNP assembly factor functionally linked to Rsa1p for binding to C/D snoRNP core proteins.**

## INTRODUCTION

Ribonucleoprotein particles (RNPs) participate in various fundamental functions for cell growth and development. In Eukarya, snoRNPs (small nucleolar RNPs) play essential role during processing of ribosomal RNAs (32). These particles are divided into two classes, each mediating one type of chemical modification on RNA: box C/D snoRNPs catalyze ribose 2'-O-methylation, and box H/ACA snoRNPs catalyze conversion of uridine to pseudouridine. Furthermore, some snoRNPs like C/D snoRNPs U3 and U8 in mammals (16), and H/ACA snoRNPs U17 in mammals and its yeast homolog snR30 (2, 36) are not involved in those chemical modifications but in RNA endonucleolytic cleavages during pre-rRNA processing.

Each box C/D snoRNPs arises by the association of a box C/D snoRNA (small nucleolar RNA) with a common set of four core proteins called Snu13p, Nop1p, Nop56p, and Nop58p in yeast *Saccharomyces cerevisiae* (respectively, 15.5K, Fibrillarine, NOP56, and

NOP58 in human) (28, 29, 51). Snu13p and 15.5K belong to the L7Ae family of RNA binding proteins whose members are components of several essential RNPs. Snu13p and 15.5K are present in C/D snoRNPs (51), in the RNP assembled on the snoRNA U3-specific boxes B and C (8), and in the U4 small nuclear ribonucleoprotein particle (U4 snRNP) involved in pre-mRNA splicing (37, 51). Nhp2 is the L7Ae protein present in H/ACA snoRNPs (18, 49) and in the H/ACA RNP of the vertebrate telomerase complex. Cytoplasmic SECIS mRNPs, which are involved in selenocystein incorporation in vertebrates proteins, contain the L7Ae protein SBP2 (1, 7, 9, 25, 51). The L7Ae family of proteins share a common RNA binding domain that specifically recognizes the structural RNA motifs Kink-turn (K-turn) and K-loop (23, 48). Initial binding of L7Ae proteins on these RNA motifs nucleates assembly of RNP by facilitating specific recruitment of the other core proteins (8, 46, 50).

Biogenesis of nuclear L7Ae RNPs is a complex process starting from transcription of RNA precursors and their maturation, to synthesis and nuclear import of RNP core proteins, which are orderly assembled on RNA during RNA processing, and finished by the delivery of mature RNPs to functional sites (32, 53). For snRNPs the process includes assembly steps in the cytoplasm for review, (39). Several studies highlighted the involvement of *trans*-acting factors during RNPs assembly. In the case of snRNPs formation, the SMN (Survival Motor Neuron) multiprotein complex controls cytoplasmic assembly steps of the particles. Indeed, SMN complex specifically recognizes snRNA sequences, and mediates the recruitment and the binding of a set of Sm proteins to form the Sm-core RNPs (24, 40). The Naf1 protein facilitates the H/ACA snoRNP co-transcriptional assembly by interacting with the Cbf5 core protein and with the RNA polymerase II CTD (10, 11, 14, 52). Recently, we and others showed that assembly of the L7Ae RNPs requires a set of conserved factors (3, 33). These factors will be written here as a couple formed by the yeast protein and human protein (yeast/human). This machinery includes Rsa1p/NUFIP, the Hsp82/Hsp90 co-chaperone R2TP complex composed of the AAA+ hexameric helicases Rvb1p/TIP49 and Rvb2p/TIP48, Tah1p/hSPAGH, and Pih1p/hPIH1(NOP17). The Hsp82/Hsp90 chaperon interacts with the R2TP complex through binding to the adaptor Tah1 (3, 33, 54).

Rsa1p and NUFIP are believed to play a central role in the RNP assembly process by interacting simultaneously with Pih1p/hPIH1, Rvb1p/hTIP49 and Rvb2p/hTIP48, and with the L7Ae family members Snu13p/15.5K, Nhp2p/hNHP2, and SBP2 (3). In particular, we demonstrated that a conserved stretch of 32 amino acids designated yPEP in Rsa1p and hPEP in NUFIP is dedicated to the interaction with proteins Snu13p/15.5K and Nhp2p/hNHP2 (3), and residues of the protein-protein interface were identified (Rothé, Back *et al.*, unpublished data). Moreover, based on NMR data for yPEP and molecular modeling, we proposed a 3D model of the interaction in which Rsa1p/NUFIP binds Snu13p/15.5K opposite to the RNA binding region suggesting that the assembly factor acts as an external scaffold during RNP assembly. In yeast, the other box C/D core protein Nop58 has been shown to be contacted by protein Pih1 (15). In a  $\Delta$ PIH1 mutant lacking Pih1p, the C/D

core proteins Snu13, Nop1, Nop56, and Nop58 accumulate into the cytosol and the nucleus (15), and the amounts of co-immunoprecipitated core proteins are reduced (55). Furthermore, Pih1p is closely associated with the low molecular weight Tah1 protein (13, 54, 55), which contains a TPR domain and acts as co-factor of the Hsp82 chaperon (54) (Back *et al.*, unpublished data).

This machinery dedicated to RNP assembly is based on a complex interaction network including activities of helicases and chaperons. Data obtained in human indicate that Hsp90, in association with the R2TP co-chaperon complex, participates to the RNP assembly process by controlling stability of the RNP core proteins (15.5K, NOP58, NHP2, SBP2) (3, 55). Furthermore, it has been proposed that NUFIP and the TIP48/TIP49 helicases facilitate interactions between C/D snoRNP components in a mechanism regulated by the ATPase activity of TIP48/TIP49 (3, 33, 34). Nevertheless, the precise mechanism by which the machinery assembles RNPs remains poorly understood. While protein-protein interactions were mapped (3, 13, 55), the *in vivo* existence of a complex formed by association of the full set of the identified factors is not settled. Moreover, data on interactions between R2TP and intermediates of the snoRNP biogenesis are lacking. Furthermore, the list of assembly factors could still be incomplete and additional factors may participate in the RNP assembly.

In this work, we set out to characterize complexes containing TAP-tagged Rsa1 and Pih1 proteins. Proteomic analysis of immunoprecipitated complexes indicates that Rsa1p is not stably associated with R2TP, but co-purifies with the Hit1 protein. Here we show that Hit1p is physically and functionally linked to Rsa1p, and that it represents an additional factor of the C/D snoRNP assembly machinery. Moreover, we purified protein complexes containing the core protein Snu13 associated with the Rsa1p-Hit1p heterodimer. We show that this heterotrimer interacts with the other core box C/D snoRNP protein Nop58.

## MATERIAL & METHODS

**Plasmid constructions.** Plasmids were manipulated by standard techniques. Constructs used for co-expression studies in *E. coli* were obtained by cloning PCR-amplified fragments between the *Nde*I and *Bam*HI sites of plasmids (pnEA-3CH and pnCS) (12). The primers used

for PCR amplification introduced a *NdeI* site at the 5' end and a *BamHI* site at the 3' end of the fragment (Rothé, Back *et al.*, unpublished data). The bacterial expression vector pGEX-6P-1::*SNU13* encoding the GST-Snu13p fusion protein was already described (31). The recombinant pACT2 plasmids expressing Gal4-AD-Rsa1p, Gal4-AD-Rsa1<sub>Cter</sub> and used for yeast two hybrid (Y2H) assays were already described (3). Recombinant plasmids pAS2 and pACT2 expressing respectively Gal4-BD-Hit1p and Gal4-AD-HA-Hit1p were obtained by the same procedure. Plasmid expressing GFP-Rsa1p was provided by E. Bertrand (IGM, Montpellier, France).

**Production of recombinant proteins.** The recombinant pGEX-6P-1::*SNU13* plasmid was used to produce Snu13p fused to the glutathione S-transferase (GST) (31) in the *E. coli* CodonPlus strain (Novagen®). The recombinant GST-Snu13 protein was purified from cellular extract under native conditions, using Glutathione-Sepharose 4B as recommended by the manufacturer (GE Healthcare®). The Snu13 protein was recovered after overnight cleavage on the beads using PreScission protease (GE Healthcare®). The cleaved recombinant Snu13 protein had two additional amino acid residues at its N-terminus (G and P), compared to the native protein.

**Yeast strains and measurements of RNA levels.** Wild-type, knock-out, and TAP-tagged strains used in this study were provided by Euroscarf®. Standard techniques were used for growth and handling of *S. cerevisiae* strains. Total RNAs were extracted from exponentially growing cells (OD<sub>600</sub> ~0.8). Northern-blot analyses were carried out according to standard procedures on 5 µg of total RNAs using as probes 5' <sup>32</sup>P-radiolabeled oligodesoxyribo nucleotides complementary to RNAs of interest.

**Synthetic lethal screen by diploid strains sporulation and dissection.** Diploid strains were obtained by cross-breeding on YPD plates of simple knock-out haploid strains By4741 (MAT<sub>a</sub>) and By4742 (MAT<sub>α</sub>) from Euroscarf®. Diploid cells were resuspended in 2 mL of liquid sporulation medium (1% potassium acetate; 0.05% zinc acetate). Sporulation cultures were incubated on a roller wheel for 5 days at 25°C, then for 3 days at 30°C. Cells were finally dissected with a micromanipulator (Singer®) and

separated on YPD plates. Strain genotypes were controlled by standard techniques including genomic DNA extraction and PCR analysis.

**Purification of multiprotein complexes by protein A selection affinity.** *S. cerevisiae* strains SC4186 (TAP-*RSAl*), SC3760 (TAP-*HIT1*), and SC3165 (TAP-*PIH1*) expressing TAP-tagged proteins, as well as the untagged strain BY4741, were grown at 30°C in YPD to OD<sub>600</sub>~0.8-1. Protein A purifications were performed in breaking buffer (100 mM NaCl, 8% glycerol, 0.1 mM DTT, 100mM Tris-HCl, pH8) from cells representing 2500 units of OD<sub>600</sub> of the culture (43). Total cellular extracts were incubated for 12 hours at 4°C with IgG-Sepharose beads (Sigma-Aldrich®). Beads were then washed in Mobicol "classic" columns (MoBiTec®) with washing buffer (100 mM NaCl, 0.1% IGEPAL, 0.1mM DTT, 20 mM Tris-HCl, pH 7.4). After TEV cleavage, eluted proteins were separated by SDS-PAGE and stained with Coomassie Blue.

**Mass Spectrometry-based Proteomic analysis.** In gel digestion of SDS-PAGE gel slices was performed as previously described (35). The resulting peptide extracts were analyzed by nanoLC-MS/MS on a nanoACQUITY Ultra-Performance-LC (UPLC, Waters, Milford, UK) coupled to a SYNAPT HDMS G1 hybrid quadrupole-time-of-flight mass spectrometer (Waters, Milford, UK) equipped with a Z-spray ion source and a lock mass system. Peptide mixtures were loaded on a Symmetry C18 trap precolumn (180 µm inner diameter × 20 mm, particle size 5 µm; Waters, Milford, UK) and peptides were separated on a ACQUITY UPLC® BEH130 C18 column, 75 µm x 200 mm, 1.7 µm particle size (Waters, Milford, UK). The solvent system consisted of 0.1 % formic acid in water (solvent A) and 0.1 % formic acid in acetonitrile (solvent B). Trapping was performed during 3 min at 5 µL/min with 99 % of solvent A and 1 % of solvent B. Elution was performed at a flow rate of 400 nL/min, using 1-50 % gradient (solvent B) over 35 min at 45°C followed by 65 % (solvent B) over 5 min. Parameters of the mass spectrometer were set as follows: the capillary voltage was set to 3500 V and the cone voltage to 35 V. Mass calibration of the TOF was achieved using [Glu1]-fibrinopeptide B (GFP) over the m/z range 50-2000 in the positive ion mode. Online correction of this calibration was achieved using lock-mass on product ions derived from the [Glu1]-fibrinopeptide B (GFP).

The ion  $(M+2H)^{2+}$  at  $m/z$  785.8426 is used to calibrate MS data and the fragment ion  $(M+H)^+$  at  $m/z$  684.3469 is used to calibrate MS/MS data during the analysis. For tandem MS experiments, the system was operated with automatic switching between MS and MS/MS modes. The 4 most abundant peptides, preferably doubly and triply charged ions, were selected on each MS spectrum for further isolation and CID fragmentation. MS/MS fragmentation was performed in the collision cell using argon as collision gas. The complete system was fully controlled by MassLynx 4.1 (SCN 566, Waters, Milford, UK). Raw data collected during nanoLC-MS/MS analyses were processed and converted with ProteinLynx Browser 2.3 (23, Waters, Milford, UK) into \*.pkl peak list format. The MS/MS data were submitted to Mascot search engine (Matrix Science, London, UK) installed in its 2.2.04 version on a local server. Searches were performed against a composite target-decoy database including both forward (target) and reversed (decoy) SwissProt protein sequences (21 January 2010, 5 1028424 total entries). Searches were performed with a tolerance on mass measurements of 0.2 Da for precursor and 25 ppm for fragment ions. Carbamidomethylation of cysteine residues, oxidation of methionine residues and acetylation of protein N-term residues were searched as variable modifications. Up to one missed cleavage was allowed. Scaffold V3\_00\_01 (Proteome Science, Portland, USA) was used for identification validation and false positive rate estimation for protein identification. For each sample, peptides were filtered out according to the cutoff set for proteins hits with 2 or more peptides (Mascot Ion Score >20 and Mascot Ion Score - Mascot Identity Score >0) and no false positive protein was identified. Further data treatment were performed with Microsoft Excel.

**Noncovalent Mass Spectrometry.** Before noncovalent mass spectrometry experiments, protein buffer was exchanged against 500 mM or 1 M ammonium acetate buffer pH 7.5 using microcentrifuge gel-filtration columns (Zeba 0.5 mL, Thermo Scientific, Rockford, IL). Protein concentrations were determined spectrophotometrically. Mass spectrometry experiments were performed on a hybrid quadrupole/ion mobility separator/time-of-flight instrument (Synapt HDMS G2, Waters, Manchester, UK) equipped with an automated chip-based nanoelectrospray source (Triversa Nanomate,

Advion Biosciences, Ithaca, NY) operating in the positive ion mode. External calibration was performed with the multiply charged ions produced by 2  $\mu$ M horse heart myoglobin diluted in 1:1 (v/v) water/acetonitrile acidified with 1 % (v/v) formic acid. Prior to noncovalent analysis, each sample was analyzed in classical denaturing conditions in order to assess its purity and homogeneity. Mass measurements under denaturing conditions were carried out by diluting samples to 2  $\mu$ M in water/acetonitrile/formic acid (50:50:1). Accelerating voltage ( $V_c$ ) was set to 20 or 40 V while pressure in the interface region ( $P_i$ ) was 1.1 mbar. Molecular weights of  $17599.0 \pm 0.1$  Da, and  $18241.9 \pm 0.3$  Da were measured for Rsa1p<sub>230-375</sub>-Hit1p expression and were respectively in agreement with theoretical masses calculated from the amino acid sequence (GPHM-Rsa1p(230-375) and Hit1p(2-164)). Molecular weights of  $17176.8 \pm 0.1$  Da,  $18243.4 \pm 0.6$  Da and  $13860.4 \pm 0.2$  Da were measured for Snu13p-Rsa1p<sub>230-375</sub>-Hit1p expression and were respectively in agreement with theoretical masses calculated from the amino acid sequence (GPH-Snu13p(1-129), Rsa1p(230-375), and Hit1p(2-164)). Experiments under nondenaturing conditions (noncovalent conditions) were realized in either 500 mM (for Rsa1p<sub>230-375</sub>-Hit1p) or 1 M (for Snu13p-Rsa1p<sub>230-375</sub>-Hit1p) ammonium acetate buffer at pH 7.5. Optimal interface parameters ( $V_c$ ,  $P_i$ ) were used in order to combine intact noncovalent complexes detection to efficient ion desolvation and ion transmission in the mass spectrometer (45). Data analysis was performed with MassLynx 3.5 (Waters, Manchester, UK).

**RNA co-IP assays.** Co-immunoprecipitation assays were realized as previously described in (3). Briefly, yeast cells expressing the TAP-tagged proteins were grown to exponential phase before centrifugation. Cells were lysed by bead-beating in breaking buffer (100 mM NaCl, 8% glycerol, 0.1 mM DTT, 100mM Tris-HCl, pH8) and lysates were added to IgG-Sepharose beads (Sigma-Aldrich®) and incubated 2 h at 4°C. Beads were washed four times in washing buffer (100 mM NaCl, 0.1% IGEPAL, 0.1mM DTT, 20 mM Tris-HCl, pH 7.4). RNAs were extracted with phenol-chloroform and analyzed by RT-PCR. cDNAs were generated using oligonucleotides OG-Rev (5'-CAGTGCAGGGTCCGAGGTATTCTGAAAACCAAACCTTTGG-3') containing the underlined sequence



exogenous to the yeast genome, and used as templates for PCR amplification. PCR was then performed by using the 5' oligonucleotide OG-5'PCRa 5'-TCTGTGTCGACGTA CTCA-3' complementary to exon 1 of *snR17A* and *snR17B* (from nucleotides 1 to 14) or the OG-5'PCRB 5'-AGTCTTAGGTACTAGAGTTT-3' complementary to *snR17A* intron (from nucleotides 66 to 85) and the 3' oligonucleotide OG-3'PCR 5'-CA GTGCAGGGTCCGAGGTATT-3' complementary to the yeast exogenous sequence.

**Western-blot.** Crude cell extracts were prepared from pellets of  $7.10^8$  yeast cells resuspended in 300  $\mu$ L of buffer (150 mM KCl, 5 mM MgCl<sub>2</sub>, 0,05% TRITON-X100, 20 mM Tris-HCl, pH 7,5), and lysed by bead beating. The lysates were centrifuged two times at 10,000 g for 5 min. Concentration of proteins in the supernatants obtained from the various samples were estimated by Bradford assays. Equivalent amounts of protein were fractionated on 12,5 % SDS-PAGE and western blots were performed according to standard procedures using rabbit commercial anti-GFP, anti-HA, anti-PMA1 (provided by B. André, LPMC, IBMM-ULB, Bruxelles, Belgium), or anti-Tah1p-Pih1p antibodies (provided by S. Moreira LEBS, Gif-sur-Yvette, France).

**Formation of complexes between proteins co-expressed in *E. coli*.** Co-expression tests in *E. coli* were carried out using the procedure described previously (44) (Rothé, Back *et al.*, unpublished data). The recombinant plasmids pNEA-3CH expressing His<sub>6</sub>Snu13p, His<sub>6</sub>Rsa1p, His<sub>6</sub>Rsa1p<sub>230-375</sub>, or His<sub>6</sub>Tah1p, and the recombinant plasmid pncS expressing Hit1p, Rsa1p<sub>230-375</sub>, or Pih1p were simultaneously transformed into pRare2 *E. coli* host cells (Novagen). Growth was performed in 500 mL of 2X concentrated LB medium at 37°C to OD<sub>600</sub> ~0.8. Then, the temperature was shifted to 20°C. At this point, co-expression of the proteins was induced by addition of isopropyl- $\beta$ -D-thiogalactopyranoside (Euromedex) at a 1 mM final concentration. Cells were further grown overnight at 20°C. They were collected by low speed centrifugation, resuspended in buffer A (20 mM HEPES, pH 7.5; 400 mM NaCl, 5 mM  $\beta$ -mercaptoethanol), and lysed by sonication. The soluble fraction recovered by high speed centrifugation was mixed with 100  $\mu$ L of Talon resin (Clontech). The supernatant was removed after 1 h incubation, and the resin washed

extensively with buffer A. Proteins were eluted from the column using buffer A containing 500 mM imidazole. Eluted proteins were analyzed by SDS-PAGE.

**Electrophoresis Mobility Shift Assays (EMSA).** Yeast RNAs U14, variant of U3 snoRNA U3 $\Delta$ 2,3,4 (31), and tRNA<sup>His</sup> were synthesized by T7 RNA polymerase by *in vitro* transcription of PCR fragments in the conditions previously described (3, 6). RNA transcripts were radiolabeled at their 5' ends by T4 polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]ATP. 500 nM of various proteins or complexes were mixed and incubated for 20 min at 4°C with 5 fmol of 5'-end <sup>32</sup>P-radiolabeled RNA in buffer D (20 mM HEPES, pH 7.9; 150 mM KCl; 1.5 mM MgCl<sub>2</sub>; 0.2 mM EDTA; 10 % glycerol). The RNA-protein complexes formed were visualized by native gel electrophoresis.

**Yeast two hybrid assays.** For yeast two-hybrid assays (Y2H), appropriate pACTII and pAS2 plasmids were used to transform haploid test cells (strains Y190 and Y187, respectively), which were then crossed. Diploids were selected on media suitable for double selection (Leu-, Trp-) and then plated on media suitable for triple selection (Leu-, Trp-, His-). Various concentrations of 3-Amino-1,2,4-triazol (3AT) (Sigma®) were used to evaluate the strength of the interactions. Growth was assessed after three days of incubation at 30°C.

**His-tagged protein pull-down.** Complex formation was tested in a volume of 20  $\mu$ L of buffer D (20 mM HEPES, pH 7.9; 150 mM KCl; 1.5 mM MgCl<sub>2</sub>; 0.2 mM EDTA; 10 % glycerol) in the presence of recombinant proteins at a concentration of 2.5  $\mu$ M and 4  $\mu$ L of <sup>35</sup>S radiolabeled Nop58p, obtained by using the kit *E. coli* T7 S30 Extract System for Circular DNA (Promega). After 20 min at 30°C, 180  $\mu$ L of buffer (10 mM Tris-HCl, pH 8, 150 mM NaCl, 0.1 % Igepal, 5 mM imidazole) and 20  $\mu$ L of cobalt-sepharose beads (GE Healthcare) were added. The mixture was incubated on a wheel at 4°C for 90 min. Finally, the beads were washed three times binding buffer, and retention of radiolabeled protein was analyzed by gel electrophoresis SDS-PAGE 7.5 %.

## RESULTS

### **Rsa1p is not stably associated with the R2TP complex**

Aiming at characterizing the variety of complexes that are formed in the cell during C/D snoRNP biogenesis, we attempted to identify proteins that co-purify with the RNPs assembly factors Rsa1p and Pih1p. We used a procedure based on a single-step immunoglobulin G (IgG) affinity purification of TAP-tagged Rsa1 and Pih1 proteins followed by protein identification with proteomic analysis. Yeast strains expressing one of the proteins tagged at their C-terminus with the tandem affinity purification (TAP) sequence were grown. Cell extracts were prepared from cells in log-phase, during which biosynthesis of snoRNPs is very active. After capture on IgG-coupled beads and cleavage with the TEV protease, the proteins co-purified with TAP-Rsa1p or TAP-Pih1p were fractionated by SDS-PAGE (Fig. S1). Gel slices were cut systematically each 2 mm, and protein contents were identified by mass spectrometry-based proteomic analysis (Supplementary data Table S1).

Proteins of interest related to RNP assembly and identified with a significant number of peptides are listed in Table 1. As expected, proteins constitutive of the whole R2TP complex mainly co-purify with TAP-Pih1p, as Rvb1p, Rvb2p, and Tah1p are identified with a high number of peptides (16 for Rvb1p, 17 for Rvb2p, and 4 for Tah1p). In parallel, C/D core proteins Snu13p, Nop56p, Nop58p, and Nop1p were also identified in association with Pih1p but with fewer peptides. Surprisingly, Rsa1p does not co-purify with any of the R2TP components as no peptides of Rsa1p were found with TAP-Pih1p. In the TAP-Rsa1p experiment, C/D core proteins were also detected with a low number of peptides. No peptide from Pih1p and from any R2TP subunits were found, which is consistent with results obtained in the TAP-Pih1p experiment.

Altogether these observations first highlight the transient feature or low stability of the interactions between these assembly factors and core proteins. In addition, the low steady-state level of the fully assembled R2TP-Rsa1p complexes can result from a transient association of Rsa1p with R2TP, or alternatively the complex is destabilized during purification due to weak interactions.

### **Co-purification of Hit1p with Rsa1p**

As proteomic analysis identified Hit1p with a high number of peptides (7) among the proteins that co-purify with TAP-Rsa1p, we next performed the IgG affinity purification of TAP-Hit1p. Rsa1p was found to be strongly associated with Hit1p as 10 Rsa1p peptides are detected in the eluted TAP-Hit1p samples (Table 1). As in TAP-Rsa1p experiment, only fewer peptides from the C/D snoRNP core proteins Nop1p and Snu13p that co-purify with TAP-Hit1p were observed, and no peptide from the R2TP complex was detected (Table 1). These data suggest the *in vivo* existence of a stable complex containing Rsa1p and Hit1p, which is in agreement with previously published two hybrid screen results (20).

### **Rsa1p, Hit1p, and Pih1p are related to the stability of C/D snoRNAs**

The involvement of Hit1p in snoRNP biogenesis is attested by our phenotypic analyses of the *HIT1* gene null mutant ( $\Delta HIT1$ ). Indeed, this mutant shows a phenotype similar to those observed for null mutants of genes *RSA1* and *PIH1*, which respectively encode proteins Rsa1p and Pih1p involved in the assembly of the C/D snoRNPs (Fig. 1). First we showed that the disruption of genes *RSA1*, *HIT1*, or *PIH1* is correlated with a slow growth phenotype (Fig. 1A). Indeed, in standard growth conditions the  $\Delta RSA1$ ,  $\Delta HIT1$ , and  $\Delta PIH1$  null mutant strains have doubling times respectively ~58 %, ~56 %, and ~31 % lower than the WT cells. Northern blot analyses indicate that these slow growth phenotypes are associated with lower levels of C/D snoRNAs whereas the levels of H/ACA snoRNAs and UsnRNAs remain unchanged (Fig. 1B). The stability of some C/D snoRNAs, i.e. U14, snR190, snR70, is decreased to the same extent in the three mutants. The levels of the C/D snoRNAs snR45, snR63, and snR13 are also reduced in the  $\Delta PIH1$  mutant, but these RNAs are barely detected in the  $\Delta RSA1$  and  $\Delta HIT1$  mutants. As the amount of U3 snoRNA was only found slightly decreased in the various strains, we analyzed a truncated version of the U3 snoRNA (U3 $\Delta$ 2,3,4) that accumulates precursors of U3 RNPs when the encoding gene is transfected in yeast cells (3, 27). We already reported the higher accumulation of U3 $\Delta$ 2,3,4 precursors in cells lacking Rsa1p (Fig. 1C) or Pih1p (3). Such accumulation is also observed in

$\Delta HIT1$  cells (Fig. 1C), indicating that Hit1p is also related to the processing of U3.

Given that snoRNPs participate in ribosome biogenesis, we further analyzed the phenotypes of the  $\Delta RSA1$ ,  $\Delta PIH1$ , and  $\Delta HIT1$  null mutant strains, and detected by Northern blot that these strains accumulate the 35S ribosomal RNA precursor and several intermediates of its processing pathway (Fig. 1D): intermediate 32S which results from a lower efficiency of cleavage at site A3, and intermediates 23S and 20S of the 18S maturation pathway accumulated due to a lower efficiency of cleavages at sites A1 and A3.

Altogether,  $\Delta HIT1$  and  $\Delta RSA1$  strains display similar marked phenotypes strongly suggesting that both proteins participate in a common step of C/D snoRNP biogenesis. Interestingly, under the tested conditions, absence of Tah1p does not strongly affect the cell growth (Fig. 1A), the snoRNA stability (Fig. 1B), nor the rRNA processing (Fig. 1D).

### **Like other assembly factors, Hit1p is associated *in vivo* with precursors of U3 snoRNA**

We next tested the association of Rsa1p, Hit1p, and the components of the R2TP complex, with intermediates formed in the U3 snoRNA processing pathway. In yeast, RNA polymerase II produces a precursor of U3 (pre-U3) containing one intron and an extended 3'-end (27). As described in Fig. 2A, unspliced pre-U3 and spliced pre-U3 could be detected by RT-PCR using a reverse-primer hybridizing in the 3'-extension. We used such detection system to test for the presence of U3 precursors among the RNAs that are co-purified with the TAP-tagged assembly factors. First, we observed that Nop1p, Nop56p, and Nop58p are stably associated with unspliced and spliced pre-U3 (Fig. 2B, lanes 2-4), demonstrating that the core proteins are recruited early during U3 snoRNP biogenesis. In addition, Rsa1p, Hit1p, Pih1p, Rvb1p, and Rvb2p, which are not present in the mature U3 particle, are also associated with the two U3 precursors (Fig. 2C, lanes 2-6). We concluded that Hit1p is present with Rsa1p and with the R2TP assembly machinery on the pre-RNPs. Nevertheless, Tah1p was not found associated with pre-U3s (Fig. 2D, lane 2). In accordance with the absence of apparent phenotype for the  $\Delta TAH1$  mutant strain (Fig. 1), this result seemed to minimize the role of Tah1p in snoRNP biogenesis. However, as shown by W. Houry

laboratory (55), the functional importance of Tah1p appears during the stationary growth phase, during which it stabilizes Pih1p (Fig. 2E). Thus, we performed TAP purification from cells harvested in the stationary growth phase (Fig. 2D). Our results clearly show the implication of Tah1p in U3 biogenesis since pre-U3 species are selectively co-precipitated with Tah1p in an extract prepared from cells in stationary phase (Fig. 2D, lane 4). Altogether, these data indicate that Tah1p could be selectively involved with the Rsa1p-Hit1p-R2TP assembly machinery under stress conditions.

### ***RSA1* and *HIT1* exhibit synthetic lethality with *PIH1***

To further increase our knowledge of the functional link existing between the assembly factors, we tested synthetic lethality between pairs of null alleles in the non-essential genes *RSA1*, *HIT1*, *PIH1*, and *TAH1* (Fig. 3). We crossed single knock-out haploid strains to obtain heterozygous diploid strains. Tetrads of haploid cells were produced by meiosis during sporulation, dissected, and growth of the haploid cells were tested on YPD media plates. The  $\Delta RSA1$ - $\Delta HIT1$  double knock-out haploid strain did not reveal any cumulative effect in the phenotypes of the single knock-out strains, as illustrated by a comparable growth defect for the  $\Delta RSA1$ - $\Delta HIT1$ ,  $\Delta RSA1$ , and  $\Delta HIT1$  strains (Fig. 4A). This result supports the observation that the lack of Rsa1p or Hit1p leads to the same phenotypic effects (Fig. 1), and that these proteins are closely linked in the same function. This function is strongly connected to Pih1p as a lethal phenotype is observed when cells combine null mutations of *RSA1* and *PIH1* genes, or null mutations of *HIT1* and *PIH1* genes.

### **The stability of Rsa1p relies on presence of Hit1p**

The strong functional connection that we have established between Rsa1p and Hit1p is based on a control of the stability of Rsa1p by Hit1p. We evaluated the levels of the fusion protein GFP-Rsa1p when the encoding gene is expressed in the  $\Delta RSA1$  strain and in the double knock-out  $\Delta RSA1$ - $\Delta HIT1$  strain (Fig. 4B). The fusion protein is functional as ectopic expression in the  $\Delta RSA1$  strains restores a wild type growth phenotype (data not shown). Western-blot analysis with anti-GFP antibodies indicates that the level of GFP-Rsa1p is drastically reduced in extracts prepared from cells lacking Hit1p (Fig.

4B, compare lane 2 with lane 3). The role of Hit1p for the *in vivo* stability of Rsa1p is confirmed by restoration of Rsa1p amounts upon ectopic expression of Hit1p in strain  $\Delta$ RSA1- $\Delta$ HIT1 (Fig. 4B, lane 4). On the contrary, absence of Rsa1p has no strong effect on the level of Hit1p fused to the epitope derived from the Influenza protein hemagglutinin (HA) (Fig. 4B, lanes 5 to 7).

### **A COOH-terminal fragment of Rsa1p directly interacts with Snu13p and Hit1p**

In a previous work, we showed that ectopic expression in the  $\Delta$ RSA1 strain of the Rsa1p COOH-terminal fragment, extending from Pro230 to Lys381 and named Rsa1p<sub>230-381</sub> allows full recovery of the wild type growth phenotype (3). We also demonstrated that this fragment interacts in a yeast two hybrid (Y2H) assay with Pih1p (Boulon et al., 2008), and that it carries the conserved PEP domain (residues 230 to 263), which constitutes the Snu13p-binding domain of Rsa1p (3) (Rothé, Back *et al.*, unpublished data). Here we show that Rsa1p<sub>230-381</sub> interacts also with Hit1p in a Y2H assay (Fig. 5A).

To better characterize the interaction network between Snu13p, Rsa1p, Pih1p, and Hit1p, we co-expressed His<sub>6</sub>-tagged Snu13p or Rsa1p with one or two of their potential partners in *E. coli*. Co-expressed proteins were co-purified from bacterial extracts by Immobilized Metal Ion Affinity Chromatography, fractionated by SDS-PAGE and analysed by mass spectrometry. The mass spectrometric analysis revealed that the last 6 amino acids of Rsa1p<sub>230-381</sub> are often degraded (data not shown). This observation led us to clone and work with a new fragment deleted of these residues and named Rsa1p<sub>230-375</sub>. Proteomic analysis of the detected bands identified Hit1p as associated with both His<sub>6</sub>Rsa1p and His<sub>6</sub>Rsa1p<sub>230-375</sub> (Fig. 5B). These data confirm the Y2H interactions, and reflect direct physical association between Rsa1p and Hit1p. In addition, we observed that Rsa1p and Hit1p also co-purify with His<sub>6</sub>Snu13p (Fig. 5C), indicating formation of a complex involving these 3 partners in *E. coli*.

To further characterize binary (Rsa1p<sub>230-375</sub> and Hit1p) or ternary (Snu13p, Rsa1p<sub>230-375</sub>, and Hit1p) complexes and to unambiguously assess their binding stoichiometries, His<sub>6</sub>-tagged complexes were cleaved by TEV protease, purified and analyzed by noncovalent mass spectrometry. Electrospray ionization mass spectra obtained in nondenaturing conditions of

Rsa1p<sub>230-375</sub>-Hit1p confirmed the presence of a 1/1 complex with a measured mass of  $35\,906 \pm 1$  Da (expected mass of the complex  $35\,905$  Da) in addition to free proteins Rsa1p<sub>230-375</sub> ( $17\,600 \pm 1$  Da ; expected mass:  $17\,599$  Da) and Hit1p ( $18\,305 \pm 1$  Da in agreement with Hit1p sequence after deletion of the N-terminus methionine and presence of 1 Zn<sup>2+</sup> ion ;  $18\,245 + 63 = 18\,308$  Da) (Fig. 5C). Similarly, noncovalent mass spectrometry analysis of the complex involving Snu13p, Rsa1p<sub>230-375</sub>, and Hit1p confirmed the presence of a 1/1/1 complex with a measured mass of  $49\,342 \pm 2$  Da (expected mass of the complex  $49\,344$  Da) in addition to free Snu13p ( $13\,861 \pm 1$  Da ; expected mass:  $13\,860$  Da) and the 1/1 heterodimer Rsa1p<sub>230-375</sub>-Hit1p ( $35\,483 \pm 5$  Da ; expected mass :  $35\,484$  Da) (Fig. 5D).

Finally, we observed by EMSA that incubation of the purified protein complexes with radiolabeled snoRNAs containing the K-turn RNA motif leads to formation of a single high molecular weight complex displaying a slower mobility than the Snu13p-Rsa1p-RNA complex (Fig. 5D). Hence, presence of Hit1p does not impair the capacity of the Snu13p-Rsa1p heterodimer to interact with RNA.

### **Core protein Nop58p binds heterodimers Rsa1p-Hit1p and Tah1p-Pih1p**

As presented in table 1, only small amounts of C/D core proteins were found in association with assembly factors, suggesting that the latter are transiently associated to pre-RNPs in normal growth conditions. It was previously established that Pih1p interacts with Nop58p (15), indicating that the interaction between this assembly factor and this core protein is intrinsically stable. In our precedent work, we showed a Y2H interaction between Rsa1p and the Nop58 core protein (3). Here, by using the purified recombinant His<sub>6</sub>-tagged complexes in pull-down assays with the <sup>35</sup>S radiolabelled Nop58p produced in *E. coli* S30 extract, we observed that Nop58 co-purifies as expected with His<sub>6</sub>Tah1p-Pih1p (Fig. 6, lane 2), and also with His<sub>6</sub>Snu13p-Rsa1p<sub>230-375</sub> (Fig. 6, lane 3), His<sub>6</sub>Rsa1p<sub>230-375</sub>-Hit1p (Fig. 6, lane 4), and His<sub>6</sub>Snu13p-Rsa1p<sub>230-375</sub>-Hit1p (Fig. 6, lane 5), but not in absence of the His<sub>6</sub>-tagged complexes (Fig. 6, lane 1). These *in vitro* binding assays confirm the direct association of Rsa1p and Nop58p. Hence, Nop58p associates directly not only with Pih1p but also with Rsa1p in absence of RNA. The presence of Tah1p bound to Pih1p, or Hit1p bound to Rsa1p does not impair the interactions Rsa1p-Nop58 or Pih1p-Nop58.

## DISCUSSION

### *The function of Rsa1p relies on Hit1p*

In this work, we have identified Hit1p (High Temperature growth, YJR055W) as a novel protein which function is related to the assembly of C/D snoRNPs. At first, we found that Hit1p and the C/D snoRNP assembly factor Rsa1p co-purify of yeast extracts (Table 1), confirming the interaction between the two proteins in a global yeast two hybrid screen (20). Here, we show that Hit1p has features in common with Rsa1p and the other R2TP assembly factors: it is associated *in vivo* with U3 snoRNA precursors, and is required for accumulation of the C/D snoRNAs. Our data indicate that the presence of Hit1p on the pre-snoRNPs relies on a direct interaction with Rsa1p. Moreover, we show that at least one of the major function of Hit1p is to stabilize Rsa1p in yeast (Fig. 6). The loss of such activity in the single-gene knock-out mutant strain  $\Delta HIT1$  gives rise to a phenotype similar to the  $\Delta RSA1$  strain, and deletion of the two *HIT1* and *RSA1* genes in a same strain does not lead to cumulative effects of the mutations (Fig. 1). As Hit1p was not found to interact with other characterized C/D snoRNPs assembly factors, nor with any of their core proteins, we propose that Hit1p intervenes during the assembly of C/D snoRNP through its stability effect on Rsa1p.

The specificity of Hit1p for Rsa1p relies on a direct association. Indeed, we demonstrate that the two proteins form an heterodimer when they are co-expressed as recombinant proteins in *E. coli*. The production of Rsa1p is not efficient in *E. coli* as most of the proteins accumulate in the insoluble fraction. The solubility is markedly enhanced when Rsa1p is co-expressed with its interacting proteins Snu13 (Rothé, Back *et al*, unpublished data) or Hit1 (Fig. 6B). This is reminiscent of the feature of the Pih1p C/D snoRNP assembly factor which is efficiently produced and purified as a recombinant protein only in the presence of its interaction partner Tah1p (13, 55). Ectopic expression of various fragments of Rsa1p in strain  $\Delta RSA1$  indicated that the last 152 amino acids of the COOH-terminal end (Rsa1p<sub>230-381</sub>) represent the minimal fragment able to restore normal growth ((3), and data not shown). All the physical interactions that we have characterized in this work were obtained with the C-terminal part of Rsa1p which constitutes a platform for several non-exclusive interactions. Indeed, we show that in presence of Hit1p, the C/D snoRNAs binding

protein Snu13p can still interact with the PEP motif located at the N-terminal end of Rsa1p<sub>230-375</sub>, and that the formed complex Snu13p-Rsa1p<sub>230-375</sub>-Hit1p can interact with Nop58, one of the core box C/D snoRNP-associated proteins.

### *Features of the Hit1 protein*

The Hit1 protein contains a zinc finger domain, which represents the prototype of the HIT-zinc finger domain (Zf-HIT) family (22). The 3D solution structure of the Zf-HIT domain present in human protein ZNHIT2 (protein FON) classified it as a new member of the treble clef domain (17). The function of the domain remains obscure and we have no evidence whether it is required for the association of Hit1p with Rsa1p. Interestingly, the Zf-HIT domain is found in nuclear proteins involved in gene regulation and chromatin remodelling, and is also present in protein Bcd1 (Box CD 1 protein), which controls the box C/D snoRNAs stability in yeast (19, 41). The human homologue of Bcd1p was shown to be involved in the C/D RNP assembly process by interacting with NUFIP, the functional homologue of Rsa1p (33). Conservation of a similar domain in the two assembly factors Hit1p and Bcd1p raises the question of whether the two domains have similar or related functions.

TRIP3 (Thyroid Hormone Receptor Interacting protein 3), the human homologue of Hit1p, is involved in multiple hormone-dependent regulation processes, e.g. the regulation of retinoic acids receptor in neuroblastoma differentiation (30), the regulation of the HNF-4 $\alpha$  (Hepatocyte Nuclear Factor 4 $\alpha$ ) transcription activity involved in glucose metabolism (21), and the regulation of PPAR $\gamma$ -mediated adipocyte differentiation (26). The precise role played by TRIP3 in these mechanisms is currently unclear. Work performed by the laboratory of E. Bertrand shows that TRIP3 participate in C/D snoRNPs biogenesis (personal communication). Furthermore, a recent study revealed a decrease of TRIP3 expression in renal cell carcinoma, and proposed TRIP3 as a new tumor suppressor gene (4). Thus, this property could rely on its capacity to control cell growth through its implication in various essential processes like nuclear-receptor dependant transcription and/or by promoting snoRNP biogenesis and ribosome synthesis.

### *Rsa1p-Hit1p represents a module for box C/D snoRNP assembly*

Based on our data of the co-purification assays showing no accumulation of complexes composed of Rsa1p, Hit1p, Rvb1p, Rvb2p, Tah1p, and Pih1p (Table 1), we propose that in yeast, these assembly factors mainly exist as two sub-complexes. One module would be composed of the heterodimer Rsa1p-Hit1p, and a second one would correspond to the R2TP complex (Rvb1p-Rvb2p-Tah1p-Pih1p). Nevertheless, as we showed that Rsa1p and Pih1p interact in Y2H assays (3), such interaction may promote transient association between the two modules during snoRNPs assembly. The importance of Rsa1p and Pih1p for the stability of C/D snoRNA *in vivo* is well established (3, 55). In addition, lethality of the  $\Delta RSA1\Delta PIH1$  mutant (Fig. 3) indicates that Rsa1p and Pih1 represent the cornerstones of the assembly machinery. Although Nop58p was shown to interact with Pih1p, our Y2H assays (3) and our *in vitro* binding assay (Fig. 6) show an association of Nop58 with the purified Rsa1p<sub>230-375</sub>-Hit1p and the Snu13p-Rsa1p<sub>230-375</sub>-Hit1p complex. Thus, at some stage of the C/D snoRNP biogenesis, the stable association of Nop58p with Pih1p could be destabilized in favour of an interaction with a pre-RNP composed of Snu13p-Rsa1p-Hit1p. Proteins of the two modules are present on pre-U3 snoRNPs (Fig. 2). Hence at this stage, the two modules can transiently cooperate for RNP formation: Rsa1p can facilitate recruitment of the Nop proteins on the RNA-Snu13p pre-RNP (3, 33, 34), and R2TP can operate a conformation remodelling required for the progression of the RNP assembly by bringing helicase activity of the Rvb1p-Rvb2p proteins (3, 34).

The precise events occurring during early steps of snoRNP assembly are not known. The initial interaction established between the box C/D RNA motif and proteins Snu13p/15.5K in Eukaryotes or L7Ae in Archaea, is required for subsequent assembly of the other proteins (5, 38, 42, 47, 50). An unresolved question is whether the K-turn is initially bound by the free form of Snu13p and next by the Rsa1p-Hit1p pair or directly by a Snu13p-Rsa1p-Hit1p complex. This last scenario is possible because such heterotrimeric complex can be purified (Fig. 5C) and because association of Rsa1p and Hit1p with Snu13p does not impair the RNA binding capacity of Snu13p (Fig. 5D). In addition, given that Snu13p does not possess nuclear localization signal (NLS), one possibility

would be that Rsa1p or Hit1p serves as protein cargo for the import of Snu13p in the nucleus. Interestingly, a NLS signal is present in the C-terminal region of Rsa1p. Hence, Snu13p-Rsa1p-Hit1p may represent a module for nuclear import and RNA binding.

### FIGURE LEGENDS

**Table 1:** Identification of Hit1p as an interacting partner of Rsa1p. List of identified proteins by proteomic analysis after tandem affinity experiments of TAP-Rsa1p, TAP-Hit1p, TAP-Pih1p, and an untagged negative control (No TAP).

**Figure 1:** Phenotypic analysis of the single null mutant strains.

(A) Growth of the strains By4741 WT, and of its mutants  $\Delta RSA1$ ,  $\Delta HIT1$ ,  $\Delta PIH$ , and  $\Delta TAH1$  was monitored overtime by measuring the optical density of the cultures at 600 nm ( $OD_{600}$ ). The doubling times were calculated (G), and the growth defect of the mutant strains were expressed as a percentage of the doubling time of the wild-type strain designated. (B) Rsa1p, Hit1p and Pih1p are required for accumulation of C/D snoRNAs. Northern blot analysis of snoRNAs steady-state levels in the null mutant strains. Extracted RNAs were separated by 7 % acrylamide/urea gels and transferred on Zeta-Probe membrane (Biorad®). Various RNAs (C/D snoRNAs, H/ACA snoRNAs and snRNAs) were detected by antisense oligonucleotide probes.

(C) Analysis by Northern blot of the processing of a modified U3 $\Delta$ 2,3,4 gene expressed in strains lacking Rsa1p ( $\Delta RSA1$ ) and in strain lacking Hit1p ( $\Delta HIT1$ ). This U3 gene accumulates pre-U3 $\Delta$ 2,3,4 precursors.

(D) Rsa1p, Hit1p, and Pih1p are required for pre-rRNA processing. Analysis of pre-rRNA processing by Northern blot. Cells were harvested in log phase ( $OD_{600} = 0.8-1$ ) and total RNAs were extracted. Equal amounts of RNA (5  $\mu$ g) were fractionated on 1.2 % glyoxal/DMSO agarose gels and transferred to a Zeta-Probe membrane (Biorad®). The membranes were consecutively hybridized with radio-labeled antisense oligonucleotides probes indicated on the scheme showing various intermediates of the pre-rRNA processing pathway (on the right).

**Figure 2:** The core proteins and assembly factors are associated with U3 precursors.

(A) Scheme of the pre-U3 precursor with its intron and extended 3' end. Positions of the cleavages occurring in this extension are indicated. Synthesis of cDNAs were primed by oligonucleotide Rev (broken arrow) interacting partially with the U3 3' end extension. Regions complementary to the primers PCRa, PCRb, and 3'PCR used for the PCR amplifications are indicated by dotted arrows.

(B, C, D) Extracts prepared from various TAP-tagged strains in log phase: TAP-*NOPI*, TAP-*NOP56* and TAP-*NOP58* (B), TAP-*RSAl*, TAP-*HIT1*, TAP-*PIH1*, TAP-*RVB1* and TAP-*RVB2* (C), TAP-*TAH1* both in log- and stationary phase (D), were purified on IgG sepharose beads and analyzed by RT-PCR with primers specific for the unspliced (unspl.) and spliced (spl.) U3 precursors carrying the full size of the 3' extension.

An untagged strain was used as negative control. Inputs correspond to signals obtained from 10 % of cellular extracts used for IPs.

(E) Tah1p is required for accumulation of Pih1p in stationary phase. Analysis by Western-blot of Pih1p and Tah1p levels in WT,  $\Delta$ *PIH1* and  $\Delta$ *TAH1* strains.

**Figure 3:** Strains lacking both Rsa1p and Pih1p, or Hit1p and Pih1p are not viable. Diploid strains generated from crossbreeding between single null mutants  $\Delta$ *RSAl*,  $\Delta$ *PIH1*, or  $\Delta$ *HIT1* of haploids strains By4741 (MAT $\alpha$ ) and By4742 (MAT $\alpha$ ) were placed under starvation for meiosis and spore formation. The obtained tetrads were dissected with a micromanipulator and haploid spores were deposited on YPD plates and incubated 1 week at 30 °C.

**Figure 4:** Hit1p is required for accumulation of Rsa1p

(A) Growth defect of the single and double null mutants for the *RSAl* and *HIT1* genes was analyzed as in Fig. 1A.

(B) Western blot analysis of equal amounts of total cell lysates from the single and double null mutants transformed with a HA-Hit1p or GFP-Rsa1p expression vector. The tagged proteins were detected by blotting the membrane with  $\alpha$ -HA and  $\alpha$ -GFP antibodies. Protein PMA1 used as a loading control was detected using specific  $\alpha$ -PMA1 antibodies.

**Figure 5:** Rsa1p and Hit1p form a stable heterodimer.

(A) Hit1p interacts in Y2H assay with Rsa1p. Rsa1p and the COOH-terminal fragment (Rsa1p<sub>230-381</sub>; amino acids 230 to 381) fused to the Gal4 activation domain interact with Hit1p fused to the Gal4 DNA binding domain. The interaction allows transcription activation of the *TRP1* reporter gene. Diploids carrying the plasmids expressing the two protein partners were grown on medium lacking Leu, His, and Trp.

(B, C) Co-expression in *E. coli* and co-purification assays. His<sub>6</sub>-tagged Rsa1p and Rsa1p<sub>230-375</sub> proteins were co-expressed with Hit1p (B). His<sub>6</sub>-tagged protein Snu13 was co-expressed with Rsa1p<sub>230-375</sub> and Hit1p (C). His-pull down assays from the crude cellular extracts were performed by IMAC. The proteins retained on cobalt resin were resolved by SDS-PAGE and stained with Coomassie blue (panels on the left). The ladder (in kDa) is indicated on the right. The retained complexes were eluted by site-specific cleavage of the His<sub>6</sub> tag with PreScission protease and the protein analyzed by mass spectrometry.

(D, E) Noncovalent mass spectrometry detection of the 1/1 Rsa1p<sub>230-375</sub>-Hit1p and 1/1/1 Snu13p-Rsa1p<sub>230-375</sub>-Hit1p complexes. NanoESI mass spectra were obtained under non-denaturing conditions for Rsa1p<sub>230-375</sub>-Hit1p (10  $\mu$ M in 500 mM ammonium acetate pH 7.5) (V<sub>c</sub> = 140 V; P<sub>i</sub> = 6 mbar) (D) and for Snu13p-Rsa1p<sub>230-375</sub>-Hit1p (10  $\mu$ M in 1M ammonium acetate pH 7.5) (V<sub>c</sub> = 80 V; P<sub>i</sub> = 6 mbar) (E).

(F) The trimeric complex His<sub>6</sub>Snu13p-Rsa1p<sub>230-375</sub>-Hit1p interacts *in vitro* with RNAs containing a K-turn RNA motif. Formation of the proteins-RNA complexes was assessed EMSA. Radiolabeled U3 $\Delta$ 2.3.4 and U14 RNAs containing K-turn RNA motif bound by Snu13p were incubated with various combinations of recombinant proteins and purified complexes obtained by co-expression in *E. coli*: His<sub>6</sub>Snu13p, His<sub>6</sub>Snu13p-Rsa1p<sub>230-375</sub>, or the trimeric complex His<sub>6</sub>Snu13p-Rsa1p<sub>230-375</sub>-Hit1p. The formed complexes were resolved by electrophoresis on 6 % non denaturing polyacrylamide gels.

**Figure 6:** Protein Nop58 interacts with purified recombinant complexes His<sub>6</sub>Tah1p-Pih1p and His<sub>6</sub>Rsa1p-Hit1p in His pull-down experiments. Nop58p was translated *in vitro* in bacterial S30 lysate in presence of <sup>3</sup>H methionine and assayed for binding to purified His<sub>6</sub>Tah1p-Pih1p, His<sub>6</sub>Rsa1p<sub>230-375</sub>-Hit1p, and His<sub>6</sub>Snu13p-

Rsalp<sub>230-375</sub>-Hit1p. Input represents 10% of total labeled protein used in interaction assay.

## REFERENCES

- Allmang, C., P. Carbon, and A. Krol. 2002. The SBP2 and 15.5 kD/Snu13p proteins share the same RNA binding domain: identification of SBP2 amino acids important to SECIS RNA binding. *RNA* **8**:1308-18.
- Atzorn, V., P. Fragapane, and T. Kiss. 2004. U17/snr30 is a ubiquitous snoRNA with two conserved sequence motifs essential for 18S rRNA production. *Mol Cell Biol* **24**:1769-78.
- Boulon, S., N. Marmier-Gourrier, B. Pradet-Balade, L. Wurth, C. Verheggen, B. E. Jady, B. Rothe, C. Pescia, M. C. Robert, T. Kiss, B. Bardoni, A. Krol, C. Branlant, C. Allmang, E. Bertrand, and B. Charpentier. 2008. The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J Cell Biol* **180**:579-95.
- Brito, G. C., A. A. Fachel, A. L. Vettore, G. M. Vignal, E. R. Gimba, F. S. Campos, M. A. Barcinski, S. Verjovski-Almeida, and E. M. Reis. 2008. Identification of protein-coding and intronic noncoding RNAs down-regulated in clear cell renal carcinoma. *Mol Carcinog* **47**:757-67.
- Cahill, N. M., K. Friend, W. Speckmann, Z. H. Li, R. M. Terns, M. P. Terns, and J. A. Steitz. 2002. Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snoRNP. *EMBO J* **21**:3816-28.
- Charpentier, B., S. Muller, and C. Branlant. 2005. Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res* **33**:3133-44.
- Clery, A., V. Bourguignon-Igel, C. Allmang, A. Krol, and C. Branlant. 2007. An improved definition of the RNA-binding specificity of SECIS-binding protein 2, an essential component of the selenocysteine incorporation machinery. *Nucleic Acids Res* **35**:1868-84.
- Clery, A., V. Senty-Segault, F. Leclerc, H. A. Raue, and C. Branlant. 2007. Analysis of sequence and structural features that identify the B/C motif of U3 small nucleolar RNA as the recognition site for the Snu13p-Rrp9p protein pair. *Mol Cell Biol* **27**:1191-206.
- Copeland, P. R., J. E. Fletcher, B. A. Carlson, D. L. Hatfield, and D. M. Driscoll. 2000. A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J* **19**:306-14.
- Darzacq, X., N. Kittur, S. Roy, Y. Shav-Tal, R. H. Singer, and U. T. Meier. 2006. Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *J Cell Biol* **173**:207-18.
- Dez, C., J. Noaillac-Depeyre, M. Caizergues-Ferrer, and Y. Henry. 2002. Naf1p, an essential nucleoplasmic factor specifically required for accumulation of box H/ACA small nucleolar RNPs. *Mol Cell Biol* **22**:7053-65.
- Diebold, M. L., S. Fribourg, M. Koch, T. Metzger, and C. Romier. 2011. Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. *J Struct Biol* **175**:178-88.
- Eckert, K., J. M. Saliou, L. Monlezun, A. Vigouroux, N. Atmane, C. Caillat, S. Quevillon-Cheruel, K. Masion, M. Nicaise, S. Lazereg, A. Van Dorsselaer, S. Sanglier-Cianferani, P. Meyer, and S. Morera. 2010. The Pih1-Tah1 cochaperone complex inhibits Hsp90 molecular chaperone ATPase activity. *J Biol Chem* **285**:31304-12.
- Fatica, A., M. Dlakic, and D. Tollervey. 2002. Naf1 p is a box H/ACA snoRNP assembly factor. *RNA* **8**:1502-14.
- Gonzales, F. A., N. I. Zanchin, J. S. Luz, and C. C. Oliveira. 2005. Characterization of *Saccharomyces cerevisiae* Nop17p, a novel Nop58p-interacting protein that is involved in Pre-rRNA processing. *J Mol Biol* **346**:437-55.
- Granneman, S., J. Vogelzangs, R. Luhrmann, W. J. van Venrooij, G. J. Pruijn, and N. J. Watkins. 2004. Role of pre-rRNA base pairing and 80S complex formation in subnucleolar localization of the U3 snoRNP. *Mol Cell Biol* **24**:8600-10.
- He, F., T. Umehara, K. Tsuda, M. Inoue, T. Kigawa, T. Matsuda, T. Yabuki, M. Aoki, E. Seki, T. Terada, M. Shirouzu, A. Tanaka, S. Sugano, Y. Muto, and S. Yokoyama. 2007. Solution structure of the zinc finger HIT domain in protein FON. *Protein Sci* **16**:1577-87.
- Henras, A., Y. Henry, C. Bousquet-Antonelli, J. Noaillac-Depeyre, J. P. Gelugne, and M. Caizergues-Ferrer. 1998. Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *EMBO J* **17**:7078-90.
- Hiley, S. L., T. Babak, and T. R. Hughes. 2005. Global analysis of yeast RNA processing identifies new targets of RNase III and uncovers a link between tRNA 5' end processing and tRNA splicing. *Nucleic Acids Res* **33**:3048-56.
- Ito, T., T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, and Y. Sakaki. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A* **98**:4569-74.
- Iwahashi, H., K. Yamagata, I. Yoshiuchi, J. Terasaki, Q. Yang, K. Fukui, A. Ihara, Q. Zhu, T. Asakura, Y. Cao, A. Imagawa, M. Namba, T. Hanafusa, J. Miyagawa, and Y. Matsuzawa. 2002. Thyroid hormone receptor interacting protein 3 (trip3) is a novel coactivator of hepatocyte nuclear factor-4alpha. *Diabetes* **51**:910-4.
- Kawakami, K., B. K. Shafer, D. J. Garfinkel, J. N. Strathern, and Y. Nakamura. 1992. Ty element-induced temperature-sensitive mutations of *Saccharomyces cerevisiae*. *Genetics* **131**:821-32.
- Klein, D. J., T. M. Schmeing, P. B. Moore, and T. A. Steitz. 2001. The kink-turn: a new RNA secondary structure motif. *EMBO J* **20**:4214-21.
- Kolb, S. J., D. J. Battle, and G. Dreyfuss. 2007. Molecular functions of the SMN complex. *J Child Neurol* **22**:990-4.
- Koonin, E. V., P. Bork, and C. Sander. 1994. A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins and a ribosome modification enzyme? *Nucleic Acids Res* **22**:2166-7.
- Koppen, A., R. Houtman, D. Pijnenburg, E. H. Jeninga, R. Ruijtenbeek, and E. Kalkhoven. 2009. Nuclear receptor-coregulator interaction profiling identifies TRIP3 as a novel peroxisome proliferator-activated receptor gamma cofactor. *Mol Cell Proteomics* **8**:2212-26.
- Kufel, J., C. Allmang, G. Chanfreau, E. Petfalski, D. L. Lafontaine, and D. Tollervey. 2000. Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. *Mol Cell Biol* **20**:5415-24.
- Lafontaine, D. L., and D. Tollervey. 1999. Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. *RNA* **5**:455-67.
- Lafontaine, D. L., and D. Tollervey. 2000. Synthesis and assembly of the box C+D small nucleolar RNPs. *Mol Cell Biol* **20**:2650-9.
- Lovat, P. E., M. Annicchiarico-Petruzzelli, M. Corazzari, M. G. Dobson, A. J. Malcolm, A. D. Pearson, G. Melino, and C. P. Redfern. 1999. Differential effects of retinoic acid isomers on the expression of nuclear receptor co-regulators in neuroblastoma. *FEBS Lett* **445**:415-9.
- Marmier-Gourrier, N., A. Clery, V. Senty-Segault, B. Charpentier, F. Schlotter, F. Leclerc, R. Fournier, and C. Branlant. 2003. A structural, phylogenetic, and



- functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA. *RNA* **9**:821-38.
32. **Matera, A. G., R. M. Terns, and M. P. Terns.** 2007. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol* **8**:209-20.
  33. **McKeegan, K. S., C. M. Debieux, S. Boulon, E. Bertrand, and N. J. Watkins.** 2007. A dynamic scaffold of pre-snoRNP factors facilitates human box C/D snoRNP assembly. *Mol Cell Biol* **27**:6782-93.
  34. **McKeegan, K. S., C. M. Debieux, and N. J. Watkins.** 2009. Evidence that the AAA+ proteins TIP48 and TIP49 bridge interactions between 15.5K and the related NOP56 and NOP58 proteins during box C/D snoRNP biogenesis. *Mol Cell Biol* **29**:4971-81.
  35. **Miguet, L., G. Bechade, L. Fornecker, E. Zink, C. Felden, C. Gervais, R. Herbrecht, A. Van Dorselaer, L. Mauvieux, and S. Sanglier-Cianferani.** 2009. Proteomic analysis of malignant B-cell derived microparticles reveals CD148 as a potentially useful antigenic biomarker for mantle cell lymphoma diagnosis. *J Proteome Res* **8**:3346-54.
  36. **Morrissey, J. P., and D. Tollervey.** 1993. Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol Cell Biol* **13**:2469-77.
  37. **Nottrott, S., K. Hartmuth, P. Fabrizio, H. Urlaub, I. Vidovic, R. Ficner, and R. Luhrmann.** 1999. Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snoRNP protein with the 5' stem-loop of U4 snRNA. *EMBO J* **18**:6119-33.
  38. **Omer, A. D., S. Ziesche, H. Ehardt, and P. P. Dennis.** 2002. In vitro reconstitution and activity of a C/D box methylation guide ribonucleoprotein complex. *Proc Natl Acad Sci U S A* **99**:5289-94.
  39. **Patel, S. B., and M. Bellini.** 2008. The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res* **36**:6482-93.
  40. **Pellizzoni, L.** 2007. Chaperoning ribonucleoprotein biogenesis in health and disease. *EMBO Rep* **8**:340-5.
  41. **Peng, W. T., M. D. Robinson, S. Mnaimneh, N. J. Krogan, G. Cagney, Q. Morris, A. P. Davierwala, J. Grigull, X. Yang, W. Zhang, N. Mitsakakis, O. W. Ryan, N. Datta, V. Jojic, C. Pal, V. Canadien, D. Richards, B. Beattie, L. F. Wu, S. J. Altschuler, S. Roweis, B. J. Frey, A. Emili, J. F. Greenblatt, and T. R. Hughes.** 2003. A panoramic view of yeast noncoding RNA processing. *Cell* **113**:919-33.
  42. **Rashid, R., M. Aittaleb, Q. Chen, K. Spiegel, B. Demeler, and H. Li.** 2003. Functional requirement for symmetric assembly of archaeal box C/D small ribonucleoprotein particles. *J Mol Biol* **333**:295-306.
  43. **Rigaut, G., A. Shevchenko, B. Rutz, M. Wilm, M. Mann, and B. Seraphin.** 1999. A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**:1030-2.
  44. **Romier, C., M. Ben Jelloul, S. Albeck, G. Buchwald, D. Busso, P. H. Celie, E. Christodoulou, V. De Marco, S. van Gerwen, P. Knipscheer, J. H. Lebbink, V. Notenboom, A. Poterszman, N. Rochel, S. X. Cohen, T. Unger, J. L. Sussman, D. Moras, T. K. Sixma, and A. Perrakis.** 2006. Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. *Acta Crystallogr D Biol Crystallogr* **62**:1232-42.
  45. **Sanglier, S., C. Atmanene, G. Chevreux, and A. V. Dorselaer.** 2008. Nondenaturing mass spectrometry to study noncovalent protein/protein and protein/ligand complexes: technical aspects and application to the determination of binding stoichiometries. *Methods Mol Biol* **484**:217-43.
  46. **Schultz, A., S. Nottrott, N. J. Watkins, and R. Luhrmann.** 2006. Protein-protein and protein-RNA contacts both contribute to the 15.5K-mediated assembly of the U4/U6 snRNP and the box C/D snoRNPs. *Mol Cell Biol* **26**:5146-54.
  47. **Tran, E. J., X. Zhang, and E. S. Maxwell.** 2003. Efficient RNA 2'-O-methylation requires juxtaposed and symmetrically assembled archaeal box C/D and C'/D' RNPs. *EMBO J* **22**:3930-40.
  48. **Vidovic, I., S. Nottrott, K. Hartmuth, R. Luhrmann, and R. Ficner.** 2000. Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol Cell* **6**:1331-42.
  49. **Wang, C., and U. T. Meier.** 2004. Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *EMBO J* **23**:1857-67.
  50. **Watkins, N. J., A. Dickmanns, and R. Luhrmann.** 2002. Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Mol Cell Biol* **22**:8342-52.
  51. **Watkins, N. J., V. Segault, B. Charpentier, S. Nottrott, P. Fabrizio, A. Bachi, M. Wilm, M. Rosbash, C. Branlant, and R. Luhrmann.** 2000. A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP. *Cell* **103**:457-66.
  52. **Yang, P. K., G. Rotondo, T. Porras, P. Legrain, and G. Chanfreau.** 2002. The Shq1p.Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis. *J Biol Chem* **277**:45235-42.
  53. **Yong, J., L. Wan, and G. Dreyfuss.** 2004. Why do cells need an assembly machine for RNA-protein complexes? *Trends Cell Biol* **14**:226-32.
  54. **Zhao, R., and W. A. Houry.** 2005. Hsp90: a chaperone for protein folding and gene regulation. *Biochem Cell Biol* **83**:703-10.
  55. **Zhao, R., Y. Kakahara, A. Gribun, J. Huen, G. Yang, M. Khanna, M. Costanzo, R. L. Brost, C. Boone, T. R. Hughes, C. M. Yip, and W. A. Houry.** 2008. Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation. *J Cell Biol* **180**:563-78.

## FUNDING

This work was supported by the Agence Nationale de la Recherche [ANR-06-BLAN-0208], the ARC, the french Ministère de l'Enseignement Supérieur et de la Recherche [doctoral fellowships to B.R. and R.B.]; the Centre National de la Recherche Scientifique; the Pôle de Recherche Scientifique et Technologique "Bioingénierie" of Région lorraine, the University of Strasbourg (UdS), the Fondation pour la Recherche Médicale and the Region Alsace. J.M.S. acknowledges the Sidaction foundation financial support.

	Protein name	Accession number	Molecular weight (Da)	No TAP		TAP-PIH1		TAP-RSA1		TAP-HIT1	
				Unique pep (% cover)	Spectral Count	Unique pep (% cover)	Spectral Count	Unique pep (% cover)	Spectral Count	Unique pep (% cover)	Spectral Count
	Rsa1p	Q08932	43 981	-	-	-	-	10 (35,4)	81	10 (32,0)	77
	Hit1p	P46973	18 358	-	-	2 (20,7)	2	7 (54,9)	31	7 (54,9)	49
Box C/D core protein	Snu13p	P39990	13 551	-	-	2 (36,5)	6	2 (36,5)	9	2 (36,5)	8
	Nop1p	P15646	34 195	2 (10,4)	4	4 (21,4)	5	5 (24,5)	9	6 (27,8)	9
	Nop56p	Q12460	56 849	-	-	-	-	5 (20,4)	7	-	-
	Nop58p	A6ZPE5	57 176	-	-	2 (6,8)	3	-	-	-	-
R2TP protein	Pih1p	P38768	39 503	-	-	13 (51,7)	58	-	-	-	-
	Tah1p	P25638	12 493	-	-	4 (69,4)	33	-	-	-	-
	Rvb1p	Q03940	50 436	-	-	16 (47,7)	25	-	-	-	-
	Rvb2p	Q12464	51 595	-	-	17 (46,9)	33	-	-	-	-

Table 1

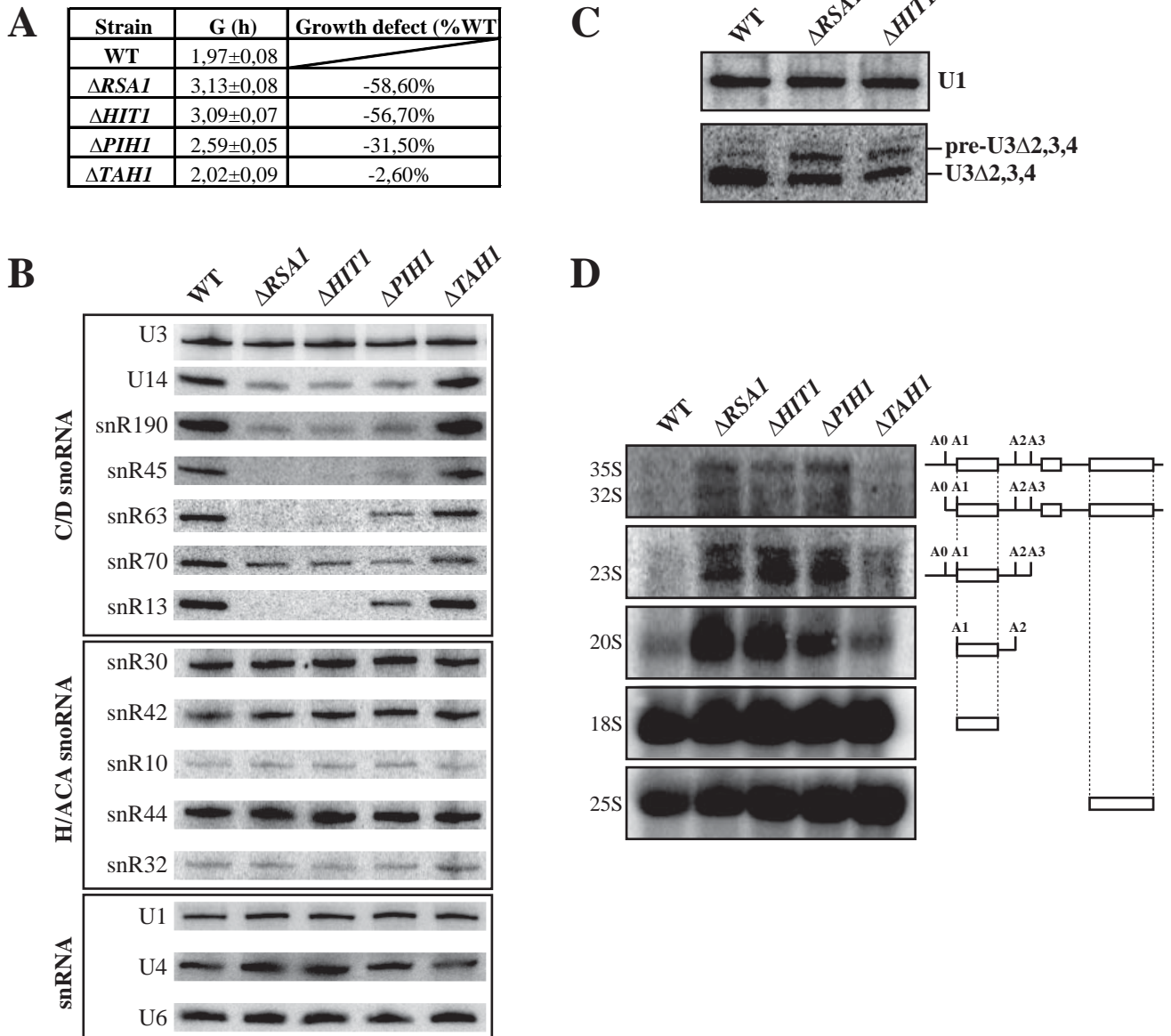
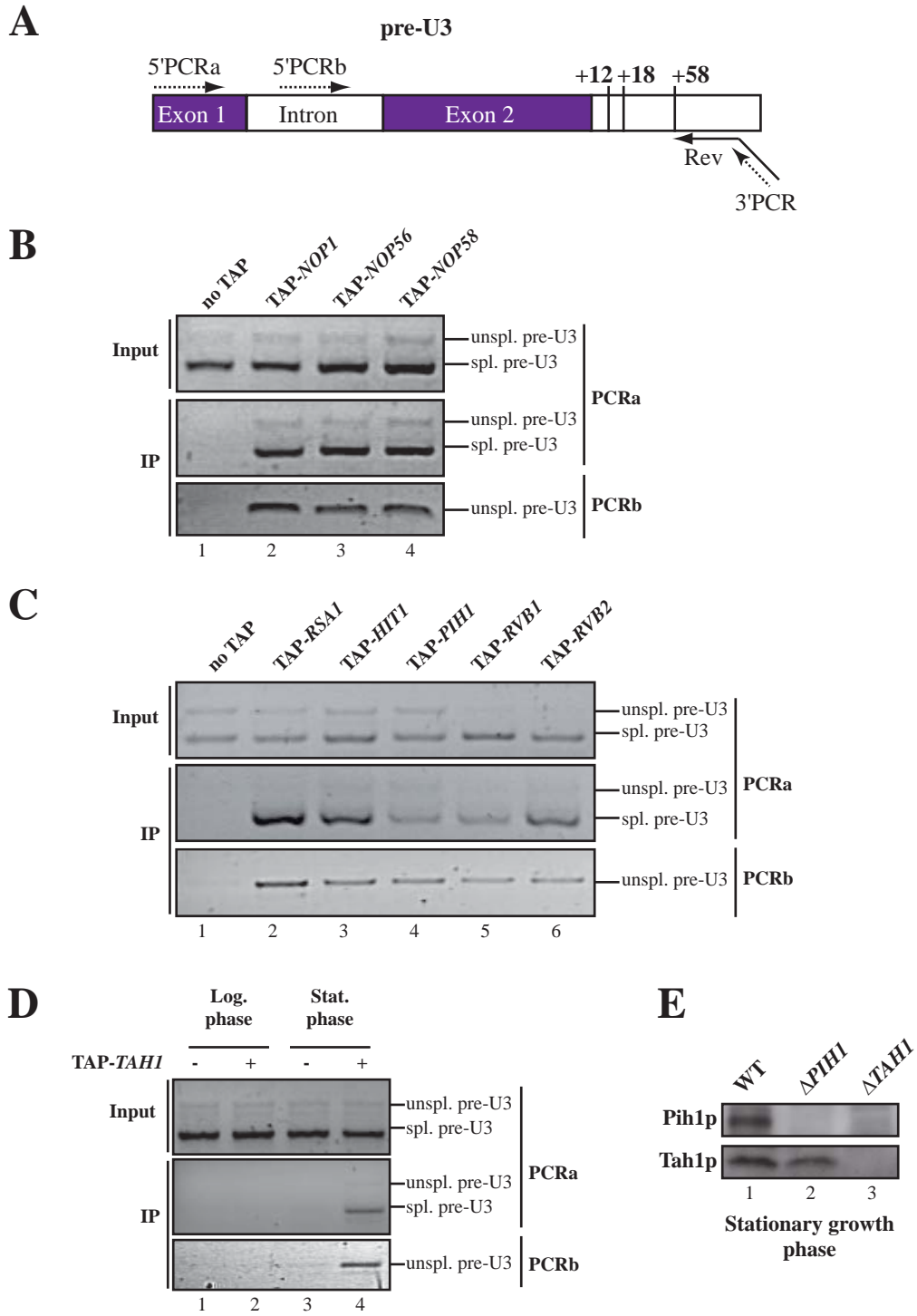
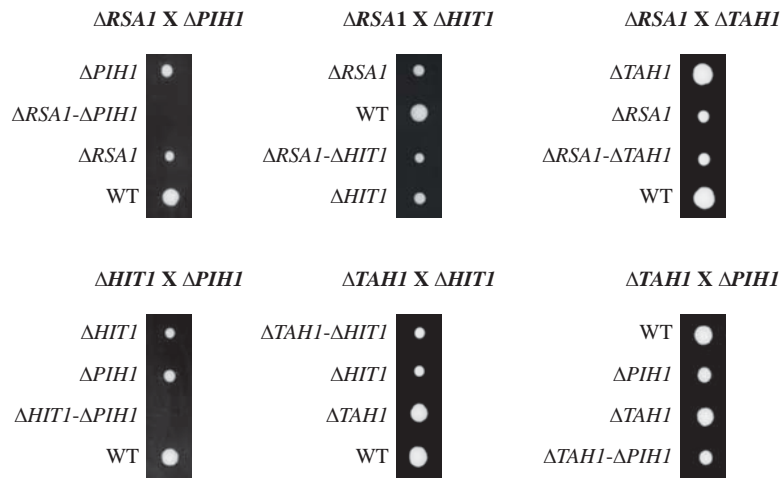


Figure 1



**Figure 2**

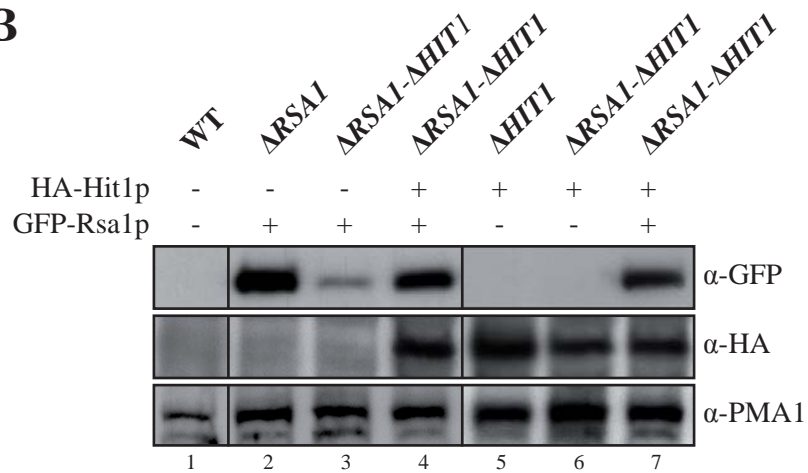


**Figure 3**

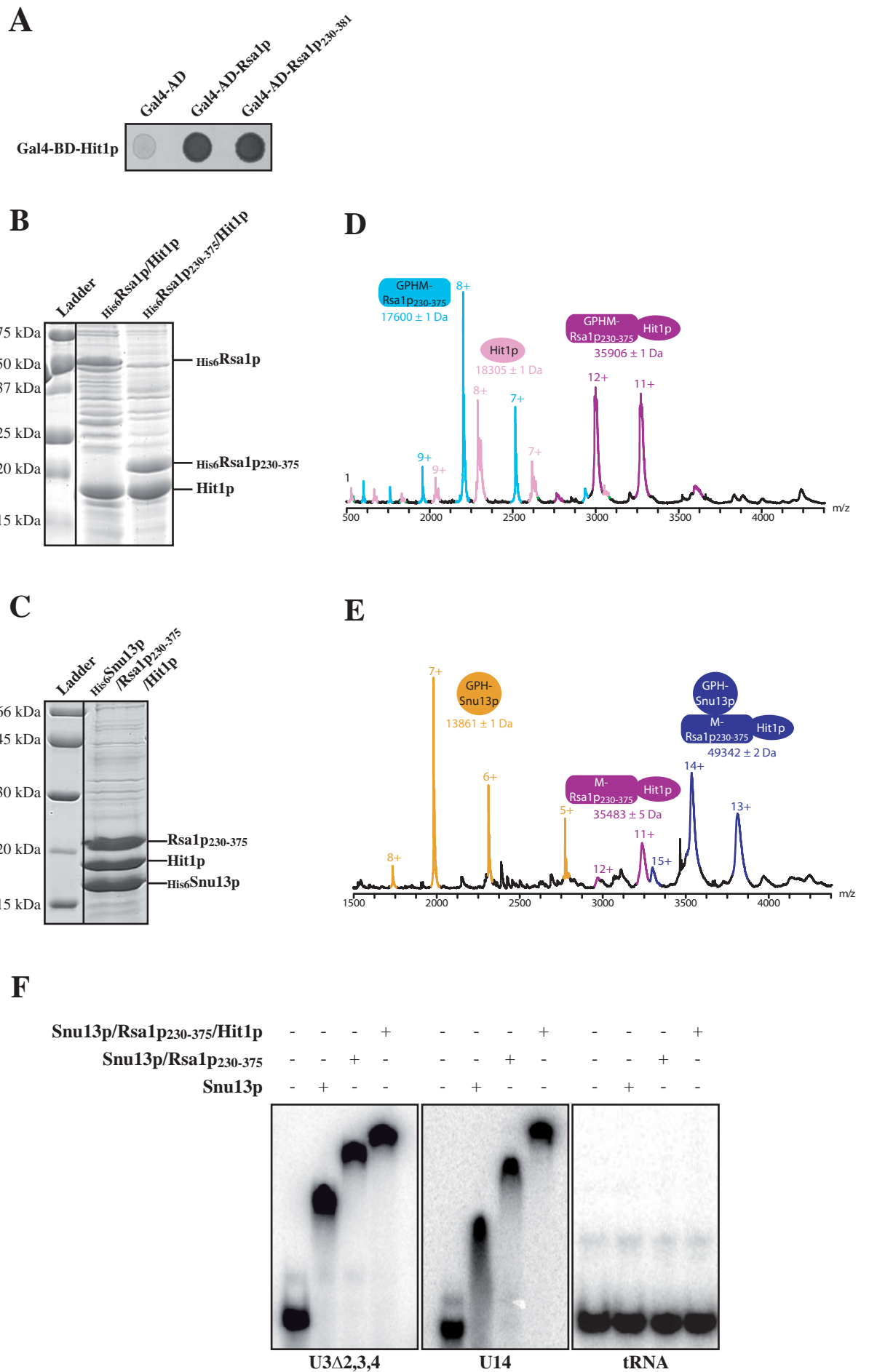
**A**

Strain	G (h)	Growth defect (%WT)
WT	1,97±0,08	
<i>ΔRSAI</i>	3,13±0,08	58,60%
<i>ΔHITI</i>	3,09±0,07	56,70%
<i>ΔRSAI-ΔHITI</i>	3,21±0,08	62,70%

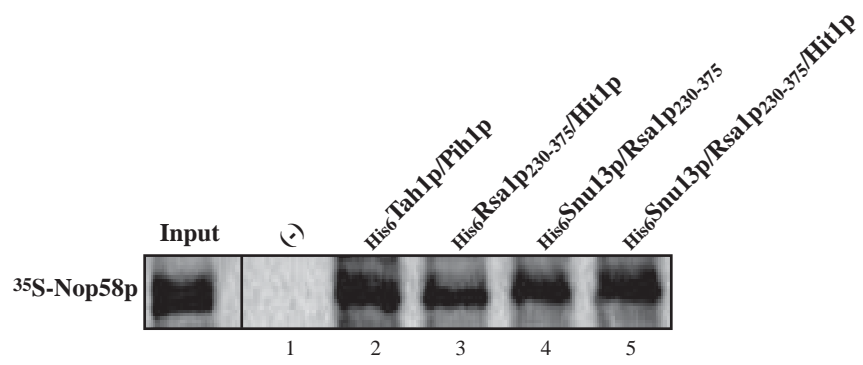
**B**



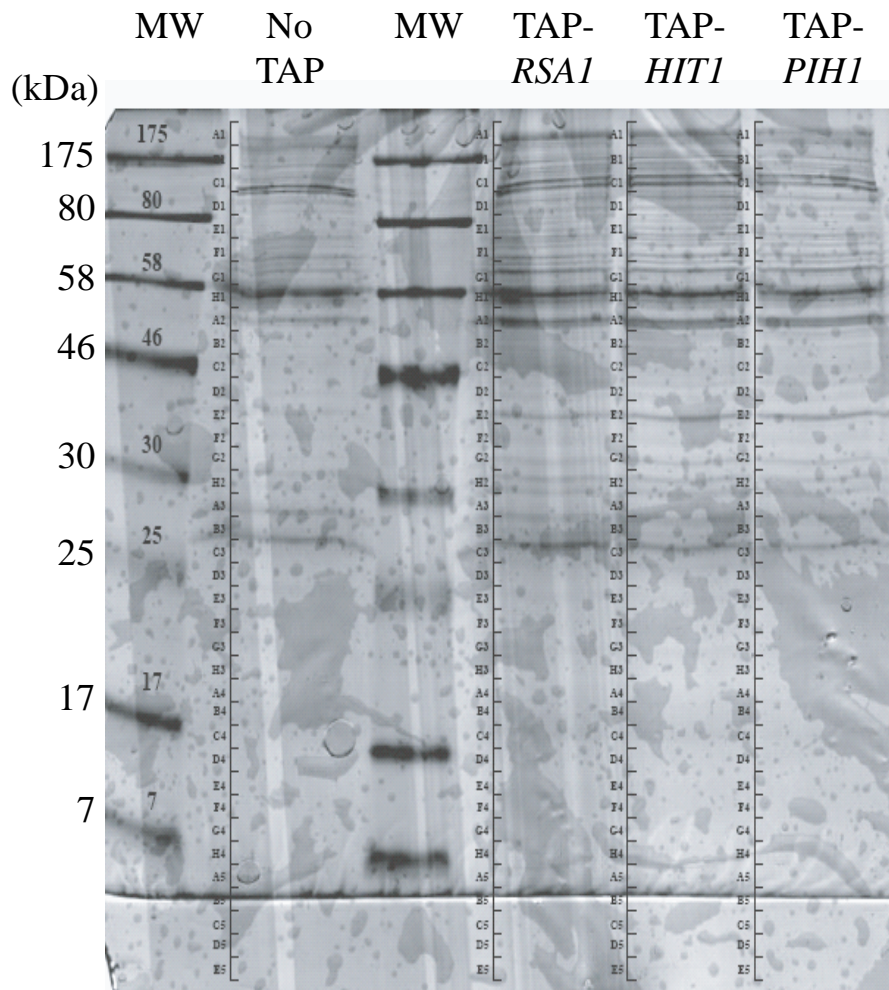
**Figure 4**



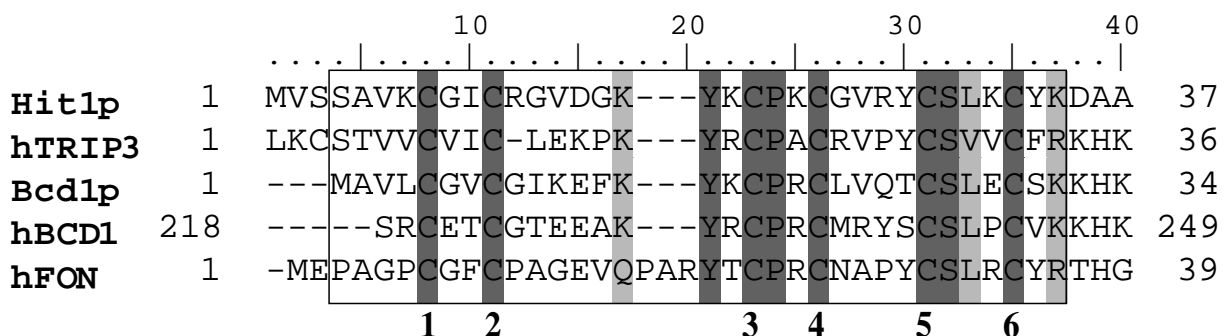
**Figure 5**



**Figure 6**



**Figure S1:** SDS-PAGE analysis of proteins contained in the extracts prepared from tandem affinity experiments of TAP-Rsa1p, TAP-Hit1p, TAP-Pih1p, and an untagged negative control (No TAP). Proteins were separated by 12,5% SDS-PAGE and stained with colloidal Coomassie Blue. Protein markers were fractionated on the left of the gel and their molecular weights in kDa are given.



**Figure S2:** Zinc finger HIT domain sequence. Multiple amino acid sequence alignments were constructed using CLUSTALW. Identical residues are highlighted in dark grey and similar residues in light grey. Zf-HIT domain is framed and zinc-finger forming cysteines are numbered below the alignment.



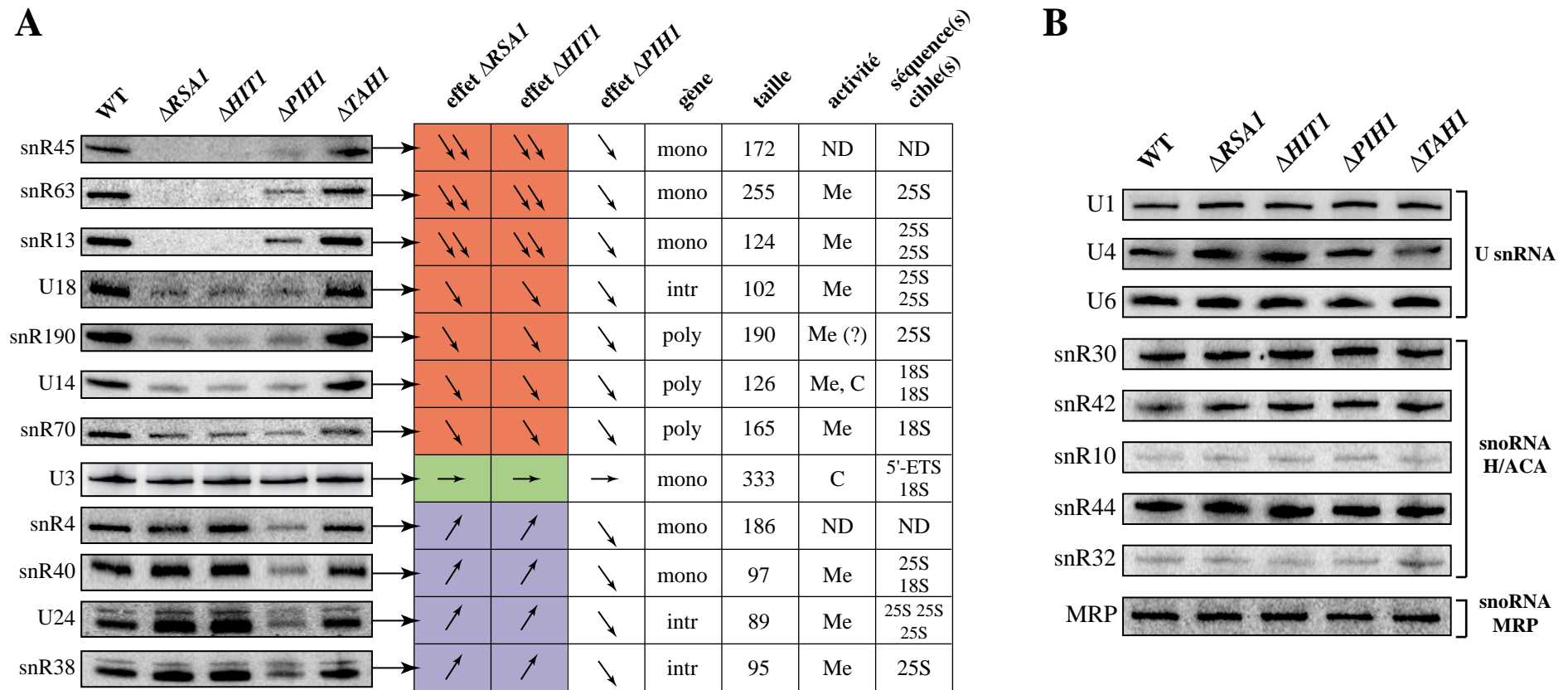


## 2.2. Données complémentaires

Dans l'article *Functional link between proteins Rsa1 and Hit1 during box C/D snoRNPs biogenesis in yeast*, nous avons montré que l'absence des facteurs d'assemblage Rsa1p, Hit1p et Pih1p conduit à une diminution spécifique du taux de snoRNA à boîtes C/D. En plus des données déjà présentées, nous avons réalisé des expériences de Northern-blot supplémentaires en ciblant d'autres snoRNA. De façon surprenante, les données obtenues indiquent que certains snoRNA à boîtes C/D ne sont pas sensibles à la délétion des gènes *RSA1* et *HIT1*, suggérant que le processus d'assemblage des RNP pourrait présenter des variations en fonction de certains sous-type d'ARN. Afin de compléter les résultats de co-immunosélection entre les facteurs d'assemblage et les précurseurs du snoRNA U3, nous avons également réalisé des expériences d'immunoprécipitation de la chromatine. La nature des interactions établies entre les protéines Rsa1p, Pih1p et Nop58p, a été précisée en utilisant diverses approches de double-hybride, de co-expression chez *E. coli* et de pull-down. Enfin, en collaboration avec C. Saveanu à l'institut Pasteur, des cribles d'interactions génétiques à partir des souches de levures invalidées pour les gènes *RSA1*, *PIH1* et *TAH1* ont été réalisés.

### 2.2.1. Rsa1p et Hit1p semblent spécifiques d'une sous-population de snoRNA à boîtes C/D

L'une des caractéristiques principales des facteurs d'assemblage des snoRNP à boîtes C/D est d'exercer une influence sur le taux de snoRNA. En effet, chez la levure, les protéines Rsa1p, Hit1p, Pih1p, Rvb1p et Rvb2p se sont toutes avérées requises pour l'accumulation de ces ARN (Rothé *et al.*, Publication n°3 ; (King *et al.*, 2001; Zhao *et al.*, 2008)). Cependant, en élargissant l'analyse phénotypique des souches de levure  $\Delta RSA1$ ,  $\Delta HIT1$ ,  $\Delta PIH1$  et  $\Delta TAH1$ , à d'autres snoRNA à boîtes C/D, nous avons constaté que les différents facteurs n'exercent pas tous la même influence (**Fig. 65**). En effet, hormis le snoRNA U3, relativement stable dans les différents contextes génétiques, deux principales catégories se distinguent parmi les snoRNA à boîtes C/D (**Fig. 65A**) : (i) des snoRNA dont le taux est diminué en absence de Rsa1p et Hit1p (*i.e.* snR45, snR63, snR13, U18, snR190, U14, snR70), (ii) des snoRNA dont le taux est augmenté en absence de ces deux facteurs (*i.e.* snR4, snR40, U24, snR38). Ce résultat pose donc la question de savoir si Rsa1p et Hit1p sont réellement impliquées dans la biogenèse de l'ensemble des snoRNA à boîtes C/D ? A titre de comparaison, la délétion du gène *PIH1* se



**Figure 65 : L'accumulation de certains snoRNA à boîtes C/D est indépendante de Rsa1p et Hit1p.** Analyse du taux de plusieurs snoRNA à boîtes C/D (A) et de divers autres petits ARN nucléaires et nucléolaires (B) dans les souches de levure WT,  $\Delta$ RSAl,  $\Delta$ HIT1,  $\Delta$ PIH1 et  $\Delta$ TAH1. Les différents ARN ont été détectés en utilisant l'approche de Northern-blot. Le snRNA U1 a été utilisé comme contrôle interne. Les flèches descendantes symbolisent une diminution du taux de snoRNA. Les flèches horizontales symbolisent un taux de snoRNA constant. Les flèches ascendantes symbolisent une augmentation du taux de snoRNA. Le fond rouge représente la population des snoRNA affectés par l'absence de Rsa1p et Hit1p. Le fond vert symbolise le snoRNA U3, peu sensible à l'absence des facteurs d'assemblage. Le fond bleu représente les snoRNA dont le taux augmente en absence de Rsa1p et Hit1p. mono : monocistronique; poly : polycistronique; intr : intronique; ND : Non Déterminé; Me : activité de 2'-O-méthylation; C : Implication dans les étapes de clivage endonucléolytiques; (?) : activité prédite mais non démontrée expérimentalement.

répercute de façon légèrement plus faible sur le taux de snoRNA, mais elle affecte de façon homogène l'ensemble des ARN classés dans ces deux groupes.

Par ailleurs, nous avons également vérifié le taux d'autres ARN non codants dans les différentes souches étudiées (**Fig. 65B**). L'absence des protéines Rsa1p, Hit1p, Pih1p et Tah1p n'affecte ni le taux de snoRNA à boîtes H/ACA (Rothé *et al.*, Publication n°3), ni celui du snoRNA de la RNase MRP. En revanche, parmi les UsnRNA, l'ARN U4 est légèrement plus abondant dans les souches  $\Delta$ RSA1 et  $\Delta$ HIT1 (**Fig. 65B**) (Boulon *et al.*, 2008). Cette tendance, également observée pour les snoRNA à boîtes C/D snR4, snR40, U24 et snR38, pourrait résulter d'un phénomène de balance. En effet, la protéine Snu13p étant commune aux snoRNA à boîtes C/D et au snRNA U4, un défaut d'assemblage des RNP dépendantes du duo Rsa1p-Hit1p pourrait augmenter la concentration en protéines constitutives disponibles et conduire de ce fait, à une production accrue des autres particules.

Afin de comprendre sur quelles caractéristiques fonctionnelles et/ou structurales repose la distinction entre les deux groupes de snoRNA à boîtes C/D, nous avons comparé leur nature génétique, leur taille ou encore leurs paramètres d'activité (cibles, type d'activité) (**Fig. 65A**), mais aucun de ces critères ne s'est révélé spécifique de l'un des deux groupes. Une possibilité est que ces différences de sensibilité reposent sur des éléments de séquence et/ou de structure secondaire. Une perspective de ce travail sera donc de réaliser une analyse de séquence comparative entre les deux sous-groupes de snoRNA. Une telle démarche pourrait notamment être facilitée par l'utilisation de puces à ARN (Hiley *et al.*, 2005; Peng *et al.*, 2003), permettant l'acquisition de plus larges quantités de données.

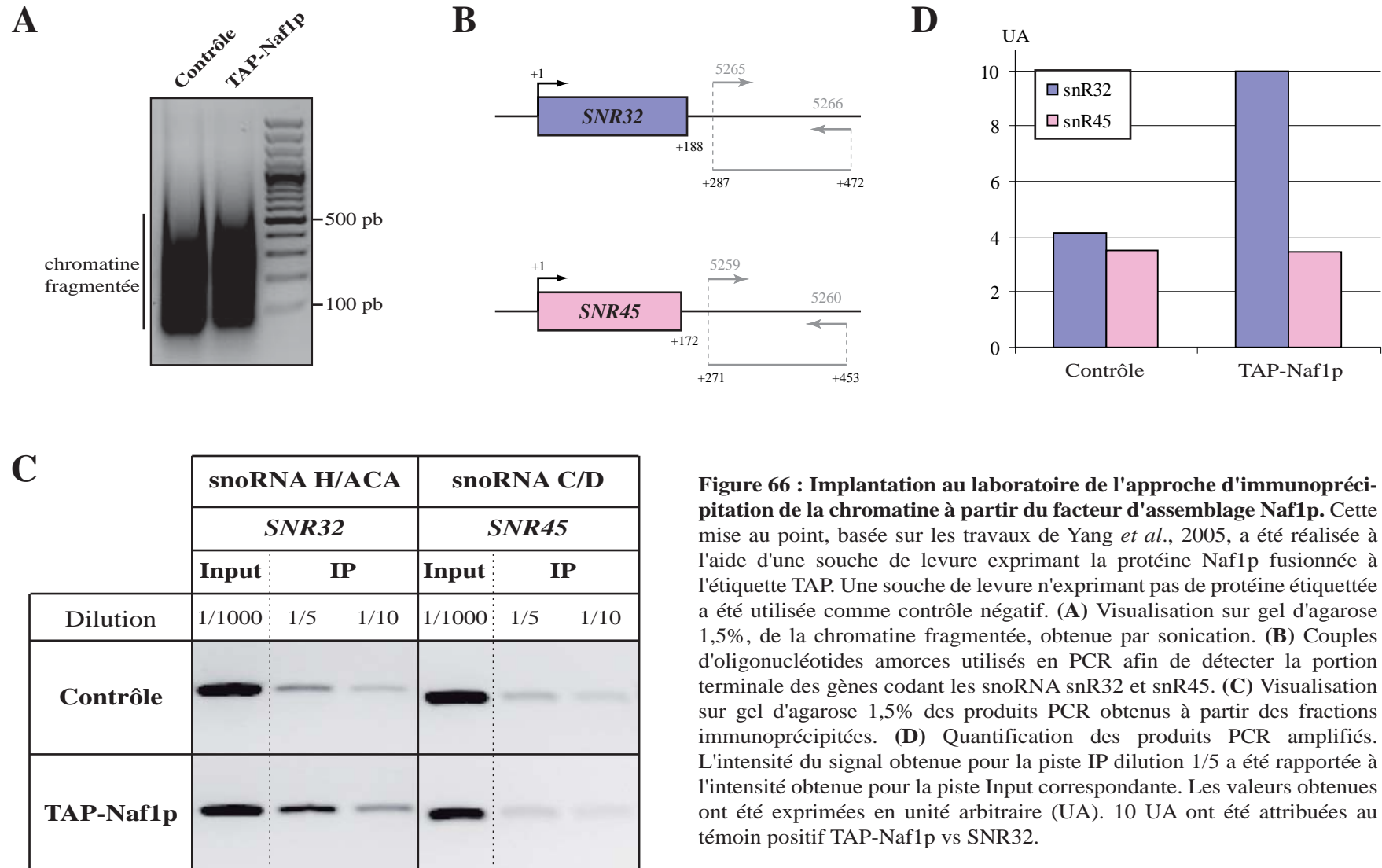
### 2.2.2. Etude des liens entre transcription et assemblage des snoRNP

En utilisant une approche d'immunosélection, couplée à une analyse par RT-PCR, nous avons montré que les protéines Nop1p, Nop56p et Nop58p, ainsi que la majorité des facteurs d'assemblage sont associés aux précurseurs du snoRNA U3 (Rothé *et al.*, Publication n°3). Ces données confirment donc l'intervention transitoire de la machinerie d'assemblage au cours de la biogenèse des snoRNP, en revanche, elles ne permettent pas de conclure quant à la nature co-transcriptionnelle de ce processus. Afin d'étudier plus spécifiquement s'il existe des liens entre transcription et assemblage des snoRNP à boîtes C/D, nous avons utilisé

l'approche d'immunoprécipitation de la chromatine. Cette dernière, mise à profit pour l'étude des snoRNP à boîtes H/ACA, avait déjà permis de mettre en évidence l'intervention co-transcriptionnelle du facteur d'assemblage Naf1p/NAF1 (Ballarino *et al.*, 2005; Darzacq *et al.*, 2006; Yang *et al.*, 2005) (cf. Introduction, Chapitre 3, 2.3.1).

Dans le but de mettre en place cette technique au laboratoire, nous avons cherché, dans un premier temps, à reproduire les résultats obtenus pour le facteur Naf1p de levure (**Fig. 66**). Ces expériences ont été réalisées à partir de souches de levure exprimant les protéines d'intérêt fusionnées à l'étiquette TAP et cultivées en phase exponentielle de croissance. Une seule étape de purification, effectuée à l'aide de l'épitope ProtA, a été réalisée, et une souche n'exprimant aucune protéine de fusion, a été utilisée comme témoin négatif. Après un traitement au formaldéhyde, permettant de ponter covalamment les protéines à l'ADN, l'ARN et à d'autres protéines (Orlando, 2000), les cellules ont été lysées et les extraits cellulaires ont été soumis à une exposition contrôlée aux ultras-sons. Cette étape nous a permis de fragmenter la chromatine, afin d'obtenir des fragments de taille variant entre 100 et 500 pb environ (**Fig. 66A**). Après l'immunosélection des protéines d'intérêt et l'élimination des pontages covalents, les fragments d'ADN co-immunosélectionnés ont fait l'objet d'une analyse ciblée par PCR (**Fig. 66B**). Dans un premier temps, nous avons analysé la portion 3' des gènes codant le snoRNA à boîtes H/ACA snR32 et le snoRNA à boîtes C/D snR45, ces deux *loci* chromosomiques constituant respectivement un témoin positif et un témoin négatif pour la fixation du facteur Naf1p (Yang *et al.*, 2005). Les produits obtenus par PCR ont été déposés sur gel d'agarose et visualisés aux UV en présence de BET (**Fig. 66C**). Cette approche très sensible est associée à un bruit de fond relativement élevé. Il a donc été nécessaire de tester différentes dilution des produits immunoprécipités, afin de discriminer les signaux positifs du bruit de fond. Dans le but d'éviter les erreurs de quantification dues au problème de saturation de la PCR, les inputs (*i.e.* fraction de l'échantillon prélevée avant l'immunoprécipitation) ont également fait l'objet d'une dilution. Comme présenté dans la **Figure 66C**, un enrichissement significatif de la protéine Naf1p sur le *locus SNR32* est constaté. La quantification des signaux obtenus confirme en effet un enrichissement d'un facteur 2,5 par rapport au signal obtenu pour le *locus SNR45*. Des résultats similaires ont été obtenus en incorporant de l' $[\alpha\text{-}^{32}\text{P}]\text{ATP}$  dans les produits PCR (données non montrées).

Une fois le protocole expérimental mis au point, nous avons réalisé cette analyse pour les protéines Nop58p, Rsa1p et Hit1p, les membres du complexes R2TP, les chaperonnes Hsp82p



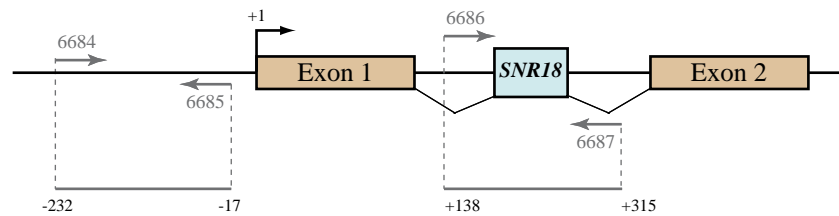
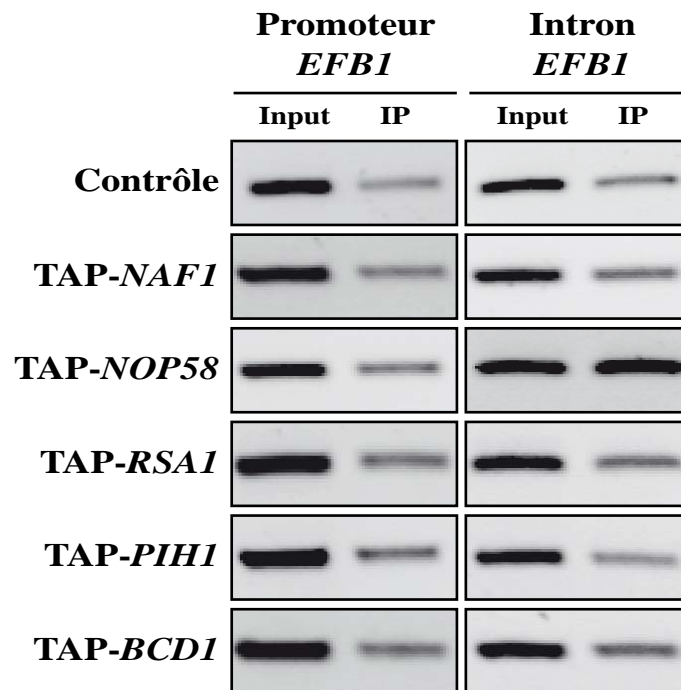
**Figure 66 : Implantation au laboratoire de l'approche d'immunoprécipitation de la chromatine à partir du facteur d'assemblage Naf1p.** Cette mise au point, basée sur les travaux de Yang *et al.*, 2005, a été réalisée à l'aide d'une souche de levure exprimant la protéine Naf1p fusionnée à l'étiquette TAP. Une souche de levure n'exprimant pas de protéine étiquetée a été utilisée comme contrôle négatif. **(A)** Visualisation sur gel d'agarose 1,5%, de la chromatine fragmentée, obtenue par sonication. **(B)** Couples d'oligonucléotides amorces utilisés en PCR afin de détecter la portion terminale des gènes codant les snoRNA snR32 et snR45. **(C)** Visualisation sur gel d'agarose 1,5% des produits PCR obtenus à partir des fractions immunoprécipitées. **(D)** Quantification des produits PCR amplifiés. L'intensité du signal obtenue pour la piste IP dilution 1/5 a été rapportée à l'intensité obtenue pour la piste Input correspondante. Les valeurs obtenues ont été exprimées en unité arbitraire (UA). 10 UA ont été attribuées au témoin positif TAP-Naf1p vs SNR32.

et Hsc82p, ainsi que le facteur Bcd1p (**Fig. 67**). Trois gènes monocistroniques ont été analysés dans un premier temps : les *loci* *SNR32* et *SNR45*, ainsi que le gène codant le snoRNA U3A (*SNR17A*) (**Fig. 67A**). Comme nous l'avons vu précédemment, la protéine Naf1p est détectée spécifiquement sur le *locus* H/ACA *SNR32* (**Fig. 67B**). En revanche, aucun enrichissement spécifique n'a été observé dans le cas des autres facteurs d'assemblage et de la protéine Nop58p. Ce constat a été effectué plusieurs fois, sur la base d'expériences réalisées à partir de larges gammes de dilution des produits immunoprécipités.

Afin d'élargir notre investigation au cas des snoRNA introniques, nous avons également analysé le *locus* *SNR18* (**Fig. 68**). Ce dernier, hébergé au sein de l'intron du gène *EFB1*, code le snoRNA U18, dont le processus de maturation a déjà fait l'objet d'études approfondies (cf. Introduction, Chapitre 1, 2.3.2 & Chapitre 3, 3.5). Ainsi, en utilisant l'approche de ChIP, l'équipe d'I. Bozzoni a notamment démontré que les protéines Nop1p, Nop56p et Nop58p sont recrutées de façon co-transcriptionnelle sur le snoRNA U18 (Vincenti *et al.*, 2007). Nous avons donc reproduit leur démarche pour la protéine Nop58p et pour certains facteurs d'assemblage. Deux régions du gène *EFB1* ont été analysées : le promoteur et la séquence intronique hébergeant le snoRNA (**Fig. 68A**). Comme attendu, la protéine Nop58p est détectée spécifiquement au niveau de la région intronique, en revanche, aucun enrichissement n'a été observé pour les différents facteurs d'assemblage, et ce en dépit de leur importance pour l'accumulation du snoRNA U18 (**Fig. 68B**).

L'ensemble de ces données suggère donc deux faits importants. D'une part, l'absence d'association des facteurs d'assemblage avec les *loci* étudiés suggère que ces derniers ne sont pas associés de façon stable avec l'ARN au cours du processus de transcription. D'autre part, la détection sélective de la protéine Nop58p sur le *locus* *SNR18* révèle que le caractère précoce du processus d'assemblage pourrait être une spécificité des snoRNA introniques. Cependant, une proposition alternative est que l'impossibilité de détecter des signaux quantifiables à partir des *loci* cistroniques découle des caractéristiques structurales des snoRNA à boîtes C/D. A la différence des snoRNA à boîtes H/ACA, la structuration des snoRNA à boîtes C/D repose essentiellement sur les séquences 5' et 3', contenant respectivement les boîtes C et D (cf. Introduction, Chapitre 2, 3.1.1.1). De ce fait, le recrutement de la protéine Snu13p, qui permet la nucléation de la particule, ne peut s'effectuer que lorsque l'intégralité de la séquence du snoRNA a été transcrite. Etant donné la petite taille des snoRNA à boîtes C/D et de leurs précurseurs, la fenêtre de détection par



**A****Gène *EFB1*****B**

**Figure 68 : La protéine Nop58p et les facteurs d'assemblage Rsa1p, Pih1p et Bcd1p ne sont pas détectés sur le snoRNA intronique U18 en cours de transcription.** (A) Couples d'oligonucléotides amorces utilisés en PCR afin de détecter différentes régions du gène *EFB1* hébergeant le snoRNA U18 (alias snR18). (B) Visualisation sur gel d'agarose 1,5% des produits PCR obtenus à partir des fractions immunoprécipitées. Ces analyses ont été réalisées à partir de souches de levures exprimant les différentes protéines d'intérêt fusionnées à l'étiquette TAP. Les produits d'Input ont été dilués au 1/500<sup>ème</sup>, tandis que les produits d'IP ont été dilués au 1/5<sup>ème</sup>.



l'approche de ChIP semble donc relativement étroite et pourrait s'avérer limitante dans le cadre de l'étude des transcrits cistroniques. Inversement, dans le cas du snoRNA intronique U18, la génération d'un long précurseur correspondant au pré-ARNm *EFB1*, pourrait favoriser le repliement de l'ARN, ainsi que le recrutement et l'assemblage des protéines relativement tôt avant la fin de la transcription. Afin de tester cette hypothèse, une perspective intéressante serait de construire un gène de snoRNA à boîtes C/D monocistronique artificiel, dans lequel la taille de la région correspondante à la séquence 3' du pré-snoRNA serait allongée d'un facteur variable. L'analyse d'un tel *locus* par l'approche de ChIP pourrait ainsi nous permettre de déterminer, indépendamment du problème cinétique évoqué, si l'assemblage des RNP s'effectue de façon précoce au niveau des sites de transcription.

### **2.2.3. Cartographie des domaines d'interaction protéiques impliqués dans le mécanisme d'assemblage des snoRNP**

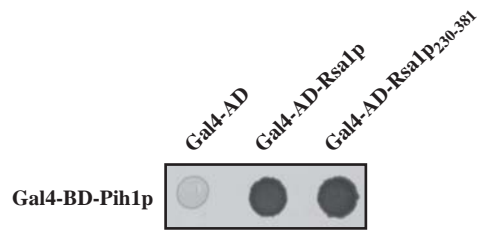
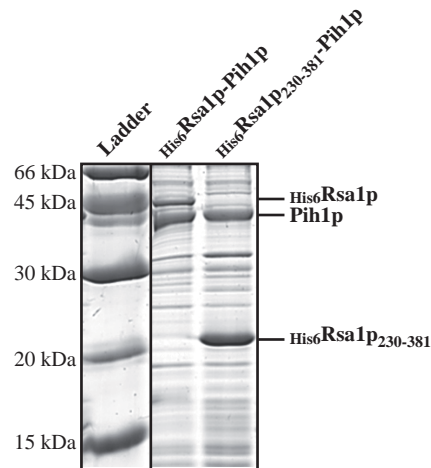
Les données acquises jusqu'ici suggèrent que Nop58p interagit successivement avec Pih1p et Rsa1p au cours de la biogenèse des snoRNP à boîtes C/D. En effet, par l'intermédiaire d'une interaction avec Pih1p, Nop58p pourrait être prise en charge de façon précoce par le complexe R2TP, avant d'être transférée vers la snoRNP par le biais de contacts avec Rsa1p. De plus, il semble que l'interaction transitoire de Rsa1p avec le facteur Pih1p puisse jouer un rôle important dans ces événements, en facilitant le recrutement du complexe R2TP sur les pré-snoRNP. Afin de mieux comprendre les bases moléculaires de l'assemblage des snoRNP à boîtes C/D, nous avons donc entrepris de caractériser le jeu d'interactions entre Rsa1p, Pih1p et Nop58p, qui constitue vraisemblablement un élément clé du processus d'assemblage. La protéine Hit1p, partenaire essentiel du facteur Rsa1p, a également été intégrée à cette étude.

Un obstacle majeur rencontré lors de notre démarche a été la difficulté à produire les différentes protéines recombinantes sous une forme soluble chez *E. coli*. Comme nous l'avons déjà mentionné, l'approche de co-expression, mise en place au laboratoire par R. Back et X. Manival, a constitué un apport considérable dans le cadre des études réalisées *in vitro* (Rothé, Back *et al.*, Publication n°2 ; Rothé *et al.*, Publication n°3). Cette technique, nous a en effet permis de tester de nombreuses interactions entre des partenaires peu solubles lorsqu'ils sont exprimés individuellement (Travail de thèse R. Back). Les protéines Rsa1p<sub>230-381</sub> et Pih1p ont notamment pu être produites en quantité importante, grâce à l'effet co-solubilisateur de leurs

partenaires respectifs Hit1p et Tah1p (Rothé *et al.*, Publication n°3). La protéine Nop58p a quant à elle été obtenue sous une forme radiomarquée, par transcription/traduction dans un extrait bactérien S30 (Rothé *et al.*, Publication n°3).

### 2.2.3.1. Rsa1p interagit avec Pih1p par l'intermédiaire de son domaine C-terminal fonctionnel

L'interaction entre Rsa1p et Pih1p a été démontrée par double-hybride et par l'approche de pull-down *in vitro* (Boulon *et al.*, 2008). Comme nous l'avons déjà mentionné, la partie C-terminale de la protéine Rsa1p, comprenant les acides aminés 230 à 381, constitue le fragment minimal fonctionnel *in vivo* (cf. Introduction, Chapitre 3, 3.6.6.1). Ce fragment, baptisé Rsa1p<sub>230-381</sub> (alias N3C1) contient le domaine PEP, responsable de l'interaction avec la protéine Snu13p, et semble suffisant pour assurer le déroulement des événements d'assemblage. Afin de savoir si cette propriété implique également les déterminants d'interaction avec Pih1p, nous avons étudié la capacité du fragment Rsa1p<sub>230-381</sub> à interagir avec la protéine Pih1p (**Fig. 69**). Par l'approche de double-hybride, nous avons tout d'abord montré que le domaine Rsa1p<sub>230-381</sub> possède la même capacité à interagir avec Pih1p que la protéine Rsa1p de pleine taille (**Fig. 69A**). Cependant, cette interaction demeure de faible intensité, puisqu'elle n'est détectée qu'à de faibles concentrations en 3-AT (< 5mM). Afin de s'assurer de la nature directe de cette interaction, R. Back a entrepris la purification du complexe Pih1p-Rsa1p<sub>230-381</sub>, obtenu par co-expression des deux protéines chez *E. coli* (**Fig. 69B**). De façon intéressante, la purification des protéines étiquetées His<sub>6</sub>Rsa1p et His<sub>6</sub>Rsa1p<sub>230-381</sub> sur billes de cobalt-Sepharose conduit à une rétention similaire de la protéine Pih1p. Ce résultat indique donc que le domaine fonctionnel de Rsa1p contient des déterminants suffisants pour assurer l'interaction avec Pih1p. En revanche, à l'image des résultats obtenus par double-hybride, les complexes Rsa1p-Pih1p et Rsa1p<sub>230-381</sub>-Pih1p se sont révélés relativement labiles, puisque leur production à plus grande échelle, incluant notamment une étape de chromatographie par exclusion de taille, n'a pas abouti.

**A****B**

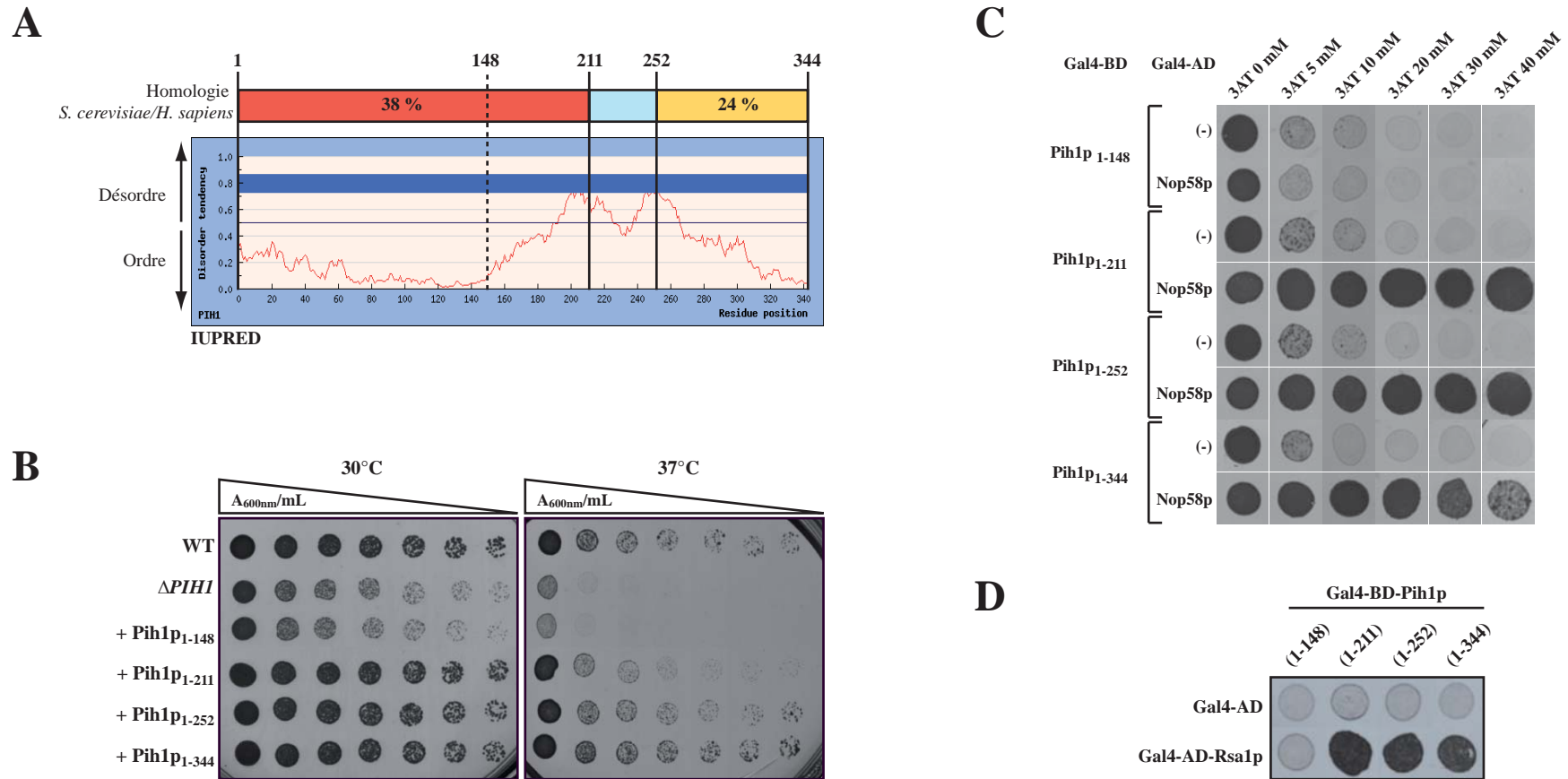
**Figure 69 : Le domaine C-terminal fonctionnel de Rsa1p interagit physiquement avec la protéine Pih1p.** (A) Test d'interaction par double-hybride entre Rsa1p ou son domaine C-terminal fonctionnel, et le facteur Pih1p. Rsa1p et Rsa1p<sub>230-381</sub> sont exprimées à partir du vecteur pACT2 (Gal4-AD). Pih1p est exprimée à partir du vecteur pGBKT7 (Gal4-BD). Les tests d'interaction ont été réalisés par croisement des souches CG1945 (pACT2) et Y187 (pGBKT7), et étalement sur milieu sélectif sans histidine et sans 3-amino-triazol. (B) Test d'interaction entre les protéines Rsa1p, Rsa1p<sub>230-381</sub> et Pih1p, par co-expression chez *E. coli* et co-purification. Les protéines étiquetées His<sub>6</sub>Rsa1p et His<sub>6</sub>Rsa1p<sub>230-381</sub> ont été retenues sur billes de cobalt-Sépha-rose. La rétention de la protéine Pih1p a été vérifiée par SDS-PAGE 12,5% et coloration au bleu de Coomassie.

### 2.2.3.2. Le domaine fonctionnel de Pih1p est responsable de l'interaction avec Nop58p et Rsa1p

Etant donné le rôle majeur joué par le facteur Pih1p dans les processus d'assemblage des snoRNP à boîtes C/D, nous avons réalisé la caractérisation fonctionnelle de ses différents domaines (**Fig. 70**). Sur la base de son homologie de séquence avec la protéine PIH1 humaine, nous avons défini les bornes de trois principaux domaines (**Fig. 70A**) : (i) la région N-terminale, composée des 211 premiers acides aminés, qui présente 38% de similitude (*i.e.* séquences similaires et identiques) avec la protéine humaine, (ii) la région comprise entre les résidus 212 et 252, qui correspond à un domaine spécifique de la levure, (iii) la région C-terminale, qui s'étend des acides aminés 253 à 344, et qui présente 24% de similitude avec la protéine humaine. Dans un premier temps, nous avons porté notre attention sur la région N-terminale de la protéine en clonant les fragments Pih1p<sub>1-212</sub> et Pih1p<sub>1-252</sub>. De plus, en utilisant le profil de structuration obtenu grâce au serveur IUPRED, nous avons également cloné le fragment court Pih1p<sub>1-148</sub> potentiellement compatible avec des études *in vitro*.

Dans un premier temps, nous avons testé la capacité de ces différents fragments à compléter le défaut de croissance d'une souche de levure invalidée pour le gène *PIH1* (**Fig. 70B**). La protéine Pih1p entière (344 acides aminés) et ses différents fragments ont été exprimés dans les cellules de levures par l'intermédiaire du vecteur pG1. La croissance des cellules a été évaluée sur milieu solide par un test en goutte à 30°C et à 37°C. Les résultats obtenus aux deux températures indiquent que les fragments Pih1p<sub>1-211</sub> et Pih1p<sub>1-252</sub> présentent une capacité de complémentarité similaire à celle de la protéine entière, tandis que les cellules exprimant le fragment court Pih1p<sub>1-148</sub> sont associées à un défaut de croissance équivalent à la souche non complétée. Ces données indiquent donc que la portion N-terminale (1-211) conservée de Pih1p est suffisante pour assurer la fonction de la protéine dans les conditions de croissance utilisées.

Afin de déterminer si la fonctionnalité Pih1p repose sur ses propriétés d'interaction avec Nop58p et/ou Rsa1p, nous avons réalisé des tests de double-hybride à partir des différents fragments de la protéine. De façon intéressante, nous avons observé pour chacun d'entre eux, une stricte corrélation entre leur propriété de complémentarité et leur capacité à interagir à la fois avec Nop58p (**Fig. 70C**) et Rsa1p (**Fig. 70D**). En effet, contrairement au fragment Pih1p<sub>1-148</sub>, les fragments fonctionnels Pih1p<sub>1-212</sub> et Pih1p<sub>1-252</sub> conservent la capacité d'interagir avec



**Figure 70 : Analyse fonctionnelle de la protéine Pih1p.** (A) Représentation schématique de la protéine Pih1p et de ses différents domaines. Les bornes des domaines ont été définies sur la base de l'homologie de séquence avec la protéine humaine PIH1 et du profil de prédiction de structuration IUPRED (Dosztányi et al., 2005). La portion représentée en bleu clair correspond à un domaine d'insertion spécifique des levures. Le % de similitude (i.e. % de résidus similaires et identiques) entre la protéine de levure et la protéine humaine est indiqué pour chaque fragment homologue. (B) Test de fonctionnalité des différents fragments de la protéine Pih1p. Chaque fragment est exprimé à partir du vecteur d'expression pG1 dans la souche de levure Y190- $\Delta$ PIH1. La croissance des cellules a été évaluée sur milieu sélectif à 30°C et 37°C, par dilution en cascade (facteur : 1/2) à partir d'une suspension de cellules à 0,5 unités  $A_{600nm}/mL$ . (C) Test d'interaction par double-hybride entre les différents fragments de Pih1p et la protéine Nop58p. Pih1p et ses fragments sont exprimés à partir du vecteur pGBKT7 (Gal4-BD). Nop58p est exprimée à partir du vecteur pACT2 (Gal4-AD). Les tests d'interaction ont été réalisés par croisement des souches Y190 (pACT2) et Y187 (pGBKT7), et étalement sur milieu sélectif sans histidine, en présence d'une gamme de concentration croissante en 3-aminotriazol. (D) Test d'interaction par double-hybride entre les différents fragments de Pih1p et la protéine Rsa1p. Rsa1p est exprimée à partir du vecteur pACT2 (Gal4-BD). Les tests d'interaction ont été réalisés par croisement des souches CG1945 (pACT2) et Y187 (pGBKT7), et étalement sur milieu sélectif sans histidine, et sans 3-aminotriazol.

ces deux partenaires. Comme évoqué précédemment, les interactions avec Rsa1p, observées sans 3-AT, demeurent relativement faibles. En revanche, les interactions avec Nop58p, résistantes à de hautes concentrations en 3-AT, présentent une grande stabilité. Il semble par ailleurs que l'absence du domaine C-terminal (acides aminés 253 à 344) renforce cette interaction, puisque la force de l'interaction entre Pih1p<sub>1-344</sub> et Nop58p est inférieure à celle observée pour les fragments plus courts Pih1p<sub>1-212</sub> et Pih1p<sub>1-252</sub> (**Fig. 70C**). Des expériences supplémentaires nous ont permis de montrer que ce domaine C-terminal constitue un domaine d'interaction dédié au facteur Tah1p. Ces résultats, qui ont fait l'objet de la rédaction d'un article, seront présentés en détail dans le chapitre suivant (Back *et al.*, Publication n°4). Enfin, l'interaction entre Pih1p et Nop58p fait actuellement l'objet d'études *in vitro* au laboratoire par le groupe de X. Manival. La purification d'un complexe entre Pih1p<sub>1-250</sub> et Nop58p<sub>159-412</sub> est en cours d'optimisation, dans l'optique de réaliser à plus long terme, des études structurales.

### 2.2.3.3. Le domaine fonctionnel de Hit1p interagit avec le domaine fonctionnel de Rsa1p

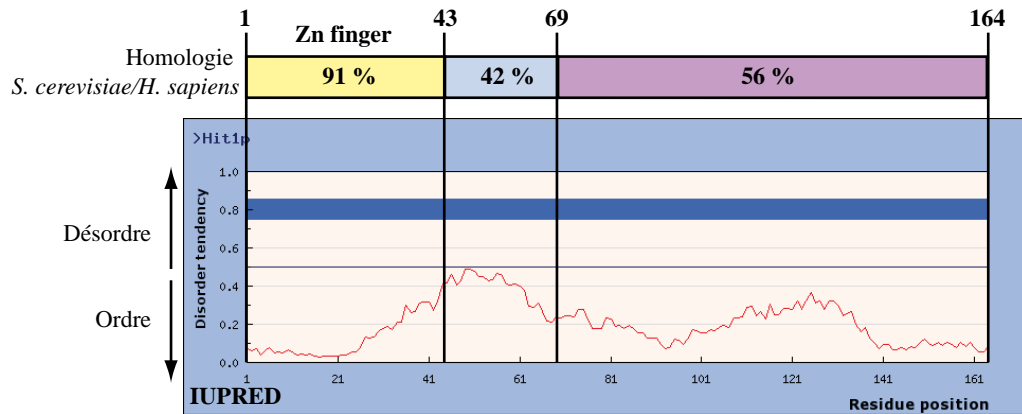
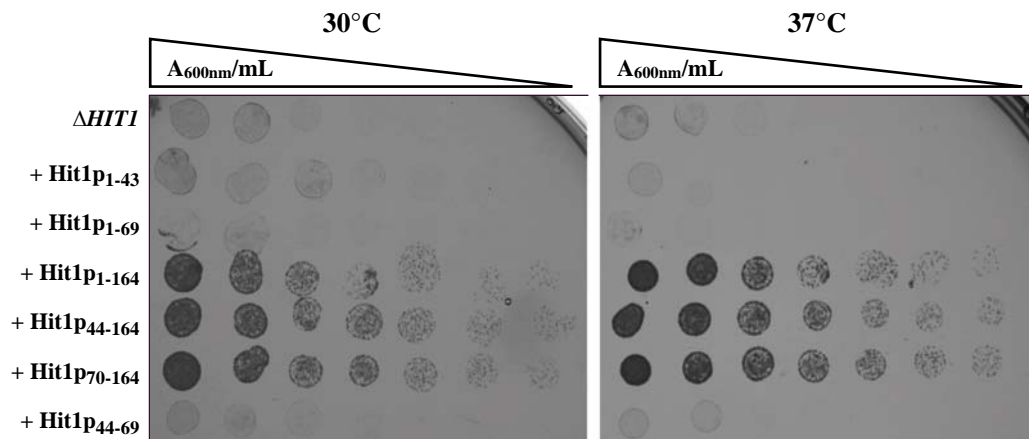
Les résultats présentés dans l'article *Functional link between proteins Rsa1 and Hit1 in yeast box C/D snoRNPs biogenesis* ont permis de mettre à jour le rôle de la protéine Hit1p dans le processus de biogenèse des snoRNP à boîtes C/D (Rothé *et al.*, Publication n°3). Ce nouveau facteur, requis pour assurer la stabilité des snoRNA, est en effet détecté sur les précurseurs du snoRNA U3. Son rôle semble essentiellement lié à celui du facteur Rsa1p, puisque ces deux protéines sont fortement associées *in vivo* et que Hit1p exerce un contrôle sur le taux de Rsa1p. L'interaction entre ces deux partenaires, initialement décrite par double-hybride (Ito 2001), a été confirmée *in vitro* par l'approche de purification de protéines co-exprimées chez *E. coli* (Rothé *et al.*, Publication n°3). En utilisant une approche de protéolyse ménagée, couplée à une analyse par spectrométrie de masse, R. Back, M. Quinternet, X. Manival et J.M. Saliou ont réussi à isoler un complexe minimal Rsa1p<sub>317-352</sub>-Hit1p<sub>70-164</sub>, qui fait actuellement l'objet d'études structurales par RMN (Travail de thèse R. Back). Afin d'obtenir plus d'informations sur la protéine Hit1p, nous avons donc entrepris de cloner et de caractériser la fonctionnalité de différents domaines, comme cela a déjà été réalisé pour Rsa1p (Boulon *et al.*, 2008) et pour Pih1p (**Fig. 70**). Cette partie du travail, réalisée sous mon encadrement, a été prise en charge par D. Tiotiu lors de son arrivée au laboratoire. Sur la base de l'homologie avec la protéine

humaine TRIP3, du profil de prédiction de structuration et des résultats de protéolyse ménagée, plusieurs fragments de la protéine ont été définis (**Fig. 71A**) et leur capacité à compléter le défaut de croissance  $\Delta HIT1$  a été évalué (**Fig. 71B**). En comparaison avec la protéine entière (164 acides aminés), le fragment minimum Hit1p<sub>70-164</sub> s'est avéré suffisant pour rétablir la croissance normale des cellules à 30°C et à 37°C. De façon intéressante, ce résultat révèle donc une corrélation entre la fonctionnalité de la protéine et sa capacité à interagir avec Rsa1p. De façon plus surprenante, il indique que le domaine à doigt de zinc, qui fait l'objet d'une forte conservation, n'est pas requis pour la fonction de la protéine dans ces conditions.

#### 2.2.3.4. Les interactions Pih1p-Nop58p et Rsa1p-Nop58p ne semblent pas mutuellement exclusives

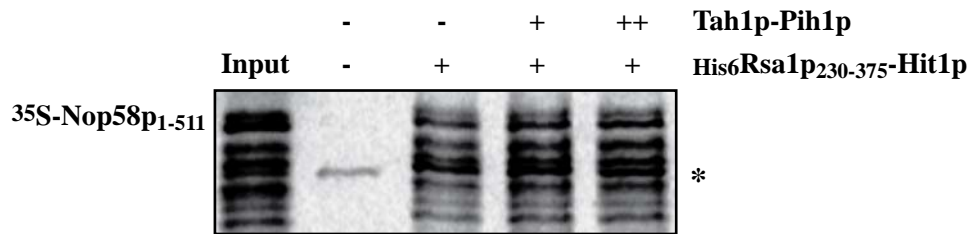
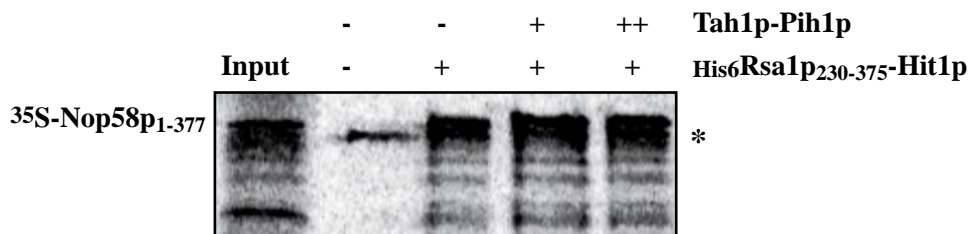
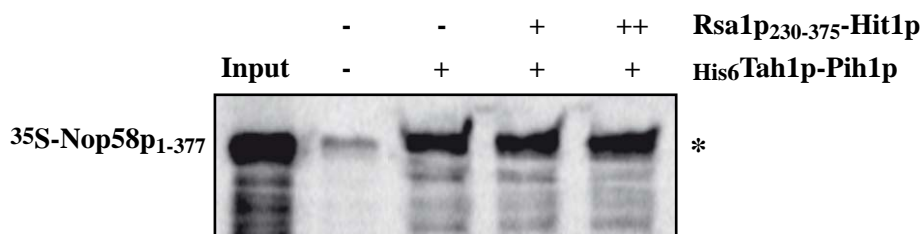
A ce stade, le réseau d'interaction reliant les facteurs d'assemblage Rsa1p, Hit1p, Pih1p, Tah1p, ainsi que la protéine Nop58p, apparaît relativement bien détaillé. Cependant, la plupart de ces interactions ont été étudiées à l'état isolé et peu d'informations sont disponibles quant à l'influence qu'elles exercent les unes sur les autres. Afin d'apporter les premiers éléments de réponse, nous avons donc utilisé les complexes Rsa1p<sub>230-375</sub>-Hit1p et Tah1p-Pih1p obtenus par co-expression chez *E. coli*, et nous avons réalisé des tests de compétition pour l'interaction avec la protéine Nop58p radiomarquée (**Fig. 72**). Nous avons reproduit l'expérience de His-pull down à partir du complexe His<sub>6</sub>Rsa1p<sub>230-375</sub>-Hit1p et de la protéine Nop58p (Rothé *et al.*, Publication n°3), en absence ou en présence du duplex Tah1p-Pih1p (**Fig. 72A**). Lorsque ce dernier est ajouté en quantité équimolaire ou en excès molaire 2X la rétention de la protéine Nop58p par le complexe His<sub>6</sub>Rsa1p<sub>230-375</sub>-Hit1p n'est pas affectée. Des résultats identiques ont également été obtenus en utilisant le fragment Nop58p<sub>1-377</sub>, correspondant à la protéine Nop58p<sub>1-377</sub>, tronquée de sa partie C-terminale (**Fig. 72B**). De même, lorsque le *pull-down* est réalisé à partir du complexe His<sub>6</sub>Tah1p-Pih1p, la présence du complexe Rsa1p<sub>230-375</sub>-Hit1p n'interfère pas avec la rétention de la protéine Nop58p<sub>1-377</sub> sur les billes (**Fig. 72C**).

Ces résultats nous conduisent donc à établir deux constats majeurs : d'une part la région C-terminale (378-511) de Nop58p n'est pas nécessaire pour interagir avec les complexes Rsa1p<sub>230-375</sub>-Hit1p et Tah1p-Pih1p, d'autre part, ces interactions ne semblent interférer ni

**A****B**

**Figure 71 : Analyse fonctionnelle de la protéine Hit1p.** (A) Représentation schématique de la protéine Hit1p et de ses différents domaines. Les bornes des domaines ont été définies sur la base de l'homologie de séquence avec la protéine humaine TRIP3 et du profil de prédiction de structuration IUPRED (Dosztányi *et al.*, 2005). Le % de similitude (i.e. % de résidus similaires et identiques) entre la protéine de levure et la protéine humaine est indiqué pour chaque fragment homologue. (B) Test de fonctionnalité des différents fragments de la protéine Hit1p. Chaque fragment est exprimé à partir du vecteur d'expression p413 dans la souche de levure By4741- $\Delta HIT1$ . La croissance des cellules a été évaluée sur milieu sélectif à 30°C et 37°C, par dilution en cascade (facteur : 1/2) à partir d'une suspension de cellules à 0,5 unités A<sub>600nm</sub>/mL.



**A****B****C**

**Figure 72 : Les sous-complexes Rsa1p<sub>230-375</sub>-Hit1p et Pih1p-Tah1p interagissent avec Nop58p de façon non compétitive.** (A) His-pull down réalisé sur billes de Cobalt-Sepharose, à partir du complexe His6Rsa1p<sub>230-375</sub>-Hit1p, en présence de la protéine Nop58p<sub>1-511</sub> radiomarquée et du complexe Tah1p-Pih1p. (B) His-pull down réalisé sur billes de Cobalt-Sepharose, à partir du complexe His6Rsa1p<sub>230-375</sub>-Hit1p, en présence de la protéine Nop58p<sub>1-377</sub> radiomarquée et du complexe Tah1p-Pih1p. (C) His-pull down réalisé sur billes de Cobalt-Sepharose, à partir du complexe His6Tah1p-Pih1p, en présence de la protéine Nop58p<sub>1-377</sub> radiomarquée et du complexe Rsa1p<sub>230-375</sub>-Hit1p. Les complexes binaires ont été obtenus par co-expression chez *E. coli*. Pour chaque expérience, ces complexes ont été incubés en proportion équimolaire (+/+) ou avec un excès 2X du complexe non étiqueté (+/+). La protéine Nop58p<sub>1-511</sub> entière ou sa forme tronquée Nop58p<sub>1-377</sub> ont été obtenues par transcription/traduction en lysat bactérien S30, en présence de méthionine marquée au soufre 35 (35S). Le symbole\* représente un contaminant protéique radiomarqué généré dans l'extrait S30 lors la réaction de transcription/traduction.

négativement, ni positivement l'une avec l'autre, dans les conditions expérimentales utilisées. Cette dernière observation suggère donc que les interactions de Nop58p avec ces deux complexes pourraient coexister au cours du processus d'assemblage. Ces données doivent cependant être considérées à titre d'indications préliminaires et ce travail devra être poursuivi par des investigations plus approfondies. Il sera notamment important de reproduire les tests de compétitions en y incluant le pré-complexe snoRNA-Snu13p et en augmentant l'excès molaire du complexe non étiqueté. De plus l'approche de co-expression pourra être utilisée afin d'évaluer la possibilité d'isoler des complexes macromoléculaires contenant ces différentes protéines.

## **2.2.4. Cartographie des interactions génétiques à partir des souches de levures $\Delta RSA1$ , $\Delta PIH1$ et $\Delta TAH1$**

### **2.2.4.1. Description de l'approche GIM**

En utilisant une approche de génétique de la levure, nous avons mis en évidence que les doubles délétions  $\Delta RSA1\Delta PIH1$  et  $\Delta HIT1\Delta PIH1$  confèrent un phénotype synthétique léthal (Rothé *et al.*, Publication n°3). Ce résultat confirme l'importance du duo Rsa1p-Pih1p pour les processus d'assemblage. De plus, il confirme également le rôle capital joué par Hit1p pour maintenir un taux normal de protéine Rsa1p, puisque la délétion du gène *HIT1* confère aux cellules un phénotype de type  $\Delta RSA1$ .

Afin de poursuivre ces analyses à plus grande échelle, nous avons débuté une collaboration avec C. Saveanu de l'Institut Pasteur à Paris. Son équipe a en effet mis au point une approche de crible d'interactions génétiques permettant d'obtenir des données quantitatives sur les effets liés au cumul de plusieurs délétions (Decourty *et al.*, 2008). Cette technique, baptisée GIM pour *Genetic Interaction Mapping*, nécessite de croiser la souche haploïde d'intérêt simplement délétée (MAT $\alpha$ , *NatMX4* = gène de résistance à la nourseothricine) avec une banque de souches haploïdes mutantes (KO ou DAMP) (MAT $\alpha$ , *KanMX4* = gène de résistance à la kanamycine). Les cellules diploïdes obtenues sont ensuite sélectionnées sur milieu sélectif et placées en conditions de sporulation. Les souches haploïdes doublement mutées sont alors sélectionnées grâce à leur double résistance à la nourseothricine et à la kanamycine. Lors de cette étape, les cellules sont cultivées en milieu liquide jusqu'à l'obtention de 18 doublements de la population, dans le but de générer un effet de compétition entre les différentes souches.

A terme, les *loci* issus de la banque sont analysés spécifiquement par PCR et hybridation sur puce à ADN. La même démarche est réalisée en parallèle par croisement de la banque de mutants avec une souche délétée pour un gène de référence. L'intensité des signaux mesurés pour chaque *loci* peut ainsi être comparée entre la population du double mutant d'intérêt et celle de référence. Les résultats sont alors exprimés en  $\log_2(Q/R)$  où Q représente l'intensité de signal obtenue pour la population d'intérêt et R représente celle mesurée pour la population de référence. Ce paramètre permet d'évaluer le *fitness* de chaque souche (= valeur sélective correspondant à la capacité d'une cellule à générer des descendants viables et fertiles). Il fournit en effet une information sur la représentativité des différents mutants au sein de la population totale et permet d'estimer l'impact de chaque double délétion sur la capacité d'adaptation des cellules. Lorsque la valeur  $\log_2(Q/R)$  est inférieure à 0, le double mutant d'intérêt est sous-représenté par rapport au double mutant de référence ( $Q < R$ ), l'effet délétère des deux mutations est alors cumulatif. Inversement, une valeur de  $\log_2(Q/R)$  supérieure à 0, correspond à un enrichissement du double mutant d'intérêt ( $Q > R$ ), on conclut alors à un effet épistatique. Dans ce deuxième cas, soit l'effet délétère de l'une des mutations est supprimé ou compensé par la seconde mutation, soit la double mutation est dénuée d'effet cumulatif car les produits des deux gènes d'intérêt interviennent dans la même voie métabolique. Enfin, la validation de ces résultats nécessite de reproduire la même approche avec différents mutants haploïde de référence, afin de s'assurer de leur neutralité par rapport au processus étudié.

Nous avons donc entrepris d'appliquer l'approche GIM pour les différentes souches de levures d'intérêt. Ces dernières ont été préparées à partir des souches BY4742 (MAT $\alpha$ , *KanMX4*) en échangeant, par recombinaison homologue, la cassette de résistance originale *KanMX4* par la cassette *NatMX4* (cf. Matériel & Méthode). Le crible d'interaction génétique a ensuite été réalisé par C. Saveanu à l'Institut Pasteur. A l'heure actuelle, les premiers résultats ont été obtenus pour les souches  $\Delta RSA1$ ,  $\Delta PIH1$  et  $\Delta TAH1$  (**Tableaux 6, 7 et 8**).

#### **2.2.4.2. Description du réseau d'interactions génétiques impliquant l'ORF *RSA1***

Concernant l'ORF *RSA1* (**Tableau 6**), nous avons constaté un lien génétique important avec l'ORF *RVB1* (-2,26) qui souligne vraisemblablement leur rôle complémentaire dans le processus d'assemblage des snoRNP. De façon plus surprenante, un effet opposé est observé

ORF	RSA1#1	RSA1#2	RSA1 moyenne	co-IP
YOL081W - IRA2	-4.872164699	-4.431864941	-4.65201482	
YDR495C - VPS3	-2.918245947	NA	-2.918245947	
YJR087W	-3.570387082	-1.988417148	-2.779402115	
YPR122W - AXL1	-3.250423202	-2.281673252	-2.766048227	
YKL177W	-3.574967387	-1.955423295	-2.765195341	
YFR048W - RMD8	-4.148579398	-1.366158077	-2.757368738	
YPL043W - NOP4 - DAMP	-3.645119964	-1.563693085	-2.604406524	
YGL024W	-2.588989481	NA	-2.588989481	
YBR193C - MED8 - DAMP	-1.966654933	-3.050500225	-2.508577579	
YKL048C - ELM1	-2.619560882	-2.271559048	-2.445559965	
YGR190C - DAMP	-2.236743157	-2.653450162	-2.445096659	
YLR175W - CBF5 - DAMP	-2.538065292	-2.339587061	-2.438826176	
YOL080C - REX4	-1.560914604	-3.118618777	-2.33976669	
YKL104C - GFA1 - DAMP	-1.963099656	-2.619291944	-2.2911958	
YDR190C - RVB1 - DAMP	-2.19540722	-2.323348751	-2.259377986	P
YJR117W - STE24	-1.904835987	-2.586076565	-2.245456276	
YDR456W - NHX1	-2.22608266	NA	-2.22608266	
YJL157C - FAR1	-3.523848804	-0.851817816	-2.18783331	
YCL036W - GFD2	-1.848365371	-2.444631434	-2.146498402	
YML097C - VPS9	-2.116065661	NA	-2.116065661	
YBR078W - ECM33	-1.425323719	-2.736598592	-2.080961155	
YNL299W - TRF5	-2.067254234	-2.060755554	-2.064004894	
YDL001W - RMD1	-0.474019337	-3.592222846	-2.033121092	
YMR043W - MCM1 - DAMP	-1.972165805	-1.944080291	-1.958123048	
YOL090W - MSH2	-1.620087769	-2.276363135	-1.948225452	
YIR033W - MGA2	-1.881347242	-2.007816687	-1.944581965	
YAL002W - VPS8	-1.942136678	NA	-1.942136678	
YNL298W - CLA4	-1.340751251	-2.507811519	-1.924281385	
YJR007W - SUI2 - DAMP	-3.06627187	-0.714975726	-1.890623798	
YBR290W - BSD2	-1.170458714	-2.573538692	-1.871998703	
YOL041C - NOP12	-1.90658736	-1.808564627	-1.857575994	
YMR274C - RCE1	-1.749860186	-1.952300543	-1.851080364	
YKR081C - RPF2 - DAMP	-2.355512338	-1.322556585	-1.839034462	
YMR136W - GAT2	-1.791758153	-1.876963666	-1.834360909	
YOL085C	-1.900521121	-1.712228229	-1.806374675	
YOL091W - SPO21	-1.481686715	-2.075894518	-1.778790616	
YJR139C - HOM6	-2.176867741	-1.3759474	-1.77640757	R H
YMR022W - UBC7	-1.729651357	-1.790181415	-1.759916386	
YPR108W - RPN7 - DAMP	-2.32059369	-1.194942129	-1.75776791	
YJL063C - MRPL8	-2.140775944	-1.354844299	-1.747810122	
YDR027C - VPS54	-1.743637957	NA	-1.743637957	
YAL056W - GPB2	-1.387530035	-2.067331323	-1.727430679	
YJL024C - APS3	-1.56442039	-1.8743163	-1.719368345	
YOL098C	0.24743915	-3.672734493	-1.712647672	
YOL050C	-1.690932904	-1.705024818	-1.697978861	
YKR024C - DBP7	-1.911151576	-1.470415898	-1.690783737	
YMR264W - CUE1	-1.39476746	-1.963029609	-1.678898535	
YPR029C - APL4	-1.44868267	-1.908699419	-1.678691044	
YDL097C - RPN6 - DAMP	-2.040320274	-1.316654826	-1.67848755	H
YKL058W - TOA2 - DAMP	-1.669541665	NA	-1.669541665	
YFR004W - RPN11 - DAMP	-2.003951244	-1.298695693	-1.651323468	
YNL062C - GCD10 - DAMP	-1.418805138	-1.882861382	-1.65083326	
YIL141W	-1.619006809	-1.661425761	-1.640216285	
YGR178C - PBP1	-1.332672124	-1.946386861	-1.639529493	
YOL082W - ATG19	-1.07048195	-2.193207661	-1.631844806	
YLR426W - TDA5	-1.420370377	-1.83626539	-1.628317884	
YDL226C - GCS1	-1.593229735	-1.659808076	-1.626518905	
YBR036C - CSG2	-1.592327546	-1.658419764	-1.625373655	
YOL135C - MED7 - DAMP	-1.164272639	-2.073314683	-1.618793661	
YBR288C - APM3	-2.08407316	-1.124517503	-1.604295332	
YDR496C - PUF6	-2.587500606	-0.612604181	-1.600052393	
YER083C - GET2	-1.470036902	-1.70863339	-1.589335146	
YJL001W - PRE3 - DAMP	-2.120756985	-1.045061364	-1.582909174	
YBR200W - BEM1	NA	-1.581016732	-1.581016732	
YBR253W - SRB6 - DAMP	-1.901725582	-1.256464772	-1.579095177	
YOL069W - NUF2 - DAMP	-1.436962315	-1.717221858	-1.577092086	
YGL025C - PGD1	-1.576647457	NA	-1.576647457	
YDL029W - ARP2 - DAMP	-1.504297683	-1.605368891	-1.554833287	P
YLR025W - SNF7	-1.553267722	NA	-1.553267722	
YKR020W - VPS51	-2.294439741	-0.760207707	-1.527323724	
YGR245C - SDA1 - DAMP	-1.735833765	-1.271019629	-1.503426697	
YFR052W - RPN12 - DAMP	-1.952950718	-1.05039378	-1.501672249	
YHR011W - DIA4	-1.485733755	-1.515380706	-1.500557231	

Tableau 6a : Résultats du crible GIM réalisé à partir de la souche de levure BY4742- $\Delta$ RSA1. (cf. légende Tableau 6c)

ORF	RSA1#1	RSA1#2	RSA1 moyenne	co-IP
YOR039W - CKB2	1.483227582	1.524699069	1.503963325	P
YFR021W - ATG18	0.698175044	2.313567046	1.505871045	
YGR158C - MTR3 - DAMP	1.214662778	1.802614821	1.5086388	
YMR194W - RPL36A	0.155027612	2.87033728	1.512682446	
YEL059W - HHY1	-0.18511799	3.223892532	1.519387271	
YKR026C - GCN3	1.27328943	1.77268798	1.522988705	R P
YCL061C - MRC1	2.027453892	1.020152034	1.523802963	
YKL166C - TPK3	1.198162119	1.853675529	1.525918824	
YDL080C - THI3	0.923810932	2.156700567	1.540255749	
YGL016W - KAP122	0.721682076	2.367248717	1.544465397	
YLR435W - TSR2	2.719804663	0.389543819	1.554674241	
YOL012C - HTZ1	1.888988377	1.221231176	1.555109777	
YDR268W - MSW1	2.547935079	0.57008725	1.559011164	
YDL229W - SSB1	1.349581391	1.768738651	1.559160021	H
YLR106C - MDN1 - DAMP	1.182209239	1.936666088	1.559437663	
YDL213C - NOP6	0.957900772	2.161524252	1.559712512	
YCL037C - SRO9	0.919231008	2.204676322	1.561953665	R
YPR070W - MED1	2.101953219	1.047248865	1.574601042	
YDR035W - ARO3	0.822634657	2.333753131	1.578193894	
YOR006C - TSR3	1.845540688	1.314272057	1.579906373	
YDL111C - RRP42 - DAMP	0.945003042	2.2178204	1.581411721	P
YPL036W - PMA2	0.650290028	2.513190755	1.581740391	
YJL177W - RPL17B	1.258051031	1.905516611	1.581783821	
YER131W - RPS26B	1.475514865	1.698592339	1.587053602	
YPR200C - ARR2	1.127490916	2.049450876	1.588470896	
YHR118C - ORC6 - DAMP	-0.06384908	3.245429786	1.590790353	
YOR123C - LEO1	1.587236864	1.602998376	1.59511762	
YGR102C - GTF1	-0.672728152	3.872560297	1.599916073	
YFR009W - GCN20	1.081547153	2.127246593	1.604396873	
YKR090W - PXL1	0.369500419	2.841312214	1.605406316	
YLR337C - VRP1	2.404786704	0.815710173	1.610248438	
YDR447C - RPS17B	1.437172353	1.78375809	1.610465222	
YML123C - PHO84	0.49527437	2.735152738	1.615213554	
YLR325C - RPL38	2.321234506	0.915529767	1.618382136	P
YJR140C - HIR3	1.404287436	1.833252387	1.618769912	
YLR441C - RPS1A	1.501038778	1.740592704	1.620815741	H
YFL013C - IES1	0.696071763	2.548777639	1.622424701	
YOR245C - DGA1	0.656091943	2.589566471	1.622829207	
YBR275C - RIF1	0.284115429	2.961701677	1.622908553	
YDR058C - TGL2	0.779337277	2.468969166	1.624153222	
YDR271C	-0.450455504	3.70820791	1.628876203	
YER092W - IES5	1.678330258	1.582457832	1.630394045	
YLR406C - RPL31B	0.68705685	2.576925828	1.631991339	
YLR197W - NOP56 - DAMP	1.195806806	2.086447979	1.641127393	R
YJL148W - RPA34	2.209762427	1.111940627	1.660851527	
YDL063C	1.396509955	1.9389392	1.667724577	
YOR038C - HIR2	1.578670166	1.756832008	1.667751087	
YNL021W - HDA1	1.463393153	1.878312427	1.67085279	
YDR157W	0.74262053	2.601042078	1.671831304	
YNL315C - ATP11	1.035040086	2.315292643	1.675166364	
YLR024C - UBR2	-0.17168496	3.537523741	1.682919391	
YBR215W - HPC2	2.17902949	1.186883255	1.682956372	
YML032C - RAD52	0.857018043	2.51094059	1.683979316	
YGR150C - CCM1	0.358458544	3.031558606	1.695008575	
YDL002C - NHP10	1.369624474	2.023199162	1.696411818	
YCL029C - BIK1	-1.220682938	4.621495233	1.700406148	P
YNL250W - RAD50	0.281154531	3.137331871	1.709243201	
YOL139C - CDC33 - DAMP	2.145806152	1.276745092	1.711275622	
YER169W - RPH1	1.549001927	1.912776561	1.730889244	
YBR263W - SHM1	1.24506133	2.267379011	1.756220171	R H P
YJL111W - CCT7 - DAMP	1.732542858	1.795185772	1.763864315	
YOL067C - RTG1	0.574899699	2.964373519	1.769636609	
YER048C - CAJ1	2.432693511	1.106820415	1.769756963	
YLR335W - NUP2	1.609881548	1.931775056	1.770828302	
YOL076W - MDM20	1.67843272	1.867532255	1.772982487	H
YGL151W - NUT1	0.65998471	2.913510827	1.786747769	
YGR118W - RPS23A	1.829073082	1.757780715	1.793426899	
YMR016C - SOK2	2.346617999	1.287095401	1.8168567	
YKR073C	2.158443523	1.496226058	1.827334791	
YGR214W - RPS0A	1.998668494	1.658560754	1.828614624	
YBR135W - CKS1 - DAMP	0.070569659	3.620721539	1.845645599	
YOR150W - MRPL23	0.028884912	3.671335831	1.850110372	
YDR324C - UTP4 - DAMP	0.693769323	3.01984243	1.856805877	

Tableau 6b : Résultats du crible GIM réalisé à partir de la souche de levure BY4742- $\Delta$ RSA1 (suite). (cf. légende Tableau 6c)

ORF	RSA1#1	RSA1#2	RSA1 moyenne	co-IP
YBL078C - ATG8	0.451026424	3.274025711	1.862526068	
YOR167C - RPS28A	1.382124525	2.349331314	1.865727919	
YDR510W - SMT3 - DAMP	-0.266391001	4.014862507	1.874235753	
YHR091C - MSR1	-0.1627325	3.945381976	1.891324738	
YDL065C - PEX19	1.711329332	2.074952978	1.893141155	
YLR367W - RPS22B	1.750323227	2.046212192	1.89826771	
YER178W - PDA1	0.870435612	2.965181668	1.91780864	R H P
YOR061W - CKA2	2.01196341	1.835027173	1.923495291	R H P
YDR295C - HDA2	0.94355563	2.90896497	1.9262603	
YPR110C - RPC40 - DAMP	0.88660461	2.999487405	1.943046007	R H P
YLR193C - UPS1	1.1750358	2.731249186	1.953142493	
YDR156W - RPA14	0.303099949	3.620809756	1.961954852	
YKR057W - RPS21A	1.732316517	2.219086691	1.975701604	H
YGR095C - RRP46 - DAMP	1.859254334	2.093988761	1.976621548	
YKL076C - PSY1	1.728795095	2.232472892	1.980633994	
YPL055C - LGE1	1.213438926	2.830075459	2.021757193	
YOR073W - SGO1	1.019370687	3.054470276	2.036920481	
YGL149W	1.676038878	2.397995898	2.037017388	
YDR315C - IPK1	1.790157655	2.298525381	2.044341518	
YBL091C - MAP2	1.100294955	3.006245772	2.053270363	R H P
YBL007C - SLA1	1.600259052	2.540666214	2.070462633	
YNR046W - TRM112 - DAMP	2.607767838	1.536683093	2.072225466	
YMR230W - RPS10B	1.839278553	2.351745497	2.095512025	R H P
YNL113W - RPC19 - DAMP	2.063151917	2.12831509	2.095733504	
YLR366W	2.083586093	2.110863363	2.097224728	
YNL206C - RTT106	1.846058265	2.366358781	2.106208523	
YPR057W - BRR1	2.036319014	2.205636348	2.120977681	
YLR312W-A - MRPL15	1.458467435	2.861610247	2.160038841	
YGR064W	1.173474533	3.199309974	2.186392254	
YPL148C - PPT2	0.995298512	3.454585499	2.224942006	
YLR287C-A - RPS30A	2.062981829	2.419855392	2.24141861	
YLR410W - VIP1	1.555370943	2.94097026	2.248170601	
YLR388W - RPS29A	2.376642649	2.126479746	2.251561198	
YNR004W - SWM2	1.545731727	2.978612341	2.262172034	
YER110C - KAP123	2.050501931	2.498961496	2.274731713	R H P
YJR118C - ILM1	1.671860101	2.892560619	2.28221036	
YHR148W - IMP3 - DAMP	1.37923017	3.197639932	2.288435051	
YBR195C - MSI1	2.113805274	2.488362933	2.301084104	
YGR063C - SPT4	1.663671753	2.989135143	2.326403448	R H
YNL232W - CSL4 - DAMP	1.069168081	3.626336712	2.347752396	
YML101C-A	1.55874378	3.202212854	2.380478317	
YML014W - TRM9	NA	2.389123004	2.389123004	
YGL029W - CGR1	1.842799627	2.941500655	2.392150141	
YGR055W - MUP1	1.938785589	2.875679364	2.407232476	
YOR191W - ULS1	2.028920113	2.83777693	2.433348521	
YIL027C - KRE27	2.027107189	2.973999234	2.500553211	
YLR200W - YKE2	1.091708516	3.920947246	2.506327881	
YJR050W - ISY1	1.242679467	3.775174421	2.508926944	
YDL182W - LYS20	0.882399779	4.172292399	2.527346089	
YBR226C	0.713807355	4.341565154	2.527686254	
YMR179W - SPT21	1.873843248	3.268361361	2.571102305	
YLR386W - VAC14	1.401651302	3.760137548	2.580894425	
YDR225W - HTA1	1.162356467	4.108419516	2.635387992	
YML102W - CAC2	1.355023146	3.92325129	2.639137218	
YGL244W - RTF1	2.195242208	3.121264676	2.658253442	
YPR132W - RPS23B	2.44740752	2.97400508	2.7107063	
YML026C - RPS18B	2.437016189	3.174008637	2.805512413	
YKL075C	2.082789032	3.610444182	2.846616607	
YOR359W - VTS1	1.835520617	3.945987974	2.890754296	
YPR133C - SPN1 - DAMP	2.005792355	3.883661352	2.944726853	
YBR268W - MRPL37	2.22359704	3.800642555	3.012119797	
YGR081C - SLX9	2.203744715	3.915859556	3.059802135	
YER161C - SPT2	0.628516103	5.573884311	3.101200207	
YML009C-A	1.561418036	4.801629759	3.181523898	
YML009W-B	2.426477076	4.525946546	3.476211811	

**Tableau 6c : Résultats du crible GIM réalisé à partir de la souche de levure BY4742-ARSAI (suite).** La première colonne indique le nom systématique et le nom standard de l'ORF. DAMP : *Decreased Abundance by mRNA Perturbation* (= insertion dans la région 3'-UTR permettant d'atténuer la production d'une protéine essentielle à la viabilité cellulaire). Les colonnes 2 et 3 indiquent les résultats de Log<sub>2</sub>(Q/R) obtenus lors de deux expériences distinctes. NA = absence de donnée. La colonne 4 représente la moyenne des valeurs de Log<sub>2</sub>(Q/R). La dernière colonne croise les résultats du crible GIM avec ceux obtenus par IP-protéomique (Rothé *et al.*, Publication n°3), en indiquant les protéines qui ont été co-purifiées spécifiquement avec Rsa1p (R), Hit1p (H) ou Pih1p (P). Seuls les résultats de Log<sub>2</sub>(Q/R) compris dans l'intervalle [-1,5; 1,5] sont représentés. Les composants des snoRNP sont symbolisés en rouge, les facteurs d'assemblage en bleu, les facteurs impliqués dans la maturation des ARN en vert, et les protéines impliquées dans la biogenèse des ribosomes en violet. Ces données ont été fournies par C. Saveanu (Institut Pasteur).

avec l'ORF *NOP56* (+1,64) et révèle soit un effet suppresseur, soit un effet non cumulatif des deux délétions. D'autres effets épistatiques majeurs ont également été observés avec des facteurs impliqués dans la terminaison de la transcription et la maturation de l'extrémité 3' des snoRNA (cf. Introduction, Chapitre 1, 2.4.2.3) (**Fig. 12**). C'est notamment le cas des ORF *MTR3* (+1,51), *RRP42* (+1,58), *RRP46* (+1,97), *CSL4* (+2,35) codant des composants de l'exosome et des ORF *LEO1* (+1,59) et *RTF1* (+2,66) qui codent des membres du complexe PAF1. Etant donné l'implication de ces différents facteurs et de Rsa1p dans des étapes distinctes de la biogenèse des snoRNP, il est envisageable que les effets observés s'apparentent à des effets suppresseurs. Il est ainsi envisageable qu'un ralentissement du processus de maturation de l'extrémité 3' permette de compenser une défaillance du système d'assemblage des snoRNP. Comme nous l'avons déjà mentionné précédemment, il est en effet probable que la cinétique de transcription et de maturation du snoRNA influence le déroulement du processus d'assemblage des particules (cf. Résultats & Discussion, Partie 3, 2.2.2). L'étude des relations entre ces différents processus constitue donc une perspective d'étude importante. Il serait notamment intéressant de reproduire les expériences de CHIP dans des souches de levures délétées de certains composants de la machinerie de terminaison de la transcription (*e.g.* Leo1p, Rtf1p). En plus de ces effets épistatiques, le gène *RSA1* est également lié aux ORF *TRF5* (-2,1) et *REX4* (-2,4), dont les produits interviennent également dans la maturation des ARN. La protéine Trf5p est une poly(A) polymérase appartenant au complexe TRAMP5, qui exerce une fonction de contrôle qualité sur certains ARN précurseurs, dont les pré-snoRNA (Egecioglu *et al.*, 2006; Grzechnik & Kufel, 2008; Houseley & Tollervy, 2006). La protéine Rex4p est quant à elle une exonucléase 3'-5', potentiellement impliquée dans la maturation du pré-ARNr (Eppens *et al.*, 2002; Faber *et al.*, 2004). Cette dernière n'a encore jamais été décrite dans le cadre de la maturation nucléolytique des snoRNA. Par la suite, il sera donc nécessaire de caractériser les snoRNA à boîtes C/D s'accumulant dans une souche de levure  $\Delta$ *REX4*, afin de vérifier son implication potentielle dans la maturation nucléolytique de ces ARN.

Outre la biogenèse des snoRNP à boîtes C/D, la délétion de l'ORF *RSA1* est corrélée à de multiples effets épistatiques impliquant des gènes codant des facteurs de remodelage de la chromatine comme des membres des complexes de remodelage HIR (*HIR2*, *HIR3*, *HPC26*), INO80 (*IES1*, *IES5*, *NHP10*), et CAF-1 (*MSI1*, *CAC2*), une histone déacétylase (*HDA1*) ou encore des protéines histones (*HTZI*, *HTAI*). Concernant cette catégorie, le résultat le plus marquant est certainement le lien établi entre *RSA1* et l'ORF *RTT106* (+2,1). Cette dernière

code en effet la protéine Rtt106p, une chaperonne d'histones, qui agit en coopération avec les complexes HIR et CAF-1 (Fillingham *et al.*, 2009; Huang *et al.*, 2005; Huang *et al.*, 2007; Lambert *et al.*, 2010) et qui interagit par double-hybride avec le facteur Bcd1p (cf. Résultats & Discussion, Partie 6, 2.2.3). Nous reviendrons en détail sur ce lien dans le dernier chapitre de la partie Résultats & Discussion.

Enfin, parmi les autres résultats notoires, un lien marqué avec l'ORF *NOP4* (-2,6) a été observée. Cette dernière code la protéine Nop4p, qui participe à la biogenèse de la grande sous-unité ribosomique et qui a également été impliquée dans l'adressage nucléolaire de la snoRNP U3 (Qiu *et al.*, 2008) (**Fig. 50**). *RSI1* établit également des liens de natures diverses avec plusieurs ORF codant des protéines ribosomiques ou des composants du protéasome 26S (*i.e.* *RPN6*, *RPN7*, *RP11*, *RPN12*), ainsi qu'avec les ORF *SMT3* (+1,87) et *ULS1* (+2,43), dont les produits sont respectivement impliqués dans la sumoylation des protéines et dans leur adressage au protéasome. Un autre lien épistatique a aussi été observé avec l'ORF *RIF1* (+1,63), dont le produit, intervient dans le contrôle de la taille des télomères. Ce résultat est particulièrement intéressant, puisque la délétion de l'ORF *RIF1*, comme celle des ORF *RSI1* et *HIT1*, conduit à une augmentation de la taille des télomères (Gatbonton *et al.*, 2006). Enfin, *RSI1* est épistatique avec les ORF *CKA2* (+1,92) et *CKB2* (+1,50) qui codent des sous-unités de la kinase CK2 retrouvées parmi les facteurs co-purifiés à partir des protéines TAP-Rsa1p, TAP-Hit1p et TAP-Pih1p (Rothé *et al.*, Publication n°3) (**Tableau 6**).

#### 2.2.4.3. Description du réseau d'interactions génétiques impliquant l'ORF *PIH1*

Concernant l'ORF *PIH1* (**Tableau 7**), nous avons pu constater un lien génétique important avec les ORF *RVB1* (-2,46), *RVB2* (-2,77), *NOP58* (-2,60) et *NOP56* (+1,95). Cette donnée confirme le lien fonctionnel unissant Pih1p, Rvb1p et Rvb2p et souligne la complémentarité du rôle joué par ces facteurs, vraisemblablement à travers leur appartenance au complexe R2TP et leur rôle dans la formation des snoRNP. L'effet antagoniste observé avec les ORF *NOP58* et *NOP56* témoigne à nouveau d'une différence notable entre les deux protéines paralogues. Une atténuation de l'expression de Nop58p, couplée à l'absence du facteur Pih1p, dont elle dépend, pourrait conduire à une faible incorporation de cette protéine au sein des snoRNP et à une diminution critique du taux de snoRNA à boîtes C/D. En revanche, comme nous l'avons déjà observé pour Rsa1p, l'absence de Pih1p n'est pas aggravée par une



ORF	PIH1#1	PIH1#2	PIH1 moyenne	co-IP
YJR105W - ADO1	-4.295843137	NA	-4.295843137	R H
YIR023W - DAL81	-4.344337745	-2.190522248	-3.267429996	
YNL120C	-3.115859891	NA	-3.115859891	
YJR139C - HOM6	-3.141130188	-2.969199594	-3.055164891	R H
YHR111W - UBA4	-2.964411779	NA	-2.964411779	
YPL235W - RVB2 - DAMP	-2.768112181	NA	-2.768112181	P
YOR310C - NOP58 - DAMP	-2.608446466	-2.595607501	-2.602026984	P
YDR190C - RVB1 - DAMP	-2.961206411	-1.954554761	-2.457880586	P
YKL149C - DBR1	-2.02705115	-2.576774518	-2.301912834	
YOR001W - RRP6	-2.538436295	-2.065193068	-2.301814682	
YER165W - PAB1 - DAMP	-2.268870337	-2.302360902	-2.285615619	
YNL119W - NCS2	-3.444891546	-1.033884428	-2.239387987	
YLR039C - RIC1	-1.96905884	NA	-1.96905884	
YIL008W - URM1	-2.95231639	-0.974568415	-1.963442403	
YLR374C	-1.865163286	-2.001379199	-1.933271242	
YKL177W	-1.272910868	-2.582977629	-1.927944249	
YAL011W - SWC3	-2.802792481	-1.004680334	-1.903736408	
YCR016W	-1.419082519	-2.357415319	-1.888248919	
YLL024C - SSA2	-2.423167361	-1.351531308	-1.887349335	
YLR384C - IKI3	-1.862658424	NA	-1.862658424	
YDL190C - UFD2	-1.50524106	-2.202412733	-1.853826897	
YNL175C - NOP13	-2.220925839	-1.464811025	-1.842868432	
YDR457W - TOM1	-2.656486655	-1.029201922	-1.842844288	
YER048C - CAJ1	-1.82224012	-1.843677837	-1.832958979	
YPR119W - CLB2	-1.170959923	-2.454280856	-1.81262039	
YIL103W - DPH1	-1.494742485	-2.001069248	-1.747905866	
YLR024C - UBR2	-1.72301901	NA	-1.72301901	
YPL131W - RPL5 - DAMP	-1.096570012	-2.284661604	-1.690615808	
YPL043W - NOP4 - DAMP	-1.173481165	-2.195701757	-1.684591461	
YFR048W - RMD8	-1.148523698	-2.219920968	-1.684222333	
YBL104C - SEA4	-1.564414952	-1.798764805	-1.681589878	
YMR043W - MCM1 - DAMP	-1.281926486	-2.075669605	-1.678798045	
YKR073C	-2.570257029	-0.785681279	-1.677969154	
YGL150C - INO80 - DAMP	-1.842091187	-1.505367196	-1.673729191	
YBR200W - BEM1	0.029627928	-3.314524537	-1.642448305	
YGL078C - DBP3	-1.628165475	NA	-1.628165475	H
YNL038W - GPI15 - DAMP	-2.049462807	-1.185110557	-1.617286682	
YER111C - SWI4	-1.183157271	-2.015890185	-1.599523728	
YJR066W - TOR1	-1.433636976	-1.743706465	-1.58867172	
YJL081C - ARP4 - DAMP	-2.27790944	-0.898117812	-1.588013626	
YER177W - BMH1	-2.098477192	-1.043572738	-1.571024965	
YJL141C - YAK1	-1.58076854	-1.558227327	-1.569497933	
YJL157C - FAR1	-0.949939649	-2.146626448	-1.548283048	
YKL135C - APL2	-1.171466099	-1.916534844	-1.544000471	
YDR008C	-2.585465867	-0.499217254	-1.542341561	
YMR100W - MUB1	-1.992787052	-1.09061876	-1.541702906	
YOR189W - IES4	-1.589970178	-1.469220112	-1.529595145	
YDL208W - NHP2 - DAMP	-1.605296242	-1.452407536	-1.528851889	
YPL142C - DAMP	-1.197628925	-1.823136144	-1.510382535	
YKL064W - MNR2	-1.172670944	-1.845457632	-1.509064288	
YLR170C - APS1	-0.968919506	-2.034322771	-1.501621138	
YNL229C - URE2	1.045450861	1.96338939	1.504420126	
YNL016W - PUB1	1.109300071	1.909353099	1.509326585	R H
YML022W - APT1	0.881853829	2.142523373	1.512188601	
YIL040W - APQ12	2.064229078	0.968220209	1.516224643	
YDL002C - NHP10	0.70495431	2.332822456	1.518888383	
YDR194C - MSS116	0.963292804	2.076984483	1.520138644	
YLR295C - ATP14	1.080262891	1.96247363	1.52136826	
YOR077W - RTS2 - DAMP	1.123260562	1.921229857	1.52224521	
YGR158C - MTR3 - DAMP	1.03743722	2.009030187	1.523233703	
YNL284C - MRPL10	0.695538895	2.377455058	1.536496976	
YMR064W - AEP1	0.929956171	2.144228137	1.537092154	
YGL167C - PMR1	1.541922927	1.540654057	1.541288492	
YML101C-A	0.435481386	2.657981628	1.546731507	
YML028W - TSA1	1.013117957	2.091089923	1.55210394	
YCR079W - PTC6	1.593115006	1.517791123	1.555453064	
YDL213C - NOP6	1.071332565	2.042142988	1.556737776	
YLR337C - VRP1	1.022208338	2.101273512	1.561740925	
YPR047W - MSF1	0.804709134	2.348274	1.576491567	
YEL003W - GIM4	1.062009643	2.094274446	1.578142045	
YDR471W - RPL27B	0.802222402	2.356417026	1.579319714	
YGL011C - SCL1 - DAMP	0.486249657	2.673611059	1.579930358	
YPR057W - BRR1	1.314998619	1.854726929	1.584862774	
YBR009C - HHF1	0.973184385	2.217984985	1.595584685	

Tableau 7a : Résultats du crible GIM réalisé à partir de la souche de levure BY4742-APIHI. (cf. légende Tableau 7c)

ORF	PIH1#1	PIH1#2	PIH1 moyenne	co-IP
YOR362C - PRE10 - DAMP	0.725754594	2.482594124	1.604174359	R H P
YKL186C - MTR2 - DAMP	1.057718441	2.155628091	1.606673266	
YKL169C	0.396601341	2.823514038	1.610057689	
YJL121C - RPE1	1.569905912	1.660765618	1.615335765	
YMR216C - SKY1	0.981376934	2.249914354	1.615645644	
YJR050W - ISY1	1.064401776	2.175764539	1.620083158	
YGR188C - BUB1	1.950266404	1.292399255	1.62133283	
YNL075W - IMP4 - DAMP	0.883371366	2.372647351	1.628009359	
YPL013C - MRPS16	0.893302134	2.404933822	1.649117978	
YOR334W - MRS2	1.200819714	2.108130938	1.654475326	
YMR194W - RPL36A	1.467430499	1.842766678	1.655098588	
YDL082W - RPL13A	1.477394342	1.842264905	1.659829623	
YGR214W - RPS0A	1.245225402	2.085716537	1.66547097	
YMR193C-A	1.50811068	1.835827771	1.671969225	
YDR156W - RPA14	1.041219076	2.306618265	1.673918671	
YLR418C - CDC73	1.017944672	2.345363236	1.681653954	
YHR004C - NEM1	0.979274866	2.398549885	1.688912375	
YER092W - IES5	0.981498088	2.417328302	1.699413195	
YGL172W - NUP49 - DAMP	1.003958345	2.40975319	1.706855768	
YDR075W - PPH3	1.506907826	1.917553118	1.712230472	
YPL213W - LEA1	0.549603588	2.883635409	1.716619498	
YPL205C	1.305321394	2.16787603	1.736598712	
YNR051C - BRE5	0.66907545	2.820741647	1.744908548	
YBR011C - IPP1 - DAMP	1.299804352	2.1930433	1.746423826	R H P
YDR500C - RPL37B	1.543957913	1.956272623	1.750115268	
YNL096C - RPS7B	1.521236333	1.982861612	1.752048973	
YJR122W - IBA57	0.890570392	2.615640601	1.753105496	
YDR392W - SPT3	1.132482503	2.378554985	1.755518744	
YBR193C - MED8 - DAMP	1.385726226	2.127823791	1.756775009	
YPL148C - PPT2	0.969017529	2.554366072	1.7616918	
YER007C-A - TMA20	1.196390196	2.330608723	1.76349946	
YOL006C - TOP1	1.245675034	2.29066211	1.768168572	
YJR145C - RPS4A	1.071332238	2.479583925	1.775458081	
YNL215W - IES2	1.194178746	2.361130256	1.777654501	
YNL067W - RPL9B	1.601536283	1.96665297	1.784094626	H
YGL115W - SNF4	1.311840051	2.259465671	1.785652861	H
YNL003C - PET8	0.941697277	2.670685715	1.806191496	
YLR018C - POM34	1.208034802	2.426304721	1.817169761	
YKR048C - NAP1	0.973781971	2.661690076	1.817736024	R H P
YOL139C - CDC33 - DAMP	1.175814568	2.483513279	1.829663924	
YKL081W - TEF4	1.341081878	2.332835845	1.836958861	R H P
YKL018W - SWD2 - DAMP	0.927257012	2.794369728	1.86081337	
YHR148W - IMP3 - DAMP	1.081512782	2.649685796	1.865599289	
YLR192C - HCR1	1.925379127	1.837025135	1.881202131	
YGR074W - SMD1 - DAMP	1.215322192	2.570418882	1.892870537	
YKL134C - OCT1	1.370204319	2.418240662	1.894222491	
YNL136W - EAF7	1.368950974	2.429443523	1.899197249	
YKR055W - RHO4	1.469420271	2.333761688	1.90159098	
YDR447C - RPS17B	1.117419384	2.693058094	1.905238739	
YNR004W - SWM2	1.340249474	2.481317129	1.910783301	
YKL156W - RPS27A	1.344851391	2.48875804	1.916804715	P
YNR047W - FPK1	1.01839383	2.819847129	1.91912048	
YNL238W - KEX2	1.290552364	2.56014822	1.925350292	
YOR340C - RPA43 - DAMP	1.284017951	2.569359546	1.926688748	
YAL009W - SPO7	1.274958536	2.58816592	1.931562228	
YNL199C - GCR2	0.880630681	2.984377145	1.932503913	
YLR185W - RPL37A	1.477058692	2.392651828	1.93485526	
YLR197W - NOP56 - DAMP	1.538885003	2.366961354	1.952923178	R
YKR057W - RPS21A	0.819599089	3.135088898	1.977343994	H
YHL033C - RPL8A	1.108140196	2.850796087	1.979468142	
YBR221C - PDB1	1.056013567	2.90769592	1.981854743	
YNL040W	1.407556977	2.564794759	1.986175868	
YIR005W - IST3	1.264582751	2.717380709	1.99098173	
YMR039C - SUB1	1.443516607	2.54104994	1.992283274	
YNL015W - PBI2	1.207124573	2.77919762	1.993161097	
YKL074C - MUD2	1.370503064	2.618232191	1.994367627	R
YIL020C - HIS6	1.208272608	2.782873625	1.995573116	
YGR085C - RPL11B	1.661265871	2.337382741	1.999324306	
YJR104C - SOD1	1.267525801	2.796578834	2.032052318	
YNL198C	0.654766224	3.418254844	2.036510534	
YMR230W - RPS10B	1.400816361	2.693807843	2.047312102	R H P
YGR055W - MUP1	1.339907355	2.756485276	2.048196316	
YLR366W	1.424177584	2.68997886	2.057078222	
YGL014W - PUF4	1.570650269	2.550494691	2.06057248	

Tableau 7b : Résultats du crible GIM réalisé à partir de la souche de levure BY4742- $\Delta$ PIH1 (suite). (cf. légende Tableau 7c)

ORF	PIH1#1	PIH1#2	PIH1 moyenne	co-IP
YOR360C - PDE2	1,156703478	2,964508765	2,060606121	
YIL142W - CCT2 - DAMP	1,458679397	2,667173461	2,062926429	
YIL153W - RRD1	1,431641735	2,729707198	2,080674466	
YPR123C	1,424974327	2,759943866	2,092459097	
YBR179C - FZO1	1,484139326	2,742113755	2,11312654	
YMR073C - IRC21	1,259480214	2,970871503	2,115175858	
YDL165W - CDC36 - DAMP	1,501906351	2,73481075	2,11835855	
YPL226W - NEW1	1,681492227	2,57763285	2,129562538	H P
YOR211C - MGM1	1,76623206	2,49901062	2,13262134	
YOR191W - ULS1	1,391165981	2,893041042	2,142103512	
YLR177W	1,285326793	3,009435726	2,147381259	
YGR081C - SLX9	0,876832464	3,448089234	2,162460849	
YGR118W - RPS23A	1,358272002	2,967326741	2,162799371	
YLR200W - YKE2	1,598388811	2,766097768	2,182243289	
YLR287C-A - RPS30A	1,219456704	3,150504492	2,184980598	
YOR167C - RPS28A	1,53463462	2,842435687	2,188535153	
YOR182C - RPS30B	1,443112313	2,957720534	2,200416423	
YGL168W - HUR1	1,583342332	2,825508242	2,204425287	
YLR388W - RPS29A	1,503301314	2,927621277	2,215461295	
YDL063C	1,917129332	2,565432038	2,241280685	
YBR174C	1,352716372	3,176630867	2,26467362	
YBR253W - SRB6 - DAMP	1,659192164	2,876336797	2,26776448	
YLR367W - RPS22B	1,534537541	3,050638917	2,292588229	
YNL255C - GIS2	0,995011386	3,612495922	2,303753654	
YGL149W	1,150267762	3,525244653	2,337756208	
YDL198C - GGC1	1,260346426	3,440454812	2,350400619	
YIL124C - LSM1	1,845694954	2,8608597	2,353277327	
YPR040W - TIP41	1,345671362	3,362425656	2,354048509	
YER131W - RPS26B	1,61838705	3,138642871	2,37851496	
YML026C - RPS18B	2,044040331	2,725538072	2,384789201	
YDL076C - RXT3	1,409954394	3,376348307	2,39315135	
YCR084C - TUP1	1,774635331	3,012216941	2,393426136	R H P
YLR407W	1,224100826	3,565295385	2,394698106	
YER146W - LSM5 - DAMP	1,642471507	3,162370763	2,402421135	
YBR167C - POP7 - DAMP	1,350878647	3,458034108	2,404456377	
YDR115W	1,329408868	3,481340704	2,405374786	
YER151C - UBP3	1,402362752	3,425800315	2,414081534	
YER178W - PDA1	1,179154417	3,678581224	2,42886782	R H P
YLR441C - RPS1A	1,254629619	3,633998782	2,444314201	
YNL170W	1,800664431	3,095749843	2,448207137	
YER169W - RPH1	1,784233887	3,121468177	2,452851032	
YDR324C - UTP4 - DAMP	1,544599365	3,391164858	2,467882112	
YLR438C-A - LSM3 - DAMP	2,17147551	2,767462153	2,469468831	
YIL041W - NSP1 - DAMP	1,664080282	3,334406234	2,499243258	
YNL005C - MRP7	1,673257849	3,343084368	2,508171108	
YBL022C - PIM1	1,592364477	3,699508762	2,645936619	
YER112W - LSM4 - DAMP	2,047970876	3,267463925	2,657717401	
YCL037C - SRO9	1,55066373	3,812426426	2,681545078	R
YAR003W - SWD1	1,796354213	3,635154136	2,715754175	
YLL045C - RPL8B	1,166483324	4,291511496	2,72899741	
YLR015W - BRE2	1,690132013	3,781100055	2,735616034	
YPR124W - CTR1	1,795843217	3,80295646	2,799399839	
YMR123W - PKR1	2,015921188	3,61784248	2,816881834	
YLR369W - SSQ1	1,87561278	3,770527583	2,823070181	
YKL121W - DGR2	1,443260508	4,279715296	2,861487902	
YBR175W - SWD3	1,934463126	3,804109864	2,869286495	
YMR038C - CCS1	1,903641748	3,991612384	2,947627066	
YDR378C - LSM6	1,695560798	4,330007691	3,012784245	
YDR469W - SDC1	1,668447819	4,527430487	3,097939153	
YNL162W - RPL42A	1,397478175	5,181310174	3,289394174	
YPL055C - LGE1	2,164008729	4,808247876	3,486128302	
YGL244W - RTF1	1,951866666	5,842753499	3,897310083	
YGR092W - DBF2	1,938295626	6,518285698	4,228290662	

**Tableau 7c : Résultats du crible GIM réalisé à partir de la souche de levure BY4742- $\Delta$ PIH1 (suite).** La première colonne indique le nom systématique et le nom standard de l'ORF. DAMP : *Decreased Abundance by mRNA Perturbation* (= insertion dans la région 3'-UTR permettant d'atténuer la production d'une protéine essentielle à la viabilité cellulaire). Les colonnes 2 et 3 indiquent les résultats de Log<sub>2</sub>(Q/R) obtenus lors de deux expériences distinctes. NA = absence de donnée. La colonne 4 représente la moyenne des valeurs de Log<sub>2</sub>(Q/R). La dernière colonne croise les résultats du crible GIM avec ceux obtenus par IP-protéomique (Rothé *et al.*, Publication n°3), en indiquant les protéines qui ont été co-purifiées spécifiquement avec Rsa1p (R), Hit1p (H) ou Pih1p (P). Seuls les résultats de Log<sub>2</sub>(Q/R) compris dans l'intervalle [-1,5; 1,5] sont représentés. Les composants des snoRNP sont symbolisés en rouge, les facteurs d'assemblage en bleu, les facteurs impliqués dans la maturation des ARN en vert, et les protéines impliquées dans la biogenèse des ribosomes en violet. Ces données ont été fournies par C. Saveanu (Institut Pasteur).

diminution du taux de Nop56p. Etant donné que l'assemblage des particules C'/D', contenant Nop56p, semble généralement dépendant de la formation des particules C/D qui contiennent Nop58p (Qu *et al.*, 2011) (cf. Introduction, Chapitre 3, 3.1.1.3.2) (**Fig. 36D**), il est envisageable que les effets épistatiques observés avec l'ORF *NOP56* résultent de phénomènes non cumulatifs. En effet, dans le cas d'un défaut d'assemblage des particules C/D, engendré par l'absence de certains facteurs (*i.e.* Rsa1p, Pih1p), il est probable qu'un défaut d'assemblage de Nop56p sur le motif C'/D' soit engendré de façon consécutive et automatique. Dans ce cas, le taux de protéine Nop56p influencerait relativement peu sur la suite du processus de biogenèse et sa diminution ne conduirait pas à un effet délétère supplémentaire.

Comme *RSAl*, l'ORF *PIHl* présente également de forts liens épistatiques avec des gènes codant des facteurs de terminaison de la transcription des snoRNA. C'est en particulier le cas des ORF *RTF1* (+3,90) et *CDC73* (+1,68) codant des membres du complexe PAF1 et du gène *SWD2* (+1,86), dont le produit est un constituant du complexe APT (**Fig. 12**). De plus, l'ORF *PIHl* est également liée aux ORF *DBR1* (-2,30) et *RRP6* (-2,30), qui codent respectivement l'enzyme de débranchement du lasso intronique Dbr1p (cf. Introduction, Chapitre 1, 2.3.2) et la sous-unité catalytique de l'exosome Rrp6p (cf. Introduction, Chapitre 1, 2.4.2). Ces observations illustrent donc à nouveau les liens étroits entre l'assemblage des RNP et les processus de terminaison de la transcription et de maturation des snoRNA. A ce titre, nous avons également constaté un effet épistatique important entre *PIHl* et les ORF *LSM1* (+2,35), *LSM3* (+2,47), *LSM4* (+2,66), *LSM5* (+2,40) et *LSM6* (+3), codant les protéines Lsm (cf. Introduction, Chapitre 1, 2.4.2). Ce résultat surprenant suggère qu'une altération de la formation du complexe Lsm pourrait compenser ou supprimer l'effet délétère dû à la délétion de Pih1p. En l'absence de données complémentaires, il apparaît difficile de formuler une interprétation solide de ce phénomène. Néanmoins il est possible qu'une mauvaise fixation du complexe Lsm sur les pré-snoRNA, facilite le remodelage des pré-snoRNP et la transition vers l'état mature, en dépit de l'affaiblissement de la composante catalytique de la machinerie d'assemblage (R2TP→R2T). De façon intéressante, des éléments de réflexions similaires avaient déjà été formulés au sujet de la snoRNP U3 de levure puisque l'association des protéines Lsm sur le snoRNA semble incompatible avec la fixation stable des protéines constitutives de la RNP (Kufel *et al.*, 2000) (cf. Introduction, Chapitre 3, 3.3.1). Nous rediscuterons de ces aspects dans le Chapitre 5 de la partie Résultat & Discussion, qui traite plus spécifiquement de l'assemblage de la snoRNP U3.

L'ensemble de ces résultats, pour certains confirmés par les données de cribles d'interactions génétiques haut-débits disponibles dans la littérature (Collins *et al.*, 2007; Wilmes *et al.*, 2008), ouvre donc un nouveau champ d'investigations. Les événements de couplage entre la maturation du snoRNA et l'assemblage de la RNP sont en effet, encore relativement peu connus et l'étude des liens physiques entre les machineries d'assemblage et de maturation des snoRNA constituent une perspective intéressante. A ce titre, quelques données ont déjà été obtenues par l'approche de double-hybride, puisque Pih1p interagit avec la protéine Lsm8p (Fromont-Racine *et al.*, 1997), tandis que Rsa1p établit des interactions conservées avec Lsm4p et avec l'hyperméthylase Tgs1p (Thèse S. Boulon).

Parmi les autres *loci* génétiquement liés à l'ORF *PIH1*, de nombreuses fonctions cellulaires sont représentées. On retrouve notamment des gènes codant des protéines ribosomiques, des facteurs impliqués dans la biogenèse des ribosomes, des gènes dont les produits sont impliqués dans l'ubiquitinylation et la dégradation des protéines, ainsi que de nombreux gènes codant des protéines impliquées dans le remodelage de la chromatine. Cette dernière catégorie est particulièrement bien représentée, puisque *PIH1* interagit génétiquement avec les membres du complexe COMPASS de méthylation de l'histone H3 (*SWD1*, *SWD2*, *SWD3*, *BRE2*), avec les membres du complexe de remodelage INO80 (*INO80*, *IES2*, *IES4*, *IES5*, *NHP10*), avec la sous-unité *EAF7* du complexe NuA4 d'acétylation des histones, avec la sous-unité *RXT3* du complexe RPD3L de déacétylation des histones, avec la sous-unité *HHF1* de l'histone H4, ainsi qu'avec le répresseur transcriptionnel *TUP1*, dont le produit, qui interagit avec les histones H3 et H4, a été co-purifié avec Rsa1p, Hit1p et Pih1p (Rothé *et al.*, Publication n°3) (**Tableau 7**). La mise à jour de ces multiples interactions génétiques suggère donc l'existence de relations fonctionnelles privilégiées entre Pih1p et les machineries de remodelage de la chromatine. Une perspective intéressante serait : (i) d'établir si ces différents liens sont dépendants du complexe R2TP, comme cela a été démontré pour les complexes INO80 et SWR-C (Zhao *et al.*, 2005) (cf. Introduction, Chapitre 3, 3.6.3), et (ii) si de telles fonctions peuvent être impliquées dans des processus de régulation de la transcription des snoRNA. Nous reviendrons sur ce dernier point dans le chapitre 6 de la partie Résultats & Discussion.

Une interaction génétique avec les ORF *TOR1* (-1,59) et *TIP41* (+2,35) a également été détectée. Les produits de ces deux gènes appartiennent à la voie de signalisation TOR qui a récemment été impliquée dans le processus de régulation de la transcription du pré-ARNr 35S (Mayer & Grummt, 2006). L'éventualité d'un lien fonctionnel entre Pih1p et le complexe

TOR, pourrait donc constituer une piste de réflexion sur le rôle du facteur d'assemblage au sein du nucléole. A la différence des autres protéines comme Rsa1p ou Hit1p, Pih1p s'accumule en effet au sein de ce domaine nucléaire, où elle pourrait participer de façon directe à la production des ARNr (Gonzales *et al.*, 2005) (cf. Introduction, Chapitre 3, 3.6.2).

Enfin, parmi les données les plus marquantes, nous avons également retenu le fort défaut synthétique de croissance observé pour la double délétion  $\Delta PIH1-\Delta ADO1$  (-4,30). L'ORF *ADO1* code en effet, une adénosine kinase impliquée dans l'utilisation de la S-adénosyl méthionine (SAM) (Lecoq *et al.*, 2001). Ce résultat apparaît d'autant plus intéressant que la SAM est le co-facteur indispensable à l'activité de 2'-O-méthylation de l'enzyme Nop1p/Fibrillarine (cf. Introduction, Chapitre 3, 3.1.1.2.1), et que la protéine Ado1p a été co-purifiée avec les protéines Rsa1p et Hit1p (Rothé *et al.*, Publication n°3) (**Tableau 7**).

#### **2.2.4.4. Description du réseau d'interaction génétique impliquant l'ORF *TAHI***

Les résultats obtenus pour l'ORF *TAHI* sont relativement restreints et la seule donnée notoire concerne le défaut synthétique de croissance observé avec l'ORF *RVB2-DAMP* (-2,7) (**Tableau 8**). Cette interaction confirme le lien mis à jour entre Tah1p et l'hélicase au sein du complexe R2TP (Zhao *et al.*, 2005). D'autre part, elle révèle que la fonction de Tah1p, pourrait dépendre d'avantage de Rvb2p, que de Rvb1p. De façon plus générale, le faible nombre d'interactants génétiques relevés pour *TAHI*, semble à l'image de la grande neutralité mise à jour pour le phénotype  $\Delta TAHI$  (Rothé *et al.*, Publication n°3). Etant donné, que le rôle du facteur Tah1p semble principalement lié à la phase stationnaire de croissance ((Zhao *et al.*, 2008); Rothé *et al.*, Publication n°3), il est cependant possible que l'approche GIM, basée sur la culture des cellules en phase exponentielle de croissance, ne constitue pas l'approche la plus adaptée afin d'apprécier pleinement les effets liés à la délétion du gène *TAHI*.

ORF	TAH1#1	TAH1#2	TAH1 moyenne	co-IP
YPL235W - RVB2 - DAMP	-3.128849384	-2.258607642	-2.693728513	
YOR014W - RTS1	NA	-2.32002248	-2.32002248	
YDR245W - MNN10	NA	-1.762143726	-1.762143726	
YPL262W - FUM1	-1.608919127	-1.887201235	-1.748060181	
YFL022C - FRS2 - DAMP	-2.49718587	-0.994179585	-1.745682727	
YOR309C	-1.726696356	-1.645706331	-1.686201343	
YOR203W - DAMP	NA	-1.638313705	-1.638313705	
YIL128W - MET18	-2.428454809	-0.83967397	-1.634064389	
YER025W - GCD11 - DAMP	-2.765597302	-0.491623033	-1.628610168	R H P
YLR242C - ARV1	-1.721175246	-1.515365688	-1.618270467	
YOR106W - VAM3	-2.095411527	-1.13052378	-1.612967654	
YKR072C - SIS2	-1.422877382	-1.777806877	-1.60034213	
YGL212W - VAM7	-1.399938198	-1.731699794	-1.565818996	
YKR007W - MEH1	NA	1.632167851	1.632167851	

**Tableau 8 : Résultats du crible GIM réalisé à partir de la souche de levure BY4742- $\Delta$ TAH1.** La première colonne indique le nom systématique et le nom standard de l'ORF. DAPM : *Decreased Abundance by mRNA Perturbation* (= insertion dans la région 3'-UTR permettant d'atténuer la production d'une protéine essentielle à la viabilité cellulaire). Les colonnes 2 et 3 indiquent les résultats de Log<sub>2</sub>(Q/R) obtenus lors de deux expériences distinctes. NA = absence de donnée. La colonne 4 représente la moyenne des valeurs de Log<sub>2</sub>(Q/R). La dernière colonne croise les résultats du crible GIM avec ceux obtenus par IP-protéomique (Rothé *et al.*, Publication n°3), en indiquant les protéines qui ont été co-purifiées spécifiquement avec Rsa1p (R), Hit1p (H) ou Pih1p (P). Seuls les résultats de Log<sub>2</sub>(Q/R) compris dans l'intervalle [-1,5; 1,5] sont représentés. Les composants des snoRNP sont symbolisés en rouge, les facteurs d'assemblage en bleu, les facteurs impliqués dans la maturation des ARN en vert, et les protéines impliquées dans la biogenèse des ribosomes en violet. Ces données ont été fournies par C. Saveanu (Institut Pasteur).

### 3. Conclusion & Perspectives

#### 3.1. Bilan

Les résultats présentés dans cette partie ont permis de mieux cerner les composants de la machinerie d'assemblage et d'en définir les principales sous-unités fonctionnelles. L'approche d'analyse protéomique nous a tout d'abord permis d'identifier la protéine Hit1p comme nouveau facteur d'assemblage des snoRNP à boîtes C/D. Cette dernière, qui interagit physiquement avec le facteur Rsa1p, est en effet requise pour assurer sa stabilité *in vivo* et contribue de façon majeure à la biogenèse des snoRNP. Comme c'est le cas pour les autres facteurs, son implication dans les processus d'assemblage pourrait être conservée chez les eucaryotes, puisque son homologue chez l'homme, la protéine TRIP3, interagit avec NUFIP (E. Bertrand, Communication personnelle). D'autre part, nos résultats indiquent également que le duo Rsa1p-Hit1p et le complexe R2TP ne font pas l'objet d'une association stable *in vivo* et qu'ils constituent deux entités distinctes de la machinerie d'assemblage. Des différences d'ordre phénotypiques ont notamment été constatées entre les souches de levures invalidées pour les gènes d'intérêt, puisque Pih1p semble exercer une influence sur la stabilité de l'ensemble des snoRNA à boîtes C/D, tandis que Rsa1p et Hit1p pourraient être spécifiques d'une partie seulement de ces ARN. Des différences d'ordre génétiques ont également été observées entre les souches  $\Delta RSAI$  et  $\Delta PIHI$ , puisqu'en dépit de partenaires communs, les ORF codant les deux facteurs d'assemblage établissent des liens spécifiques avec certains gènes. De forts effets épistatiques, avec les gènes codant les protéines Lsm, ont notamment été constatés pour l'ORF *PIHI*, tandis qu'une interaction spécifique entre les ORF *RSAI* et *REX4* a été mise en évidence. Par la suite, ce dernier lien devra faire l'objet d'une attention particulière. Il sera en effet nécessaire de vérifier : (i) si la protéine Rex4p intervient dans la maturation des snoRNA à boîtes C/D, (ii) si son rôle potentiel est relatif aux snoRNA dont la stabilité est influencée sélectivement par Rsa1p et Hit1p.

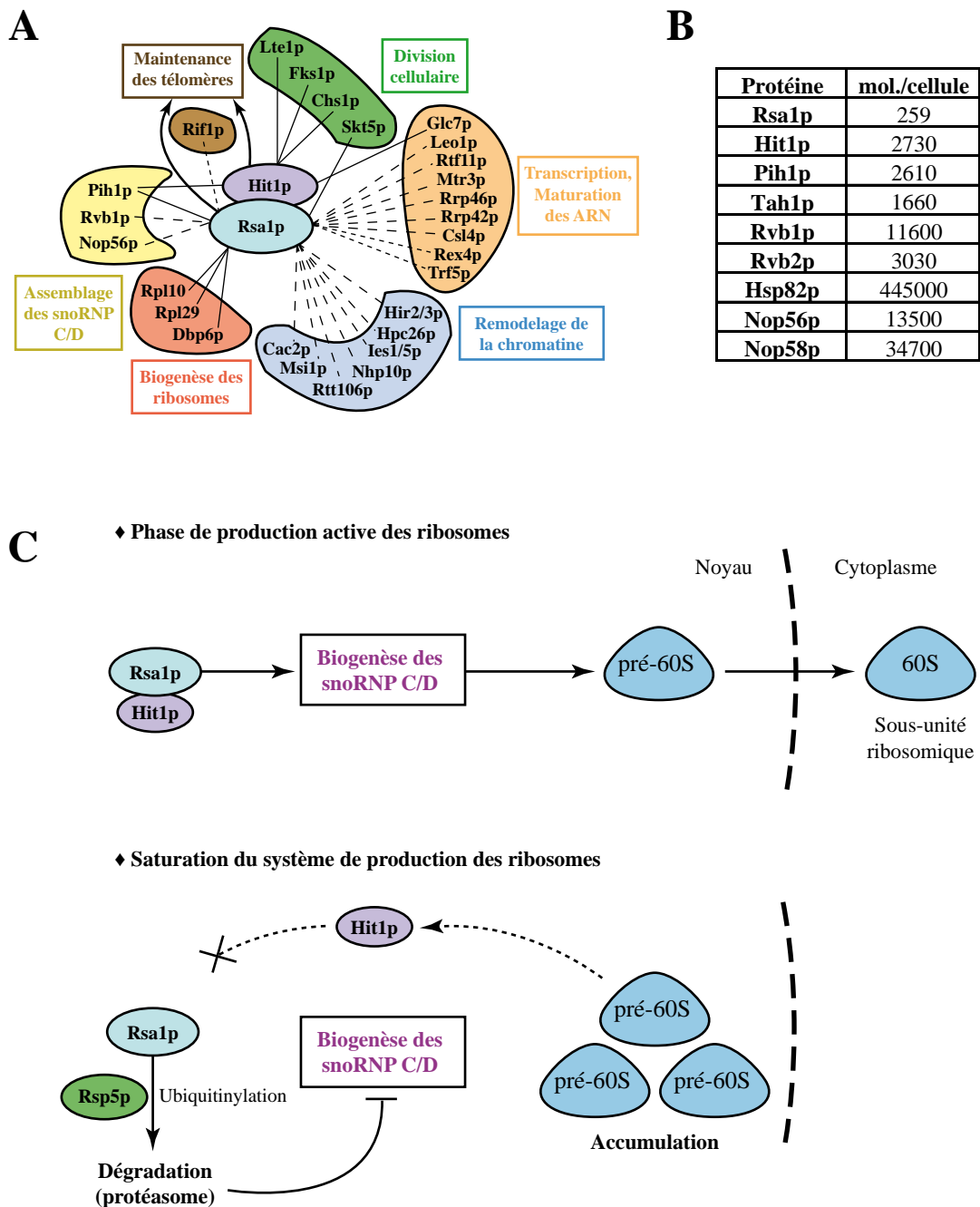
En dépit de ces différences, de nombreuses données font cependant état de l'interaction physique et fonctionnelle de ces facteurs au cours du processus d'assemblage. Rsa1p, Hit1p, les membres du complexe R2TP et les protéines constitutives des snoRNP sont en effet détectés en association avec les mêmes précurseurs du snoRNA U3, suggérant leur présence transitoire au sein des mêmes complexes lors de la biogenèse de la snoRNP. De plus, la caractérisation fonctionnelle des principaux facteurs a révélé une forte corrélation entre la



fonctionnalité de ces protéines et leurs capacités à interagir ensemble. Le domaine fonctionnel de Rsa1p concentre en effet les déterminants requis pour interagir avec Hit1p, Snu13p, Pih1p et Nop58p, tandis que le domaine fonctionnel de Pih1p est responsable des interactions avec Nop58p et Rsa1p. Ces observations, ainsi que la mise à jour de phénotypes synthétiques létaux entre *RSA1*, *HIT1* et *PIH1*, confirment donc l'importance de ce réseau d'interactions et suggère que l'association transitoire du complexe R2TP avec le duplex Rsa1p-Hit1p constitue un événement primordial pour l'assemblage des snoRNP. Comme nous l'avons abordé dans le chapitre précédent, l'utilisation de certains mutants protéiques, visant à ralentir la formation des snoRNP matures, devra être mise à profit, afin d'accumuler des intermédiaires d'assemblage dans les cellules et de réaliser une caractérisation plus approfondie de leur composition.

### 3.2. Rsa1p, Hit1p et régulation métabolique ?

Une perspective majeure de ce travail concerne la relation fonctionnelle entre les facteurs Rsa1p et Hit1p. Alors que Rsa1p favorise la nucléation des pré-snoRNP en établissant des contacts avec de nombreux partenaires, les données actuelles indiquent que le spectre d'interaction de Hit1p pourrait se restreindre à Rsa1p. Son rôle n'en est cependant pas moins important, puisque son expression apparaît essentielle à l'accumulation de Rsa1p et que l'absence de l'un ou l'autre des facteurs conduit aux mêmes effets phénotypiques. De façon intéressante, cette redondance phénotypique est observée pour d'autres paramètres cellulaires, comme la taille des télomères qui augmente à la fois dans les souches  $\Delta RSA1$  et  $\Delta HIT1$  (Gatbonton *et al.*, 2006) (**Fig. 73A**). Rsa1p et Hit1p semblent également reliées à des fonctions relevant de la division cellulaires, puisque *HIT1* est synthétique létal avec le gène *CHS1*, codant une chitine synthase, tandis que *RSA1* est synthétique létal avec le gène *SKT5*, qui code une autre protéine relative à cette fonction (Lesage *et al.*, 2005; Tong *et al.*, 2004) (**Fig. 73A**). Ce constat suggère donc que la fonction de ces deux protéines est intrinsèquement liée. Hit1p pourrait en effet agir en tant que co-facteur capable de potentialiser la fonction de la protéine Rsa1p, notamment en la protégeant de la dégradation. Afin de confirmer cette hypothèse, il sera cependant nécessaire de vérifier si un défaut d'interaction entre les deux partenaires conduit aux mêmes effets que l'absence de Hit1p. Dans cette optique, la structure du complexe Rsa1p<sub>317-352</sub>-Hit1p<sub>70-164</sub>, actuellement en cours de résolution au laboratoire, apportera les informations structurales nécessaires à la définition de variants d'interaction.



**Figure 73 : Implication potentielle de Rsa1p et Hit1p dans des événements de régulation.** (A) Bilan des principaux liens génétiques mis à jour entre Rsa1p, Hit1p et différents processus cellulaires (données issues de ce travail et de la base de données Saccharomyces Genome Database; Cherry *et al.*, 1997; www.yeastgenome.org/). Les traits en pointillés fins représentent des défauts synthétiques de croissance, les traits en pointillés larges, des liens épistatiques, et les traits pleins, des liens synthétiques létaux. Les flèches matérialisent l'influence des protéines Rsa1p et Hit1p sur la taille des télomères. (B) Abondance des différents facteurs d'intérêt, estimée en nombre de molécules par cellule, sur la base de l'immunodétection de protéines étiquetées (Ghaemmaghami *et al.*, 2003). (C) Exemple de couplage potentiel entre assemblage des snoRNP C/D et biogenèse des ribosomes, impliquant Rsa1p et Hit1p.

De plus, il sera également important de vérifier si Rsa1p fait l'objet d'une dégradation spécifique par le système ubiquitine-protéasome. Sur la base d'un crible, réalisé à partir d'un système de puce à protéine, Rsa1p a en effet été identifiée comme l'un des substrats privilégiés de l'ubiquitine-ligase E3 Rsp5p (Gupta *et al.*, 2007). Cette dernière cible le facteur d'assemblage en interagissant avec une séquence de type PPXY (*i.e.* P<sub>95</sub>P<sub>96</sub>N<sub>97</sub>Y<sub>98</sub>) située dans sa région N-terminale et pourrait ainsi favoriser sa dégradation par le protéasome. Afin de confirmer l'implication de Rsp5p dans la régulation du taux de Rsa1p, il sera donc nécessaire de vérifier si l'absence du domaine N-terminal et si la mutation du motif PPNY protègent Rsa1p de la dégradation dans une souche  $\Delta HIT1$ . De plus, l'implication du protéasome pourra également être vérifiée grâce à l'utilisation de peptides inhibiteurs (*e.g.* MG132), dans des souches mutantes (*e.g.*  $\Delta ERG6$  ou  $\Delta PDR5$ ) dont la perméabilité membranaire favorise l'entrée des peptides inhibiteurs dans la cellule (Liu *et al.*, 2007a).

Enfin, il sera nécessaire de vérifier l'importance fonctionnelle de ce phénomène, en particulier en termes de régulation. En effet, de part son rôle central dans l'assemblage des snoRNP, la protéine Rsa1p pourrait constituer un candidat de choix pour exercer un contrôle sur le taux d'assemblage des snoRNP. Cette éventualité paraît d'autant plus vraisemblable qu'en comparaison avec les autres protéines du système d'assemblage, la très faible abondance de Rsa1p dans les cellules de levure pourrait en faire le facteur limitant de ce processus (Ghaemmaghami *et al.*, 2003) (**Fig. 73B**). Une hypothèse est que la dépendance de Rsa1p vis-à-vis de Hit1p permet d'assurer une réponse rapide en cas de saturation du système de production des ribosomes. Dans un contexte d'accumulation de complexes pré-ribosomiques au sein du noyau, la protéine Hit1p pourrait par exemple être titrée par un partenaire d'interaction, avec pour conséquence une dégradation rapide du facteur Rsa1p et un ajustement coordonné de la production des snoRNP à boîtes C/D (**Fig. 73C**). La prise en compte d'une telle possibilité apparaît d'autant plus intéressante que Rsa1p s'est elle-même avérée requise pour l'export optimal de la sous-unité pré-60S (Gadal *et al.*, 2001) (cf. Introduction, Chapitre 3, 3.6.6.1). Une perspective serait donc d'évaluer le taux de Rsa1p dans des cellules de levure affectées transitoirement pour les étapes tardives de biogenèse des ribosomes. A ce titre, l'utilisation de souches exprimant des formes variantes de la protéine Nmd3p, qui contrôle le processus d'export de la sous-unité pré-60S pourra notamment être envisagée (Gadal *et al.*, 2001; Hedges *et al.*, 2005) (cf. Introduction, Chapitre 1, 3.4) (**Fig. 3 et 20**). Une telle démarche pourrait notamment nous conduire à aborder la question du lien entre Rsa1p et l'assemblage de la grande sous-unité ribosomique (Kressler *et al.*, 1999). Il

serait en effet intéressant de définir si le rôle initialement attribué à ce facteur, rend compte de son implication physique dans ce processus, ou, s'il découle d'effets indirects, liés à des phénomènes de co-régulation avec le processus de biogenèse des snoRNP ? Par ailleurs, ces expériences devront être conduites en présence de la protéine Hit1p entière, mais également à partir de la forme Hit1p<sub>44-164</sub>, afin de vérifier si le domaine à doigt de zinc (1-43) est requis pour la fonctionnalité de la protéine dans de telles conditions.

Ces différents éléments de réflexion nous amènent donc à réévaluer l'impact des conditions expérimentales utilisées au cours de notre étude. En effet, la majorité de nos résultats ont été obtenus dans des conditions de croissance optimales (milieu riche, 30°C, phase exponentielle de croissance) et renseignent sur le fonctionnement de la machinerie d'assemblage au cours de la phase de production active des ribosomes. Il n'en demeure pas moins, que certains éléments de cette machinerie, comme le domaine N-terminal de Rsa1p ou encore le domaine à doigt de zinc de Hit1p, qui s'avèrent non essentiels dans de telles conditions, pourraient jouer un rôle majeur dans des circonstances alternatives. La prise en compte des variations environnementales, par l'induction de différents stress et par l'étude des cellules en phase stationnaire de croissance, constituent donc une perspective importante. Comme nous l'avons évoqué dans l'introduction, de nombreuses études font en effet état de liens étroits entre le contexte environnemental, la production des ribosomes et la régulation de la croissance cellulaire (Mayer & Grummt, 2006). Les multiples relations établies entre les composants de la machinerie d'assemblage et d'autres processus majeurs, comme les étapes tardives de la biogenèse des ribosomes (Rsa1p), la division cellulaire (Rsa1p, Hit1p), la synthèse des télomères (Rsa1p, Hit1p) ou encore le remodelage de la chromatine (R2TP), nous poussent donc à reconsidérer l'envergure fonctionnelle de ces facteurs. De part leur rôle clé dans l'assemblage des snoRNP, ces derniers constituent en effet des candidats de choix pour l'intégration du processus biogenèse des snoRNP dans le cadre des régulations métaboliques globales.

## CHAPITRE 4

### Etude du rôle de Hsp82p et de sa co-chaperonne Tah1p dans le processus de biogenèse des snoRNP à boîtes C/D

#### 1. Présentation des travaux

L'implication des protéines HSP90 dans la biogenèse des RNP de la famille L7Ae a été démontrée chez l'homme (Boulon *et al.*, 2008) (cf. Résultats & Discussion, Chapitre 1). Ces chaperonnes contrôlent en effet la stabilité de certaines protéines constitutives comme NOP58, 15.5K ou NHP2, et exercent une influence majeure sur la production des particules et sur le taux des ARN non codants qui les composent. Chez la levure, cette fonction ne semble en revanche pas strictement conservée, puisque l'accumulation des protéines constitutives des snoRNP à boîtes C/D s'est avérée indépendante des chaperonnes, et que ces dernières n'exercent qu'une influence mineure sur le taux des snoRNA durant la phase stationnaire de croissance (Zhao *et al.*, 2008). En dépit de ces différences, le mode d'intervention des protéines HSP90 dans ces processus pourrait cependant être conservé. Il semble en effet reposer sur la présence, au sein du complexe R2TP, d'une co-chaperonne à domaine TPR, la protéine Tah1p/SPAGH, capable d'interagir spécifiquement avec les chaperonnes HSP90 (Boulon *et al.*, 2008; Zhao *et al.*, 2005; Zhao *et al.*, 2008) (cf. Introduction, Chapitre 3, 3.6.3).

Afin de mieux comprendre les bases moléculaires du rôle spécifique joué par HSP90 dans la biogenèse des RNP de la famille L7Ae, des études structurales par RMN ont été entreprises au laboratoire par R. Back. Menés en collaboration avec l'équipe de F. Allain de l'ETH à Zurich, ces travaux ont abouti à la résolution de la structure en solution de la protéine Tah1p libre et de son complexe avec l'extrémité C-terminale de Hsp82p. Les résultats obtenus ont ainsi permis de définir avec précision deux modules distincts au sein de la protéine Tah1p : un domaine TPR atypique, constitué de 5 hélices antiparallèles et responsable de l'interaction avec Hsp82p, ainsi qu'un domaine C-terminal non structuré, nécessaire et suffisant pour assurer l'interaction avec Pih1p. De plus, la définition des déterminants d'interaction entre Tah1p et la chaperonne, nous ont permis de définir différents mutants et de tester leur effet *in vivo*. En parallèle de l'emploi de la geldanamycine, un inhibiteur général des chaperonnes

HSP90, l'utilisation de formes variantes de la protéine Tah1p, a donc constitué un outil de choix afin d'étudier spécifiquement le rôle des chaperonnes dans la biogenèse des snoRNP à boîtes C/D. Cette approche, nous a ainsi conduit à deux constats majeurs : (i) chez la levure, Tah1p et Hsp82p ne semblent pas requises pour la production des snoRNP à boîtes C/D dans des conditions standard de croissance, (ii) en phase stationnaire de croissance, Tah1p contrôle la stabilité de Pih1p indépendamment de son interaction avec Hsp82p.

## **2. Résultats**

L'ensemble de ce travail a abouti à la rédaction d'un article qui sera soumis dans la revue *Structure*. Des résultats d'analyses fonctionnelles seront également présentés en complément de l'article.

### **2.1. Publication n°4**

Ma contribution personnelle à cet article concerne les études réalisées par l'approche de double-hybride : la caractérisation des domaines d'interactions entre les différentes protéines et la caractérisation des variants de l'interaction entre Tah1p et Hsp82p.

---

---

Publication n°4

**The NMR high resolution structures of yeast Tah1  
in a free form and in complex with the Hsp90 C-  
terminal tail explains the specific recognition of  
Hsp90 by the R2TP complex**

Régis Back, Cyril Dominguez, Benjamin Rothé, Claude Bobo, Chrystel Beaufils, Solange Moréra, Philippe Meyer, Bruno Charpentier, Christiane Branlant, Frédéric Allain and Xavier Manival

**En préparation pour *Structure***

---

---





# The NMR high resolution structures of yeast Tah1 in a free form and in complex with the Hsp90 C-terminal tail explains the specific recognition of Hsp90 by the R2TP complex

Régis Back<sup>1</sup>, Cyril Dominguez<sup>2,+</sup>, Benjamin Rothé<sup>1</sup>, Claude Bobo<sup>1</sup>, Chrystel Beaufiles<sup>3</sup>, Solange Moréra<sup>4</sup>, Philippe Meyer<sup>4,\*</sup>, Bruno Charpentier<sup>1,§</sup>, Christiane Branlant<sup>1</sup>, Frédéric Allain<sup>2,§</sup>, and Xavier Manival<sup>1,§</sup>

(1) AREMS (ARN, RNP, structure-fonction-maturation, Enzymologie Moléculaire et Structurale), Nancy Université, UMR 7214 CNRS-UHP, Bâtiment Biopôle, Faculté de médecine, 9 Avenue de la Forêt de Haye, BP184, 54506 Vandoeuvre-lès-Nancy Cedex, France.

(2) IMB (Institute of Molecular Biology and Biophysics), ETH, Zurich, Suisse.

(3) LCPM, Groupe Synthèse Organique et Biostructures, UMR CNRS-INPL 7568, 1, rue Grandville, BP 20451, F- 54001 Nancy Cedex, France.

(4) LEBS (Laboratoire d'Enzymologie et Biochimie Structurales) UPR CNRS 3082, Bât. 34, 91198 Gif-sur-Yvette Cedex, France.

+ Present address: Henry Wellcome Laboratories of Structural Biology, Department of Biochemistry, University of Leicester, Leicester LE1 9HN, United Kingdom.

\* Present address: Laboratoire de Biologie Moléculaire et Cellulaire des Eucaryotes - Institut de Biologie Physico-Chimique, CNRS-UPMC, 15 rue Pierre et Marie Curie 75005 Paris, France.

§ To whom correspondence should be addressed

[xavier.manival@maem.uhp-nancy.fr](mailto:xavier.manival@maem.uhp-nancy.fr)

[allain@mol.biol.ethz.ch](mailto:allain@mol.biol.ethz.ch)

[bruno.charpentier@maem.uhp-nancy.fr](mailto:bruno.charpentier@maem.uhp-nancy.fr)

## SUMMARY

**Hsp90 is a ubiquitous molecular chaperone conserved in bacteria and eukarya essential for cell viability. Tah1, known to be a cofactor of Hsp90, was proposed to contain two TPR motifs forming a single TPR domain interacting with the C-terminus of Hsp90. In yeast, Pih1, Rvb1, and Rvb2 form a multiprotein complex called R2TP involved in box C/D snoRNP and RNA polymerases assemblies. Here, we report here the solution structures of Tah1 free and its ADTEMEEVD peptide bound form. Tah1 structure shows the presence of a 1-92 N-terminal domain including two TPR motifs followed by a C helix and, a unstructured 93-111 C-terminal domain. These two domains are responsible for the binding of Hsp90 C-ter and Pih1 C-ter respectively. The binding of Tah1 to C-terminal Hsp90 is mediated by the EMEEVD extended motif of the chaperone. Based on these results, a role of Tah1 for the Hsp90 recognition by R2TP complex is proposed.**

## INTRODUCTION

The highly conserved eukaryotic Hsp90 (90 kDa heat-shock protein) protein is an ubiquitous ATP-dependent molecular chaperone which is essential for cell viability (Borkovich et al., 1989). In addition to its major activity as a chaperone protein in conditions of stress, Hsp90 is required for conformational maturation of signal transduction proteins, which links the Hsp90 activity to cell cycle regulation and also to cancer progression (Caplan et al., 2007; Maloney and Workman, 2002; Pearl and Prodromou, 2006; Whitesell and Lindquist, 2005). This specialized role of Hsp90 depends upon its specific recognition of a set of protein substrates

designated as client proteins. They includes: nuclear hormone receptors (androgen and estrogen receptors), other transcription factors (p53, HIF-1), protein kinases (ERBB2, B-RAF, C-RAF, and CDK4), and a group of non-signaling proteins, one of them, hTERT, being a major component of telomerase (Caplan et al., 2007; Maloney and Workman, 2002; Pearl and Prodromou, 2006). Furthermore, new very important roles of Hsp90 were recently discovered. They concern the assembly of several essential eukaryotic ubiquitous cellular machineries: i) all RNPs containing an RNA-binding protein of the L7Ae family, namely, the C/D box snoRNPs that guide and catalyse 2'-O methylations in pre-ribosomal RNA and are

required for pre-rRNA cleavages, the H/ACA snoRNPs which catalyze pseudouridine formation in the pre-rRNA, the spliceosomal U4 snRNP, the telomerase RNP responsible for telomere synthesis, and selenoprotein mRNPs (Boulon et al., 2008; Zhao et al., 2008), ii) the nuclear RNA polymerases (Boulon et al., 2010) and, iii) the phosphatidylinositol 3-kinase-related protein kinases cell-cycle control family, including proteins mTOR (mammalian Target Of Rapamycin) and SMG1 (Suppressor of Morphogenetic effect on Genitalia-1) (Horejsi et al., 2010). How Hsp90 is involved in these new highly important functions still has to be determined.

Progresses in the understanding of other previously identified Hsp90 functions are useful to understand its new functions, we will summarize this knowledge. Hsp90 acts at a late stage of protein folding. It interacts with its client proteins when they are in a near-native state. As a result, the client proteins acquire activity (Pearl and Prodromou, 2001). Hsp90 is also involved in the degradation of trapped misfolded proteins via the ubiquitin-proteasome pathway (Connell et al., 2001). The Hsp90 chaperone activity depends upon the ATPase activity carried by the N-terminal and middle domains of Hsp90 (Richter et al., 2008), and inhibition of this activity by geldanamycin, blocks cell proliferation (Schneider et al., 1996; Whitesell and Cook, 1996; Whitesell et al., 1994). Interestingly, geldanamycin was also shown to block the newly discovered functions of Hsp90 in snoRNP and RNA polymerase assembly (Boulon et al., 2008; Zhao et al., 2008). The Hsp90 C-terminal domain allows Hsp90 dimerization (Minami et al., 1994). In its dimeric form Hsp90 acts as a molecular clamp whose closure is mediated by transient ATP hydrolysis (Ali et al., 2006).

To achieve its specific protein maturation activity, Hsp90 cooperates with another molecular chaperone, Hsp70, and numerous co-factors called co-chaperones. Through modulation of ATP affinity or ATP hydrolysis, these co-chaperones participate in the regulation of Hsp90 ATPase cycle (Panaretou et al., 2002; Prodromou et al., 1999; Richter et al., 2003; Siligardi et al., 2004; Siligardi et al., 2002), and in the specific recognition of the protein substrates (Echtenkamp and Freeman, 2011; Grammatikakis et al., 1999; Riggs et al., 2004; Roe et al., 2004). Many co-chaperones contain a tetratricopeptide repeat domain (TPR domains) which mediates their association with the 20

amino acid C-terminal sequence of Hsp90. This Hsp90 terminal sequence contains the MEEVD motif required for TPR recognition (Prodromou et al., 2000; Scheufler et al., 2000; Young et al., 1998). In addition, the Hsp90 C-terminal domain and/or the middle domain participate to the recognition of the client proteins (Harris et al., 2004; Meyer et al., 2003; Panaretou et al., 1999; Prodromou et al., 1997; Prodromou et al., 1999; Stebbins et al., 1997; Street et al., 2011).

Numerous 3D structures of individual domains and full-length Hsp90 (*Escherichia coli* HtpG and yeast Hsp82 proteins) have been solved at high resolution (Ali et al., 2006; Harris et al., 2004; Meyer et al., 2003; Prodromou et al., 1997; Shiao et al., 2006). The apo-enzyme is in an open V shape, while the AMP-PNP bound enzyme is in a closed twisted conformation. This closed conformation is stabilized by interaction with co-chaperones, and contrary to the expectation, recent data revealed that in its closed conformation, the Hsp90 dimer would not enclose its client protein but provides a bipartite binding surface whose formation and disruption are coupled to the chaperone ATPase cycle (Ali et al., 2006; Shiao et al., 2006; Vaughan et al., 2006).

The crystal structures established for Hsp90 and Hsp70 C-terminal domains in complex with partner TPRs brought information on determinants ensuring their specific recognitions (Scheufler et al., 2000). For instance, Hsp70 and Hsp90 recognize distinct TPR domains in the Hop (HSP-Organizing Protein) co-factor (TPR1 and TPR2A, respectively). TPR domains contain 3 up to 16 TPR motifs, that consist in a 34 amino acid sequence (D'Andrea and Regan, 2003) with only a few conserved amino acids (Gly/Ala8, Ala 20, Tyr24, and Pro32) (Blatch and Lassel, 1999; D'Andrea and Regan, 2003; Lamb et al., 1995; Prasad et al.). However, based on the numerous crystal structures that were established (Das et al., 1998; Scheufler et al., 2000; Sinars et al., 2003; Taylor et al., 2001; Ward et al., 2002; Zhang et al., 2005), TPR domains share a common helix-turn-helix structure, that packs in a parallel fashion resulting in a spiral of repeating anti-parallel  $\alpha$  helices. Their essential two-carboxylate clamp interacts with the C-terminal EEVD motif present in both Hsp70 and Hsp90. (Russell et al., 1999; Scheufler et al., 2000). Discrimination of Hsp90 or Hsp70 by the co-chaperone depends upon the identity of the amino

acid located upstream of the EEVD sequence (Scheufler et al., 2000).

To assume its newly discovered functions in macro-complex assembly, Hsp90 associates with one peculiar co-chaperone, the R2TP complex. This complex contains the two AAA+ ATPases/Helicases Rvb1 and Rvb2 (TIP49 and TIP48 in human, respectively), proteins Pih1 (Protein Interacting with Hsp90) and Tah1 [TPR containing protein associated with Hsp90] in yeast, or its Spaghetti counterpart in vertebrates (Boulon, 2008). Protein Pih1 is unstable. However, it forms a stable heterodimeric complex with Tah1, and this complex can down regulate the yeast Hsp90 ATPase activity cycle (Eckert et al., 2010). As a result, the dimeric clamp of Hsp90 is open, allowing the recognition and loading of the client proteins. This R2TP property is shared by members of the early co-chaperone family (Roe et al., 2004; Siligardi et al., 2002; Southworth and Agard, 2011). Protein Tah1 was proposed to contain two TPR motifs forming a single TPR domain. This TPR domain would explain how Tah1 tethers Hsp90 to the R2TP complex. Amino acid sequence comparison suggests that the Tah1 TPR domain shares similarities with the Hop TPR2B domain (Millson et al., 2008). A manuscript in press (ref) describes an NMR analysis of the complex formed between yeast Tah1 and an MEEVD peptide corresponding to the C-terminal region of Hsp90 (ref). However, due to the low number of intermolecular NOE restraints obtained by the authors, they had to model the peptide structure using a limited number of NMR data, and the validity of the proposed structure is questionable. Furthermore, the structure of free Tah1 was not studied.

For a better understanding of the mechanism of recognition of yeast Hsp90 by the R2TP complex, here, we use a large panel of NMR approaches to establish the high resolution structures of both free Tah1 and the complex formed by Tah1 and the ADTEMEEVD amino-terminal sequence of Hsp90. Our data disqualify the modeled interaction which is in press. It highlights how the Hsp90 terminal peptide is bound to Tah1 and explains the specificity of the interaction. We also study the Tah1- Pih1 interaction *in cellulose*. Altogether, our data bring consistent information on how Hsp90 recognizes the R2TP complex.

## EXPERIMENTAL PROCEDURES

**Biological Material used in this study.** The *Saccharomyces cerevisiae* strains Y190 and Y187 and plasmids pACTII::PIH1, PACTII::HSP82 (WT and mutants) and pACTII::TAH1, pGBKT7::TAH1 (WT, domains and mutants) and pGBKT7::PIH1 (WT and domains) were used for Y2H assays (supplemental experimental procedures). To express Tah1-His6 the cDNA encoding full length Tah1 was cloned into plasmid pKHS (Eckert et al., 2010). *Escherichia coli* BL21 (DE3) (Novagen®) cells complemented with the pRare2 plasmid were used as the host cells for Tah1-His6 production. The nonapeptide ADTEMEEVD was purchased from GenCust® (Luxembourg).

**Production and purification of the Tah1 recombinant protein.** Recombinant Tah1-His<sub>6</sub> protein was purified as described previously (Eckert et al., 2010). The bound proteins were eluted with an imidazole gradient and dialyzed in 20 mM Tris HCl (pH 7.5), 150 mM NaCl. A second binding on Ni-Sepharose High Performance was performed and the Tah1 part of the bound Tah1-His<sub>6</sub> protein fusion was recovered by PreScission proteolysis (5-10 U per mg of protein at 4 °C overnight). Residual impurities and the protease were removed by exclusion size chromatography using a Superdex 75 prep grade column (GE Healthcare®) equilibrated with the NMR buffer (10 mM NaH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub> (pH 7.2), 150 mM NaCl). The protein was stored in this buffer at -20°C at a 1 mM concentration.

**NMR measurements.** NMR spectra used to build the 3D structure of free Tah1p were recorded at 288K on a BRUKER DRX 600 US equipped with a triple resonance (<sup>1</sup>H, <sup>15</sup>N, <sup>13</sup>C) TCI Z 5mm cryoprobe using a <sup>15</sup>N-<sup>13</sup>C labeled Tah1p sample at a concentration of ~1 mM in the NMR buffer. Data were processed by Topspin (Bruker®) and analyzed with the a Sparky 3.113 software (Goddard and Kneller, 2002). Sequence-specific backbone, side chain and aromatic proton assignments were achieved using 2D and 3D classical experiments (supplemental experimental procedures). NMR measurements for assignment of the Tah1: peptide complex were performed in the same conditions as above using Bruker AVIII-500 MHz, AVIII-700 MHz, and Avance-900 MHz spectrometers. The

samples were prepared at a ~ 1 mM concentration using  $^{15}\text{N}$ - $^{13}\text{C}$  labeled Tah1 and unlabeled yeast Hsp90 peptide at a 1:1 molar ratio. We assigned the bound Tah1 backbone, side chains and aromatic protons using 2D and 3D experiments described in supplemental experimental procedures. We assigned protons of the bound Hsp90 peptide in the Tah1: peptide complex by using 2D F1-filtered, F2-filtered TOCSY and 2D F1-filtered, F2-filtered NOESY experiments in  $\text{H}_2\text{O}$  or 100%  $\text{D}_2\text{O}$ . The intermolecular NOEs were established from 3D  $^{15}\text{N}$  F1-filtered, F2-edited NOESY-HSQC, 3D  $^{13}\text{C}$  F1-filtered, F3-edited NOESY-HSQC experiments in  $\text{H}_2\text{O}$  or 100%  $\text{D}_2\text{O}$ , and 2D  $^{13}\text{C}$  F1-filtered, F2-edited NOESY performed on the  $^{15}\text{N}$ - $^{13}\text{C}$  Tah1: peptide complex.

**Structure calculation and refinement.** The automated peak picking, NOE assignment and structure calculation of free and bound Tah1 were performed using the Atnos/Candid software (Herrmann et al., 2002a, b). We used: i) 3D  $^{15}\text{N}$  and  $^{13}\text{C}$  NOESY-HSQC, and 2D homonuclear NOE spectra recorded in  $\text{D}_2\text{O}$  for structure analysis of free Tah1, ii) 3D  $^{15}\text{N}$  and  $^{13}\text{C}$  NOESY-HSQC experiments for structure determination of the bound protein, and iii) 2D  $^{13}\text{C}$  F1-filtered F2-filtered NOESY in  $\text{H}_2\text{O}$  or 100%  $\text{D}_2\text{O}$  for structure analysis of the bound peptide. Seven iterations were performed and 250 independent structures were calculated at each iteration step. We used the CYANA software to calculate the structures of the Tah1-Hsp90 peptide complex, calculation was based on a list of automatically assigned intra-molecular NOE distance constraints that we established for the bound Tah1 and bound peptide using the Atnos/Candid software and manually assigned intermolecular distance constraints for the complex (Guntert, 2004). The 50 structures with the lowest energy were refined using the AMBER 7.0 software with simulated annealing protocol (Case et al., 2005). We analyzed the 20 best final structures based on energy, NOE violations, and structural quality using the PROCHECK software (Laskowski et al., 1996).

**Yeast two-hybrid assays (Y2H).** Recombinant plasmids obtained after cloning the various protein coding sequences into plasmids pACTII (*LEU2*) and pGBKT7 (*TRP1*), were used to transform haploid cells (strains Y190 and Y187, respectively). After mating, the diploid cells were selected on media suitable for double

selection ( $\text{Leu}^-$ ,  $\text{Trp}^-$ ), and interaction tests were performed after plating the cells on media suitable for triple selection ( $\text{Leu}^-$ ,  $\text{Trp}^-$ ,  $\text{His}^-$ ). Various concentrations of 3-amino-1,2,4-triazol (3AT) were used to evaluate the strength of the interactions. Growth was assessed after three days of incubation at 30°C.

## RESULTS

### Solution structure of free Tah1 reveals a folded 5-helix-extended TPR domain and a flexible C-terminus

To study the structure of free Tah1 by NMR, we produced a  $^{15}\text{N}$   $^{13}\text{C}$  labeled Tah1 protein carrying a His6 Tag sequence which is eliminated during purification. Backbone atoms were assigned using conventional methods. Figure 1 shows the superposition of the energetically lowest conformers that best satisfied the experimental restraints (Fig. 1A-1B, Table 1). It reveals the presence of a well define N-terminal TPR domain (N-domain) flanked by a flexible short C-terminal domain (C-domain). As expected from previous predictions (for review see (Blatch and Lassle, 1999; D'Andrea and Regan, 2003), the N-terminal residues 1-67 are folded into two tandem TPR motifs (TPR1 and TPR2) composed of a pair of antiparallel  $\alpha$ -helices: residues 4-17 and 20-34, 38-51 and 54-67, form the first TPR1 and second TPR2 motifs respectively) (Fig. 1). Sequence analysis of figure 1C confirmed that Tah1 is part of the [TPR domain containing protein] family but its homology is restrained to the first 70 amino acids. Among all sequences of eukaryal co-chaperone TPR domains, the strictly conserved residues Ala23, Tyr27, Pro35 and the less conserved residues Gly11, Leu14, and Leu30 of Tah1 are the footprint of TPR motifs (Blatch and Lassle, 1999; Lamb et al., 1995; Prasad et al., 2010). In contrast to the first TPR motif of Tah1, the second one is more degenerate (Fig. 1C). TPR2 contains a unique strictly conserved residue, Asn43, another carboxylate clamp residue. The inter-helical angles calculated by MOLMOL software for the 4 helices of the TPR1 and TPR2 motifs are comprised between 155 and 168° {Koradi, 1996, MOLMOL: a program for display and analysis of macromolecular structures}. They are highly similar to those calculated for other Hsp co-chaperones. Interestingly, TPR2, and especially its 2B helix, is tightly associated with a fifth helix, the capping helix or C helix (residues 74 to 90) through multiple hydrophobic contacts (Fig. 1). This helix

does not display the hallmarks of TPR N-terminal helix (Fig. 1C), suggesting that it does not belong to a 3<sup>rd</sup> TPR motif whose second helix would not be folded in ligand free Tah1 structure. Moreover, this C helix forms an unusual angle with the 2B helix (145°). However, like helix 1 of TPR1, the presence of Lys79 and Arg83 conserved residues in helix C suggests that these two residues might form a carboxylate clamp (Fig. 1C). Moreover, some of the conformers of the family of structures present a 3<sub>10</sub> helix encompassing residues 71 to 73 (Fig. 1A). The folding of the Tah1 5 helices for the family of 20 conformers is well defined with a RMSD for the backbone of 0.44 ± 0.08 Å (Table 1). This stable structure is driven by the amphipathic properties of the helices. As usual, the two helices of each TPR motif interact through their hydrophobic sides, while their opposite hydrophilic sides are exposed to solvent {Lamb, 1995, Tetratricopeptide repeat interactions: to TPR or not to TPR?}{D'Andrea, 2003, TPR proteins: the versatile helix}{Blatch, 1999, The tetratricopeptide repeat: a structural motif mediating protein-protein interactions}. The additional C helix is also amphipathic and participates to the formation of a compact helix-turn-helix arrangement characteristic of TPR domains, with a concave inner surface and a right handed super helix twist of the entire structure. Finally, C helix can be considered as the capping helix of a very short TPR domain to form an extended TPR domain.

The 18 amino acid C-domain (positions 93 to 111) located immediately downstream from C helix displays the characteristic features of an unstructured region, with random coil chemical shifts and reduced <sup>1</sup>H-<sup>15</sup>N heteronuclear NOE values indicative of elevated motion relative to the TPR domain (supplemental Figure S1A). This strong difference between the N- and C-domains reinforces the idea that C helix is a capping helix. Another difference of the N- and C-domains, concerns their surface charges which are predominantly negative for the C-domain and positive for the N-domain (supplemental Figure S1B). Altogether, these strong differences suggested very distinctive protein binding properties of the two domains and different targets.

### **The two domains of Tah1 interact with different protein partners**

In the R2TP-Hsp90 complex, Tah1 is bound to both Pih1 and Hsp90 (Zhao et al., 2005). The N-

terminal TPR domain was expected to bind Hsp90. Interestingly, a recombinant protein which contained the two N-terminal TPR motifs without C helix failed to interact with Hsp90 {Zhao, 2008, Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation}. Having clearly delineated the extended TPR (TPR domain+helix C) and C-domains in free Tah1, we used yeast two-hybrid assays (Y2H) (Fig. 2A) to test their capacities to bind to Hsp90 and Pih1. To this end, the nucleotide sequences encoding, i) full-length Tah1, ii) the C-domain, and iii) the extended TPR domain were individually cloned into the Y2H plasmid pGBKT7 and the nucleotide sequences encoding Pih1 and Hsp90 were inserted into the partner Y2H plasmid pACTII. Various concentrations of 3-amino-1,2,4-triazol (3AT), were used to evaluate the strength of the interactions in yeast cells (Fig. 2A). In these assays, the full length Tah1 and extended TPR domain were found to establish very strong (stable at 10 mM 3AT) and almost identical interactions with Hsp90, strongly indicating that the extended TPR domain represents the Hsp90 recognition domain of Tah1. This domain did not interact with Pih1 in the assays (Fig. 2A), while the C-domain interacted with a strong affinity with this protein, as illustrated by detection of an interaction which is stable at a 10 mM 3AT concentration (Fig. 1A). Therefore, the very short unstructured C-domain of Tah1, without the presence of helix C, is sufficient to bind Pih1 *in cellulo* and the interaction is as strong as to the full length Tah1.

### **The 253-344 C-terminal fragment of Pih1 interacts with the flexible C-domain of Tah1**

Recent *in vitro* mass spectrometry analysis (Eckert et al., 2010) showed that the Pih1 fragment extending from position 199 to 344 binds to Tah1 *in vitro*. To further delineate the Tah1 binding site in Pih1, we again used Y2H assays. Alignment of the yeast and human Pih1 amino acid sequences, revealed an inserted 42 amino acid sequence in yeast Tah1. It separates an N-terminal 211 amino acid sequence showing 38% similarity in the two proteins and a C-terminal 91 amino acid sequence displaying 24% of similarity in the two proteins (Fig. 2B). The 1-252 Pih1 N-domain including the yeast specific sequence did not interact with Tah1 in Y2H assays, while an efficient interaction was detected with the Pih1 253-344 C-domain (Fig. 2B). Therefore, our data indicate that the 91 amino

acid C-domain of Pih1 is sufficient for *in vivo* interaction with both full length Tah1 and the flexible C-domain of Tah1.

### **The NMR structure of bound Tah1 reveals a strong interaction of the Hsp90 C-terminal sequence with the Tah1 extended TPR domain**

As the 9 amino acid C-terminal segment of yeast Hsp90 (ADTEMEEVD) had been shown to be sufficient for a specific interaction with Tah1 (Millson et al., 2008), we used this segment to study the Tah1-Hsp90 interaction at the atomic level. In order to define the best conditions for study of the complex, we monitored by NMR HSQC experiments the titration of <sup>15</sup>N labeled Tah1 protein by peptide until the 1:2 stoichiometry (Supplemental Figure S2A). Above a 1:1 mole ratio we observed a broadening of some peaks and an absence of chemical shift variations. In the studied complex formed at a 1:1 molar ratio, only few Tah1 residues displayed high significant chemical shifts (Supplemental Figure S2B). They corresponded to the conserved residues of the carboxylate clamp (Asn12, Asn43, and Arg83) and residues Tyr27 and Lys50. None residues in the C-terminal domain showed a significant change in chemical shift consistent with the fact that this unstructured region do not participate to the interaction. The fact that the HSQC spectra of Tah1 in the presence and absence of peptide are globally similar strongly suggests that the peptide does not influence the structure of Tah1. Then, we established the 3D structure of this complex by using <sup>15</sup>N-<sup>13</sup>C labeled Tah1 protein and an unlabeled C-ter Hsp90 peptide. Resonance assignments for the bound peptide were obtained by specific filtered-filtered NMR experiments. To determine the solution structure of the Tah1-Hsp90 peptide complex, we used a total of 2,771 NOE-derived distance restraints (Table 1). The interaction between Tah1 and the pentapeptide is characterized by a lot of intermolecular NOEs (56 restraints) determined using filtered-edited NMR experiments and the structure calculations of the peptide relied on 56 upper distance limit values derived from the <sup>1</sup>H-<sup>1</sup>H NOE intensities. According to the 20 lowest energy structures established with the AMBER software, in its bound form, Tah1 adopts the same fold and organization as in its free form (Fig. 3 and Supplemental Figure S2C). In particular, it maintains its peculiarity to contain 5 helices and the relative orientation of these helices as well as the conformation of their linking segments are

almost identical in the free and bound states. The calculated RMS for residues 4-70 between the free and bound TPR structures was 0.943 Å (Supplemental Figure S2C). The C-terminal domain remains unfolded in the bound form and comparative analysis of Tah1 flexibility in the free and bound forms by <sup>1</sup>H-<sup>15</sup>N heteronuclear NOE measurements revealed no significant difference (Supplemental Figure S1A).

According to the 20 energetically lowest calculated conformers (Fig. 3), the N-terminus of the bound peptide (A<sup>701</sup>DT<sup>703</sup>) adopts a rather flexible conformation, while the E<sup>704</sup>MEEVD<sup>709</sup> motif has a constrained conformation (Fig. 3A). The E<sup>707</sup>VD<sup>709</sup> segment adopts an extended conformation similar to that observed for Hsp peptides bound to HOP-, CHIP-, and Tom71-peptides (Li et al., 2009; Scheufler et al., 2000; Zhang et al., 2005). As observed in CHIP-Hsp90 complex (Zhang et al., 2005), the main chain of the M<sup>705</sup>E<sup>706</sup>E<sup>707</sup> triplet performs a turn by formation of H-bonds between the two carbonyl groups of M<sup>705</sup> and E<sup>707</sup> and the side chain of Arg83 which brings the side chain of M<sup>705</sup> close to C helix (Fig. 4A). Simultaneously, the backbone atoms of E<sup>704</sup>M<sup>705</sup>E<sup>706</sup> make a <sub>3</sub><sub>10</sub> helix enabling the N-terminal extremity of the peptide to exit the protein channel in a perpendicular orientation (Fig. 3C-D). This peculiar structure of the EMEEVD peptide bound to Tah1 is required for its accommodate in the groove of the Tah1 TPR domain (Fig. 3B and C).

The highly asymmetric charge distribution between the peptide and Tah1 (Supplemental Figure S1C-D) results in the formation of several salt bridges. The interaction is reinforced by formation of several intermolecular H-bonds and hydrophobic interactions explaining the strength of the Tah1-Hsp90 and Tah1TPR-Hsp90 interactions observed *in vivo* (Fig. 4B). Each of the residues in the EMEEVD peptide contributes to the interaction. The terminal D<sup>709</sup> amino acid plays a key role as expected. Its side chain forms H-bonds with Asn12 and Asn43 and salt bridges with Lys8 residues. The natural carboxylate terminal group of D709 establishes salt bridges with Lys79 and Arg83 residues. Its amide establishes an H-bond with the Asn43 carbonyl side chain (Fig. 4). A similar network of interactions was already observed in complexes formed by the HOP, PP5 and Tom 71 TPRs and C-terminal peptides of chaperon proteins (Li et al., 2009; Scheufler et al., 2000; Zhang et al., 2005). The side chain of residue V<sup>708</sup> is located in

a hydrophobic pocket formed by the side chains of Lys8 and Lys16 and the aromatic cycle of Phe15. Its carbonyl group establishes H-bond with the amino group of Lys79 (Fig. 4). Carboxylate groups of the side chains of residues E<sup>706</sup> and E<sup>707</sup> establish electrostatic interactions with the NH<sub>3</sub><sup>+</sup> groups of Lys50 and Lys16 respectively. The divalent sulfur group of M<sup>705</sup> specifically interacts with the  $\pi$  electron of the aromatic carbon cycle (sp<sup>2</sup>) of Tyr82, resulting in localization of its methyl group into a hydrophobic pocket formed by the methylene group of Tyr82 and the carbon chain of Lys79 (Fig. 4). Finally, the carboxylate side chain and the main chain carbonyl of E<sup>704</sup> form a salt bridge and an H-bond with the amino group of Lys79. Lastly, we do not observe interaction between A<sup>701</sup>, D<sup>702</sup>, and T<sup>703</sup> residues of Hsp90 and Tah1. Therefore, both the canonical TPR motif and C helix involved in interaction with the EMEEVD Hsp90 peptide. The two contacting patches in helices 1A and 2A of canonical part of the TPR contribute to the formation of 1 hydrophobic interaction, 4 salt bridges and 3 H-bonds, while the contacting large patch in C helix contributes to the formation of 3 salt bridges, 5 H-bonds, and 1 hydrophobic interaction, strongly suggesting that it plays a crucial role in the stability of the interaction.

### **Residues in the Tah1 C helix are essential for the Tah1-Hsp90 interaction *in vivo***

To confirm the functional importance of C helix in the interaction of Tah1 with Hsp90, we generated individual substitutions by alanine on the Lys79, Tyr82 and Arg83 residues in Tah1 and their partner amino acids E<sup>706</sup>, E<sup>707</sup>, V<sup>708</sup> and D<sup>709</sup> in Hsp90. Then, we tested the effects on the Hsp90-Tah1 interaction by Y2H assays (Fig 5). As a control, Arg77, Glu85 in the Tah1 C helix, and D<sup>702</sup>, T<sup>703</sup> in Hsp90 which are not predicted to be involved in the protein-peptide interaction were also substituted for Ala. We controlled the expression of mutated Tah1 and Hsp90 proteins by testing their interactions with Pih1, using Y2H assays. As expected the Lys79Ala, Tyr82Ala, and Arg83Ala substitutions in C helix had no effect on the Tah1p-Pih1 interaction (Fig. 5A), however they abolished the interaction with Hsp90. By contrast, control substitutions of Arg77 or Glu85 had no effect on the interaction. Concerning Hsp90, the D<sup>709</sup>A substitution completely abolished its interaction with Tah1, confirming the crucial role of D<sup>709</sup>. The E<sup>706</sup>A and V<sup>708</sup>A substitutions had negative although less

pronounced effects on the interaction. The E<sup>707</sup>A mutation had a more limited negative effect, but its implication is evidenced by a significant decrease of the 3-AT resistance. As expected, mutations of amino acids D<sup>702</sup> and T<sup>703</sup> had no marked effect. Therefore, we confirmed the functional importance of the Tah1 C helix *in vivo* interaction of Tah1-Hsp90 interaction.

## **DISCUSSION**

In this article, we established the solution structures of Tah1 free and bound to the C-terminal Hsp90 peptide. We produced to do this, a soluble <sup>15</sup>N/<sup>13</sup>C labeled Tah1 protein with concentration up to 1 mM. This sample is stable for long time (exceed one month) in the NMR tube. The instability of apo-TPR domains previously observed for PP5 and Tah1 is certainly an artifact (Cliff *et al.*, 2005) {Jimenez, 2011, The structure of the minimal tetratricopeptide repeat domain protein Tah1 reveals the mechanism of its interaction with Pih1 and Hsp90}. Initially, Tah1 was described to be a TPR motif protein. On the strength of sequence alignment Zhao *et al.* have defined the TPR domain from the residue 1 to residue 75 (Zhao *et al.*, 2008). In addition, the sequence comparison performed by Millson *et al.* in 2008 failed to define the exact boundaries of the third TPR motif and especially the position of the two carboxylate clamp residues (Millson *et al.*, 2008). On the basis of the Tah1 NMR structure, our laboratory and Jiménez *et al.* have simultaneously shown that Tah1 protein is composed by two distinct domains, an extended TPR domain with a fifth C helix and an unfolded domain. The folding of the unique extended TPR domain (1-92) of Tah1 was a surprise for its small size because it encompasses only two classical TPR motifs followed by an additional C helix. Indeed, many TPR domain structures are already describing in the literature, in particular PP5 TPR domain (Das *et al.* 1998; Cliff *et al.*, 2006), the two TPR domains of HOP, TPR1 and TPR2A (Scheufler *et al.*, 2000), CHIP (Zhang *et al.*, 2005), and Tom71 (Li *et al.*, 2009) proteins, and all of them includes at least three TPR motifs by TPR domain (D'Andrea and Regan, 2003). Surprisingly, the minimum domain represented by the extended TPR domain of Tah1 is stable and able to interact with the nonapeptide of yeast C-terminal Hsp90. An additional fifth C helix (74-90), not conserved in sequence, plays both: (i) the role of the capping helix, like in other TPR

containing proteins such as PP5, necessary to solubilize hydrophobic surfaces from the last TPR motif (Das et al., 1998) and (ii) the role of the helix 3A of the third TPR motif (TPR3) by the presence of the last two residues of the dicarboxylate clamp. Now, the results explain the loss of ability of the domain 1-75 to interact with yeast Hsp90 observed by Zhao *et al.*

The superposition of Tah1 TPR domains between the free and bound structures provide a backbone RMS of 0.943Å indicating that the global folding remains unchanged upon the fixing of peptide except for the C helix (Supplemental Figure S2B). This binding induces only a slight shift of C helix toward the helix 2A shortening the channel of Tah1. This movement of the C helix is also observed in comparing the bound TPR domains of Tah1 with other TPR domains that bind Hsp70/90 like TPR2A, Tom71, and CHIP (Fig. 6). Overall, as shown by the RMS values (average of 1.1Å), the four helices of bound Tah1 TPR domain and the other TPR structures are superimposable each other except for the structure obtained by Jimenez *et al.* (Fig. 6). In terms of length of helices, only helices 1A and C vary significantly. We observe two additional turns at the N-terminus of the helix 1A from Tom71 and PP5 and one turn in less at the N-terminus of the helix 3A from Tom71, CHIP and PP5 compared to the helix 1A and C respectively from Tah1. Unfortunately, we cannot superimpose properly our structure of bound Tah1p with the one obtained by Jimenez *et al.* as demonstrated the high value for RMS of 2.5Å. Finally, the results of heteronuclear NOE for each residues of Tah1 in free form overlap well with those of the linked form, confirming that the dynamic properties of the protein Tah1 is the same with or without the peptide (Supplemental Figure S1A).

The residues Lys8, Asn12, Asn43, Lys79 and Arg83 of Tah1 form a so-called two-carboxylate clamp that provides the binding anchors for the tightly interaction with both acidic functions to EMEEVD motif of yeast Hsp90. They represent the primary recognition motif for Hsp70/90 binding on TPR domains proteins. These five strictly conserved residues represent more than half of the strictly identical residues, which suggest their importance for the function in the co-chaperone Hsp70/90 family. All these residues were mutated in Tah1 by Jimenez *et al.* and were analyzed *in vitro* using isothermal titration calorimetry (ITC) experiments. The K8A, N12A and N43A, K79A,

R83A mutants respectively reduce and abolish the binding of MEEVD peptide to Tah1 (Jimenez *et al.*, 2011). Our *in vivo* Y2H experiments on mutated Tah1 versus yeast Hsp90 showed that the two last basically residues (Lys79 and Arg83) of the carboxylate clamp in the C helix are both essentials for the interaction of the Hsp90 with the groove of TPR domain. This confirms the lost of interaction between human Hsp90 and PP5 Lys97/Ala or FKB52 Lys354/Ala mutants (corresponding to Lys79 in Tah1) (Ramsey et al., 2000; Russell et al., 1999). Russell *et al.* obtained the same result with the mutation of Lys32 and Arg101 in PP5 (corresponding to Lys8 and Arg83 respectively in Tah1) (Russell et al., 1999). In addition, the single alanine mutation of clamp residues Lys229, Asn264, Lys301 and Arg305 in TPR2A of Hop protein (Lys8, Asn43, Lys79 and Arg83 in Tah1 respectively) was analyzed by SPR-based binding assay (Brinker et al., 2002). Each point mutation of these residues severely impaired the binding to human Hsp90. In contrast, the mutation Arg74/Ala in PP5 (Lys50 in Tah1 which is not part of the carboxylate clamp) decrease, but not abrogate the interaction (Russell et al., 1999). It is interesting to note that our simple mutation of residue E<sup>706</sup>A in yeast Hsp90 which interacts with Lys50 of Tah1, decrease but not disrupt the *in vivo* complex formation (Fig 6B). Finally, in this complex, the five helices represent the minimal domain for interaction with the yeast chaperone Hsp90 compared to that observed in the classical TPR domain composed at least of three TPR motifs (D'Andrea and Regan, 2003).

Ramsey *et al.* has also shown that the simultaneous mutations of the three acid residues of the C-terminal 12-kDa fragment of human Hsp90 (EEVD) abolished the *in vitro* interaction with PP5, FKBP52 and Hop proteins (Ramsey et al., 2000). In order to refine the results, we performed Y2H experiments on mutated yeast Hsp90 versus Tah1 (Fig. 6B) to observe the effect of the each mutation on these acid residues. We observed gradual effects on the binding to EEVD core motif in the context of protein chaperone. The most pronounced effect occurs with D<sup>709</sup> substituted in alanine (D<sup>709</sup>A), the intermediate effect with the two single mutations E<sup>706</sup>A and V<sup>708</sup>A and the weakness one with E<sup>707</sup>A. Interestingly, Scheufler *et al.* indicated that the residue E<sup>707</sup> is the lower conserved of the EEVD motif among eukaryotic cytosolic Hsp90 proteins (Scheufler *et al.*, 2000). The comparison with 3D structures of known TPR/Hsp complexes



revealed that each side chain of the two conserved glutamic residues from the EEVD motif are more implicated in complex formation in the case of Tah1.

TPR1 and TPR2A domains of Hop protein specifically recognize the human C-terminal five amino acids of Hsp70 (IEEVD) and Hsp90 (MEEVD) respectively, but residues upstream the position -4 (corresponding to M<sup>705</sup> in yeast Hsp90) did not affect the binding (Scheufler et al., 2000). This constitutes a distinctive feature compared to results obtained in yeast complex by Millson *et al.* They showed that the binding affinity of Tah1 for the C-terminal MEEVD peptide increases by a factor 30 with the addition of T<sup>703</sup>E<sup>704</sup> amino acids (Millson et al., 2008). Our structural data indicated that the mode of fixation and specificity between TPR domains and chaperones is different between Hop TPR domains and Tah1/CHIP proteins (Fig. 6B-F). The structure obtain in this work indicated that the residue E<sup>704</sup> of Hsp90 interacts with Lys79 of Tah1 by salt bridge (Fig. 4). This acidic residue is conserved (over 70%) in the species of yeast. This additional interaction may explain the affinity gain obtained by Millson *et al.* A second difference concerns the hydrophobic residues I<sup>637</sup>, M<sup>728</sup> and, M<sup>705</sup> in Hsp70, Hsp90 and, yeast Hsp90 respectively that are not only necessary for the high affinity of TPR interactions but are also the primary determinants of binding specificity (Brinker et al., 2002; Scheufler et al., 2000). These residues I<sup>637</sup> and M<sup>728</sup> increase by a factor of 2 and 8 the affinity of TPR1 and TPR2 of Hop protein respectively for the pentapeptides IEEVD and MEEVD compared to EEVD motif (Scheufler et al., 2000). In our 3D structure complex, the main chain of the E<sup>704</sup>MEE<sup>707</sup> residues forms a 3<sub>10</sub> helix following by a turn, bringing the M<sup>705</sup> residue at the position -4 of peptide towards the C helix of Tah1 instead of helix 1A for TPR2A (Fig. 6B-F). It enables a strong S-CH3/ $\pi$  interaction with residue Tyr82 of Tah1 in place of a hydrophobic interaction with Tyr236 in the opposite side for the Hop-peptide complex (Scheufler et al., 2000). It is the unique interaction of this residue Tyr82 with the C-terminal tail of the chaperone. Less significantly, the M<sup>705</sup> engages a van der Waals contact with the carbon side chain of Lys79 of Tah1p which performs several other interactions with the EEVD motif (Fig. 4). An another difference is the presence of a large hydrophobic pocket in

CHIP or HOP proteins that accommodates either the methionine of Hsp90 or isoleucine residues of Hsp70 while for the Tah1, this hydrophobic pocket is reduced to residues Lys79 and Tyr82. This is in agreement with a report showing that the residue M at the position -4 of the peptide is more important for interaction with a designed TPR domain (CTPR390) than TPR2A (Cortajarena et al., 2010). Interestingly, at the position of amino acid Tyr82 in Tah1, PP5 and CHIP proteins have also an aromatic residue (Tyr100 and Phe99 respectively). By contrast, Hop and Tom71 proteins contain respectively an alanine and a leucine. The specificity for methionine residue at the -4 position of the peptide end is explained by a S-CH3 interaction with  $\pi$  electron of aromatic Tyr82 which is impossible with an isoleucine or valine. This data could also explain the results of Millson *et al.* which observed that the M<sup>705</sup>V mutation in yeast Hsp90 decreases Tah1p binding by a factor of 20 (Millson et al., 2008). Our results of Y2H assays confirm that the Tyr82 is essential for the *in vivo* interaction between Tah1 protein and yeast Hsp90 as well as the two last Lys79 and Arg83 carboxylate clamp residues (Fig. 5A).

Our Y2H interactions are tested in a homologous genetic context and can thus be mediated by a third endogenous protein. We favor the hypothesis that the Y2H interaction between Pih1 and yeast Hsp90 rely on a direct interaction arising from the N-terminal part of Pih1 and an indirect interaction mediated by the endogenous Tah1 interacting with the C-terminal part of Pih1. This interpretation completes published data (Eckert et al., 2010; Zhao et al., 2005) and strongly suggests that Tah1 mediates and reinforces the interaction between Pih1 and yeast Hsp90. If all TPR domains of co-chaperone TPR proteins bound to Hsp70/90 *in vitro*, the biological difference between these co-chaperones would reside outside of their TPR domains. The majority of TPR containing protein, which interacts with Hsp90 by its TPR domain, encompass a second domain which contains a specific activity: e.g. PP5 with a phosphatase domain, FKBP52 a peptidylprolyl isomerase domain, CHIP an ubiquitin ligase domain. Tah1 protein is different in this sense because it has not all features of Hsp90 co-chaperones: it contains an Hsp90 interaction domain, but it lacks an Hsp90 regulatory domain (Eckert *et al.*, 2010). The second domain of Tah1 (93-111) represents the fixation site of Pih1

protein which was shown to negatively modulate the Hsp90 activity. Indeed, the Tah1-Pih1 complex has shown to reduce the ATPase activity of yeast Hsp90 (Eckert *et al.*, 2010) while Tah1 alone weakly stimulates it (Millson *et al.*, 2008). Tah1 is thus an interacting mediator for yeast Hsp90 and its regulatory protein, Pih1. Finally, the Pih1-Tah1 heterodimer must be considered like a functional entity enables to recognize specifically Hsp90 and to negatively modulate its activity. The specificity of interaction between C-terminal domain of Hsp90 and TPR domain of Tah1 studied in this work would be supplemented by other direct contacts between Hsp90 and Pih1.

## CONCLUSION

This studies reported here provide the first high resolution structures of Tah1 free and in interaction with the Hsp90 C-terminal tail (ADTEMEEVD peptide). The original features of the Tah1 structure in complex with chaperone peptide are the existence of a TPR domain with only two TPR motifs followed by a C helix. This additional helix is not only a solvating/stabilizing helix but it is also directly involved in C-terminal Hsp90 recognition by the two Lys79 and Arg83 carboxylate clamp residues. Moreover, this helix also includes residue Tyr82 that performs a  $\pi$  electron of aromatic interaction with the S-CH3 of the M<sup>705</sup> of the Hsp90 peptide. This interaction explains the preference of Tah1 for a methionine versus an isoleucine or valine residues. This is the first time that this Tah1 residue is described as part of the interaction interface. Similarly, in Hsp90, the E<sup>704</sup> residue has never been described until now in this interaction. The unstructured C-terminal domain of Tah1 is sufficient to bind the Pih1 253-344 C-domain *in vivo* and this interaction is as strong as to full length Tah1 and Pih1. Finally, the role of Tah1 would be to recruit Pih1 on Hsp90 to negatively modulate its ATPase activity.

## ACCESSION NUMBERS

The coordinates for these free and bound Tah1 structures were deposited in April, 2012 with the PDB codes of XXXX and XXXX respectively. The resonance assignments have been deposited in the BMRB with accession codes XXXXX.

## ACKNOWLEDGMENTS

This work was supported by the Agence Nationale de la Recherche (ANR-06-BLAN-0208), the Centre National de la Recherche

Scientifique, and the University of Lorraine. RB and BR were the recipients of a French Ministère de l'Enseignement Supérieur et de la Recherche fellowships.

## REFERENCES

- Ali, M.M., Roe, S.M., Vaughan, C.K., Meyer, P., Panaretou, B., Piper, P.W., Prodromou, C., and Pearl, L.H. (2006). Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex. *Nature* 440, 1013-1017.
- Blatch, G.L., and Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* 21, 932-939.
- Borkovich, K.A., Farrelly, F.W., Finkelstein, D.B., Taulien, J., and Lindquist, S. (1989). hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. *Mol. Cell. Biol.* 9, 3919-3930.
- Boulon, S., Marmier-Gourrier, N., Pradet-Balade, B., Wurth, L., Verheggen, C., Jady, B.E., Rothe, B., Pescia, C., Robert, M.C., Kiss, T., *et al.* (2008). The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J. Cell. Biol.* 180, 579-595.
- Boulon, S., Pradet-Balade, B., Verheggen, C., Molle, D., Boireau, S., Georgieva, M., Azzag, K., Robert, M.C., Ahmad, Y., Neel, H., *et al.* (2010). HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol. Cell* 39, 912-924.
- Brinker, A., Scheufler, C., Von Der Mulbe, F., Fleckenstein, B., Herrmann, C., Jung, G., Moarefi, I., and Hartl, F.U. (2002). Ligand discrimination by TPR domains. Relevance and selectivity of EEVD-recognition in Hsp70 x Hop x Hsp90 complexes. *J. Biol. Chem.* 277, 19265-19275.
- Caplan, A.J., Mandal, A.K., and Theodoraki, M.A. (2007). Molecular chaperones and protein kinase quality control. *Trends Cell Biol* 17, 87-92.
- Case, D.A., Cheatham, T.E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K.M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R.J. (2005). The Amber biomolecular simulation programs. *J Comput Chem* 26, 1668-1688.
- Cliff, M.J., Williams, M.A., Brooke-Smith, J., Barford, D., and Ladbury, J.E. (2005). Molecular recognition via coupled folding and binding in a TPR domain. *J Mol Biol* 346, 717-732.
- Connell, P., Ballinger, C.A., Jiang, J., Wu, Y., Thompson, L.J., Hohfeld, J., and Patterson, C. (2001). The co-chaperone CHIP regulates protein triage decisions mediated by heat-shock proteins. *Nat Cell Biol* 3, 93-96.
- Cortajarena, A.L., Wang, J., and Regan, L. (2010). Crystal structure of a designed tetratricopeptide repeat module in complex with its peptide ligand. *FEBS J* 277, 1058-1066.

- D'Andrea, L.D., and Regan, L. (2003). TPR proteins: the versatile helix. *Trends Biochem. Sci.* *28*, 655-662.
- Das, A.K., Cohen, P.W., and Barford, D. (1998). The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *EMBO J.* *17*, 1192-1199.
- Echtenkamp, F.J., and Freeman, B.C. (2011). Expanding the cellular molecular chaperone network through the ubiquitous cochaperones. *Biochim. Biophys. Acta.*
- Eckert, K., Saliou, J.M., Monlezun, L., Vigouroux, A., Atmane, N., Caillat, C., Quevillon-Cheruel, S., Madiou, K., Nicaise, M., Lazereg, S., *et al.* (2010). The Pih1-Tah1 cochaperone complex inhibits Hsp90 molecular chaperone ATPase activity. *J. Biol. Chem.* *285*, 31304-31312.
- Goddard, T.D., and Kneller, D.G. (2002). SPARKY 3. S.F. University of California, ed.
- Gouet, P., Robert, X., and Courcelle, E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* *31*, 3320-3323.
- Grammatikakis, N., Lin, J.H., Grammatikakis, A., Tsihchlis, P.N., and Cochran, B.H. (1999). p50(cdc37) acting in concert with Hsp90 is required for Raf-1 function. *Mol. Cell. Biol.* *19*, 1661-1672.
- Guntert, P. (2004). Automated NMR structure calculation with CYANA. *Methods Mol Biol* *278*, 353-378.
- Harris, S.F., Shiau, A.K., and Agard, D.A. (2004). The crystal structure of the carboxy-terminal dimerization domain of htpG, the Escherichia coli Hsp90, reveals a potential substrate binding site. *Struct.* *12*, 1087-1097.
- Herrmann, T., Guntert, P., and Wuthrich, K. (2002a). Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* *319*, 209-227.
- Herrmann, T., Guntert, P., and Wuthrich, K. (2002b). Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J Biomol NMR* *24*, 171-189.
- Horejsi, Z., Takai, H., Adelman, C.A., Collis, S.J., Flynn, H., Maslen, S., Skehel, J.M., de Lange, T., and Boulton, S.J. (2010). CK2 phospho-dependent binding of R2TP complex to TEL2 is essential for mTOR and SMG1 stability. *Mol. Cell* *39*, 839-850.
- Lamb, J.R., Tugendreich, S., and Hieter, P. (1995). Tetratricopeptide repeat interactions: to TPR or not to TPR? *Trends Biochem. Sci.* *20*, 257-259.
- Laskowski, R.A., Rullmann, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* *8*, 477-486.
- Li, J., Qian, X., Hu, J., and Sha, B. (2009). Molecular chaperone Hsp70/Hsp90 prepares the mitochondrial outer membrane translocon receptor Tom71 for preprotein loading. *J. Biol. Chem.* *284*, 23852-23859.
- Maloney, A., and Workman, P. (2002). HSP90 as a new therapeutic target for cancer therapy: the story unfolds. *Expert Opin Biol Ther* *2*, 3-24.
- Meyer, P., Prodromou, C., Hu, B., Vaughan, C., Roe, S.M., Panaretou, B., Piper, P.W., and Pearl, L.H. (2003). Structural and functional analysis of the middle segment of hsp90: implications for ATP hydrolysis and client protein and cochaperone interactions. *Mol. Cell* *11*, 647-658.
- Millson, S.H., Vaughan, C.K., Zhai, C., Ali, M.M., Panaretou, B., Piper, P.W., Pearl, L.H., and Prodromou, C. (2008). Chaperone ligand-discrimination by the TPR-domain protein Tah1. *Biochem J* *413*, 261-268.
- Minami, Y., Kimura, Y., Kawasaki, H., Suzuki, K., and Yahara, I. (1994). The carboxy-terminal region of mammalian HSP90 is required for its dimerization and function in vivo. *Mol. Cell. Biol.* *14*, 1459-1464.
- Panaretou, B., Siligardi, G., Meyer, P., Maloney, A., Sullivan, J.K., Singh, S., Millson, S.H., Clarke, P.A., Naaby-Hansen, S., Stein, R., *et al.* (2002). Activation of the ATPase activity of hsp90 by the stress-regulated cochaperone aha1. *Mol. Cell* *10*, 1307-1318.
- Panaretou, B., Sinclair, K., Prodromou, C., Johal, J., Pearl, L., and Piper, P.W. (1999). The Hsp90 of *Candida albicans* can confer Hsp90 functions in *Saccharomyces cerevisiae*: a potential model for the processes that generate immunogenic fragments of this molecular chaperone in *C. albicans* infections. *Microbiology* *145 ( Pt 12)*, 3455-3463.
- Pearl, L.H., and Prodromou, C. (2001). Structure, function, and mechanism of the Hsp90 molecular chaperone. *Adv Protein Chem* *59*, 157-186.
- Pearl, L.H., and Prodromou, C. (2006). Structure and mechanism of the Hsp90 molecular chaperone machinery. *Annu. Rev. Biochem.* *75*, 271-294.
- Prasad, B.D., Goel, S., and Krishna, P. (2010). In silico identification of carboxylate clamp type tetratricopeptide repeat proteins in Arabidopsis and rice as putative co-chaperones of Hsp90/Hsp70. *PLoS One* *5*, e12761.
- Prodromou, C., Panaretou, B., Chohan, S., Siligardi, G., O'Brien, R., Ladbury, J.E., Roe, S.M., Piper, P.W., and Pearl, L.H. (2000). The ATPase cycle of Hsp90 drives a molecular 'clamp' via transient dimerization of the N-terminal domains. *EMBO J.* *19*, 4383-4392.
- Prodromou, C., Roe, S.M., O'Brien, R., Ladbury, J.E., Piper, P.W., and Pearl, L.H. (1997). Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. *Cell* *90*, 65-75.
- Prodromou, C., Siligardi, G., O'Brien, R., Woolfson, D.N., Regan, L., Panaretou, B., Ladbury, J.E., Piper, P.W., and Pearl, L.H. (1999). Regulation of Hsp90 ATPase activity by tetratricopeptide repeat (TPR)-domain co-chaperones. *EMBO J.* *18*, 754-762.
- Ramsey, A.J., Russell, L.C., Whitt, S.R., and Chinkers, M. (2000). Overlapping sites of tetratricopeptide repeat protein binding and chaperone activity in heat shock protein 90. *J. Biol. Chem.* *275*, 17857-17862.

- Richter, K., Muschler, P., Hainzl, O., Reinstein, J., and Buchner, J. (2003). Sti1 is a non-competitive inhibitor of the Hsp90 ATPase. Binding prevents the N-terminal dimerization reaction during the atpase cycle. *J. Biol. Chem.* 278, 10328-10333.
- Richter, K., Soroka, J., Skalniak, L., Leskovar, A., Hessling, M., Reinstein, J., and Buchner, J. (2008). Conserved conformational changes in the ATPase cycle of human Hsp90. *J. Biol. Chem.* 283, 17757-17765.
- Riggs, D.L., Cox, M.B., Cheung-Flynn, J., Prapapanich, V., Carrigan, P.E., and Smith, D.F. (2004). Functional specificity of co-chaperone interactions with Hsp90 client proteins. *Crit Rev Biochem Mol Biol* 39, 279-295.
- Roe, S.M., Ali, M.M., Meyer, P., Vaughan, C.K., Panaretou, B., Piper, P.W., Prodromou, C., and Pearl, L.H. (2004). The Mechanism of Hsp90 regulation by the protein kinase-specific cochaperone p50(cdc37). *Cell* 116, 87-98.
- Russell, L.C., Whitt, S.R., Chen, M.S., and Chinkers, M. (1999). Identification of conserved residues required for the binding of a tetratricopeptide repeat domain to heat shock protein 90. *J. Biol. Chem.* 274, 20060-20063.
- Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., Hartl, F.U., and Moarefi, I. (2000). Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell* 101, 199-210.
- Schneider, C., Sepp-Lorenzino, L., Nimmesgern, E., Ouerfelli, O., Danishefsky, S., Rosen, N., and Hartl, F.U. (1996). Pharmacologic shifting of a balance between protein refolding and degradation mediated by Hsp90. *Proc. Natl. Acad. Sci. USA* 93, 14536-14541.
- Schrodinger, L. (2010). The PyMOL Molecular Graphics System, Version 1.3r1.
- Shiau, A.K., Harris, S.F., Southworth, D.R., and Agard, D.A. (2006). Structural Analysis of E. coli hsp90 reveals dramatic nucleotide-dependent conformational rearrangements. *Cell* 127, 329-340.
- Siligardi, G., Hu, B., Panaretou, B., Piper, P.W., Pearl, L.H., and Prodromou, C. (2004). Co-chaperone regulation of conformational switching in the Hsp90 ATPase cycle. *J. Biol. Chem.* 279, 51989-51998.
- Siligardi, G., Panaretou, B., Meyer, P., Singh, S., Woolfson, D.N., Piper, P.W., Pearl, L.H., and Prodromou, C. (2002). Regulation of Hsp90 ATPase activity by the co-chaperone Cdc37p/p50cdc37. *J. Biol. Chem.* 277, 20151-20159.
- Sinars, C.R., Cheung-Flynn, J., Rimerman, R.A., Scammell, J.G., Smith, D.F., and Clardy, J. (2003). Structure of the large FK506-binding protein FKBP51, an Hsp90-binding protein and a component of steroid receptor complexes. *Proc. Natl. Acad. Sci. USA* 100, 868-873.
- Southworth, D.R., and Agard, D.A. (2011). Client-loading conformation of the Hsp90 molecular chaperone revealed in the cryo-EM structure of the human Hsp90:Hsp complex. *Mol. Cell* 42, 771-781.
- Stebbins, C.E., Russo, A.A., Schneider, C., Rosen, N., Hartl, F.U., and Pavletich, N.P. (1997). Crystal structure of an Hsp90-geldanamycin complex: targeting of a protein chaperone by an antitumor agent. *Cell* 89, 239-250.
- Street, T.O., Lavery, L.A., and Agard, D.A. (2011). Substrate binding drives large-scale conformational changes in the Hsp90 molecular chaperone. *Mol. Cell* 42, 96-105.
- Taylor, P., Dornan, J., Carrello, A., Minchin, R.F., Ratajczak, T., and Walkinshaw, M.D. (2001). Two structures of cyclophilin 40: folding and fidelity in the TPR domains. *Struct.* 9, 431-438.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673-4680.
- Vaughan, C.K., Gohlke, U., Sobott, F., Good, V.M., Ali, M.M., Prodromou, C., Robinson, C.V., Saibil, H.R., and Pearl, L.H. (2006). Structure of an Hsp90-Cdc37-Cdk4 complex. *Mol. Cell* 23, 697-707.
- Ward, B.K., Allan, R.K., Mok, D., Temple, S.E., Taylor, P., Dornan, J., Mark, P.J., Shaw, D.J., Kumar, P., Walkinshaw, M.D., and Ratajczak, T. (2002). A structure-based mutational analysis of cyclophilin 40 identifies key residues in the core tetratricopeptide repeat domain that mediate binding to Hsp90. *J. Biol. Chem.* 277, 40799-40809.
- Whitesell, L., and Cook, P. (1996). Stable and specific binding of heat shock protein 90 by geldanamycin disrupts glucocorticoid receptor function in intact cells. *Mol Endocrinol* 10, 705-712.
- Whitesell, L., and Lindquist, S.L. (2005). HSP90 and the chaperoning of cancer. *Nat Rev Cancer* 5, 761-772.
- Whitesell, L., Mimnaugh, E.G., De Costa, B., Myers, C.E., and Neckers, L.M. (1994). Inhibition of heat shock protein HSP90-pp60v-src heteroprotein complex formation by benzoquinone ansamycins: essential role for stress proteins in oncogenic transformation. *Proc. Natl. Acad. Sci. USA* 91, 8324-8328.
- Young, J.C., Obermann, W.M., and Hartl, F.U. (1998). Specific binding of tetratricopeptide repeat proteins to the C-terminal 12-kDa domain of hsp90. *J. Biol. Chem.* 273, 18007-18010.
- Zhang, M., Windheim, M., Roe, S.M., Pegg, M., Cohen, P., Prodromou, C., and Pearl, L.H. (2005). Chaperoned ubiquitylation--crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex. *Mol. Cell* 20, 525-538.
- Zhao, R., Davey, M., Hsu, Y.C., Kaplanek, P., Tong, A., Parsons, A.B., Krogan, N., Cagney, G., Mai, D., Greenblatt, J., et al. (2005). Navigating the chaperone network: an integrative map of physical and genetic

interactions mediated by the hsp90 chaperone. *Cell* 120, 715-727.

Zhao, R., Kakihara, Y., Gribun, A., Huen, J., Yang, G., Khanna, M., Costanzo, M., Brost, R.L., Boone, C., Hughes, T.R., *et al.* (2008). Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation. *J. Cell. Biol.* 180, 563-578.

## FIGURE LEGENDS

**Figure 1: Overview of the free Tah1 NMR solution structure and.** (A) Backbone representation of the overlay between the 20 energetically lowest conformers of free Tah1 protein. The first and second TPR motifs are colored in turquoise and gray blue respectively and helix C in brown. The  $3_{10}$  helix is colored in red. The names of helices, motifs, and the amino/carboxyl extremities are indicated. (B) Molecular surface views of the free Tah1 average conformer with ribbon representation transparency. The two views are rotated by 180°. The color coding is the same as in (A). Figures were generated with PyMOL software (Schrodinger, 2010). (C) Structural sequence alignment with different TPR domains of Hsp70/90 co-chaperones family in Eukaryote. All sequences are taken from GenBank and are of human origin (*Homo sapiens* or Hs) except for CHIP which is from mouse (*Mus musculus* or Mm) and Tah1 and Tom71 which are from *Saccharomyces cerevisiae* (Sc). Hsp70-binding TPR domains: TPR1 (P31948) and CHIP (Q9WUD1). Hsp90-binding TPR domains: TPR2A (P31948), TPR2B (P31948), PP5 (P53041), Tah1 (P25638), and Tom71 (P38825). The numbering corresponds to the Tah1 sequence. Identical residues are represented by white letter on blue background, and homologous residues which exhibit more than 50% of sequence homology by black letter on grey background. Residues implicated on the carboxylate clamp are indicated by "\*". Conserved residues of TPR motif consensus are shown. The secondary structural elements and helical boundaries are mapped to the Tah1 sequence on the top of alignment with color coding and nomenclature as in (A). The sequence alignment was performed using CLUSTAL W and manually optimized (Thompson et al., 1994).

**Figure 2: Interaction fragments mapping between Tah1, Pih1 and Hsp90.** (A) Yeast two-hybrid assays between Tah1 fragments used as the bait (pGBKT7-Gal4BD) and Pih1 or Hsp90

used as the preys (pACTII-Gal4AD). (B) Yeast two-hybrid assays between Pih1 fragments used as the bait (pGBKT7-Gal4BD) and Tah1 or Hsp90 used as the preys (pACTII-Gal4AD). For each assay, a control was performed with an empty pACTII vector (-). Growth was performed on medium lacking histidine but containing an increasing range of 3-Amino-1,2,4-triazol (3AT), a competitive inhibitor of the product of the HIS3 reporter gene. Domain boundaries of Tah1 (with color coding as in (A)) and Pih1 are indicated on the left of the panels.

**Figure 3: NMR solution structure of the Tah1-Hsp90 peptide complex.** (A) Ensemble of 20 lowest energy structures of Tah1 bound to C-terminal peptide of Hsp90 in ribbon representation. (B, C, and D) Molecular surface representation of the Tah1 average conformer with ribbon representation transparency and alpha carbon trace representation (symbolized by black sphere connected by a green line) of the Hsp90 C-terminal peptide in front (B), bottom (C) and side (D) views of Tah1-Hsp90 complex. The amino acids of the peptide are labeled. For clarity, the unfolded C-terminal domain of Tah1 protein is not displayed in B, C and D panels. The nomenclature and color coding are the same than the figure 1.

**Figure 4: Interaction interfaces detailed of 20 lowest energy NMR structures in Tah1-Hsp90 peptide complex.** (A) Magnified views of different interaction regions are represented around the full view of the complex in the center. The protein and the peptide are represented in ribbon and sticks respectively. The nomenclature and color coding are the same than the figure 1 except for peptide carbons which are green. The  $3_{10}$  helix of Tah1 is represented. Residues involved in the interaction are labeled and their side chains shown in stick. For clarity, some views are rotated relative to the central panel. Hydrogen bonds and salt bridges are symbolized by dotted and full lines respectively. (B) Schematic summary of the interaction network at the Tah1p/Hsp90 peptide interface. The primary sequence of C-terminal peptide of Hsp90 was developed and the amino-acids labeled. Only amino acids of the protein interacting with the peptide are indicated. Red, blue and violet lines represent salt bridges, H-bonds and Van Der Waals interactions respectively.

**Figure 5: *In cellulo* analysis of single mutations effect on the Tah1-Hsp90 interaction.** (A) Yeast two-hybrid assays between Tah1 and its variants used as the bait (pGBKT7-Gal4BD) and Hsp90 or Pih1 used as the prey (pACTII-Gal4AD). Surface views of Tah1 (beige) with trace of Hsp90 peptide (green) are indicated on the left panel. The mutated amino acids are mapped on the Tah1 surface. The nomenclature and color coding are the same than the figure 1. (B) Yeast two-hybrid assays between Hsp90 and its variants used as the prey (pACT2-Gal4AD) and Tah1 or Pih1 used as the bait (pGBKT7I-Gal4BD). The correct expression of various variants was checked by testing their capacity to interact with Pih1. For each variant, a negative control was also performed with an empty pACT2 vector (-). Growth was performed on medium lacking histidine but containing an increasing range of 3-Amino-1,2,4-triazol (3AT), a competitive inhibitor of the product of the HIS3 reporter gene. Surface views of Tah1 (beige) with trace of Hsp90 peptide (green) are indicated on the left panel. The side chains of mutated amino acids of Hsp90 peptide are represented in stick. The nomenclature and color coding are the same than the figure 3.

**Figure 6: Structural comparison of Tah1-Hsp90 C-terminal peptide complex with different TPR domains of Hsp90 co-chaperones family bound to Hsp90 peptide.** The yeast lowest energy Tah1-ADTEMEEVD peptide structure is overlaid with human PP5-MEEVD peptide (PDB code 2BUG) (A) {Cliff, 2006, Conformational diversity in the TPR domain-mediated interaction of protein phosphatase 5 with Hsp90.} human Hop TPR2A-MEEVD peptide (PDB code 1ELR) (B) (Scheufler et al., 2000), mouse CHIP-DTSRMEEVD peptide (PDB code 2C2L) (C) {Zhang, 2005, Chaperoned ubiquitylation crystal structures of the CHIP U box E3 ubiquitin ligase and a CHIP-Ubc13-Uev1a complex}, yeast Tom71-MEEVD peptide (PDB code 3FP2) (D) {Li, 2009, Molecular chaperone Hsp70/Hsp90 prepares the mitochondrial outer membrane translocon receptor Tom71 for preprotein loading}, and yeast Tah1\*-MEEVD peptide (PDB code 2L6J) (E) {Jimenez, 2011, The structure of the minimal tetratricopeptide repeat domain protein Tah1 reveals the mechanism of its interaction with Pih1 and Hsp90.} complexes. For the comparison of TPR domain in the structure of the complexes, the peptide was

drawing in transparency mode. (F) Overlay of all TPR domains in transparency mode and magnified views on backbone peptides. The alpha carbon of the residue M<sup>705</sup> is indicated by a full sphere. \* symbolized the Tah1-MEEVD complex recently published by Jimenez et al. The structure alignment was performed by PyMOL using 4 to 70 backbone residues as initial parameters {Schrodinger, 2010, The PyMOL Molecular Graphics System, Version 1.3r1}.

## SUPPLEMENTAL DATA

### Supplemental Legends

**Supplemental Figure S1: Flexibility and electrostatic potentials of Tah1 protein in free and bound forms.** (A) Superposition of heteronuclear NOE intensities for each residues of Tah1 in free (black point) and bound (red square) forms. (B) Vacuum electrostatic potentials mapped on the Tah1 molecular surface in two opposite views using VASCo Plugin software {Steinkellner, 2009, VASCo: computation and visualization of annotated protein surface contacts}. Blue, white and red regions correspond to positive, neutral and negative electrostatic potentials respectively. (C) Vacuum electrostatic potentials mapped on the bound Tah1 molecular surface complexed with the Hsp90 C-terminal peptide in trace drawing mode. The color coding is the same as in (B). (D) Vacuum electrostatic potentials mapped on the molecular surface of the Tah1-Hsp90 peptide complex. The process and the color code are like to (B).

**Supplemental Figure S2: Mapping of the Tah1-Hsp90 peptide interaction by chemical shift perturbations.** (A) <sup>15</sup>N-<sup>1</sup>H HSQC spectral overlays of free (black) and bound forms (red) of Tah1 protein [1:1 protein:peptide molar ratio]. (B). Bar graphs represent normalized amide chemical shift perturbations per residues for Tah1 in complex with Hsp90 peptide. Amide chemical shift changes were normalized using the equation:  $\Delta\delta^{\text{HN}} = \sqrt{(\Delta\text{HN})^2 + (\Delta\text{N}/6.51)^2}$ . The horizontal red dot line indicates the threshold for significant chemical shift changes (average +1 of standard deviation). Residues that had significant chemical shift changes are identified in colors (red: >0.4 ppm, orange: 0.3 to 0.4 ppm, and yellow: <0.3 ppm). (C) TPR domain structural overlays between the Tah1-ADTEMEEVD peptide complex and free Tah1. The peptide

backbone is shown in green transparency and the ribbon representation of bound and free Tah1 are colored in light and dark brown respectively. The right side view is rotated by 180°.

### Supplemental Experimental Procedures

**Biological Material used in this study.** Recombinant pACTII and pGBKT7 plasmids expressing Gal4AD and Gal4BD fusion proteins respectively were obtained by cloning various coding sequences between restriction site BamHI/XhoI and BamHI/SalI respectively. DNA inserts were obtained by PCR amplification using appropriate primers containing the restriction sites. These amplifications were realized by using genomic DNA of *Saccharomyces cerevisiae* BY4741 strain (Euroscarf). The yeast encoding amino acids sequences for the yeast two-hybrid assays include the following domains: Tah1(1-111), Tah1(1-92), Tah1(93-111), Pih1(1-344), Pih1(1-252), Pih1(253-344), Hsp82(1-709) and mutants Tah1 R77/A, Tah1 K79/A, Tah1 R83/A, Tah1 E85/A, Hsp82 D702/A, Hsp82 D703/A, Hsp82 E706/A, Hsp82 E707/A, Hsp82 V708/A, Hsp82 D709/A.

**NMR measurements.** Sequence-specific backbone assignments were achieved using 2D ( $^{15}\text{N}$ - $^1\text{H}$ ) and ( $^{13}\text{C}$ - $^1\text{H}$ ) HSQCs, 3D HNCO, 3D HN(CA)CO, 3D CBCANH, and 3D CBCA(CO)NH experiments. Side chain assignments were performed using 3D H(C)CH-TOCSY, 3D HNHA, 3D HBHANH, 3D HBHA(CO)NH, 3D NOESY-( $^{15}\text{N}$ - $^1\text{H}$ )-HSQC and 3D NOESY-( $^{13}\text{C}$ - $^1\text{H}$ )-HSQC experiments. Aromatic proton assignments were performed using 2D-TOCSY and 2D-NOESY in 100%  $\text{D}_2\text{O}$ . Relaxation measurements were performed in the same sample conditions using  $^{15}\text{N}$  labeled Tah1 protein at a ~1 mM concentration and a Bruker AVIII-600 MHz spectrometer.

We assigned the bound Tah1 backbone and side chains using 2D ( $^{15}\text{N}$ - $^1\text{H}$ ) and ( $^{13}\text{C}$ - $^1\text{H}$ ) HSQC, 3D HNCA, 3D hCcoNH, 3D HcccoNH, 3D NOESY ( $^{15}\text{N}$ - $^1\text{H}$ ) and ( $^{13}\text{C}$ - $^1\text{H}$ ) HSQC experiments, taking advantage of the backbone and side chains assignments of free Tah1. We assigned the aromatic proton of the bound Tah1 using 2D ( $^{13}\text{C}$ - $^1\text{H}$ )-HSQC in  $\text{H}_2\text{O}$  and 100%  $\text{D}_2\text{O}$  as well as 3D NOESY-( $^{13}\text{C}$ - $^1\text{H}$ )-HSQC recorded in the aromatic region.

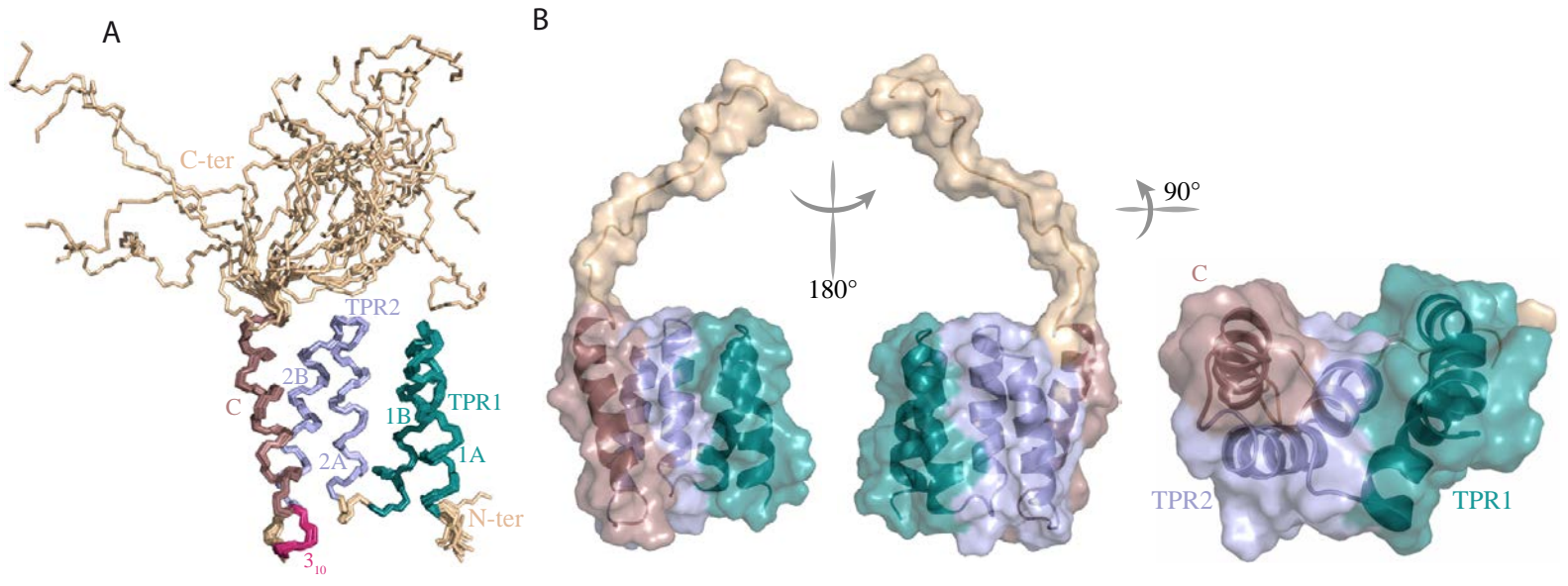


FIGURE 1.



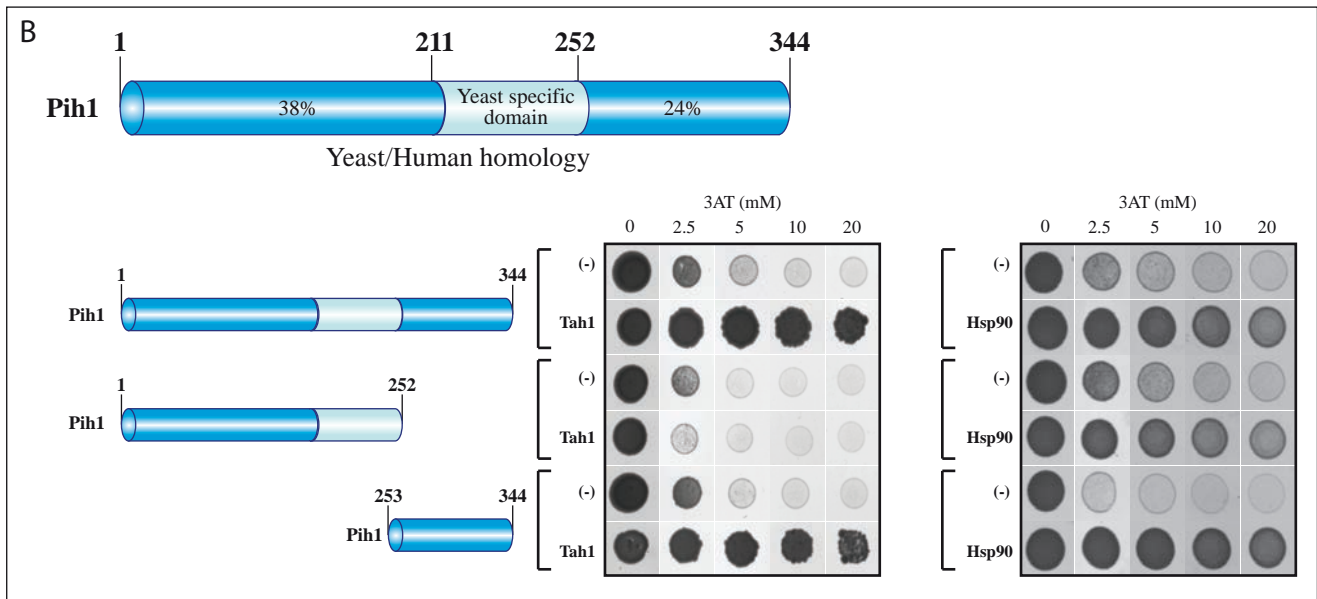
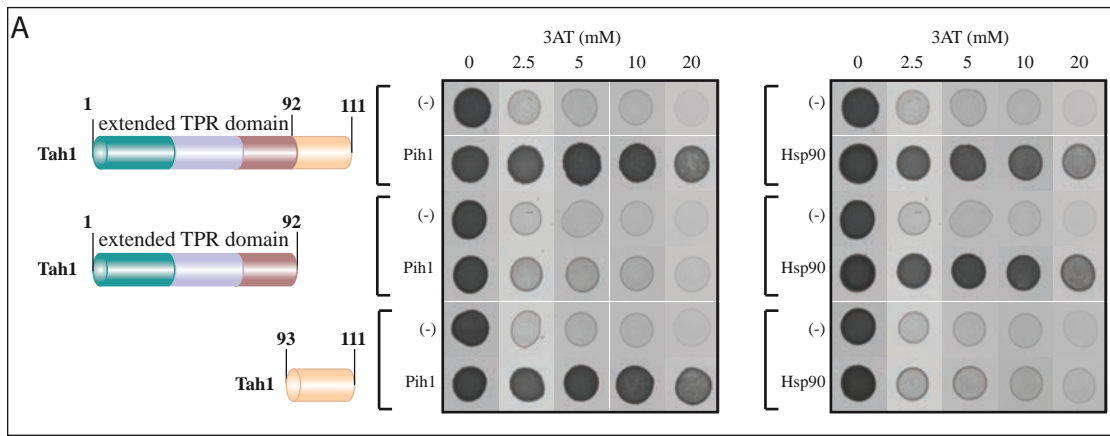


FIGURE 2 .

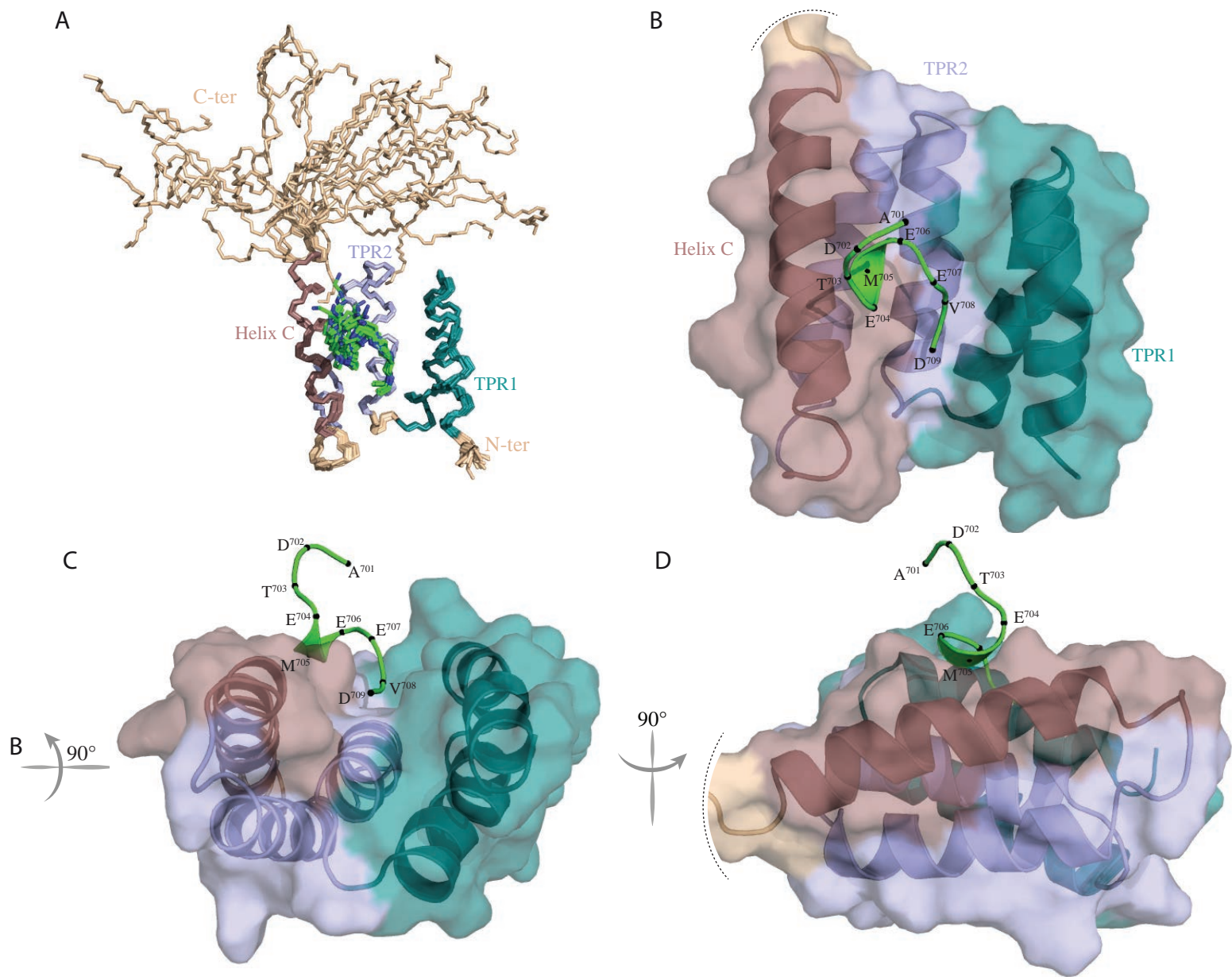
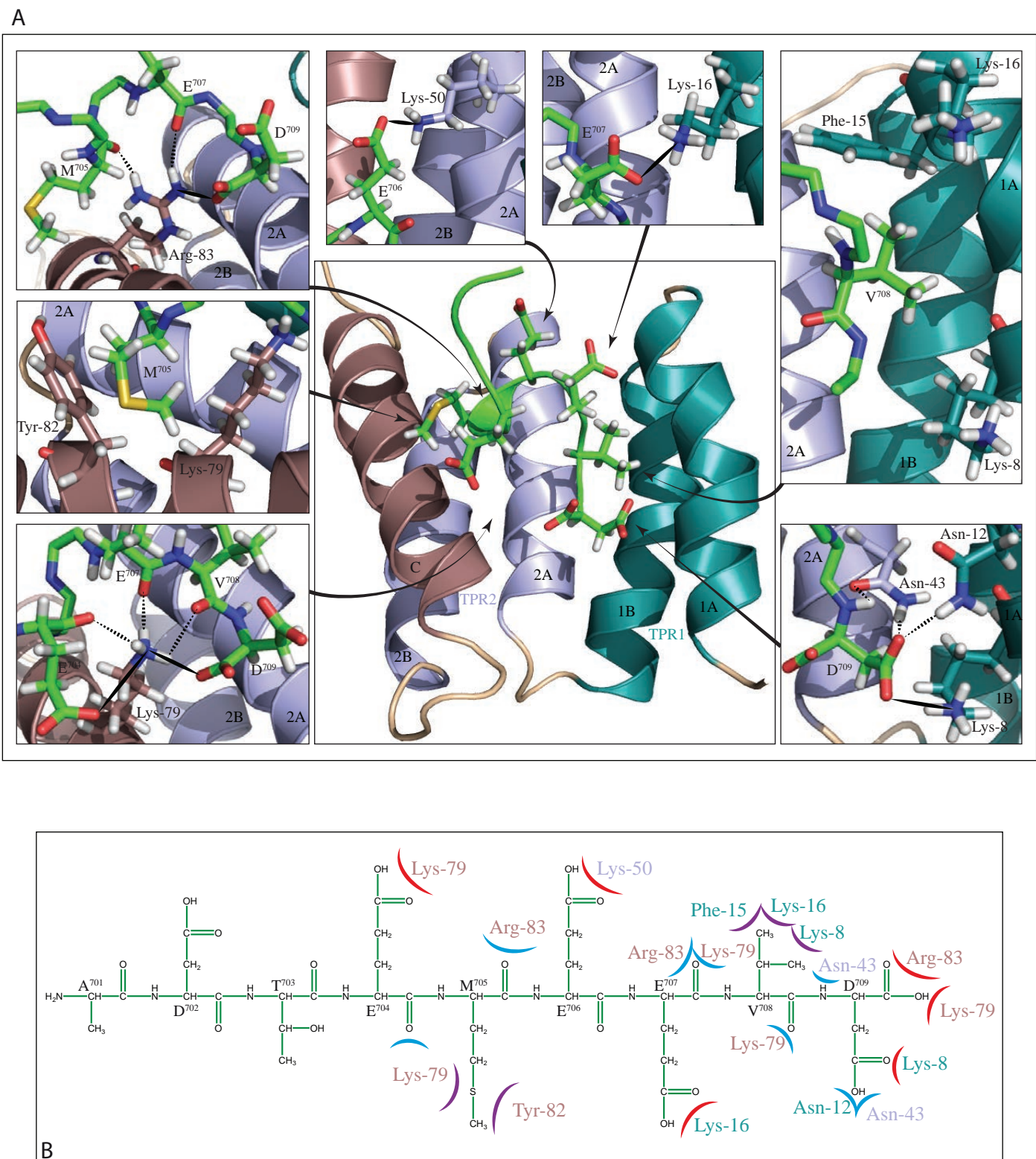
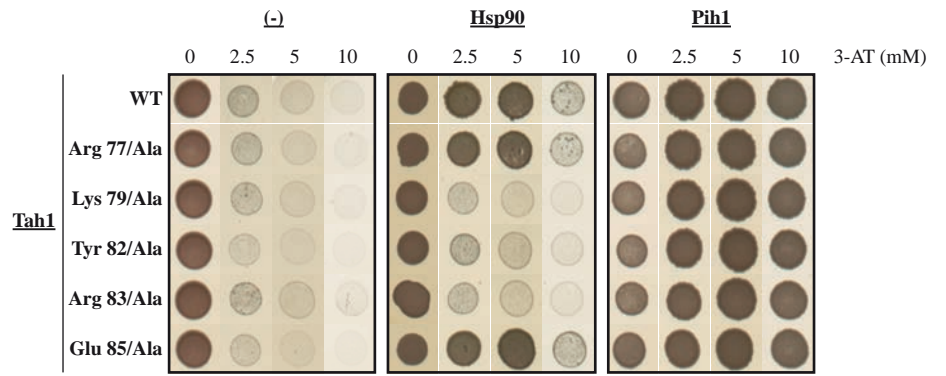
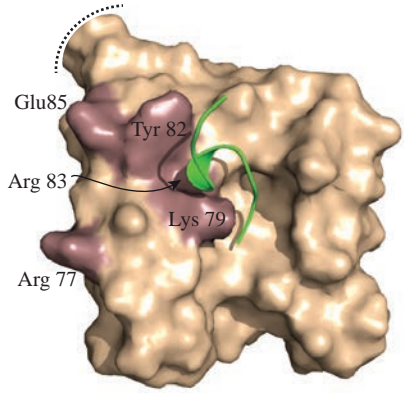


FIGURE 3.



A



B

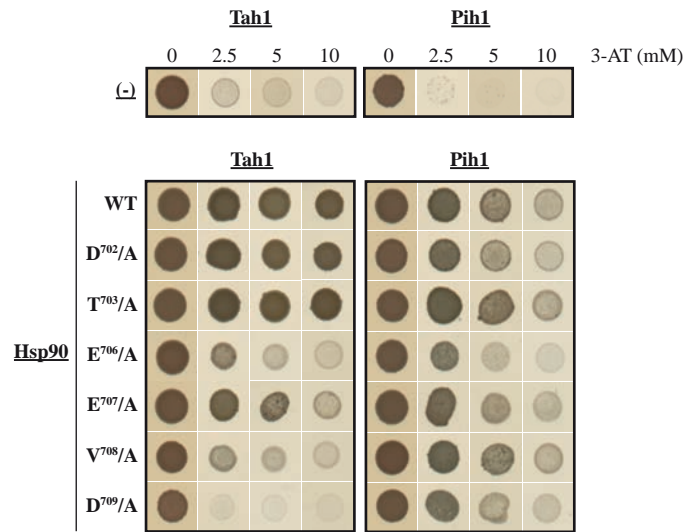
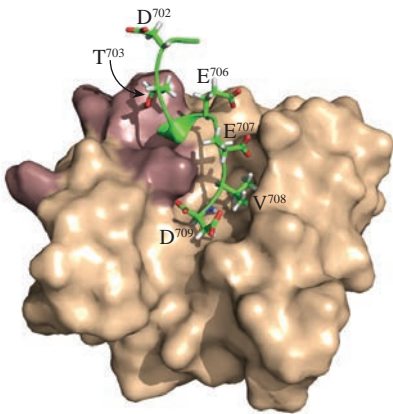


FIGURE 5 .

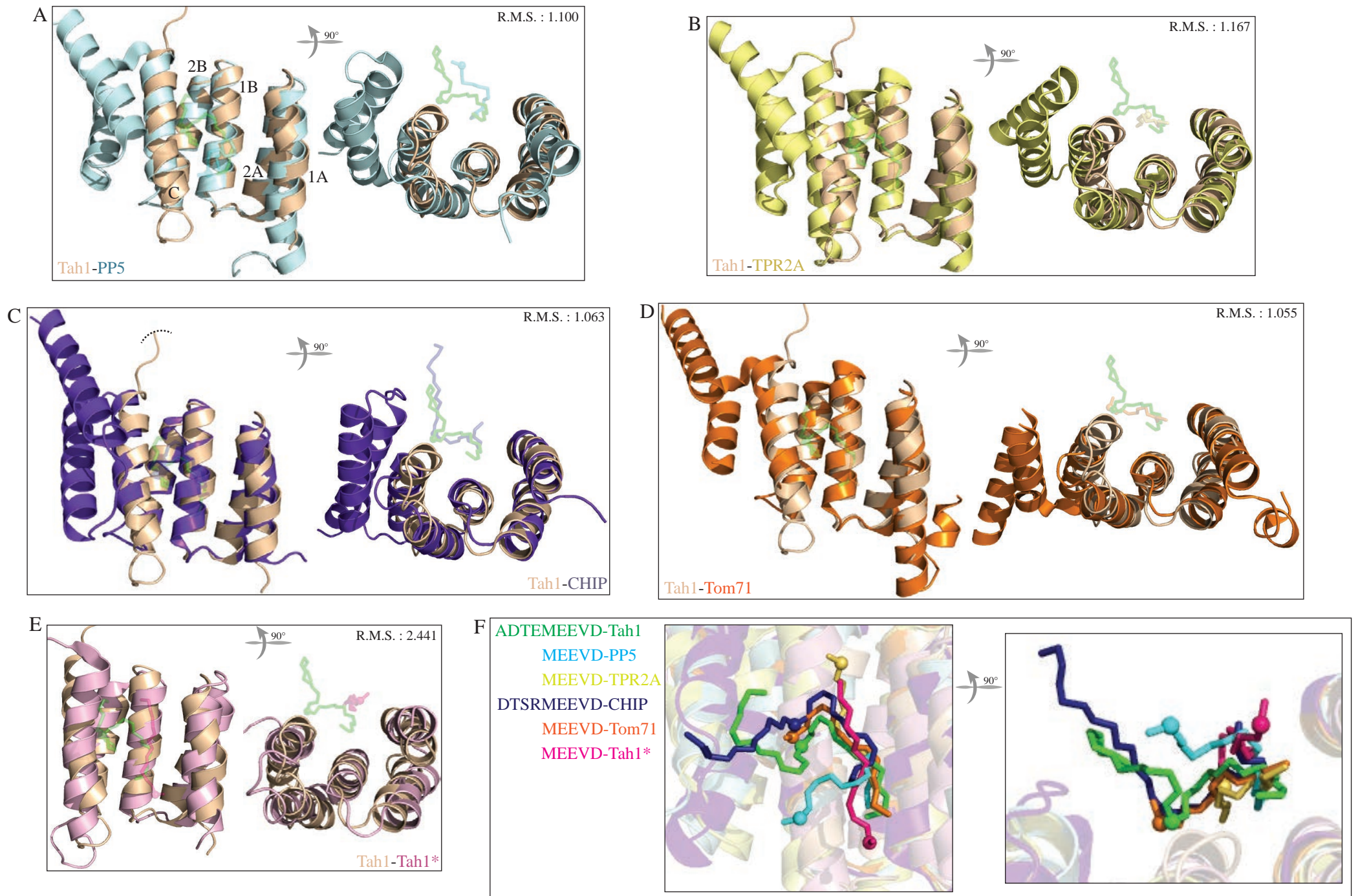


FIGURE 6.

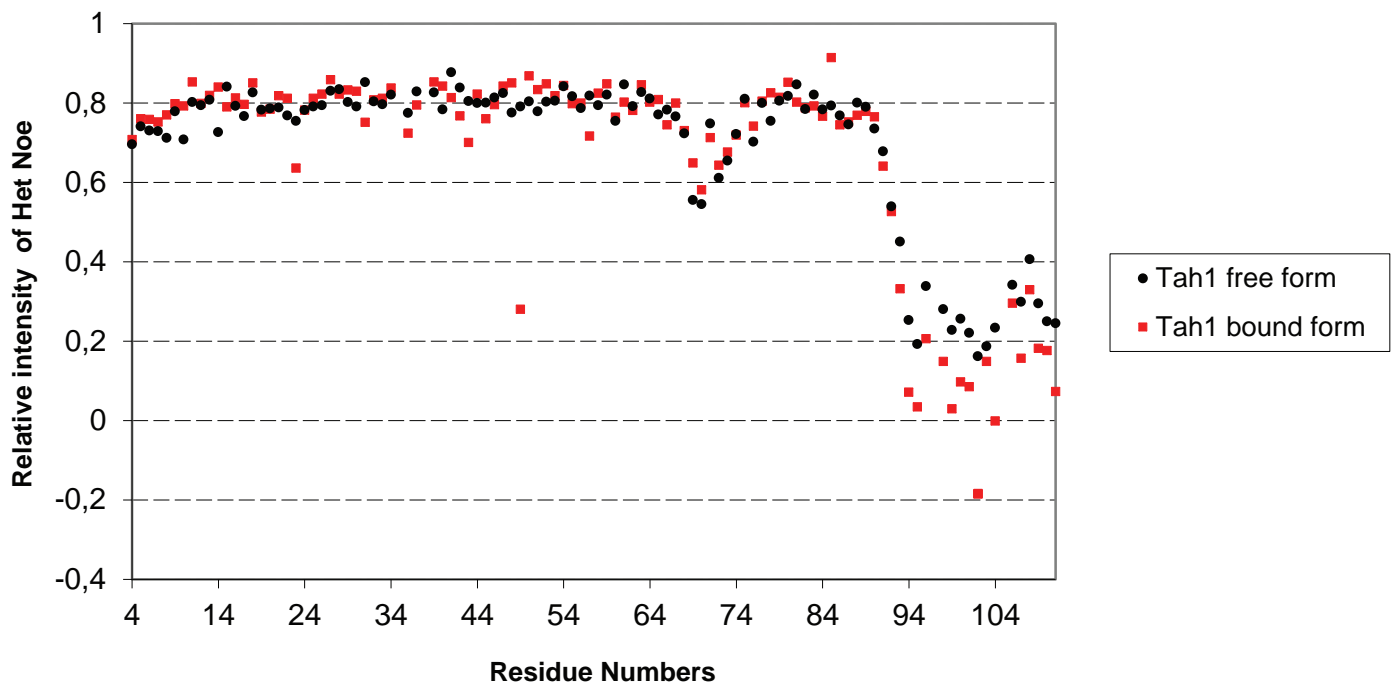
	Free form		Complex	
	Tah1 protein	Tah1 protein	Tah1 protein	Hsp82 peptide
<b>NMR distance constraints</b>				
• Total Intramolecular Noes	1786	2659		56
○ Intra-residue	482	610		19
○ Inter-residues				
▪ Sequential	448	687		25
▪ Medium range	411	776		12
▪ Long range	445	586		0
• Total Intermolecular Noes			56	
<b>Structure statistics</b>				
• RMSD to mean structure				
○ Backbone atoms	0.44 ± 0.08 (a)	0.36 ± 0.10 (a)	0.42 ± 0.10 (a+b)	0,56 ± 0.21 (b)
○ Heavy atoms	0.98 ± 0.13 (a)	0.86 ± 0.12 (a)	0.92 ± 0.12 (a+b)	1.30 ± 0.25 (b)
• RMSD from idealized covalent geometry				
○ Bond length (Å)	0.0038 ± 0.0001		0.0042 ± 0.0001	
○ Bond angle (°)	1.43 ± 0.02		1.74 ± 0.03	
• Violations				
○ Average number of violations (>0.4 Å)	0.00		9.6	
○ Highest violation (Å)	0.39		0.95	
• Energies (kcal.mol <sup>-1</sup> )				
○ Average AMBER	-4009 ± 20		-4146 ± 17.5	
○ Average constraint violations energy	96 ± 10		530 ± 18	
• Ramachandran analysis				
○ Most favoured	70.6		74.8	
○ Additionnaly allowed region	28.3		22.9	
○ Generously allowed region	1.0		2.1	
○ Disallowed region	0.1		0.2	

<sup>(a)</sup> r.m.s. deviation was calculated using residues 4-90 for the ensemble of 20 refined structures

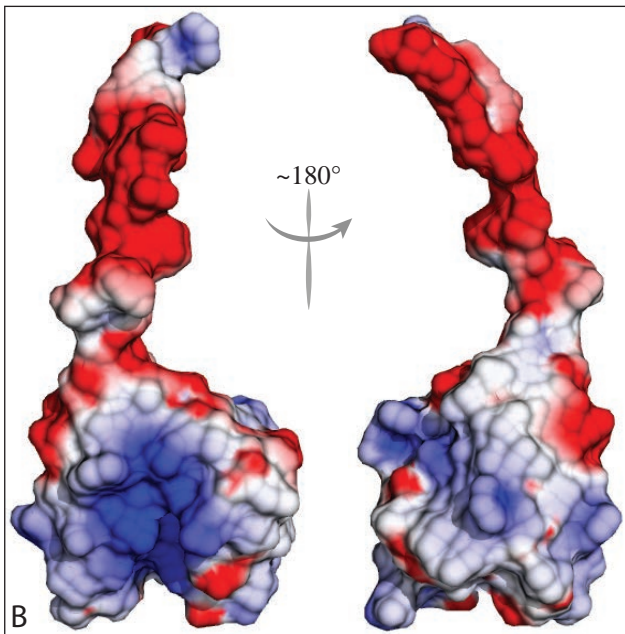
<sup>(b)</sup> r.m.s. deviation was calculated using residues 704-709 for the ensemble of 20 refined structures

Table 1: NMR structure determination statistics of the Tah1 protein and Tah1p-Hsp82 complex

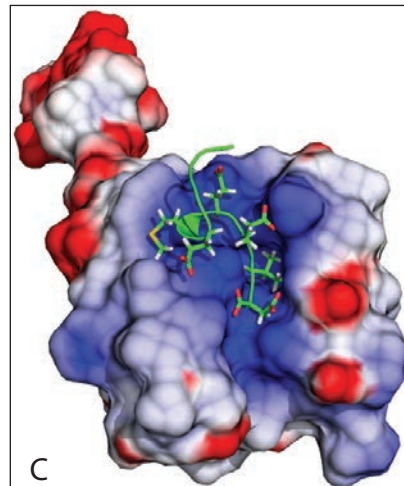
### Flexibility of yTah1p free vs bound form at 288K



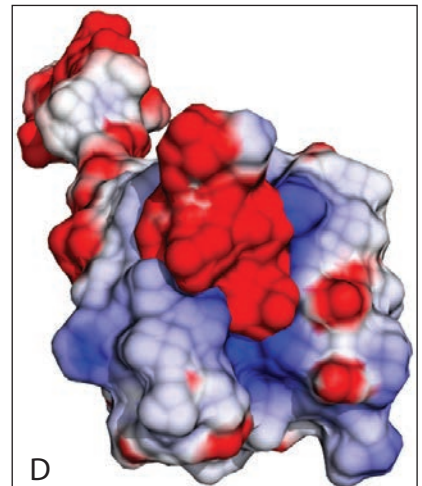
A



B



C



D

FIGURE S1.





## 2.2. Données complémentaires

Dans les chapitres précédents, nous avons étudié le rôle des différents facteurs d'assemblage dans la biogenèse des snoRNP à boîtes C/D. De façon surprenante, la caractérisation des souches de levures invalidées pour les gènes d'intérêt a révélé la totale neutralité du phénotype  $\Delta T A H 1$ . En effet, dans des conditions standard de culture de la levure *S. cerevisiae* en laboratoire, l'absence de la co-chaperonne n'a aucune incidence sur la croissance des cellules, sur le taux de snoRNA à boîtes C/D et sur la production des ARNr (Rothé *et al.*, Publication n°3). En dépit de sa présence au sein du complexe R2TP, la co-chaperonne Tah1p ne semble donc pas exercer une influence majeure sur la biogenèse des snoRNP et l'implication réelle des chaperonnes HSP90 (*i.e.* Hsp82p et Hsc82p) reste à démontrer chez la levure. L'équipe de W.A. Houry a proposé que le rôle de ces facteurs puisse prévaloir durant la phase stationnaire de croissance (Zhao *et al.*, 2008). En effet, d'après cette étude, l'accumulation de la protéine Pih1p et le taux de snoRNA sont contrôlés par Tah1p et HSP90 au cours de cette phase. Afin d'apporter de nouveaux éléments de réflexion, nous avons donc mis à profit l'apport des données structurales, en testant les conséquences d'une perte d'interaction entre Tah1p et HSP90 sur le taux de snoRNA et sur l'accumulation de la protéine Pih1p au cours de la phase stationnaire de croissance.

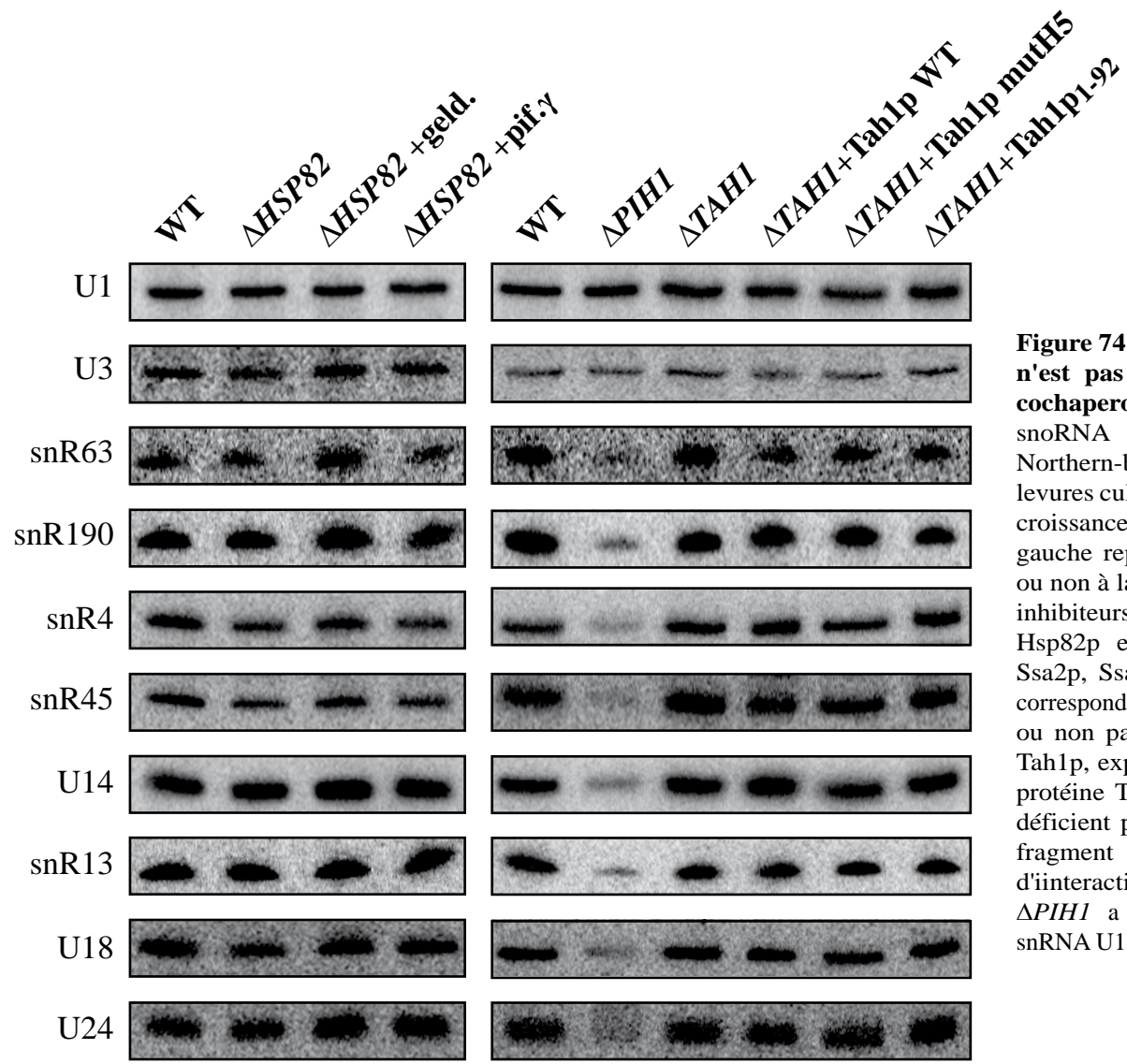
### 2.2.1. HSP90 et Tah1p ne sont pas requis pour l'accumulation des snoRNA à boîtes C/D

L'une des caractéristiques communes des facteurs d'assemblage est leur influence sur le taux de snoRNA. Les protéines Rvb1p et Rvb2p sont en effet requises pour l'accumulation des snoRNA à boîtes C/D et H/ACA (King *et al.*, 2001; Zhao *et al.*, 2008), la protéine Pih1p est quant à elle importante pour l'accumulation spécifique des snoRNA à boîtes C/D ((Zhao *et al.*, 2008); Rothé *et al.*, Publication n°3), enfin les facteurs Rsa1p et Hit1p contrôlent le taux d'une sous-population de snoRNA à boîtes C/D (Rothé *et al.*, Publication n°3) (cf. Résultats & Discussion, Chapitre 3) (**Fig. 65**). Cette propriété s'exprime de façon non ambiguë au cours de la phase exponentielle de croissance, qui correspond à une phase de haute activité métabolique, au cours de laquelle la biogenèse des ribosomes est maximale (Warner, 1999). D'après les résultats obtenus par l'équipe de W.A. Houry, il semblerait en revanche que l'influence de Tah1p et de la chaperonne s'exerce spécifiquement pendant la phase stationnaire de croissance, qui correspond à une phase de stress, au cours de laquelle la

surexpression des chaperonnes assure la maintenance de certaines fonctions (Borkovich *et al.*, 1989; Donalies & Stahl, 2001; Zhao *et al.*, 2008). Afin d'aborder cette problématique, nous avons décidé de reproduire ces expériences en analysant, au cours de la phase stationnaire de croissance, le taux de snoRNA à boîtes C/D dans un contexte d'inhibition des HSP90 ou en absence de la protéine Tah1p. De plus, afin d'affiner cette analyse, nous avons également utilisé des formes de la protéine Tah1p déficientes soit pour l'interaction avec HSP90, soit pour l'interaction avec Pih1p (Back *et al.*, Publication n°4).

Concernant l'inhibition de la fonction HSP90, nous avons utilisé une souche de levure  $\Delta HSP82$ , qui nous a été fournie par W.A. Houry (Zhao *et al.*, 2008). Dans son article, l'équipe de W.A. Houry avait préféré l'utilisation d'une souche de levure cumulant la délétion du gène *HSC82* et l'expression d'un mutant de Hsp82p. Afin de s'assurer de l'inhibition complète de la seconde isoforme de la chaperonne, la protéine Hsc82p, nous avons donc cultivé les cellules en présence de geldanamycine, un inhibiteur qui bloque son activité ATPase (Boulon *et al.*, 2008; Zhao *et al.*, 2005). De plus, nous avons réalisé un essai en présence de pifithrine  $\gamma$ , un inhibiteur spécifique des HSP70, une autre famille de chaperonnes qui participent au cycle d'activité des HSP90 (cf. Introduction, Chapitre 3, 3.6.3) (**Fig. 52**). Ces inhibiteurs ont été ajoutés dans le milieu pendant les phases de préculture et de culture des cellules afin de maintenir une inhibition constante tout au long de la production de biomasse. Comme attendu, la présence de geldanamycine s'est traduite par un défaut de croissance marqué (données non montrées). En revanche, la pifithrine  $\gamma$  qui à notre connaissance n'a jamais fait l'objet d'étude sur la physiologie des cellules de levure, a provoqué un effet beaucoup plus modéré (données non montrées). Concernant Tah1p, nous avons utilisé la souche de levure  $\Delta TAH1$ , complémentée ou non par différentes formes de la protéine exprimées à partir du vecteur p413 : (i) Tah1p WT, (ii) Tah1p mutH5 dont la mutation des résidus I76, S78, K79, Y82, R83 en alanine assure la perte d'interaction avec HSP90 et (iii) la protéine Tah1p<sub>1-92</sub> tronquée du domaine d'interaction avec Pih1p (Back *et al.*, Publication n°4).

Les cultures ont été réalisées en milieu riche (YPD) à une température de 30°C, jusqu'à obtention d'une valeur d' $A_{600nm}/mL$  supérieure à 4. En comparaison avec le snRNA U1, le taux des snoRNA a été analysé par l'approche de Northern-blot (**Fig. 74**). De façon surprenante, aucun effet n'a été constaté sur le taux de snoRNA à boîtes C/D dans les différentes souches utilisées. A titre de comparaison avec les travaux de l'équipe de W.A.



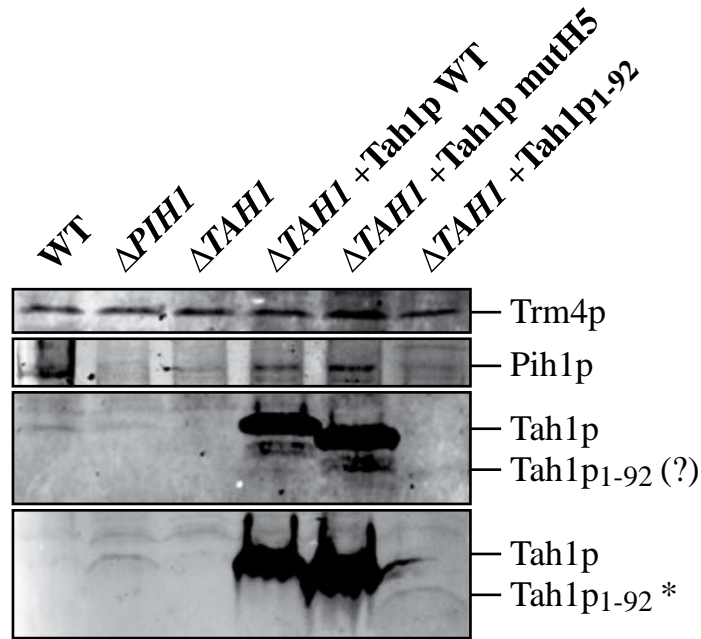
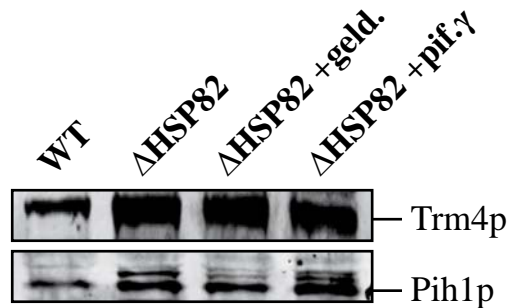
**Figure 74 : Le taux de snoRNA à boîtes C/D n'est pas dépendant de Hsp82p et de sa cochaperone Tah1p.** Le taux de plusieurs snoRNA à boîtes C/D a été analysé par Northern-blot dans différentes souches de levures cultivées jusqu'en phase stationnaire de croissance ( $A_{600nm/mL} >4$ ). Le panneau de gauche représente la souche  $\Delta HSP82$ , traitée ou non à la geldanamycine et à la pifithrin  $\gamma$ , 2 inhibiteurs respectifs des protéines HSP90 (i.e. Hsp82p et Hsc82p) et HSP70 (i.e. Ssa1p, Ssa2p, Ssa3p et Ssa4p). Le panneau de droite correspond à la souche  $\Delta TAH1$  complétementée ou non par différentes formes de la protéine Tah1p, exprimées à partir du vecteur p413 : la protéine Tah1p WT, le variant Tah1p mutH5 déficient pour l'interaction avec Hsp82p et le fragment Tah1p<sub>1-92</sub>, dépourvu du domaine d'interaction avec Pih1p. La souche de levure  $\Delta PIH1$  a été analysée à titre comparatif. Le snRNA U1 a été utilisé comme contrôle interne.

Houry, nous avons notamment analysé le taux des ARN snR4, snR13 et U24, pour lesquels une baisse de stabilité avait été conclue, cependant nos résultats ne sont pas en mesure de confirmer ces données. A la différence des autres facteurs d'assemblage, il semble donc que les protéines HSP90 et Tah1p exercent une influence mineure sur l'accumulation des snoRNA à boîtes C/D, y compris pendant la phase stationnaire de croissance.

### 2.2.2. Tah1p contrôle la stabilité de Pih1p indépendamment de HSP90

Si Tah1p n'exerce pas de contrôle apparent sur le taux de snoRNA, il a en revanche été montré que son absence conduit à une diminution drastique du taux de la protéine Pih1p durant la phase stationnaire de croissance ((Zhao *et al.*, 2008); Rothé *et al.*, Publication n°3). Par ailleurs, l'équipe de W.A. Houry a mis en évidence que cet effet n'est pas corrélé à une variation du taux d'ARNm codant Pih1p, mais plutôt à une perte de stabilité de la protéine (Zhao *et al.*, 2008). En utilisant des approches de chromatographie et de DLS, ils ont également montré qu'en présence d'ATP, HSP90 empêche l'agrégation de la protéine Pih1p recombinante et que cette propriété est améliorée en présence de Tah1p (Zhao *et al.*, 2008). Ces observations réalisées *in vitro* les ont notamment conduit à la proposer que HSP90 pourrait exercer son activité de chaperonne sur Pih1p et que cette fonction pourrait lui être spécifiée par la co-chaperonne Tah1p (Zhao *et al.*, 2008). Cette hypothèse n'a cependant pas fait l'objet d'une vérification *in vivo*, puisque l'impact réel de HSP90 sur l'accumulation de la protéine Pih1p durant la phase stationnaire de croissance n'a pas été analysé. Afin de vérifier cette proposition, nous avons utilisé les systèmes génétiques présentés dans le paragraphe précédent et nous avons contrôlé par Western-blot le taux de protéine Pih1p pendant la phase stationnaire de croissance (**Fig. 75**). Les protéines Pih1p et Tah1p ont été détectées grâce à l'utilisation d'un sérum dirigé spécifiquement contre le duplex Pih1p-Tah1p (Rothé *et al.*, Publication n°3). La protéine Trm4p a été utilisée comme contrôle interne.

Dans un premier temps, la complémentation du phénotype  $\Delta T A H 1$  par les différentes formes de la protéine Tah1p nous a permis de vérifier spécifiquement les conséquences d'une perte d'interaction avec HSP90 ou avec Pih1p (**Fig. 75A**). En comparaison avec la souche  $\Delta T A H 1$  non complétementée, la surexpression de Tah1p WT et du variant Tah1p-mutH5 conduisent à une restauration similaire du taux de protéine Pih1p. Ce résultat indique donc que le contrôle exercé par Tah1p sur son partenaire est indépendant de l'interaction avec la chaperonne. En

**A****B**

**Figure 75 : Le contrôle du taux de protéine Pih1p par Tah1p n'est pas dépendant de Hsp82p.** (A) Evaluation par Western-blot, du taux de la protéine Pih1p en présence de différentes formes de la protéine Tah1p. L'accumulation de la protéine Pih1p a été analysée dans des souches WT,  $\Delta$ PIH1 et  $\Delta$ TAH1. La souche  $\Delta$ TAH1 complémentée par différentes formes de la protéine Tah1p, exprimées à partir du vecteur p413 : la protéine Tah1p WT, le variant Tah1p mutH5 déficient pour l'interaction avec Hsp82p et le fragment Tah1p<sub>1-92</sub>, dépourvu du domaine d'interaction avec Pih1p. La protéine Trm4p a été utilisée comme contrôle interne. Le panneau inférieur correspond au redépôt des échantillons sur un gel plus concentré en polyacrylamide (15% contre 10%), afin de mieux visualiser la forme tronquée Tah1p<sub>1-92</sub>\*. Ces expériences ont été réalisées à partir de cellules cultivées jusqu'en phase stationnaires de croissance ( $A_{600\text{nm}}/\text{mL} > 5$ ). (B) Evaluation par Western-blot, du taux de protéine Pih1p en fonction de différentes conditions d'inhibition des protéines chaperonnes. L'accumulation de la protéine Pih1p a été analysée dans des souches de levure WT et  $\Delta$ HSP82, traitées ou non à la geldanamycine et à la pifithrin  $\gamma$ , 2 inhibiteurs respectifs des protéines HSP90 (i.e. Hsp82p et Hsc82p) et HSP70 (i.e. Ssa1p, Ssa2p, Ssa3p et Ssa4p). La protéine Trm4p a été utilisée comme contrôle interne.

revanche, l'expression de la protéine Tah1p<sub>1-92</sub>, tronquée du domaine d'interaction avec Pih1p, est associée à une diminution de la quantité de protéine Pih1p similaire à celle observée dans la souche  $\Delta T A H 1$ . Il est cependant difficile de conclure sur ce résultat puisqu'en comparaison avec les autres formes de Tah1p exprimées de façon ectopique, le taux de protéine Tah1p<sub>1-92</sub> est relativement faible, suggérant que son expression ou sa stabilité pourrait être altérées en l'absence du domaine C-terminal. Ces données indiquent donc que le contrôle exercé par Tah1p sur la stabilité de Pih1p au cours de la phase stationnaire est indépendant de sa capacité à interagir avec la chaperonne. Cependant, ce résultat ne permet pas d'exclure totalement l'intervention de HSP90 dans ce phénomène. En effet, comme nous l'avons montré dans l'article *The NMR high resolution structures of yeast Tah1 in a free form and in complex with the Hsp90 C-terminal tail explains the specific recognition of Hsp90 by the R2TP complex*, le domaine N-terminal de Pih1p peut interagir avec la chaperonne indépendamment de Tah1p (Back *et al.*, Publication n°4). Autrement dit, en dépit de la mutation du domaine TPR de Tah1p, le ciblage direct de Pih1p par HSP90 doit également être envisagé.

Dans un second temps, nous avons donc évalué le taux de Pih1p dans les conditions d'inhibition des chaperonnes décrites précédemment (**Fig. 75B**). Contrairement à la disparition presque totale de la protéine Pih1p observée dans le contexte génétique  $\Delta T A H 1$  ((Zhao *et al.*, 2008) ; Rothé *et al.*, Publication n°3), nous avons ainsi constaté que le taux de protéine Pih1p est insensible aux différentes conditions d'inhibition utilisées ( $\Delta H S P 8 2 \pm$  geldanamycine ou pifithrin- $\gamma$ ). Ce résultat indique donc que les protéines HSP90 n'exercent pas d'impact majeur sur le taux de Pih1p dans les cellules en phase stationnaire de croissance. De façon générale, ces différentes observations tendent donc à réfuter l'hypothèse selon laquelle Pih1p pourrait être une protéine cliente de HSP90. Il semble en effet que sa stabilité au cours de la phase stationnaire de croissance soit assurée exclusivement par Tah1p.

### 3. Conclusion & Perspectives

#### 3.1. Bilan

La structure en solution de la protéine Tah1p et de son complexe avec le peptide C-terminal ADTEMEEVD de HSP90, nous a permis de définir avec plus de précision le réseau

d'interaction unissant Tah1p, la chaperonne et Pih1p (Back *et al.*, Publication n°4). Initialement, Tah1p avait été décrite comme une protéine à deux domaines, capable d'interagir avec HSP90 par le biais de son domaine TPR (Zhao *et al.*, 2005; Zhao *et al.*, 2008) et avec Pih1p par l'intermédiaire de sa région C-terminale (Zhao *et al.*, 2005; Zhao *et al.*, 2008). Les données structurales ont permis de définir la nature exacte de ces deux domaines, puisque la région Tah1p<sub>1-92</sub> forme un domaine TPR atypique constitué de 5 hélices  $\alpha$ , tandis que le domaine d'interaction avec Pih1p s'étend des résidus 93 à 111, sous la forme d'un domaine non structuré. La partition de la protéine Pih1p en différents domaines a également permis de montrer que sa région C-terminale Pih1p<sub>253-344</sub>, constitue le domaine minimal d'interaction avec Tah1p, tandis que sa région N-terminal Pih1p<sub>1-252</sub> présente la capacité d'interagir avec HSP90 par double-hybride. A ce titre, l'équipe de W.A. Houry avait montré qu'en absence de Tah1p, la protéine Pih1p n'interagit pas de façon stable avec HSP90 *in vitro* (Zhao *et al.*, 2008). La situation semble donc légèrement différente *in cellulo*, puisque cette interaction peut être maintenue indépendamment de Tah1p, par l'intermédiaire du domaine N-terminal de Pih1p. En revanche la nature physique de cette interaction n'a pas été démontrée, et il est également envisageable qu'elle soit favorisée par l'intervention de partenaires supplémentaires, présents dans la cellule. Enfin, l'analyse de l'interface d'interaction entre Tah1p et Hsp82p nous a permis d'identifier les déterminants de cette interaction et de définir des variants délétères pour l'association des deux partenaires *in vivo*. Leur utilisation nous a permis de découpler les fonctions de Tah1p et de HSP90, afin d'étudier spécifiquement le rôle de chacune d'entre elles dans la biogenèse des snoRNP à boîtes C/D. L'expression des variants de la protéine Tah1p *in vivo* a été menée en parallèle de l'inhibition des chaperonnes HSP90, afin de comparer les effets spécifiques dus à la relation entre Tah1p et HSP90, des effets généraux liés au caractère ubiquitaire de la chaperonne. A la différence des observations réalisées chez l'homme (Boulon *et al.*, 2008), les résultats obtenus nous ont amenés à reconsidérer le rôle respectif de Tah1p et HSP90, ainsi que leur implication dans le processus d'assemblage des snoRNP à boîtes C/D chez la levure.

### **3.2. Précision des relations fonctionnelles entre Pih1p, Tah1p et HSP90**

Précédemment, l'équipe de W.A. Houry avait proposé que HSP90 puisse contrôler la stabilité de Pih1p au cours de la phase stationnaire de croissance (Zhao *et al.*, 2008). Les données structurales acquises par RMN, nous ont permis de développer des outils appropriés afin de

tester cette hypothèse. De façon inattendue, les résultats obtenus indiquent que la stabilisation de Pih1p est indépendante de l'activité des chaperonnes HSP90 et de leur capacité à interagir avec Tah1p. Ces données révèlent donc que cette dernière possède une propriété spécifique et indépendante de son rôle de co-chaperonne, qui lui permet vraisemblablement d'exercer un contrôle sur la stabilité de la protéine Pih1p durant les phases de carence nutritive. Ce résultat est particulièrement intéressant, puisque la stabilité des deux acteurs centraux du processus d'assemblage des snoRNP à boîtes C/D, Rsa1p et Pih1p, semble dépendante de leur association avec des partenaires spécifiques (*i.e.* Hit1p et Tah1p respectivement). Comme nous en avons déjà discuté dans le chapitre précédent au sujet de la relation fonctionnelle entre Rsa1p et Hit1p, il sera nécessaire de vérifier si la dépendance de Pih1p vis-à-vis de Tah1p, donne prise à des événements de régulation visant à limiter la production des snoRNP sous certaines conditions. Une perspective sera notamment de vérifier la stabilité de Pih1p dans un contexte de saturation de la production des ribosomes. A ce titre, l'utilisation de souches de levures mutées pour le gène *NMD3* et présentant un défaut de biogenèse des ribosomes, pourra être envisagée (Gadal *et al.*, 2001; Hedges *et al.*, 2005) (cf. Résultats & Discussion, Chapitre 3, 3.2).

Etant donné que Pih1p ne se comporte pas comme une protéine cliente des chaperonnes HSP90, la relation qu'elle établit avec ces dernières doit également être reconsidérée. Sur la base d'observations expérimentales, il avait été initialement proposé que Tah1p, mais aussi Pih1p, constituent deux co-chaperonnes des protéines HSP90 (Zhao *et al.*, 2005) (cf. Introduction, Chapitre 3, 3.6.3). Leur importance pour la structuration du récepteur aux glucocorticoïdes et de la tyrosine kinase vSrc, deux protéines clientes des HSP90, avait notamment été mise en évidence. Des travaux récents sont venus confirmer cette idée, puisqu'il a été démontré *in vitro* que l'hétérodimère Tah1p-Pih1p exerce une action inhibitrice sur l'activité ATPase de la chaperonne (Eckert *et al.*, 2010). Ce rôle, déjà établi pour d'autres co-facteurs comme Cdc37p ou Sti1p, constitue une propriété typique de certaines co-chaperonnes et pourrait assurer le maintien du dimère de HSP90 dans une conformation compatible avec la prise en charge des protéines clients au début de son cycle fonctionnel (Eckert *et al.*, 2010) (cf. Introduction, Chapitre 3, 3.6.3) (**Fig. 52**). Comme l'implication de Pih1p dans la biogenèse des snoRNP à boîtes C/D semble constituer une propriété distincte de son rôle de co-chaperonne, il est possible que ce facteur fasse l'objet d'une remobilisation durant des phases de stress, afin de participer au maintien de certaines fonctions dépendantes de HSP90. A ce titre, les expériences de co-immunosélection menées à partir de Pih1p ou de



Tah1p n'ont pas permis de mettre en évidence une association significative entre HSP90 et le complexe R2TP dans des conditions de croissance exponentielles ((Zhao *et al.*, 2005; Zhao *et al.*, 2008); Rothé *et al.*, Publication n°3). En revanche, une relation fonctionnelle entre ces deux entités semble s'établir dans des conditions de carence nutritive, puisque HSP90 est associée au facteur Pih1p, spécifiquement pendant la phase stationnaire de croissance (Zhao *et al.*, 2008).

### 3.3. HSP90, Tah1p et maintenance du processus d'assemblage ?

L'ensemble des données acquises à ce stade indiquent que les différents facteurs ayant fait l'objet d'études chez la levure *S. cerevisiae*, ne revêtent pas tous la même importance fonctionnelle pour la biogenèse des snoRNP à boîtes C/D. Parmi eux, les protéines Rsa1p, Pih1p, Rvb1p et Rvb2p ont clairement été caractérisées comme des facteurs d'assemblage, capables d'interagir avec les composants des snoRNP et intervenant physiquement dans la formation des particules. Dans l'état actuel de nos connaissances, la protéine Hit1p semble quant à elle jouer un rôle différent, puisque son importance pourrait être d'avantage liée au contrôle qu'elle exerce sur le facteur Rsa1p. En revanche, dans ce chapitre, nous avons montré que la protéine Tah1p et les chaperonnes HSP90 jouent un rôle secondaire, puisque leur intervention semble se limiter à des opérations de maintenance. Précédemment, l'équipe de W.A. Houry avait observé que l'absence de la protéine Tah1p et la délétion partielle des protéines HSP90 altèrent spécifiquement le taux de certains snoRNA à boîtes C/D durant la phase stationnaire de croissance (Zhao *et al.*, 2008). En dépit, de l'utilisation d'approches d'analyses similaires, il nous a été impossible de reproduire ces résultats, suggérant que les variations initialement observées pourraient être d'une amplitude relativement faible. En comparaison avec des facteurs comme Rsa1p et Pih1p, les chaperonnes HSP90 et la protéine Tah1p, n'exercent donc pas d'influence réelle sur l'accumulation des snoRNP dans la cellule, y compris dans des conditions de carence nutritive. Certaines observations attestent néanmoins de liens entre ces facteurs et les composants des snoRNP. Dans le chapitre précédent nous avons notamment montré que Tah1p est associée de façon spécifique avec les pré-snoRNA U3 au cours de la phase stationnaire de croissance (Rothé *et al.*, Publication n°3). L'importance fonctionnelle de cette association n'a pas été déterminée, cependant, il est possible qu'elle contribue à stabiliser la protéine Pih1p présente au sein des pré-snoRNP bloquée dans la voie de biogenèse des particules. D'autre part, l'équipe de W.A. Houry a

également mis en évidence l'association de la chaperonne avec la protéine Nop56p en cas de défaillance de la machinerie d'assemblage (Zhao *et al.*, 2008). En effet, dans un contexte génétique  $\Delta PIH1$ , la quantité de protéines Snu13p et Nop1p co-immunosélectionnées avec Nop56p fait l'objet d'une diminution drastique, résultant vraisemblablement d'une diminution de leur association au sein des snoRNP. En contrepartie, d'importantes quantités de HSP90 sont co-immunosélectionnées avec Nop56p, suggérant que les chaperonnes pourraient intervenir dans la prise en charge de certaines protéines en cas de saturation du système d'assemblage des RNP. Cette observation demeure cependant paradoxale, puisqu'à l'image des différents facteurs d'assemblage, HSP90 n'exerce aucune influence sur la stabilité des protéines Snu13p, Nop1p, Nop56p et Nop58p ((Zhao *et al.*, 2008) ; données non montrées). Une possibilité est donc que cette prise en charge ne soit pas requise pour assurer la stabilité des protéines, mais plutôt pour les maintenir dans une conformation compatible avec les processus d'assemblage et/ou pour prévenir de leur interaction avec des partenaires illégitimes.

De façon plus générale, il est important de noter que toutes les observations reliant Tah1p, HSP90 et les snoRNP à boîtes C/D, ont exclusivement été réalisées à partir de cellules en phase stationnaire de croissance. Chez la levure en phase exponentielle de croissance, près de 60% de l'activité transcriptionnelle est destinée à la transcription des ARNr et environ 50% de l'activité de l'ARN polymérase II est dédiée à l'expression de protéines ribosomiques (Warner, 1999). L'entrée en phase stationnaire, qui résulte de l'épuisement en éléments nutritifs et de l'acidification du milieu, conduit en revanche les levures à adopter un métabolisme ralenti. Leur multiplication est stoppée par arrêt du cycle cellulaire en phase G1, leurs parois sont modifiées et la majorité des fonctions cellulaires sont mises en état de veille. Au cours de cette phase, l'activité transcriptionnelle chute d'environ 90% (Oliver & McLaughlin, 1977; Waldron *et al.*, 1977; Warner, 1999) et la quantité de ribosomes est réduite à moins de 25% sous l'effet d'un phénomène de dégradation par autophagie (Kraft *et al.*, 2008; Warner, 1999). Le processus de biogenèse des ribosomes, qui fait l'objet d'une coordination étroite avec les besoins en protéines et avec la croissance des cellules, est également inhibé par le biais de régulations transcriptionnelles impliquant notamment la voie TOR (Lempiainen & Shore, 2009; Mayer & Grummt, 2006; Mayer *et al.*, 2004). A l'inverse de cette tendance, parmi les rares protéines encore synthétisées, certaines protéines chaperonnes, comme Hsp82p, font l'objet d'une surexpression visant à maintenir le taux et la conformation de nombreuses protéines clientes (Donalies & Stahl, 2001).

Dans un tel contexte, il est donc évident que la mesure de paramètres relatifs aux composants des snoRNP à boîtes C/D ne renseigne pas sur la biogenèse active des particules. De façon intéressante, la détection de précurseurs du snoRNA U3 au cours de cette phase (Rothé *et al.*, Publication n°3), suggère que les pré-snoRNP ne sont pas dissociées, mais qu'elles pourraient être préservées au sein du noyau dans l'attente de conditions de croissance plus favorables. En effet, la phase stationnaire correspond à un état réversible et des cellules remises en culture dans un nouveau milieu nutritif sont capables de reprendre une croissance normale en seulement quelques heures. Il est donc envisageable que les pré-snoRNP et leurs constituants fassent l'objet d'un processus de maintenance, visant à permettre une remobilisation rapide en cas de reprise du processus de biogenèse des ribosomes. L'intervention de Tah1p et des protéines HSP90 dans ce cadre est donc tout à fait envisageable et correspond à l'un des principaux attributs fonctionnels des chaperonnes. Plutôt que de jouer un rôle dans l'assemblage des RNP à proprement parler, il est en effet possible que Tah1p et HSP90 permettent de préserver l'intégrité des acteurs du processus de biogenèse en cas d'arrêt métabolique prolongé. Afin de vérifier cette hypothèse et de quantifier l'éventuel impact de ces facteurs, il serait particulièrement intéressant d'étudier les capacités de reprise de croissance des différentes souches de levure d'intérêt. En effet, si l'absence de Tah1p ou l'expression d'un variant d'interaction avec HSP90 altère ce processus de maintenance, il est possible que des cellules en phase stationnaire présentent plus de difficultés à reprendre une croissance rapide après renouvellement du milieu nutritif.

Les données présentées dans cette partie indiquent donc que le mécanisme de biogenèse des snoRNP à boîtes C/D pourrait faire l'objet de variations majeures chez les eucaryotes. Les résultats obtenus chez l'homme ont en effet révélé le rôle prépondérant joué par les protéines HSP90 au sein de la machinerie d'assemblage, puisque ces dernières sont requises pour l'accumulation de certains constituants essentiels, comme les protéines 15.5K et NOP58 (Boulon *et al.*, 2008) (**Tableau 5**). Au regard de la diversification fonctionnelle de la machinerie d'assemblage entre la levure et l'homme, il est donc possible que la fonction du duo HSP90-Tah1p ait évolué différemment en fonction des *phyla*, pour acquérir chez certains un rôle à part entière dans le mécanisme d'assemblage des RNP.



## CHAPITRE 5

### Etude de la biogenèse de la snoRNP U3

#### 1. Présentation des travaux

De part son caractère essentiel, son abondance et sa conservation chez tous les eucaryotes, la snoRNP U3 a souvent constitué un modèle de choix pour l'étude des snoRNP. Comme nous l'avons abordé en détail dans l'introduction, cette particule présente cependant de nombreuses particularités qui en font un membre singulier de la classe des snoRNP à boîtes C/D (cf. Introduction, Chapitre 2, 3.1.2). En termes de fonction tout d'abord, puisqu'elle est à l'origine de la formation du SSU-processome et qu'elle est requise pour les clivages précoces du pré-ARNr. En termes de structure également, puisqu'à la différence des autres snoRNA à boîtes C/D, le snoRNA U3 est un ARN hautement structuré, qui possède un motif C/D dégénéré appelé motif C'/D, ainsi qu'un motif en *K-turn* spécifique appelé motif B/C (cf. Introduction, Chapitre 2, 1.2.4) (**Fig. 23**). Enfin, sa composition est également particulière, puisqu'en plus des protéines Snu13p/15.5K, Nop1p/Fibrillarine, Nop56p/NOP56 et Nop58p/NOP58, elle contient la protéine Rrp9p/U3-55K, qui est recrutée spécifiquement au niveau du motif B/C et qui est requise pour la fonction de la particule. Les motifs C'/D et B/C constituent ainsi les centres de nucléation de deux RNP, composées respectivement des jeux de protéines Snu13p/15.5K, Nop58p/NOP58, Nop1p/Fibrillarine et Snu13p/15.5K, Rrp9p/U3-55K, Nop56p/NOP56, Nop1p/Fibrillarine (cf. Introduction, Chapitre 2, 3.1.2.3) (**Fig. 41**).

Depuis de nombreuses années, le laboratoire de C. Branlant étudie la snoRNP U3 sous l'angle de la relation structure-fonction. Ces travaux ont largement contribué à la caractérisation des déterminants ARN requis pour l'assemblage de cette snoRNP et pour l'acquisition de la fonctionnalité (cf. Introduction, Chapitre 2, 3.1.2). En revanche, de nombreuses questions restent encore posées au sujet des mécanismes cellulaires régissant la biogenèse de la snoRNP U3. Précédemment, nous avons montré que ce processus implique les facteurs de la machinerie d'assemblage, puisque ces derniers ont été détectés en association avec les précurseurs du snoRNA U3 chez la levure et chez l'homme ((Boulon *et al.*, 2008; Watkins *et al.*, 2004) ; Rothé *et al.*, Publication n°3). Cependant, l'accumulation du snoRNA U3 s'est révélée relativement insensible à l'absence de certains facteurs majeurs comme Rsa1p ou

Pih1p, suggérant que la production de la snoRNP U3 pourrait être peu dépendante de la machinerie d'assemblage (Rothé *et al.*, Publication n°3). A ce titre, des travaux récents de l'équipe de N.J. Watkins, ont permis de montrer que l'incorporation de la protéine U3-55K au sein de la RNP U3B/C du snoRNA U3 humain, contribue à l'assemblage subséquent de la protéine NOP58 au niveau du motif C'/D (Knox *et al.*, 2011) (cf. Introduction, Chapitre 2, 3.1.2.3). Cette propriété intrinsèque pourrait donc contribuer à assurer une forte production de la snoRNP U3 en dépit des défaillances de la machinerie d'assemblage.

Afin de mieux comprendre le rôle joué par les facteurs d'assemblage dans la biogenèse de la snoRNP U3, nous avons donc mis en place diverses approches *in vitro* afin d'étudier de façon distincte les propriétés des RNP B/C et C'/D. De façon intéressante, les résultats obtenus indiquent que l'assemblage de la protéine Rrp9p sur le motif B/C pourrait être indépendant des facteurs d'assemblage. En revanche, son recrutement sur le motif B/C ne semble pas interférer avec celui de la protéine Rsa1p, qui pourrait donc intervenir au niveau de cette RNP spécifique de la snoRNP U3. D'autre part, nous avons également montré que la protéine Rsa1p améliore la fixation de la protéine Snu13p au niveau du motif sous-optimal C'/D. Cette propriété pourrait reposer sur la capacité du facteur d'assemblage à contacter physiquement la partie 5' de l'ARN et ainsi stabiliser les pré-snoRNP. L'ensemble de ces données suggère donc que Rsa1p pourrait être recrutée à la fois au niveau du motif B/C et du motif C'/D, afin de faciliter l'assemblage des protéines Nop56p et Nop58p, respectivement sur ces deux motifs. D'autre part, en utilisant une stratégie de RT-PCR permettant la détection spécifique des différents intermédiaires de maturation du snoRNA U3, nous avons entrepris de retracer la chronologie des événements de biogenèse de la snoRNP. Contrairement au modèle proposé par l'équipe de D. Tollervey, nous avons pu mettre en évidence que l'intron, contenu dans le pré-snoRNA U3 de levure, est éliminé avant la maturation de l'extrémité 3'. De plus, nos résultats indiquent que les protéines de la RNP et les facteurs d'assemblage sont présents sur l'ensemble des pré-snoRNA U3, suggérant la concomitance des processus d'assemblage et de maturation de l'ARN. Enfin, l'utilisation de l'ARN U3 $\Delta$ 2,3,4, qui est déficient pour le recrutement de Rrp9p *in vitro*, mais néanmoins fonctionnel *in vivo*, s'est avéré sensible à un défaut d'interaction entre Rsa1p et Snu13p. Cette donnée confirme donc que Rsa1p participe à la biogenèse de la snoRNP, mais que sa contribution pourrait être masquée par l'intervention de Rrp9p.

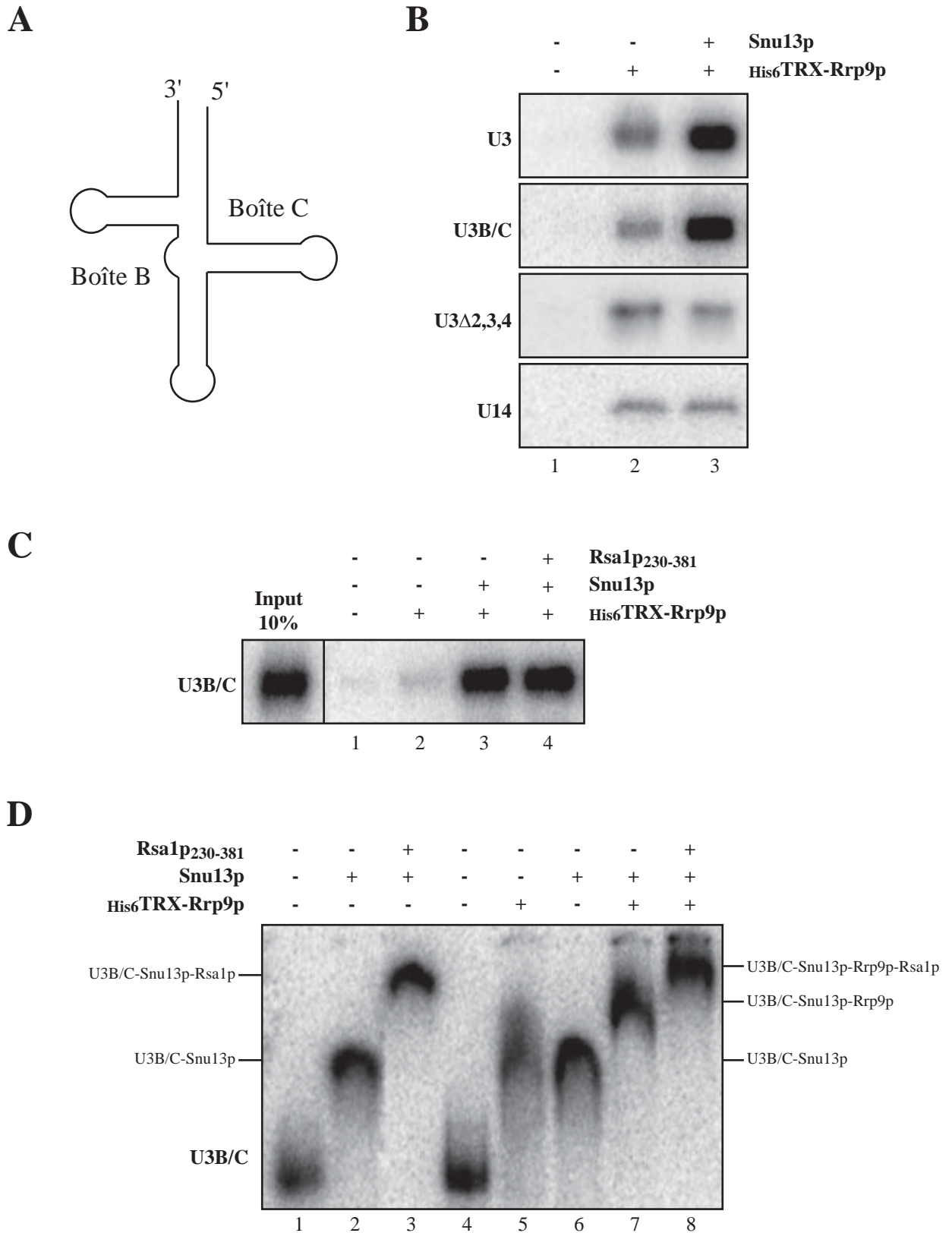
## 2. Résultats

### 2.1. Etude *in vitro* de la RNP U3B/C

Chez l'homme, les expériences de double-hybride *bridged* indiquent que NUFIP facilite l'interaction de la protéine 15.5K avec la protéine U3-55K (Boulon *et al.*, 2008) (cf. Résultats & Discussion, Partie 1, 2.2). Cette donnée suggère donc que le facteur d'assemblage pourrait être également impliqué dans la formation de la particule B/C du snoRNA U3. Afin de tester cette hypothèse chez la levure, nous avons réalisé des expériences d'assemblage *in vitro*, en utilisant la protéine Rrp9p recombinante. Cette dernière, obtenue au laboratoire par M. Estrela et X. Manival, a été produite sous forme soluble grâce à une fusion avec la thioredoxine (TRX), puis purifiée par l'intermédiaire d'une étiquette 6-Histidines.

#### 2.1.1. *In vitro*, l'assemblage de la protéine Rrp9p sur le snoRNA U3 ne requiert pas de facteur d'assemblage

Les résultats préliminaires obtenus par l'approche de double-hybride indiquaient que Rrp9p et Rsa1p n'interagissent pas de façon stable (Boulon *et al.*, 2008) (**Fig. 54**). Cependant, à l'image de Nop56p, il est possible que des contacts restreints entre les protéines de la RNP et le facteur d'assemblage favorisent la nucléation de la particule. Nous avons donc utilisé une approche de pull-down afin de tester la capacité de la protéine  $_{\text{His6}}\text{TRX-Rrp9p}$  à retenir un snoRNA radiomarqué, en présence des protéines recombinantes Snu13p et Rsa1p<sub>230-381</sub> (**Fig. 76**). Quatre ARN ont été utilisés dans ces expériences : (i) le snoRNA U3 entier, (ii) l'ARN U3B/C, tronqué du domaine 5' et du motif C'/D, qui nous a permis d'étudier spécifiquement l'assemblage de la RNP B/C (**Fig. 76A**), (iii) l'ARN U3 $\Delta$ 2,3,4 qui nous a permis de tester l'importance des hélices 2, 3 et 4 pour la fixation de Rrp9p (Marmier-Gourrier *et al.*, 2003), (iv) le snoRNA U14 qui a été utilisé afin de vérifier la spécificité du processus d'assemblage. Après incubation des différents partenaires à 30°C, la protéine  $_{\text{His6}}\text{TRX-Rrp9p}$  a été purifiée sur billes de nickel-sépharose et la rétention d'ARN radiomarqué a été vérifiée sur gel dénaturant (**Fig. 76B**).



**Figure 76 : *In vitro* l'assemblage de la protéine Rrp9p sur le motif B/C du snoRNA U3 est indépendante de la machinerie d'assemblage.** (A) Schéma de l'ARN U3B/C qui constitue l'ARN minimal permettant le recrutement de la protéine Rrp9p. Cet ARN correspond aux nucléotides 95 à 224 du snoRNA U3. Il est dépourvu du domaine 5' et du motif C'D. (B, C) Expériences de pull-down réalisées sur billes de Nickel-séparation à partir de la protéine His6TRX-Rrp9p, en présence d'un snoRNA radiomarqué et des protéines Snu13p et Rsa1p<sub>230-381</sub>. (D) Expériences de retard sur gel réalisées à partir de l'ARN U3BC radiomarqué et des protéines His6TRX-Rrp9p, Snu13p et Rsa1p<sub>230-381</sub>.



Les résultats obtenus indiquent que la protéine  $_{\text{His6}}\text{TRX-Rrp9p}$  interagit de façon faible et non spécifique avec les molécules d'ARN, puisque les différents ARN sont retenus en quantité restreinte par la protéine seule (**Fig. 76B, piste 2**). En revanche, l'ajout de la protéine Snu13p conduit à une rétention accrue et spécifique des ARN U3 et U3B/C par la protéine  $_{\text{His6}}\text{TRX-Rrp9p}$  (**Fig. 76B, piste 3**). De façon cohérente avec les données obtenues *in vivo* par A. Cléry (Cléry *et al.*, 2007b), ce résultat confirme donc l'importance de la protéine Snu13p pour initier l'assemblage de la particule B/C. De plus, il indique que les hélices 2, 3 et 4, qui ne sont pas requises pour la viabilité cellulaire (Samarsky & Fournier, 1998) et pour le recrutement de Rrp9p *in vivo* (Cléry *et al.*, 2007b), sont néanmoins requises pour son recrutement *in vitro*. Des résultats identiques ont été obtenus au laboratoire par N. Rolland et V. Senty-Ségault, grâce à une approche d'assemblage en extrait cellulaire (Travail de thèse N. Rolland). Il est probable que l'apport de ces hélices repose sur l'établissement de contacts avec la protéine Rrp9p, puisque l'équipe de D. Tollervey a montré la formation de pontages covalents entre Rrp9p et les hélices 2 et 4 par l'approche de CRAC (Granneman *et al.*, 2009) (cf. Introduction, Chapitre 2, 3.1.2.3). L'ensemble de ces données indique donc que la fixation optimale de Rrp9p sur le snoRNA U3 requiert la présence des hélices jouxtant le motif B/C, mais que leur absence peut néanmoins être compensée *in vivo*. Il est cependant important de noter que cette compensation n'est pas totale, puisque la substitution du snoRNA U3 endogène par l'ARN U3 $\Delta$ 2,3,4 ne permet pas de restaurer complètement les capacités de croissance de la souche de levure (Marmier-Gourrier *et al.*, 2003).

De façon intéressante, ces données révèlent également la possibilité de réaliser l'assemblage de cette RNP *in vitro*, sans l'assistance des facteurs d'assemblage. Afin de vérifier l'éventuelle influence de la protéine Rsa1p sur ce processus, nous avons reproduit la même expérience en présence du fragment soluble Rsa1p<sub>230-381</sub> (**Fig. 76C**). En dépit de la capacité de ce fragment à interagir avec la protéine Snu13p, aucun effet notable n'a été constaté sur la formation de la particule B/C. Ce résultat indique donc que le fragment fonctionnel de Rsa1p ne permet pas d'améliorer l'assemblage de cette RNP et que sa présence n'est néanmoins pas délétère pour ce processus.

Ces différentes données suggèrent donc que l'assemblage de la protéine Rrp9p de levure sur le snoRNA U3 pourrait s'effectuer indépendamment de la machinerie d'assemblage. A la différence du recrutement des protéines Nop56p et Nop58p sur le snoRNA U14, il semble en effet que le rôle de plateforme joué par Rsa1p ne contribue pas à la fixation de Rrp9p sur le

complexe U3B/C-Snu13p. Une démonstration complète aurait néanmoins nécessité de reproduire ces expériences avec la protéine Rsa1p entière, afin de vérifier l'influence de son domaine N-terminal. En dépit de pouvoir réaliser une telle démarche, la confirmation de cette hypothèse a cependant été apportée au laboratoire grâce à une approche de reconstitution de la particule U3B/C en extrait de cellules de levure. En utilisant ce système à partir d'une souche de levure invalidée pour le gène *RSAL*, N. Rolland et V. Senty-Ségault ont en effet montré que la protéine Rsa1p endogène n'est pas requise pour ce processus (Travail de thèse N. Rolland).

### **2.1.2. *In vitro*, Rsa1p<sub>230-381</sub> est recrutée sur la particule U3 B/C en présence de Rrp9p**

Rsa1p ne semble donc pas influencer sur le recrutement de la protéine Rrp9p sur le snoRNA U3, cependant dans le but de vérifier si les protéines His6TRX-Rrp9p et Rsa1p<sub>230-381</sub> peuvent co-exister au sein des mêmes complexes, nous avons réalisé des expériences de retard sur gel selon les conditions décrites précédemment ((Boulon *et al.*, 2008); Rothé, Back *et al.*, Publication n°2) (**Fig. 76D**). L'observation d'un retard de migration de l'ARN U3B/C radiomarqué en présence de la protéine His6TRX-Rrp9p confirme sa capacité à établir des contacts directs avec l'ARN (piste 5). Lorsque His6TRX-Rrp9p et Snu13p sont co-incubées avec l'ARN, l'apparition d'une bande de migration électrophorétique plus lente (*super-shift*) indique la formation du complexe U3B/C-Snu13p-Rrp9p précédemment étudié par l'approche de pull-down (piste 7). Enfin, l'ajout de la protéine Rsa1p<sub>230-381</sub> conduit à un retard de migration supplémentaire, légèrement supérieur à celui observé pour le complexe trimérique U3B/C-Snu13p-Rsa1p<sub>230-381</sub> (piste 8), suggérant que les protéines His6TRX-Rrp9p et Rsa1p<sub>230-381</sub> peuvent être incorporées au sein d'une même RNP. Ce résultat suggère donc que Rsa1p pourrait interagir avec la protéine Snu13p au sein de la RNP U3B/C, indépendamment de la présence de Rrp9p.

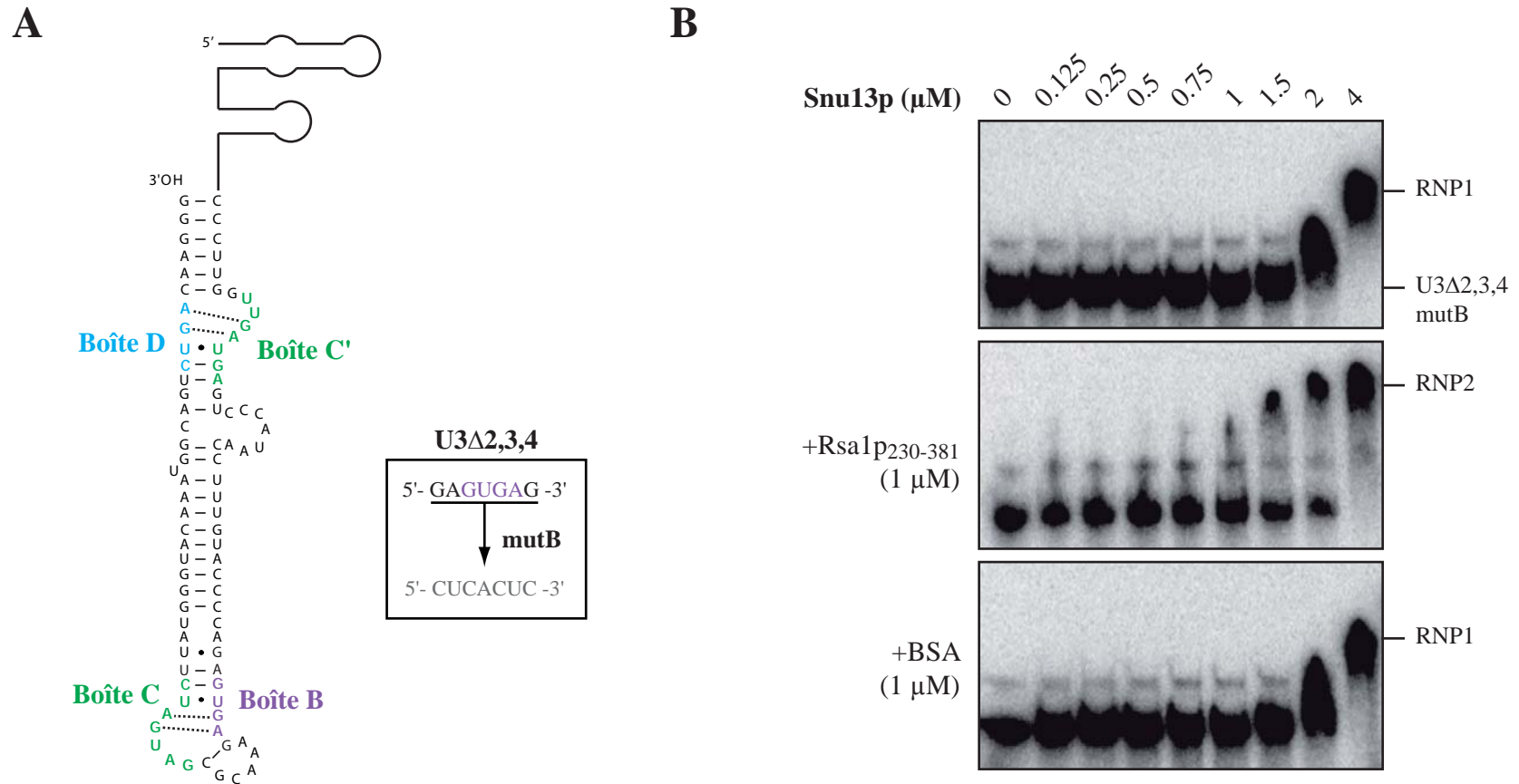
## **2.2. Etude *in vitro* de la RNP U3C'/D**

En dépit de sa singularité, l'appartenance du snoRNA U3 à la classe des snoRNA à boîtes C/D repose sur la présence d'un motif C'/D correspondant à un motif C/D non consensusel

(cf. Introduction, Chapitre 2, 1.2.4) (**Fig. 23**). Comme nous l'avons abordé dans l'introduction, le motif C'/D, qui constitue un site de faible affinité pour la protéine Snu13p/15.5K *in vitro*, est néanmoins parfaitement fonctionnel *in vivo* et s'avère requis pour le recrutement des protéines constitutives de la RNP (cf. Introduction, Chapitre 2, 3.1.2.3) (Knox *et al.*, 2011; Marmier-Gourrier *et al.*, 2003) (Travail de thèse N. Rolland). Ce constat paradoxal suggère donc que certains partenaires pourraient stabiliser la protéine Snu13p sur ce motif *in vivo*. Récemment, l'équipe de N.J. Watkins a montré que l'incorporation de la protéine humaine U3-55K au niveau du motif B/C compense ce défaut intrinsèque, en favorisant l'assemblage de la protéine NOP58 au niveau du motif C'/D (Knox *et al.*, 2011). Etant donné la capacité de Rsa1p à interagir avec Snu13p et à favoriser la nucléation des particules, nous avons cherché à évaluer son influence sur le recrutement de Snu13p au niveau du motif C'/D du snoRNA U3. Pour cela nous avons utilisé les approches de retard sur gel et d'empreintes aux RNases.

### **2.2.1. L'efficacité de fixation de Snu13p sur le motif U3C'/D est influencée par Rsa1p *in vitro***

Dans le but d'étudier spécifiquement le recrutement de la protéine Snu13p sur le motif C'/D du snoRNA U3, nous avons utilisé une forme de l'ARN U3 $\Delta$ 2,3,4 mutée au niveau de la boîte B (Marmier-Gourrier *et al.*, 2003) (**Fig. 77A**). Cet ARN, qui ne possède plus qu'un seul motif structuré en *K-turn*, a ainsi été incubé en présence d'une gamme de concentration croissante en protéine Snu13p recombinante et la formation de complexes ARN-protéine a été vérifiée par électrophorèse en gel natif (**Fig. 77B**). Comme cela a été démontré au laboratoire, l'affinité intrinsèque de la protéine Snu13p pour cet ARN est relativement faible, puisque le Kd apparent observé pour cette interaction est compris entre 2 et 4  $\mu$ M en protéine Snu13p. Afin de savoir si Rsa1p exerce une influence sur la formation de ce complexe, nous avons reproduit cette expérience en présence de la protéine Rsa1p<sub>230-381</sub> ajoutée à la concentration fixe de 1  $\mu$ M. Comme attendu, la présence de Rsa1p conduit à un retard de migration supplémentaire, qui indique la formation d'un complexe de haut poids moléculaire, contenant l'ARN, la protéine Snu13p et le fragment Rsa1p<sub>230-381</sub>. De plus, nous avons également constaté que la présence de Rsa1p améliore de façon significative la formation de la RNP, puisque le Kd apparent de cette interaction est estimé entre 1,5 et 2  $\mu$ M en protéine Snu13p.



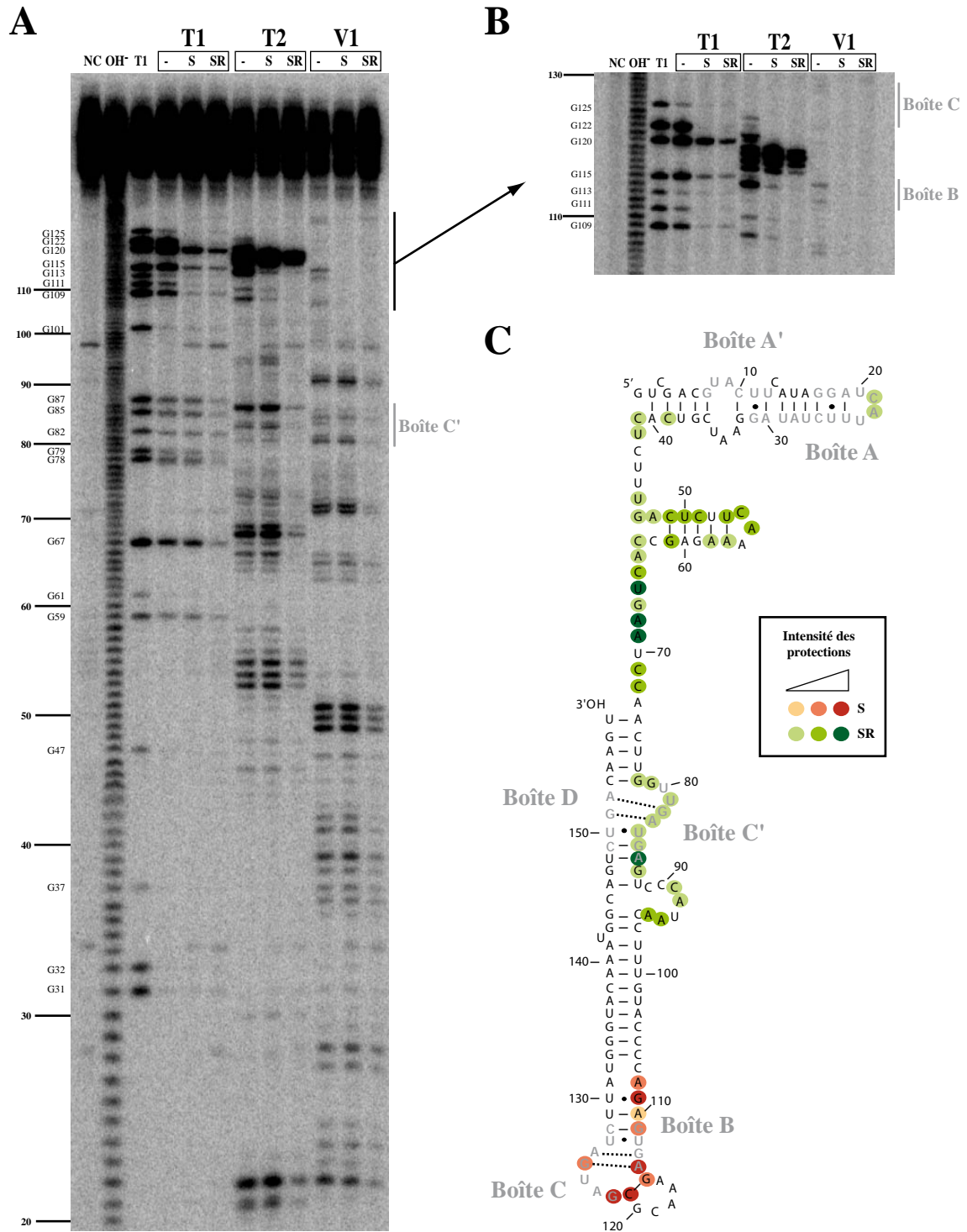
**Figure 77 : Rsa1p améliore la fixation de la protéine Snu13p sur le motif C'/D du snoRNA U3.** (A) ARN mini U3Δ2,3,4 mutB utilisé dans les expériences de retard sur gel. La boîte B a été mutée afin d'annihiler la formation du motif en K-turn B/C (Marmier-Gourrier *et al.*, 2003). Seul le motif C'/D assure le recrutement de la protéine Snu13p. (B) Expériences de retard sur gel réalisées à partir de l'ARN mini U3Δ2,3,4 mutB radiomarqué au niveau de son extrémité 5'. La protéine recombinante Snu13p a été ajoutée en concentration croissante. Les mêmes tests ont été réalisés en présence d'une concentration fixe (1 μM) du complexe de la protéine Rsa1p<sub>230-381</sub> ou de BSA. RNP1 : complexe ARN-Snu13p; RNP2 : complexe ARN-Snu13p-Rsa1p<sub>230-381</sub>.

Enfin, dans le but de confirmer la spécificité de cet effet nous avons reproduit la même expérience en remplaçant le fragment Rsa1p<sub>230-381</sub> par de la BSA acétylée en concentration équivalente. Le profil de migration électrophorétique obtenu est strictement identique à celui observée pour la protéine Snu13p seule, suggérant que l'effet positif de Rsa1p n'est pas dû à un simple effet de concentration, mais bien à une action spécifique du facteur d'assemblage.

### 2.2.2. Le domaine 5' du snoRNA U3 est important pour le recrutement de Rsa1p sur le motif U3 C'/D

D'après les résultats obtenus par l'approche de retard sur gel, il semble que Rsa1p contribue à stabiliser la protéine Snu13p sur le motif C'/D du snoRNA U3. Afin de savoir si cet effet repose sur des contacts physiques entre le facteur d'assemblage et l'ARN, nous avons réalisé des expériences d'empreintes aux RNases (**Fig. 78**). Pour cela l'ARN U3 $\Delta$ 2,3,4 radiomarqué à son extrémité 5' a été soumis à une digestion ménagée par les ribonucléases T1, T2 et V1, et les produits de clivage ont été analysés sur gel dénaturant (cf. Matériel et Méthodes). La ribonucléase T1 hydrolyse les liaisons phosphodiester en aval des résidus G non appariés, la ribonucléase T2 présente également une spécificité pour les régions en simple brin, avec une préférence pour les résidus A, enfin, la ribonucléase V1 agit sur les régions en double-brin. Ces différents traitements ont été réalisés sur l'ARN nu et sur les complexes ARN-Snu13p et ARN-Snu13p- Rsa1p<sub>230-381</sub> préformés. Dans ces expériences, la protéine Snu13p a été utilisée à la concentration limite de 1,5  $\mu$ M, afin de pouvoir distinguer l'effet positif de Rsa1p de la fixation naturelle de Snu13p (cf. section précédente). D'autre part, le motif B/C de l'ARN U3 $\Delta$ 2,3,4, pour lequel la protéine présente une très forte affinité ( $K_d < 75$  nM), a été conservé sous sa forme sauvage afin de disposer d'un contrôle positif d'interaction. Le profil de migration a été annoté grâce à la réalisation d'une échelle par hydrolyse alcaline, révélant l'ensemble des nucléotides, et grâce à une hydrolyse à la ribonucléase T1 en conditions dénaturantes, permettant d'identifier tous les résidus G.

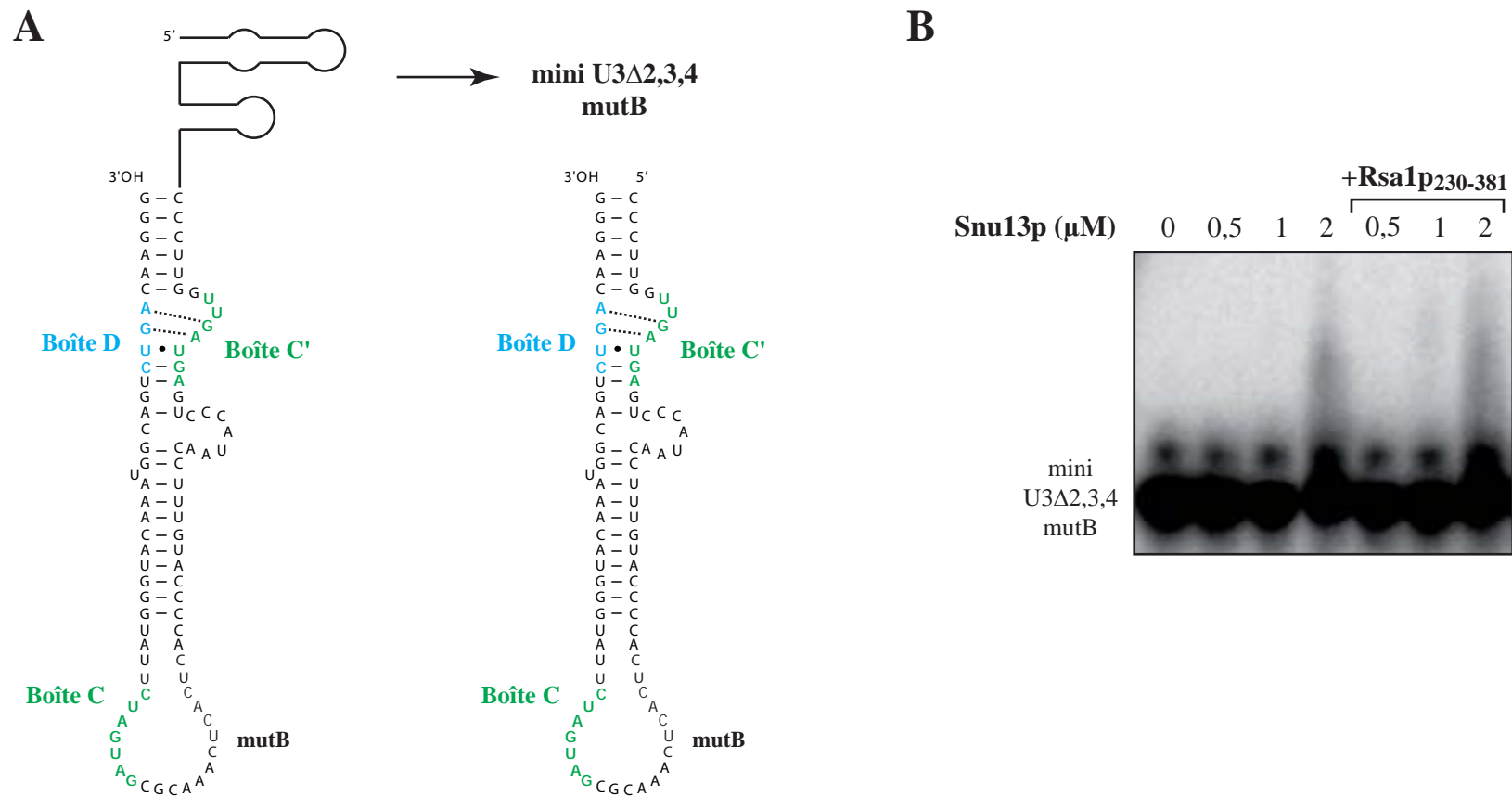
Conformément à sa grande affinité pour le motif B/C, la pré-incubation de l'ARN avec la protéine Snu13p, conduit à la disparition de nombreuses coupures aux RNases T1 et T2 au niveau des boîtes B et C de l'ARN (**Fig. 78**). Cette modification du profil d'hydrolyse constitue une marque caractéristique qui révèle la fixation de la protéine Snu13p sur cette



**Figure 78 : Le profil de digestion enzymatique de l'ARN U3 $\Delta$ 2,3,4 est modifié en présence de Rsa1p.** (A) Analyse du profil de digestion de l'ARN U3 $\Delta$ 2,3,4 par les enzymes RNases T1, T2 et V1, en présence de la protéine Snu13p seule (S) ou co-incubée avec la protéine Rsa1p<sub>230-381</sub> (SR). L'ARN est visualisé grâce à un marquage radioactif au niveau de l'extrémité 5'. Les protéines recombinantes ont été incubées à une concentration finale de 1,5  $\mu$ M. NC : Non Coupé (ARN incubé avec les protéines recombinantes sans ajout de RNases); OH<sup>-</sup> : échelle obtenue par hydrolyse alcaline; T1 : échelle des résidus G obtenue par digestion à la RNase T1 en conditions dénaturantes. (B) Vue détaillée de la région correspondant au motif B/C. (C) Localisation dans le modèle de structure 2D de l'ARN, des nucléotides protégés en présence de Snu13p (en rouge) ou en présence des protéines Snu13p et Rsa1p<sub>230-381</sub> (en vert). L'intensité de l'effet protecteur est symbolisé par l'intensité de la coloration (cf. encadré).

région et sa protection (Clery *et al.*, 2007b; Marmier-Gourrier *et al.*, 2003). Inversement, l'analyse des coupures situées dans la région de la boîte C' ne révèle aucune protection induite par la présence de Snu13p, suggérant l'absence d'interaction stable entre la protéine et le motif C'/D dans ces conditions de concentration. De façon intéressante, l'ajout de la protéine Rsa1p<sub>230-381</sub> modifie totalement ce profil, puisque de nombreuses coupures disparaissent au niveau de la boîte C' et dans les régions proximales, tandis qu'aucun effet notable n'est constaté au niveau des boîtes B et C. Au regard des résultats obtenus par la technique de retard sur gel, ces observations coïncident avec la fixation de la protéine Snu13p sur le motif C'/D de l'ARN. De plus, la disparition de nombreuses coupures au niveau de la région charnière et du domaine 5', qui ne sont pas des zones naturellement protégées par Snu13p, résulte vraisemblablement de la présence de la protéine Rsa1p<sub>230-381</sub>. Ces changements pourraient donc découler soit d'une obstruction des sites de coupures par la protéine Rsa1p<sub>230-381</sub>, soit d'une modification de la structure de l'ARN au sein du complexe.

Une hypothèse envisageable est que Rsa1p stabilise la protéine Snu13p sur le motif C'/D en établissant elle-même des contacts avec le snoRNA, au niveau de sa région charnière et de son domaine 5'. Afin de tester cette éventualité, nous avons réalisé de nouvelles expériences de retard sur gel à partir de l'ARN U3Δ2,3,4 muté au niveau de la boîte B et délété de sa portion 5' (**Fig. 79A**). Cet ARN a été incubé avec la protéine Snu13p recombinante à des concentrations de 0.5, 1 et 2 μM, en absence ou en présence de Rsa1p<sub>230-381</sub>. Alors qu'à une concentration de 2 μM en protéine Snu13p, la présence de Rsa1p<sub>230-381</sub> conduisait à un net retard de migration de l'ARN U3Δ2,3,4 mutB (**Fig. 78B**), aucun retard de migration n'est observé pour la version tronquée de l'ARN (**Fig. 79B**). Ce résultat indique donc que l'effet stabilisateur de Rsa1p est dépendant de cette région de l'ARN, avec laquelle le facteur d'assemblage pourrait établir des contacts physiques. Contrairement à la protéine Snu13p, Rsa1p ne contient aucun domaine d'interaction avec l'ARN connu, mais sa région C-terminale (230-381) présente un fort caractère basique (pI = 9,45) qui pourrait lui permettre d'établir des interactions avec les groupements phosphate de l'ARN. Ce résultat suggère donc que, par le biais de contacts aspécifiques avec la partie 5' du snoRNA U3, le facteur d'assemblage Rsa1p pourrait contribuer à stabiliser Snu13p sur le motif C'/D au cours des étapes d'assemblage. Afin de poursuivre cette étude, l'utilisation de l'approche de CRAC, mise au point par l'équipe de D. Tollervey (Granneman *et al.*, 2009), constitue une perspective intéressante qui pourrait nous permettre de confirmer ces données et de préciser les régions d'ARN potentiellement contactées par le facteur d'assemblage.



**Figure 79 : Le domaine 5' de l'ARN U3 $\Delta$ 2,3,4 est requis pour le recrutement des protéines Snu13p et Rsa1p<sub>230-381</sub> au niveau du motif C'/D.**  
**(A)** Représentation schématique de l'ARN mini U3 $\Delta$ 2,3,4 mutB dans lequel le domaine 5' et la région charnière ont été éliminés. **(B)** Expériences de retard sur gel réalisées à partir de l'ARN mini U3 $\Delta$ 2,3,4 mutB radiomarqué au niveau de son extrémité 5'. La protéine recombinante Snu13p a été ajoutée en concentration croissante. Le même test a été réalisé en présence d'une concentration fixe (1 $\mu$ M) de la protéine Rsa1p<sub>230-381</sub>.



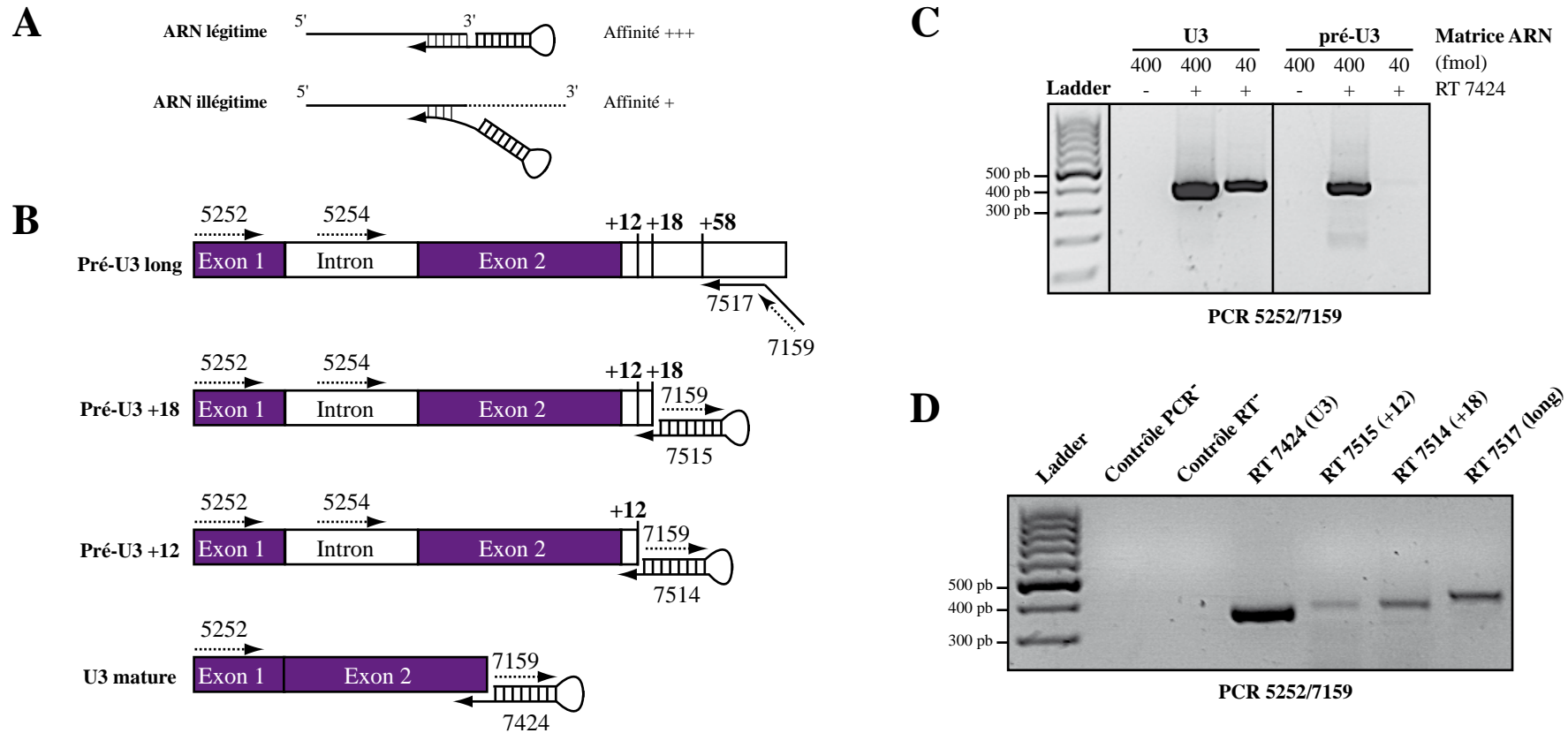
### 2.3. Etude de la voie de maturation du snoRNA U3

Outre les particularités liées à sa structure, à sa composition et à sa fonction, la snoRNP U3 présente chez la levure *S. cerevisiae* une autre caractéristique remarquable, puisque le précurseur du snoRNA contient un intron épissé par la machinerie spliceosomale (Brule *et al.*, 1995; Brule *et al.*, 1996; Fournier *et al.*, 1998; Myslinski *et al.*, 1990). La maturation du snoRNA U3 est donc un processus complexe qui implique la maturation nucléolytique de son extrémité 3', l'élimination de l'intron et l'hyperméthylation de sa coiffe (cf. Introduction, Chapitre 3, 3.4.1). La caractérisation de cette voie de maturation a partiellement été réalisée et un modèle, que nous avons détaillé dans l'introduction, a été proposé par l'équipe de D. Tollervey (Kufel *et al.*, 2000) (**Fig. 50**). De façon surprenante, ce modèle suggère que l'épissage constitue une étape tardive qui pourrait s'effectuer sur les précurseurs porteurs d'extensions de +12 et +18 nucléotides au niveau de leur extrémité 3' (*i.e.* pré-U3 +12 et +18), et qui pourrait précéder l'adressage vers les NB. De plus, il a également été proposé que l'assemblage de la RNP s'effectue de façon tardive, sous l'effet d'une compétition entre les protéines constitutives de la snoRNP et les protéines Lsm et Lhp1p, qui interagissent avec les pré-U3 +12 et +18 (cf. Introduction, Chapitre 3, 3.4.1). Or, différentes observations réalisées au cours de notre étude, nous amènent à contester ce modèle. Tout d'abord, la détection par RT-PCR de pré-snoRNA U3 non maturés en 3' (portant une extension supérieure à 58 nucléotides), nous a permis d'identifier des espèces épissées et non épissées, indiquant que l'étape d'épissage pourrait s'effectuer avant la maturation nucléolytique de l'extrémité 3' (Rothé *et al.*, Publication n°3). D'autre part, les principaux facteurs d'assemblage et les protéines Nop1p, Nop56p et Nop58p ont été détectées sur ces différents précurseurs non maturés en 3', suggérant que l'assemblage de la RNP pourrait également débiter de façon précoce (Rothé *et al.*, Publication n°3). Afin, de clarifier la chronologie de ces événements, nous avons donc mis au point une approche de RT-PCR permettant la détection spécifique des principaux intermédiaires de maturation du snoRNA U3. Cette démarche nous a notamment conduit à proposer un nouveau modèle de biogenèse pour la snoRNP U3.

### 2.3.1. L'épissage du snoRNA U3 est un processus précoce

La mise à jour de la chronologie des événements de maturation du snoRNA U3 implique la détection spécifique des différents intermédiaires produits lors de ce processus. L'utilisation de l'approche de RT-PCR constitue un outil intéressant qui permet de cloner et de séquencer les différentes espèces amplifiées. Précédemment, cette technique a été utilisée afin de détecter les précurseurs du snoRNA U3 non maturés en 3', à l'aide d'un oligonuclotide amorce s'hybridant dans l'extension 3' de l'ARN. Cependant, la détection des intermédiaires successifs (*i.e.* +58, +21, +18, +12), par l'utilisation d'amorces de transcription inverse classiques, s'hybridant dans l'extension 3' de ces molécules, ne permet pas de les distinguer des espèces les précédant dans la voie de maturation. Afin de cibler de façon sélective chaque intermédiaire, nous avons donc eu recours à l'utilisation d'oligonucléotides possédant une structure en tige-boucle interne qui favorise l'hybridation de l'amorce au niveau d'extrémités 3' définies (Chen *et al.*, 2005a) (**Fig. 80A**). Des amorces ont ainsi été sélectionnées pour la détection spécifique des extrémités 3' matures, +12 et +18 (**Fig. 80B**). Leur partie 3' est complémentaire avec l'extrémité des différentes espèces d'ARN sur une longueur de 6 à 8 nucléotides, tandis que la partie 5' de l'oligonucléotide forme une tige-boucle contraignant les amorces à s'hybrider préférentiellement sur une séquence finie. L'amplification des ADNc par PCR a ensuite été réalisée à l'aide d'une amorce 5' s'hybridant soit dans l'exon 1 (amorce 5252), soit dans l'intron (amorce 5254), et grâce à une amorce 3' qui correspond à la portion 5' de l'amorce de transcription inverse, dont la séquence est exogène au génome de la levure (amorce 7159).

Dans un premier temps, afin de vérifier la sélectivité de ce système, nous avons réalisé une RT-PCR à partir d'ARN transcrits *in vitro*, en utilisant l'amorce de transcription inverse 7424, spécifique de l'extrémité 3' du snoRNA U3 mature (**Fig. 80C**). Cette réaction a été réalisée à partir de transcrits correspondant au snoRNA U3 mature, mais aussi au précurseur non maturé en 3' (pré-U3 long). Les amorces 5252 et 7159 ont été utilisées pour l'étape de PCR. En utilisant 400 fmol d'ARN matrice, un produit d'amplification est détecté pour les deux ARN, indiquant que l'amorce en tige-boucle 7424 peut s'hybrider sur l'ARN U3 mature, mais également sur l'ARN illégitime. En revanche, lorsque la réaction est réalisée à partir de 40 fmol d'ARN, seule la forme U3 mature fait l'objet d'une amplification. Ce résultat indique



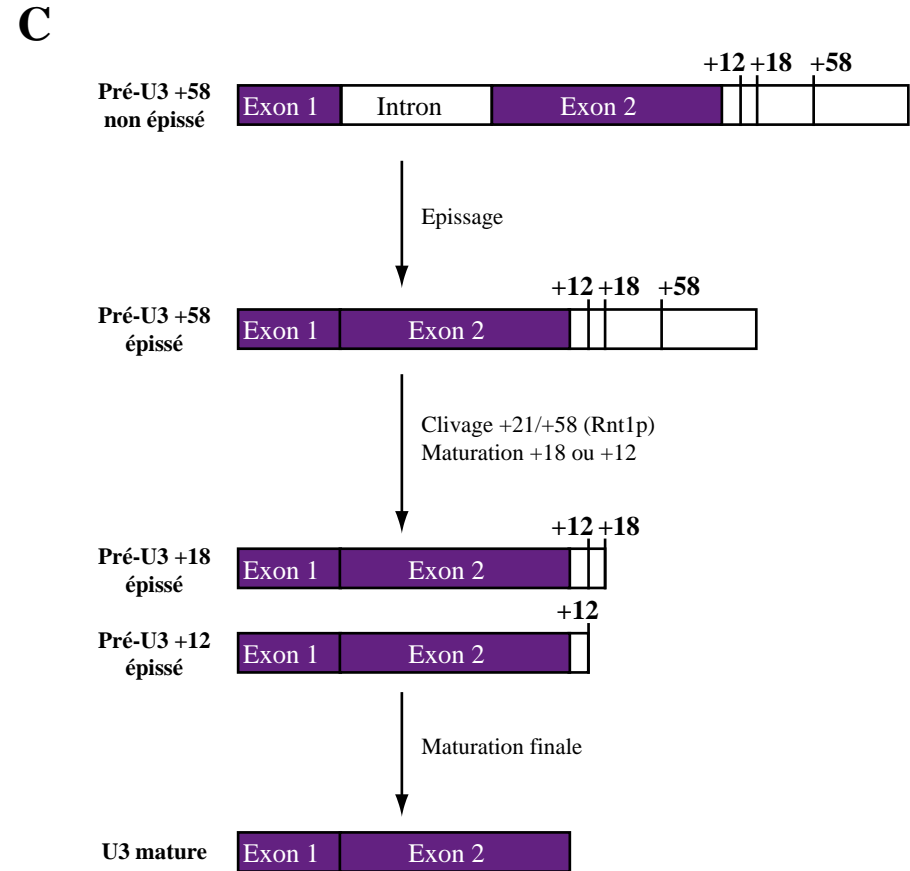
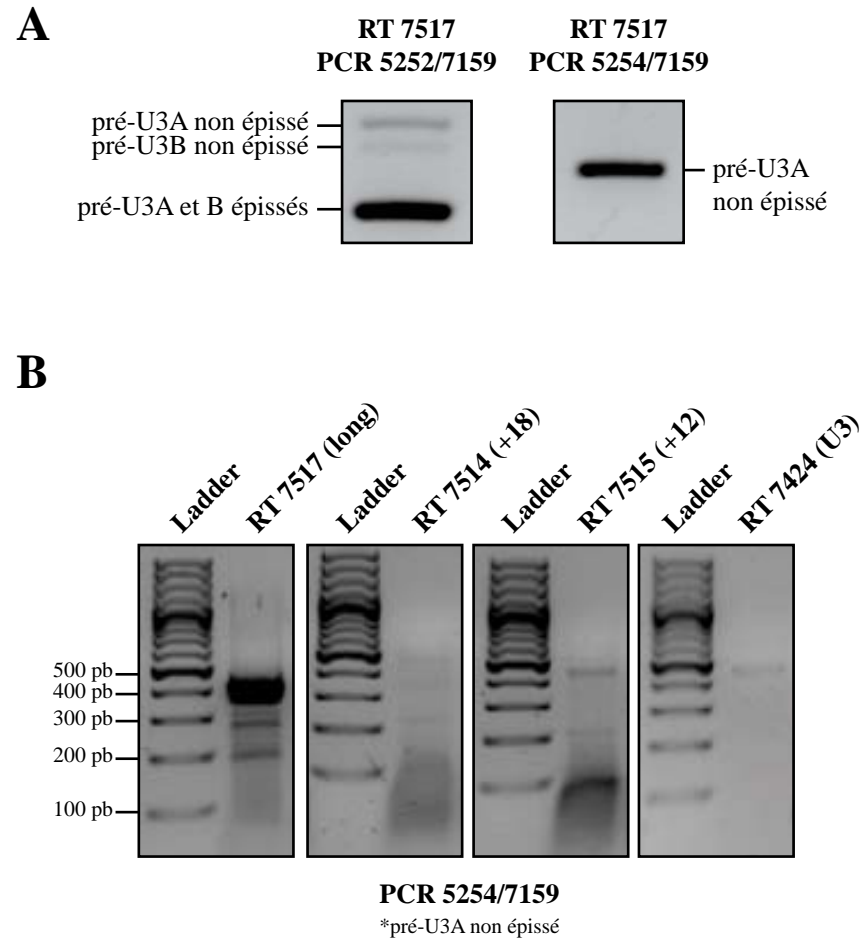
**Figure 80 : Mise au point d'une technique de détection des précurseurs du snoRNA U3 par RT-PCR.** (A) Principe de l'amorçage de la transcription inverse à partir d'oligonucléotides en tige-boucle. (B) Stratégie de détection des différentes formes de snoRNA U3 chez *S. cerevisiae*. Les oligonucléotides amorces 7517, 7515, 7514 et 7424 sont respectivement utilisés pour la transcription inverse des ARN pré-U3 long, pré-U3 +18, pré-U3 +12 et U3 mature. Certains d'entre eux contiennent une structure en tige-boucle qui favorise leur hybridation sélective au niveau d'extrémités franches. Les ADNc obtenus après transcription inverse sont analysés par PCR à l'aide des couples d'oligonucléotides amorces 5252/7159 ou 5254/7159. L'oligonucléotide 7159 équivaut à la portion 5' des différentes amorces de transcription inverse. (C) Visualisation sur gel d'agarose 2,25% des produits de RT-PCR obtenus à partir d'ARN U3 mature ou de pré-U3 transcrits *in vitro* en utilisant l'amorce de transcription inverse 7424. Afin d'évaluer la spécificité de cette amorce, différentes quantités d'ARN matrice légitime (U3 mature) et illégitime (pré-U3) ont été testées. Pour chaque expérience plusieurs contrôles ont été réalisés. (D) Visualisation sur gel d'agarose 2,25% des produits de RT-PCR obtenus à partir d'un extrait d'ARN totaux et correspondant aux principales formes de snoRNA U3. Les ARN ont été obtenus à partir d'une souche de levure WT. Les réactions de transcription inverse (RT : Reverse Transcription) ont été réalisées à partir des amorces indiquées sur la figure. La PCR a été réalisée à l'aide du couple d'oligonucléotides amorces 5252/7159. Le contrôle PCR<sup>-</sup> correspond à une réaction sans matrice. Le contrôle RT<sup>-</sup> correspond à une réaction de transcription inverse sans amorce.

donc que le système de détection choisi est sélectif lorsque l'ARN cible est en faible concentration.

Etant donné la faible abondance des précurseurs du snoRNA U3 *in vivo*, nous avons donc privilégié l'emploi de cette approche à partir d'ARN totaux extraits de la souche de levure BY4741. La détection des différentes espèces de snoRNA U3 a été réalisée à l'aide des amorces de transcription inverse présentées dans la **Figure 80B**. Dans tous les cas, l'utilisation du couple d'oligonucléotides amorces 5252/7159 pour l'étape de PCR conduit à l'amplification d'une bande majoritaire, de taille équivalente à des formes épissées (**Fig. 80D**). Cependant, comme nous l'avons observé précédemment (Rothé *et al.*, Publication n°3), la détection des pré-snoRNA U3 non maturés en 3' (pré-U3 long ciblés par l'amorce 7517) permet, lorsque la réaction de RT-PCR est réalisée à partir d'une quantité suffisante d'ARN totaux, de détecter deux fines bandes de plus haut poids moléculaire (**Fig. 81A**). Ces deux bandes, qui ont été clonées et séquencées, correspondent aux formes non épissées des pré-snoRNA U3A et U3B. L'utilisation de l'amorce 5' de PCR 5254, spécifique de l'intron du gène U3A, permet d'amplifier ce signal et de détecter spécifiquement les ARN non épissés dérivant de l'expression majoritaire du gène U3A. Ce résultat indique que l'épissage de l'intron pourrait se dérouler avant la maturation nucléolytique de l'extrémité 3'. Afin de confirmer cette hypothèse, nous avons cherché à vérifier, pour les différents intermédiaires de maturation, l'absence ou la présence de la séquence intronique en utilisant l'amorce 5' de PCR 5254 (**Fig. 81B**). Comme attendu, l'amorce 5254 permet de détecter un produit correspondant aux pré-snoRNA U3 longs non épissés, en revanche, ni le snoRNA U3 mature, ni les pré-snoRNA U3 +12 et +18, ne semblent contenir l'intron. Ce résultat indique donc que l'épissage s'effectue avant la maturation nucléolytique de l'extrémité 3' et qu'il constitue donc bien une étape précoce dans la voie de maturation de l'ARN (**Fig. 81C**).

### 2.3.2. Rsa1p est associée aux pré-snoRNA U3 +12 et +18

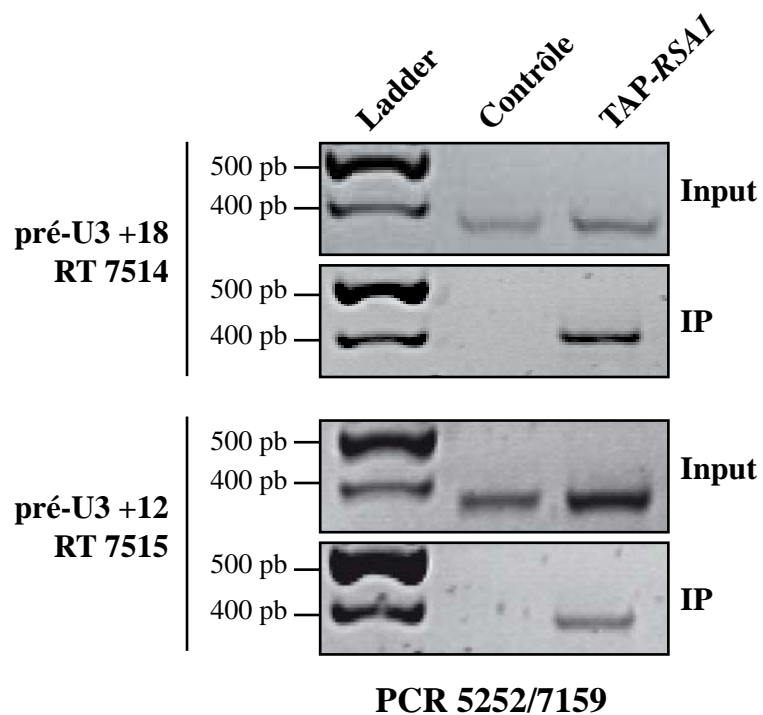
Dans le modèle proposé par l'équipe de D. Tollervey, l'assemblage des protéines Nop1p, Nop56p et Nop58p sur l'ARN, s'effectue de façon tardive, au niveau des pré-U3 +12 et +18 (Kufel *et al.*, 2000) (cf. Introduction, Chapitre 3, 3.4.1) (**Fig. 50**). Leur recrutement permettrait en effet de déplacer les protéines Lsm et Lhp1p de leur site de fixation, afin de



**Figure 81 : Mise en évidence du caractère précoce de l'épissage du snoRNA U3 chez la levure *S. cerevisiae*.** (A) Détection par RT-PCR de la séquence intronique contenue dans les pré-snoRNA U3 non maturés en 3'. (B) Analyse par RT-PCR de la présence de l'intron dans les différentes formes de pré-snoRNA U3. Chaque intermédiaire de maturation a été ciblé, lors de l'étape de transcription inverse, grâce à un oligonucléotide amorce spécifique. Les formes porteuses de l'intron ont ensuite été amplifiées par PCR, à l'aide du couple d'amorces 5254/7159. (C) Proposition d'un nouveau modèle pour la maturation du snoRNA U3 de *S. cerevisiae*. L'étape d'épissage précède la maturation nucléolytique de l'extrémité 3' du snoRNA.

faciliter l'intervention de l'exosome et la maturation nucléolytique finale de l'extrémité 3'. Cette interprétation, qui découle d'expériences d'immunoprécipitation, couplées à des analyses par Northern-blot, repose sur le fait que les protéines constitutives sont détectées uniquement sur le snoRNA U3 mature. Cependant, en couplant nos expériences d'immunoprécipitation à des analyses par RT-PCR, nous avons montré que les protéines Nop1p, Nop56p, Nop58p, ainsi que certains facteurs d'assemblage sont déjà associés sur les pré-snoRNA U3 non épissés et non maturés en 3'. Cette apparente disparité pourrait résider dans la différence de sensibilité entre les deux approches de détection utilisées, puisque l'approche de RT-PCR permet de détecter des quantités d'ARN beaucoup plus faibles que l'approche de Northern-blot. De plus, nos conditions d'immunoprécipitation ont également été optimisées afin de préserver les complexes de faible stabilité. A la différence du protocole expérimental employé par l'équipe de D. Tollervey (Kufel *et al.*, 2000), nous avons réalisé des extraits cellulaires dans un tampon sans détergeant et enrichi en glycérol (cf. Matériel & Méthode).

De façon intéressante les résultats distincts, obtenus à partir de ces deux approches, semblent cohérents avec les observations réalisées chez l'homme. En effet, en jouant sur la stringence des conditions d'immunosélection, l'équipe de N.J. Watkins avait montré que les pré-snoRNP U3 contiennent déjà les différentes protéines constitutives de la particule, mais que ces dernières sont associées de façon labile avec l'ARN. Sur la base d'observations structurales nous avons notamment proposé que cette faible stabilité résulte de l'occupation de l'hélice  $\alpha 3$  de la protéine Snu13p/15.5K par le facteur Rsa1p/NUFIP (cf. Résultats & Discussion, Chapitre 2) (**Fig. 65**). En effet, au sein des RNP matures, certains résidus de cette hélice sont vraisemblablement impliqués dans une interaction avec Nop56p/NOP56 ou Nop58p/NOP58, et leur titration par le facteur d'assemblage pourrait maintenir la particule dans un état immature. Dans un tel contexte, la présence de Rsa1p sur l'ARN semble constituer un bon marqueur de l'état d'avancement du processus d'assemblage et de l'état de stabilité des pré-snoRNP. Nous avons donc vérifié la présence de Rsa1p sur les intermédiaires de maturation +12 et +18, en couplant l'approche d'immunoprécipitation à l'analyse par RT-PCR (**Fig. 82**). De façon intéressante, la détection des ARN co-immunosélectionnés avec la protéine TAP-Rsa1p en phase exponentielle de croissance, révèle que le facteur d'assemblage est toujours associé sur les pré-snoRNA U3 +12 et +18. Ce résultat confirme donc qu'à ce stade de la maturation de l'ARN, les remodelages structuraux conduisant à l'assemblage définitif de la particule ne sont pas encore achevés. Le recrutement de la machinerie d'assemblage et des



**Figure 82 : *In vivo*, Rsa1p est associée aux pré-snoRNA U3 +12 et +18.** Les ARN co-immunoprécipités avec la protéine TAP-Rsa1p ont été analysés par RT-PCR en utilisant les amorces de transcription inverse 7514 et 7515, permettant de cibler spécifiquement les pré-snoRNA U3 +18 et +12. Le couple d'oligonucléotides amorces 5252/7159 a été utilisé pour l'étape de PCR. Les produits d'amplification ont été déposés sur gel d'agarose 2,25%. Une souche contrôle n'exprimant pas de protéine étiquetée a été utilisée comme témoin négatif lors de l'immunoprécipitation.

protéines constitutives sur l'ARN constitue donc un événement précoce qui débute avant l'épissage de l'intron (Rothé *et al.*, Publication n°3). Cependant, le départ des facteurs d'assemblage et la libération des particules correctement assemblées s'effectuent vraisemblablement de façon tardive, au cours des dernières étapes de maturation du snoRNA.

### **2.3.3. Des phénomènes de compensation impliquant Rrp9p pourraient assurer la production constante de snoRNP U3**

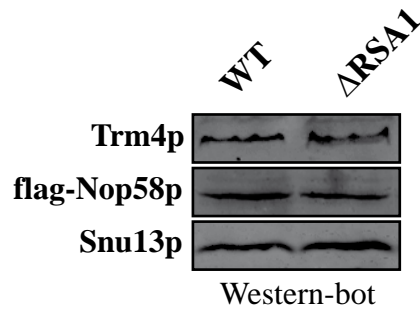
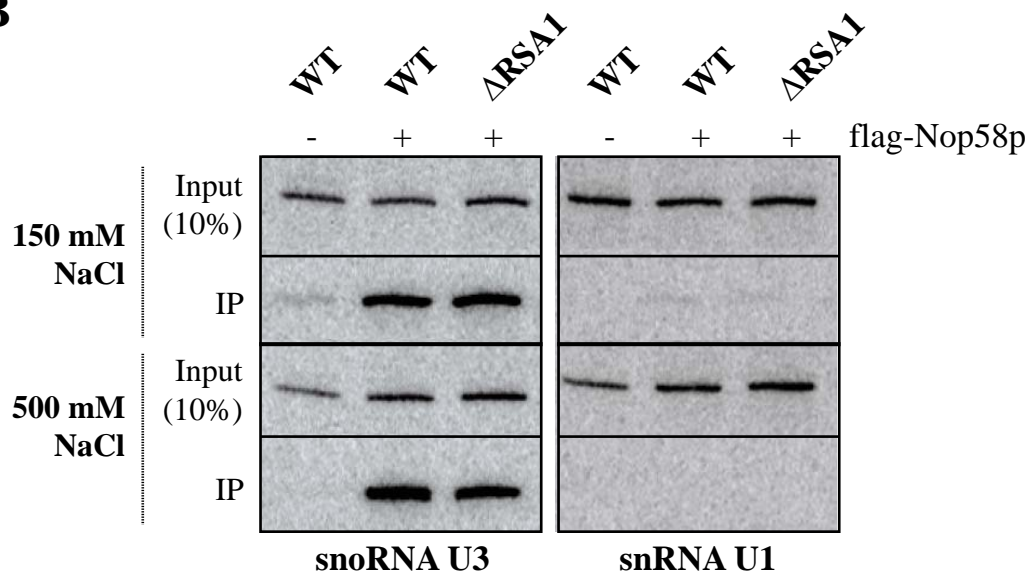
Comme nous l'avons déjà abordé, le snoRNA U3 présente une remarquable stabilité, qui, à la différence des autres snoRNA à boîtes C/D, se matérialise par un taux relativement constant dans de nombreux mutants de *S. cerevisiae*. Si les protéines de levure Snu13p et Nop58p sont essentielles pour la stabilité du snoRNA U3 (Lafontaine & Tollervey, 2000; Watkins *et al.*, 2000), ce n'est pas le cas de Nop56p (Lafontaine & Tollervey, 2000), de la Fibrillarine (Tollervey *et al.*, 1991) ou de Rrp9p (Venema *et al.*, 2000), qui n'ont pas ou peu d'influence sur son accumulation. De même alors que la machinerie d'assemblage exerce un contrôle généralement important sur le taux des snoRNA à boîtes C/D, le snoRNA U3 de levure s'est révélé peu sensible à l'absence de facteurs d'assemblage comme Rvb2p (King *et al.*, 2001), Rsa1p, ou Pih1p (Rothé *et al.*, Publication n°3). Cette grande stabilité, probablement liée à son rôle essentiel pour la production des ribosomes et pour la viabilité cellulaire, implique que la maturation du snoRNA U3 et l'assemblage des protéines Snu13p et Nop58p, sont maintenues en dépit de ces conditions défavorables. Comme nous l'avons déjà mentionné, sur la base d'observations réalisées chez l'homme, il a notamment été proposé que le recrutement de la protéine Rrp9p/U3-55K au niveau du motif B/C contribue à l'incorporation de Nop58p/NOP58 au niveau de la particule C/D (Knox *et al.*, 2011). A la différence de certaines RNP (ex : snR13, snR45, snR63), qui semblent totalement dépendantes des facteurs d'assemblage (Rothé *et al.*, Publication n°3), la snoRNP U3 pourrait donc détenir des capacités d'auto-assemblage susceptibles de compenser un défaut d'intervention de la machinerie d'assemblage.

L'approche d'assemblage de snoRNP *de novo* en extrait cellulaire, nous a permis de montrer l'importance de Rsa1p pour le recrutement de Nop56p et Nop58p sur l'ARN U14 (cf. Résultats & Discussion, Chapitre 1). Concernant la snoRNP U3, nous avons profité de la



stabilité du snoRNA U3 dans les cellules, afin de vérifier l'influence de Rsa1p directement sur l'assemblage des particules endogènes (**Fig. 83**). Nous avons ainsi immunosélectionné la protéine flag-Nop58p exprimée de façon ectopique dans des souches de levure WT et  $\Delta RSAI$  et nous avons évalué par Northern-blot la quantité de snoRNA U3 co-immunopécipité. Deux conditions d'immunosélection ont été utilisées, une condition « douce » avec un tampon à 150 mM en NaCl, ainsi qu'un tampon plus stringent, présentant une concentration de 500 mM en NaCl. Avant l'étape d'immunosélection, des contrôles par Western-blot ont été réalisés sur les extraits cellulaires, afin de s'assurer que l'absence de Rsa1p n'a pas d'effet sur le taux des protéines flag-Nop58p et Snu13p (**Fig. 83A**). En comparaison avec le taux de la méthylase Trm4p, utilisée comme contrôle interne, les résultats obtenus confirment que les deux protéines sont exprimées en quantités comparables dans les souches de levure WT et  $\Delta RSAI$ . Après immunosélection de la protéine flag-Nop58p, la quantité de snoRNA U3 a donc été évaluée dans ces deux contextes génétiques (**Fig. 83B**). Les résultats obtenus révèlent que l'absence du facteur d'assemblage n'a d'effet apparent ni sur l'association entre la protéine étiquetée et le snoRNA U3 endogène, ni sur la stabilité de cette association. Ce résultat est donc cohérent avec le maintien du taux de snoRNA U3 observé dans une souche  $\Delta RSAI$ , puisque l'association de Nop58p avec l'ARN est l'un des rares paramètres requis pour son accumulation. En dépit de l'intervention des facteurs d'assemblage au cours de la biogenèse de la snoRNP U3 (Rothé *et al.*, Publication n°3), ce résultat confirme l'existence de phénomènes de compensation assurant la formation de cette particule malgré la défaillance du système d'assemblage.

Afin de savoir si cette compensation repose précisément sur l'intervention de la protéine Rrp9p, les résultats obtenus précédemment à partir de l'ARN U3 $\Delta$ 2,3,4 constituent des éléments de réflexion intéressants. En effet, cet ARN qui ne complémente que partiellement la déplétion du snoRNA U3 *in vivo* (Marmier-Gourrier *et al.*, 2003), s'est révélé défaillant pour l'assemblage de la protéine Rrp9p *in vitro* (**Fig. 76**) (Travail de thèse N. Rolland). Or, comme nous l'avons observé dans les chapitres précédents, l'expression de l'ARN U3 $\Delta$ 2,3,4 dans les souches de levure  $\Delta RSAI$ ,  $\Delta HIT1$  et  $\Delta PIH1$  est associée à une diminution du taux de sa forme mature et à l'accumulation de ses précurseurs +12 et +18 (alias I et I') ((Boulon *et al.*, 2008); Rothé *et al.*, Publication n°3). Cet effet suggère donc qu'un recrutement sous-optimal de la protéine Rrp9p pourrait augmenter la sensibilité de l'ARN vis à vis d'une défaillance du système d'assemblage. Inversement, dans le cas du snoRNA U3 WT, l'absence des facteurs d'assemblage pourrait être compensée par la fixation optimale de Rrp9p et par son implication

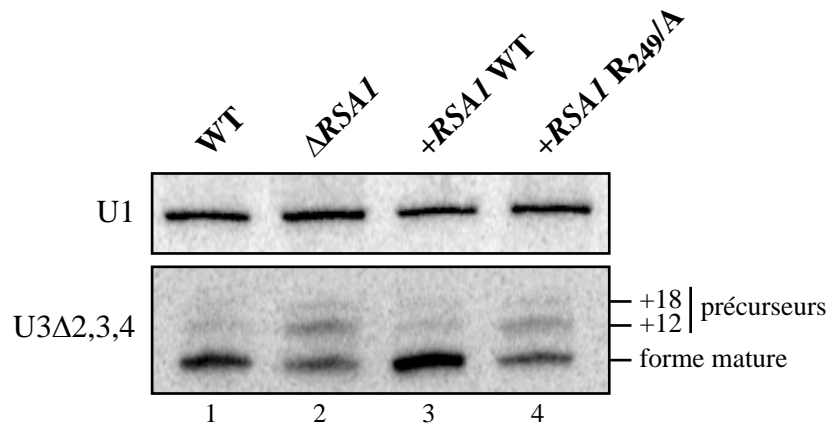
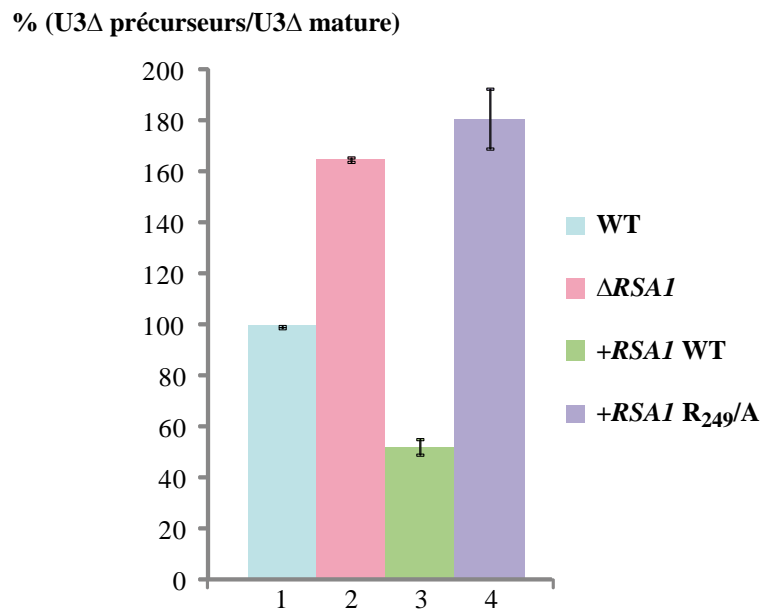
**A****B**

**Figure 83 : L'absence de Rsa1p n'inhibe pas l'incorporation de Nop58p au sein de la snoRNP U3.** (A) Analyse par Western-blot, du taux de protéines flag-Nop58p et Snu13p, dans des souches de levure WT et  $\Delta RSA1$  exprimant la protéine flag-Nop58p à partir du vecteur p413 (*HIS3*). La protéine flag-Nop58p a été détectée à l'aide d'anticorps anti-flag. La protéine Snu13p endogène a été détectée grâce à un sérum anti-Snu13p. La protéine Trm4p, utilisée comme contrôle interne, a été détectée à l'aide d'un sérum anti-Trm4p. (B) Co-immunoprécipitation du snoRNA U3 avec la protéine flag-Nop58p dans des souches WT et  $\Delta RSA1$ . L'immunosélecton a été réalisée sur des billes de sépharose couplées à des anticorps anti-flag. Deux conditions salines différentes (150 ou 500 mM NaCl) ont été utilisées pour l'immunosélecton. Les ARN co-sélectionnés ont été analysés par Northern-blot. Une souche n'exprimant pas la protéine flag-Nop58p a été utilisée en témoin. Le snRNA U1 a été utilisé comme contrôle interne.

physique dans la formation de la snoRNP. Afin de confirmer cette hypothèse, il sera nécessaire de transposer ces analyses sur le snoRNA U3 de pleine taille en utilisant des mutants ponctuels de l'interaction avec Rrp9p. Au laboratoire, les travaux d'A. Cléry ont permis de mettre en évidence que l'inversion de la paire de bases G-C dans l'hélice II du motif B/C perturbe l'association de Rrp9p avec le snoRNA U3 *in vivo* et affecte la croissance des cellules (Clery *et al.*, 2007b). L'expression du variant U3(G-C→C-G) dans une souche de levure JH84 (GAL::SNR17A, ΔSNR17B) (Clery *et al.*, 2007b; Marmier-Gourrier *et al.*, 2003) invalidée pour le gène *RSA1*, pourrait en effet nous permettre de cumuler un défaut de la machinerie d'assemblage à une altération ponctuelle et spécifique de la fixation de la protéine Rrp9p sur le motif B/C.

D'autre part, nous avons également cherché à savoir si le défaut de biogenèse de l'ARN U3Δ2,3,4 observé dans une souche de levure Δ*RSA1*, peut être reproduit sous l'effet d'une perte d'interaction entre Snu13p et Rsa1p (**Fig. 84**). Pour cela, nous avons exprimé de manière ectopique l'ARN U3Δ2,3,4 dans une souche de levure Δ*RSA1*, complétée ou non par Rsa1p WT ou par le variant Rsa1p R<sub>249</sub>/A qui n'interagit plus avec la protéine Snu13p (Rothé, Back *et al.*, Publication n°2). L'accumulation de l'ARN U3Δ2,3,4 et de ses précurseurs a été analysée par Northern-blot (**Fig. 84A**). Les résultats obtenus indiquent que le variant Rsa1p R<sub>249</sub>/A (piste 4) provoque un effet similaire à celui observé en absence de Rsa1p (piste 2), avec une diminution du taux d'ARN mature, associée à une accumulation des précurseurs +12 et +18 (**Fig. 84A et 84B**). Ce résultat illustre donc l'importance de l'interaction entre Rsa1p et Snu13p pour la biogenèse de l'ARN U3Δ2,3,4. Par la suite il sera intéressant de reproduire cette expérience dans une souche de levure JH84 exprimant le variant U3(G-C→C-G), afin d'étudier les conséquences d'une altération cumulée du recrutement de Rsa1p et de Rrp9p sur la snoRNP U3.

L'ensemble de ces données suggèrent donc que l'incorporation sous-optimale de la protéine Rrp9p, couplée à un défaut d'intervention des facteurs d'assemblage pourrait ralentir les étapes finales de la biogenèse de la snoRNP U3. En effet, Rsa1p est importante pour recruter Nop58p et Nop56p, mais également pour faciliter l'intervention des autres facteurs d'assemblage comme Pih1p (cf. Résultats & Discussion, Chapitre 2) (Rothé, Back *et al.*, Publication n°2). La perte d'interaction entre Rsa1p et Snu13p, comme l'absence de Pih1p, pourraient donc compromettre l'intervention des hélicases Rvb1p et Rvb2p, et perturber les étapes finales de remodelage des pré-snoRNP U3. Par la suite, il sera donc important de

**A****B**

**Figure 84 : L'interaction entre Rsa1p et Snu13p est requise pour la biogenèse optimale de l'ARN U3Δ2,3,4.** (A) Analyse par Northern-blot de l'efficacité de maturation du snoRNA U3Δ2,3,4 dans une souche ΔR<sub>SAI</sub> complétementée par Rsa1p WT ou par le variant R<sub>249/A</sub>. Le vecteur pG1(*TRP1*) a été utilisé pour l'expression de Rsa1p WT/mut. L'ARN U3Δ2,3,4 a été exprimé de façon ectopique à partir du vecteur pASZ11(*ADE2*). Les cellules ont été cultivées en milieu sélectif Trp<sup>-</sup>Ade<sup>-</sup> jusqu'à saturation. Le snRNA U1 a été utilisé comme contrôle interne. (B) Quantification du défaut de maturation de l'ARN U3Δ2,3,4 observé par Northern-blot. Les résultats, exprimés en % du signal obtenu pour la souche WT, rendent compte du ratio (précurseurs)/(forme mature).

déterminer par quel mécanisme la protéine Rrp9p contribue au processus d'assemblage. Il est possible que des contacts physiques avec Nop56p et/ou Nop58p au sein des pré-snoRNP contribuent au bon positionnement de ces deux protéines et à la formation du dimère de domaines *Coiled-coil*. Une autre possibilité est que Rrp9p soit directement impliquée dans le recrutement de certains composants de la machinerie d'assemblage, puisqu'une interaction entre Rrp9p et Pih1p a également été décrite par double-hybride (Boulon *et al.*, 2008) (**Tableau 5**). Une perspective intéressante serait donc de créer une souche de levure  $\Delta RSA1$  dans laquelle l'ORF *PIHI* est placée sous le contrôle d'un promoteur inductible. Une telle souche nous permettrait en effet de vérifier si l'absence transitoire des deux facteurs d'assemblage peut annuler ces phénomènes de compensation et affecter l'accumulation du snoRNA U3 WT endogène.

### 3. Conclusion & Perspectives

#### 3.1. Proposition d'un nouveau modèle pour la biogenèse de la snoRNP U3

Les données présentées au cours de cette partie constituent des éléments de réflexion nouveaux qui nous ont conduits à proposer un nouveau modèle de biogenèse pour la snoRNP U3 de levure (**Fig. 85**). Initialement, l'équipe de D. Tollervey avait proposé que l'épissage de l'intron et l'assemblage des protéines constitutives de la snoRNP se déroulent de façon tardive, à l'issue de la génération des extrémités 3' +12 et +18 par l'exosome (Kufel *et al.*, 2000) (cf. Introduction, Chapitre 3, 3.4.1) (**Fig. 50**). Cependant, de nouveaux arguments basés sur des données expérimentales contredisent cette chronologie. En utilisant une approche de RT-PCR permettant la détection spécifique de chaque intermédiaire de maturation, nous avons en effet montré que l'élimination de l'intron s'effectue avant la maturation de l'extrémité 3'. De plus, en couplant ce type d'analyse à l'approche d'immunoprécipitation, nous avons montré que les protéines Nop1p, Nop56p, Nop58p, ainsi que les principaux facteurs d'assemblage, s'associent sur le pré-snoRNA U3 avant l'étape d'épissage (Rothé *et al.*, Publication n°3). Comme nous l'avons abordé au cours du Chapitre 3 de la partie Résultats & Discussion, l'utilisation de l'approche de ChIP ne nous a pas permis de détecter ces différentes protéines au niveau du *locus SNR17A*. Pour des raisons techniques liées à la cinétique de ce processus, il n'est toutefois pas totalement exclu que cet assemblage débute

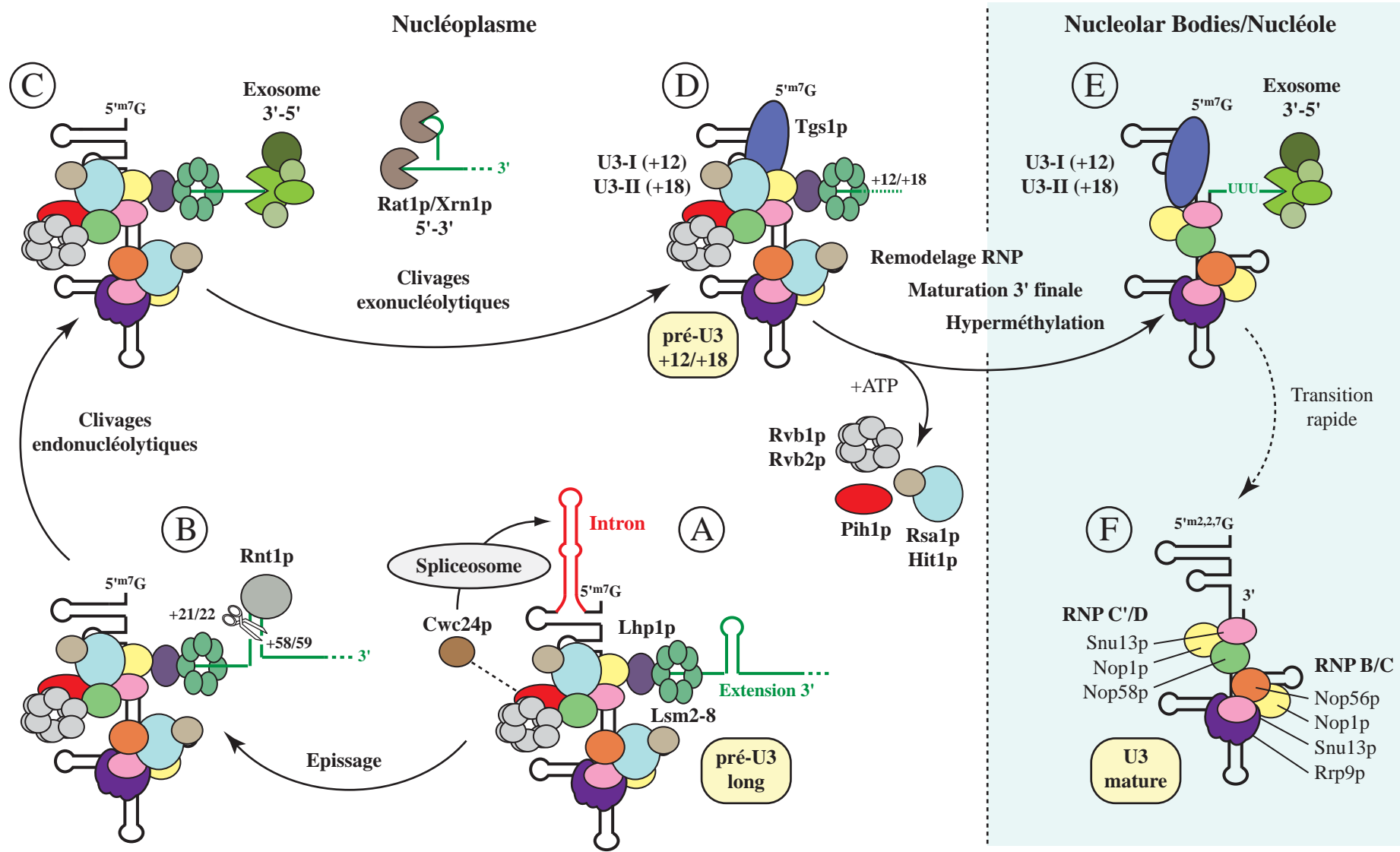


Figure 85 : Nouveau modèle de biogenèse proposé pour la snoRNP U3 chez *Saccharomyces cerevisiae*.  
 (cf. détail texte Résultats & Discussion, Chapitre 5, 3.1)

sur l'ARN naissant, comme cela semble être le cas pour les snoRNA introniques U18 et U24 (Vincenti *et al.*, 2007) (cf. Résultats & Discussion, Chapitre 3, 2.2.2).

Dans le modèle que nous proposons, la biogenèse de la snoRNP U3 débute donc de façon précoce par le recrutement des facteurs d'assemblage et des protéines constitutives sur le pré-snoRNA non épissé et non mûré en 3' (**Fig. 85**) (étape A). Seule la présence de Rrp9p sur ces intermédiaires reste à confirmer, mais étant donné son importance pour la formation de la snoRNP U3, il est fortement probable que cette dernière soit présente sur les mêmes intermédiaires que Nop56p et Nop58p. D'autre part, les protéines Lhp1p et Lsm semblent également recrutées à ce stade, puisqu'elles sont détectées sur les pré-U3 non épissés (Kufel *et al.*, 2000; Kufel *et al.*, 2003). Ces dernières contribuent vraisemblablement à stabiliser l'extrémité 3' de l'ARN afin de ralentir la dégradation par l'exosome, pendant l'assemblage de la RNP. Leur intervention pourrait par ailleurs faire l'objet d'un couplage avec la machinerie d'assemblage, puisque Rsa1p interagit avec Lsm4p par double-hybride (Thèse S. Boulon) et que Pih1p est génétiquement liée à plusieurs protéines Lsm (cf. Résultats & Discussion, Chapitre 3, 2.2.4.2).

Suite au recrutement de ces différentes protéines, la maturation du snoRNA U3 se poursuit donc par l'élimination de l'intron (étape B). Cette étape pourrait également être couplée au processus d'assemblage, puisque la protéine Cwc24p, qui est requise pour l'épissage du pré-snoRNA U3, interagit à la fois avec le facteur d'épissage Cef1p et avec le facteur d'assemblage Pih1p (Goldfeder & Oliveira, 2008) (cf. Introduction, Chapitre 3, 3.6.4). A l'issue de cette réaction, l'intervention de l'endonucléase Rnt1p et de l'exosome (étape C), permet d'aboutir à la formation des pré-U3 +12 et +18 (étape D), dont les extrémités sont protégées par les protéines Lhp1p et Lsm. La détection du facteur Rsa1p sur ces intermédiaires suggère qu'à ce stade, les RNP sont toujours dans un état pré-mature, qui pourrait correspondre à l'état de faible stabilité décrit pour les pré-snoRNP U3 chez l'homme (Watkins *et al.*, 2004). Ceci semble par ailleurs expliquer pourquoi l'équipe de D. Tollervey n'a pas réussi à observer l'association stable des protéines Nop1p, Nop56p et Nop58p avec les pré-U3 +12 et +18 (Kufel *et al.*, 2000). La transition vers les snoRNP mature pourrait en effet impliquer l'intervention des ATPases Rvb1p et Rvb2p au cours d'événements de remodelage permettant de déplacer la protéine Rsa1p et d'assurer la mise en place des interactions définitives entre les composants de la particule (cf. Résultats & Discussion, Chapitre 2) (**Fig. 64**). Au cours de ces étapes, il est possible que les différents facteurs d'assemblage, ainsi que les protéines Lhp1p et Lsm soient dissociés du pré-snoRNA U3, laissant libre cours à l'action

de l'exosome. La maturation exonucléolytique pourrait alors se poursuivre rapidement, jusqu'à l'extrémité 3' mature définie par la RNP C'/D nouvellement formée. Cette étape pourrait s'effectuer dans les NB, puisque les intermédiaires +12 et +18 sont détectés dans ces corpuscules nucléolaires (Verheggen *et al.*, 2002). Enfin, l'hyperméthylation de la coiffe, catalysée par l'enzyme Tgs1p dans les NB (Verheggen *et al.*, 2002), semble se dérouler au cours des dernières étapes de ce processus (étapes E, puis F), puisque seul le snoRNA U3 de taille mature est co-immunoprécipité à l'aide d'anticorps dirigés contre la coiffe triméthylée (Kufel *et al.*, 2000). Toutefois, il est probable que l'intervention de Tgs1p repose sur un couplage avec la machinerie d'assemblage, puisque cette dernière interagit par double-hybride avec Rsa1p (Boulon, 2005). Comme cela a été démontré chez l'homme (Pradet-Balade *et al.*, 2011; Watkins *et al.*, 2004), son recrutement pourrait donc s'effectuer au stade pré-snoRNP, avant les événements de remodelage.

Une perspective intéressante concernant ce travail serait d'étudier plus en détail les liens physiques et fonctionnels s'établissant entre les facteurs d'assemblage et les différents acteurs impliqués dans la maturation du snoRNA. A ce titre, l'utilisation des approches de double-hybride ou de co-expression de protéines chez *E. coli* constituent des outils de choix permettant de cribler de façon exhaustive les interactions potentielles entre les différentes protéines d'intérêt. De plus, comme nous l'avons déjà évoqué dans les chapitres précédant, il serait intéressant d'étudier les conséquences d'un défaut de remodelage des pré-snoRNP, sur l'association entre les différents facteurs et les pré-snoRNA U3. A ce titre, l'utilisation de variants de la protéine Nop58p ou de mutants d'activités des protéines Rvb1p et Rvb2 pourra être envisagée (cf. Résultats & Discussion, Chapitre 2).

### 3.2. Spécificités liées à l'assemblage de la snoRNP U3

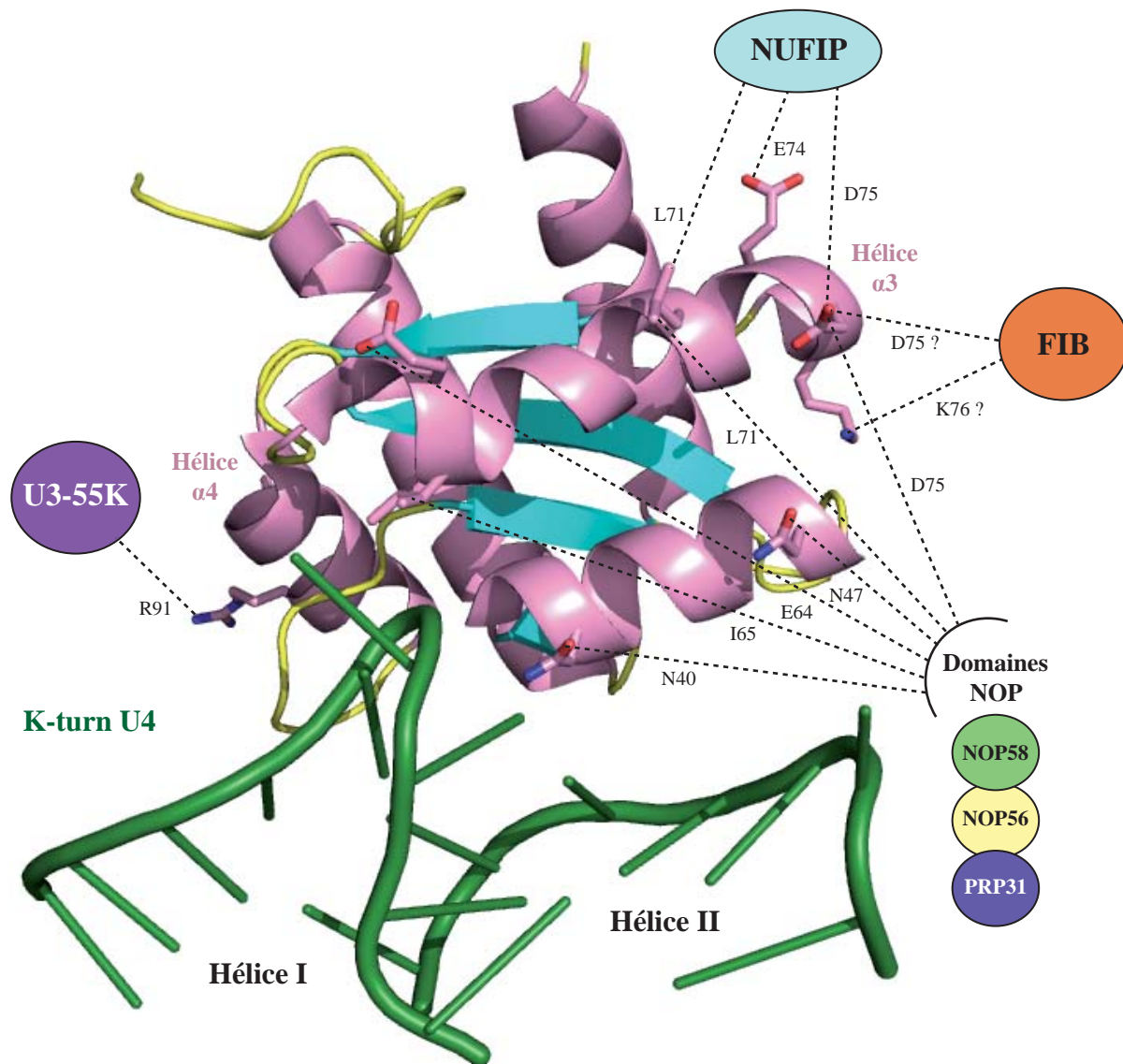
L'utilisation d'approches *in vitro*, nous a permis d'aborder certaines questions liées aux caractéristiques propres de la snoRNP U3. Nous avons ainsi pu montrer que l'assemblage de la protéine Rrp9p sur le motif B/C du snoRNA U3 est possible *in vitro* sans l'assistance de facteurs d'assemblage. Nos données indiquent notamment que Rsa1p n'exerce aucune influence sur le recrutement de Rrp9p sur la pré-snoRNP U3-Snu13p, suggérant que le facteur d'assemblage est dédié exclusivement aux protéines Nop56p et Nop58p. D'autre part, nous avons montré qu'en présence de Snu13p, Rsa1p<sub>230-381</sub> peut être recrutée indépendamment au



niveau des motifs C'/D et B/C. Concernant le motif C'/D, nous avons mis en évidence un effet stabilisateur de Rsa1p sur l'interaction entre la protéine Snu13p et l'ARN. Cette propriété semble reposer sur l'établissement de contacts physiques entre Rsa1p et la partie 5' du snoRNA et pourrait contribuer à compenser la faible affinité de Snu13p pour ce motif. A ce titre, il serait intéressant de vérifier si ce phénomène peut être étendu aux motifs C'/D' des autres snoRNA, qui constituent généralement des sites de fixation de faible affinité pour la protéine Snu13p (Qu *et al.*, 2011). En ce qui concerne le motif B/C, nous avons montré que le recrutement de la protéine Rsa1p<sub>230-381</sub> peut s'effectuer en présence de Rrp9p, suggérant que le facteur d'assemblage et la protéine de la snoRNP pourrait coexister au sein des mêmes pré-snoRNP durant la biogenèse de la particule.

La mise à jour par différentes études des résidus de la protéine Snu13p/15.5K impliqués dans des contacts avec ses nombreux partenaires, nous a permis de dresser une cartographie des sites d'interactions en surface de cette protéine (**Fig. 86**). Concernant la protéine Rrp9p/U3-55K, l'équipe de R. Lührmann a notamment mis en évidence l'importance du résidu R91 de la protéine 15.5K humaine (*i.e.* R95 chez la levure) pour le recrutement de la protéine U3-55K au sein de la snoRNP U3 (Schultz *et al.*, 2006b). Ce résultat, confirmé chez la levure par N. Rolland (Travail de thèse N. Rolland), est donc cohérent avec nos observations expérimentales puisque les sites de la protéine Snu13p/15.5K dédiés à l'interaction avec Rsa1p/NUFIP et Rrp9p/U3-55K sont situés sur des faces opposées de la protéine. La protéine Rrp9p/U3-55K semble interagir avec l'hélice  $\alpha$ 4 de la protéine Snu13p/15.5K, qui est située à proximité de l'hélice I du *K-turn*. Dans le snoRNA U3, cette région de l'ARN correspond l'embranchement des hélices 2 et 4, qui sont en contact physique avec la protéine Rrp9p (Granneman *et al.*, 2009) (**Fig. 23 et 40**). Le facteur d'assemblage Rsa1p/NUFIP contacte en revanche l'hélice  $\alpha$ 3 de Snu13p/15.5K, qui est située à l'opposé du domaine d'interaction avec l'ARN, au niveau d'une région partiellement chevauchante avec le site d'interaction des protéines à domaine NOP (cf. Résultats & Discussion, Chapitre 2) (**Fig. 62 et 63**). Cette configuration suggère donc que Rsa1p/NUFIP pourrait être recrutée au niveau des deux motifs du snoRNA U3 afin de favoriser l'assemblage respectif des protéines Nop56p/NOP56 et Nop58p/NOP58.

Sur la base d'observations expérimentales et des données structurales obtenues chez les archées, un modèle pseudo-symétrique a récemment été proposé pour l'organisation de la snoRNP U3 (Knox *et al.*, 2011) (cf. Introduction, Chapitre 2, 3.1.2.3) (**Fig. 41**). Dans ce



**Figure 86 : Cartographie des sites d'interaction potentiels de la protéine 15.5K avec différents partenaires protéiques.** Les déterminants d'interaction mis à jour par différentes études sont récapitulés sur la structure cristallographique du complexe U4-15.5K (Vidovic *et al.*, 2000; PDB: 1E7K). La zone d'interaction présumée de la protéine U3-55K a été identifiée par des expériences de pull-down (Schultz *et al.*, 2006a). Les sites d'interactions des protéines NOP56, NOP58 et Fibrillarine ont été déduits de la structure tridimensionnelle de la sRNP à boîtes C/D de *S. solfataricus* (Lin *et al.*, 2011). Les déterminants d'interaction dédiés à la protéine PRP31 ont été mis à jour dans la structure tridimensionnelle du complexe U4-15.5K-PRP31 humain (Liu *et al.*, 2007). Enfin, les résidus dédiés à l'interaction de NUFIP ont été caractérisés au cours de nos travaux. Pour chaque interaction, matérialisée par un trait pointillé, le résidu de la protéine 15.5K impliqué est mentionné.

modèle le motif C'/D est occupé par les protéines Snu13p/15.5K, Nop58p/NOP58 et Nop1p/Fibrillarine, tandis que le motif B/C est associé au lot de protéines Snu13p/15.5K, Rrp9p/U3-55K, Nop56p/NOP56 et Nop1p/Fibrillarine. De plus, comme dans les modèles proposés pour les snoRNP à boîtes C/D conventionnelles (**Fig. 36**), les deux motifs sont physiquement couplés par le biais de l'interaction entre les domaines *Coiled-coil* des deux protéines NOP. Comme nous l'avons démontré pour la snoRNP U14, il est donc envisageable que Rsa1p/NUFIP facilite l'assemblage des protéines Nop56p/NOP56 et Nop58p/NOP58 respectivement sur les motifs B/C et C'/D. Dans le cas de la snoRNP U3, cette fonction ne semble cependant pas essentielle pour la production de la particule, puisque Rsa1p/NUFIP pourrait être secondée dans ce rôle par la protéine Rrp9p/U3-55K. En effet, chez l'homme l'équipe de N.J. Watkins a démontré l'importance de cette protéine pour le recrutement de la protéine NOP58 au niveau du motif C'/D et pour l'accumulation du snoRNA U3 dans les cellules (Knox *et al.*, 2011). De plus, cette propriété semble reposer sur des contacts physiques entre les protéines à domaine NOP et la protéine U3-55K, puisque cette dernière interagit physiquement avec NOP56 et NOP58 (Knox *et al.*, 2011).

En se basant sur les données structurales obtenues expérimentalement pour les sRNP d'archées et pour certaines protéines eucaryotes, C. Charron et N. Rolland ont construit au laboratoire, un modèle tridimensionnel de la RNP U3 C'/D de levure. De plus, en utilisant l'approche de modélisation structurale par homologie de séquence, ils ont également établi un modèle tridimensionnel du domaine en *Helix-proppeler* de la protéine Rrp9p (Travail de Thèse N. Rolland). Ces modèles constituent des outils de choix dans le but d'étudier l'assemblage de la snoRNP U3 et d'en comprendre les spécificités. Leur exploitation, en combinant des approches de mutagenèse dirigée aux tests d'interactions et aux tests de fonctionnalité mis au point au laboratoire, constitue donc une perspective de premier plan. A ce titre, le criblage de variants de surface de la protéine Rrp9p a notamment été entrepris au laboratoire par G. Clerget et M. Rederstorff. Ce travail ouvre ainsi la possibilité d'identifier les zones de contact entre Rrp9p et ses différents partenaires protéiques, en particulier les facteurs d'assemblage Pih1p (Boulon *et al.*, 2008) et Bcd1p (cf. Résultats & Discussion, Partie 6), qui pourraient jouer un rôle clé dans l'implication de Rrp9p dans le processus d'assemblage de la snoRNP.

La formation de la snoRNP U3 semble donc reposer à la fois sur l'intervention de la machinerie d'assemblage et sur des propriétés intrinsèques, liées à la présence de la protéine

Rrp9p/U3-55K. Etant donné le rôle essentiel de cette RNP dans l'assemblage des structures pré-ribosomiques, il est donc probable que cette particularité constitue un système de sécurité visant à assurer une production optimale de snoRNP U3. De plus, les travaux réalisés par l'équipe de N.J. Watkins, indiquent que cette dépendance vis-à-vis de Rrp9p/U3-55K a donné prise chez l'homme à la mise en place d'un système de régulation du taux de snoRNP U3 (Knox *et al.*, 2011). Ils ont en effet observé que le taux de protéine U3-55K diminue de façon spécifique au cours de la différenciation des cellules épithéliales, conduisant vraisemblablement à un défaut d'assemblage et à la déstabilisation du snoRNA U3. Ce système, pourrait ainsi permettre d'ajuster sélectivement le taux de snoRNP U3, en fonction des besoins traductionnels de la cellule. Enfin, une autre conséquence de ce mode d'assemblage particulier, pourrait concerner la nature atypique du motif C'/D du snoRNA U3. En dépit de sa position terminale et du recrutement de Nop58p/NOP58, qui sont des attributs caractéristiques des motifs C/D, le motif C'/D du snoRNA U3 se comporte comme les motifs C'/D' retrouvés chez les autres snoRNA à boîtes C/D (cf. Introduction, Chapitre 2, 3.1.1.3.2) (Qu *et al.*, 2011). En effet, comme c'est le cas pour les motifs C'/D', le motif C'/D constitue un site sous-optimal pour la fixation de la protéine Snu13p/15.5K et son assemblage est dépendant d'un motif voisin plus consensuel, ici le motif B/C. Il est donc possible que l'implication majeure de la protéine Rrp9p/U3-55K dans l'assemblage de la particule U3 ait conduit à une diminution progressive de la pression de sélection au niveau du motif C'/D et à une divergence de séquence de la boîte C'.

## CHAPITRE 6

### Etude fonctionnelle de la protéine Bcd1p

#### 1. Présentation des travaux

Hormis le dimère Rsa1p-Hit1p et les membres du complexe R2TP, un autre facteur conservé chez les eucaryotes, la protéine Bcd1p, pourrait également être impliquée dans le processus d'assemblage des snoRNP à boîtes C/D (cf. Introduction, Chapitre 3, 4.6.5). Cette dernière, qui est essentielle à la viabilité cellulaire, s'est en effet révélée requise pour l'accumulation spécifique des snoRNA à boîtes C/D chez la levure (Hiley *et al.*, 2005; Peng *et al.*, 2003). L'analyse des ARN co-immunoprécipités avec la protéine Bcd1p, a révélé son association avec des snoRNA à boîtes C/D (Peng *et al.*, 2003), cependant comme Bcd1p n'a jamais été identifiée au sein des snoRNP matures, il est probable que cette association s'effectue de façon transitoire au cours de la biogenèse des particules. Afin de mieux comprendre le rôle joué par la protéine Bcd1p et dans le but de savoir si ce facteur s'intègre dans la machinerie d'assemblage des snoRNP, nous avons entrepris son étude par diverses approches *in vitro* et *in vivo*. Notre démarche a été centrée sur trois principaux axes d'investigation, visant à (i) évaluer sa capacité à interagir avec les snoRNA, (ii) identifier ses partenaires protéiques, et (iii) déterminer les éléments de séquences requis pour sa fonctionnalité.

Les résultats obtenus indiquent que Bcd1p possède la capacité d'interagir physiquement, mais de façon peu spécifique avec l'ARN, suggérant que son recrutement sur les snoRNA à boîtes C/D pourrait reposer sur des contacts avec d'autres protéines. A ce titre, l'approche de double-hybride a permis de mettre en évidence son interaction avec la protéine Rrp9p et avec les facteurs Pih1p et Rvb2p, tandis que l'analyse de ses partenaires *in vivo*, nous a conduit à observer son association avec les protéines Snu13p, Nop1p, Nop56p, Nop58p et avec les hélicases Rvb1p et Rvb2p. Ces résultats confirment donc l'intervention de Bcd1p au cours de la biogenèse des snoRNP et indiquent qu'elle pourrait coopérer avec certains composants de la machinerie d'assemblage. D'autre part, Bcd1p interagit également par double-hybride avec la chaperonne d'histone Rtt106p. Cette donnée est d'autant plus intéressante que les hélicases Rvb1p et Rvb2p ont récemment été co-purifiées avec Rtt106p, au sein de complexes chromatiniens (Lambert *et al.*, 2010). L'éventualité d'un lien entre la machinerie

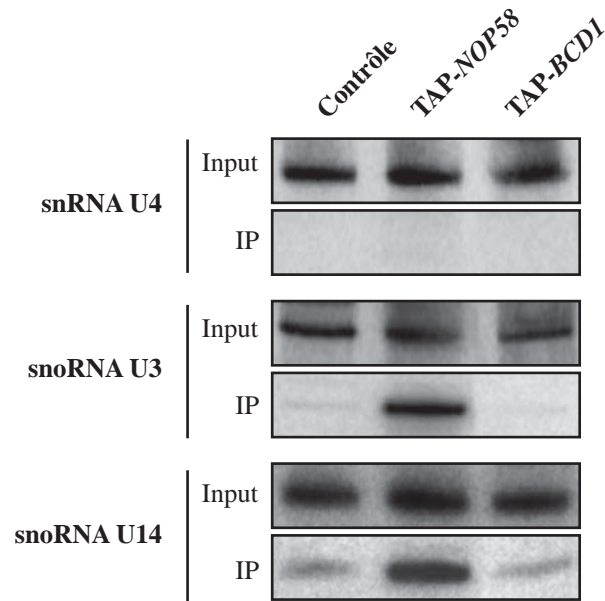
d'assemblage des snoRNP et les mécanismes de régulation transcriptionnelle devra donc être à envisagée. Enfin, Bcd1p contient dans sa région N-terminale, un domaine à doigt de zinc homologue à celui retrouvé dans la protéine Hit1p. Ce dernier ne constitue pas un domaine d'interaction avec l'ARN, mais il est cependant nécessaire à la fonctionnalité de la protéine *in vivo* et peut être remplacé par le domaine à doigt de zinc de Hit1p sans affecter la viabilité des cellules.

## 2. Résultats

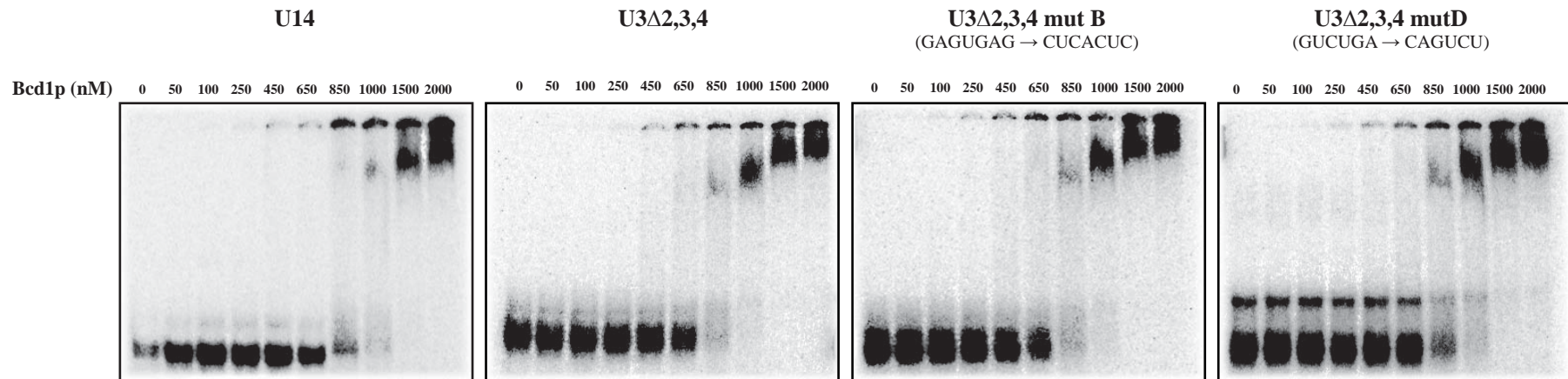
### 2.1. Etude de l'interaction entre Bcd1p et les snoRNA à boîtes C/D

#### 2.1.1. *In vivo* Bcd1p n'est pas stablement associée aux snoRNA matures

L'équipe de T.R. Hugues a mis en évidence l'association spécifique de Bcd1p avec certains snoRNA à boîtes C/D *in vivo* (Peng *et al.*, 2003). Cependant, cette observation, qui incluait notamment le snoRNA U14, a été réalisée grâce à l'utilisation de puces à ADN et ne renseignait pas sur la nature des ARN co-immunosélectionnés (précurseurs ou matures), ni sur le niveau de ces associations. Afin d'évaluer la capacité de Bcd1p à co-immunosélectionner ces molécules, nous avons donc reproduit cette approche à partir de la souche de levure TAP-Bcd1p et nous avons analysé spécifiquement la rétention des snoRNA U3 et U14 par Northern-blot (**Fig. 87A**). Par comparaison avec la protéine de fusion TAP-Nop58p, qui est fortement associée à ces deux ARN, nous n'avons constaté aucun enrichissement spécifique dû à la protéine TAP-Bcd1p. Ce résultat suggère donc que Bcd1p, comme les autres facteurs d'assemblage, n'est associé qu'à une infime proportion de ces snoRNA, qui correspond vraisemblablement à des molécules en cours de biogenèse et qui ne sont aisément détectées qu'à l'aide d'approches sensibles, comme les techniques de puce à ADN (Peng *et al.*, 2003) ou de RT-PCR (cf. Résultats & Discussion, Partie 5). Afin de compléter ce travail, il sera donc intéressant de vérifier si Bcd1p est détectée spécifiquement sur les intermédiaires du snoRNA U3, comme cela a été démontré pour les autres facteurs d'intérêt (cf. Résultats & Discussion, Chapitres 3 et 5) (Rothé *et al.*, Publication n°3).

**A**

**Figure 87 : *In vitro* Bcd1p interagit avec les snoRNA à boîtes C/D, indépendamment de la présence de motifs structurés en K-trun. (A)** Test de co-immunoprécipitation des snoRNA à boîtes C/D U3 et U14, à partir de la protéine TAP-Bcd1p exprimée dans des cellules de levure. La souche de levure TAP-NOP58 a été utilisée comme contrôle positif, tandis que la souche BY4741, n'exprimant aucune protéine étiquetée, a été utilisée comme témoin négatif. Les snoRNA ont été détectés spécifiquement par l'approche de Northern-blot. Le snRNA U4 a été utilisé comme contrôle interne et comme témoin de spécificité. **(B)** Expériences de retard sur gel réalisées à partir des ARN U14, U3 $\Delta$ 2,3,4, U3 $\Delta$ 2,3,4 mutB et U3 $\Delta$ 2,3,4 mutD radiomarqués au niveau de leur extrémité 5' (Marmier-Gourrier *et al.*, 2003). La protéine recombinante Bcd1p a été ajoutée en concentration croissante selon les gammes indiquées.

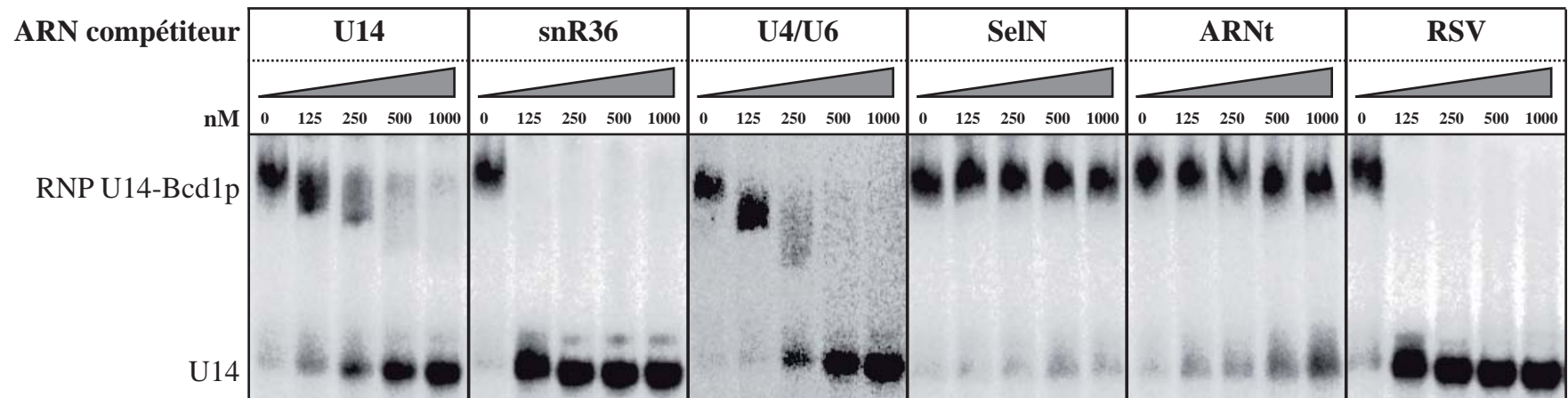
**B**

### 2.1.2. *In vitro* Bcd1p interagit de façon peu spécifique avec l'ARN

Dans le but de vérifier si les interactions détectées par l'équipe de T.R. Hugues (Peng *et al.*, 2003) reposent sur une interaction physique entre Bcd1p et les molécules d'ARN, nous avons réalisé des expériences de retard sur gel (**Fig. 87B**). La protéine Bcd1p a été produite sous forme recombinante et purifiée grâce à une étiquette 6-Histidine. Dans un premier temps, nous avons testé sa capacité à interagir avec le snoRNA U14 et avec l'ARN U3 $\Delta$ 2,3,4. Les molécules d'ARN, radiomarquées à leur extrémité 5', ont ainsi été incubées avec des quantités croissantes de protéine Bcd1p et la formation de complexes ribonucléoprotéiques a été vérifiée par électrophorèse en gel natif. La totalité de l'ARN radiomarké est déplacée dans un complexe de haut poids moléculaire lorsque la concentration en protéine Bcd1p est égale ou supérieur à 1  $\mu$ M. Le Kd apparent de la protéine pour ces deux ARN peut être estimé à environ 0,85  $\mu$ M. Ces résultats indiquent donc que la protéine Bcd1p, qui fait l'objet d'une association transitoire avec les snoRNA à boîtes C/D *in vitro*, présente une propriété d'interaction directe avec l'ARN. Etant donné que les motifs en *K-turn* constituent les éléments structuraux à l'origine de la formation des snoRNP, nous avons ensuite cherché à savoir si leur intégrité était importante pour le recrutement de la protéine (**Fig. 87B**). Pour cela, nous avons utilisé les ARN U3 $\Delta$ 2,3,4 mutB et mutD, respectivement mutés au niveau des boîtes B et D, qui abolissent la structuration des motifs en *K-turn* B/C et C'/D correspondants (Marmier-Gourrier *et al.*, 2003). Les profils de migration électrophorétique obtenus avec ces deux variants sont strictement identiques à celui observé pour l'ARN U3 $\Delta$ 2,3,4 WT, suggérant que la protéine Bcd1p ne présente pas de réelle spécificité pour ces éléments de structure.

Afin de vérifier si la protéine Bcd1p cible néanmoins de façon spécifique les snoRNA à boîtes C/D, nous avons testé son affinité pour différentes molécules d'ARN en procédant par des expériences de compétition (**Fig. 88**). En utilisant le snoRNA U14 radiomarké, nous nous sommes placés dans des conditions expérimentales permettant l'association de la totalité des molécules d'ARN avec la protéine Bcd1p. La formation de ces complexes a alors été testée en présence d'une gamme croissante d'ARN compétiteurs non radiomarqués. Dans le but de comparer le pouvoir de compétition relatif de chaque ARN, nous avons tout d'abord réalisé l'expérience en utilisant comme compétiteur le snoRNA U14 lui-même. Dans ce cas, nous avons pu observer une inhibition totale d'interaction de la protéine avec l'ARN radiomarké,



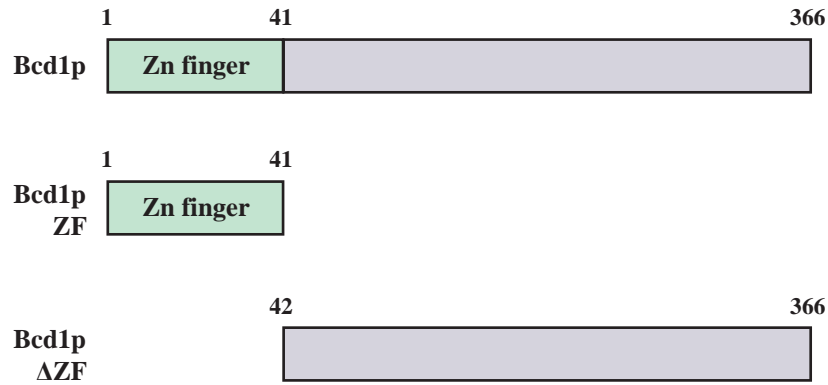
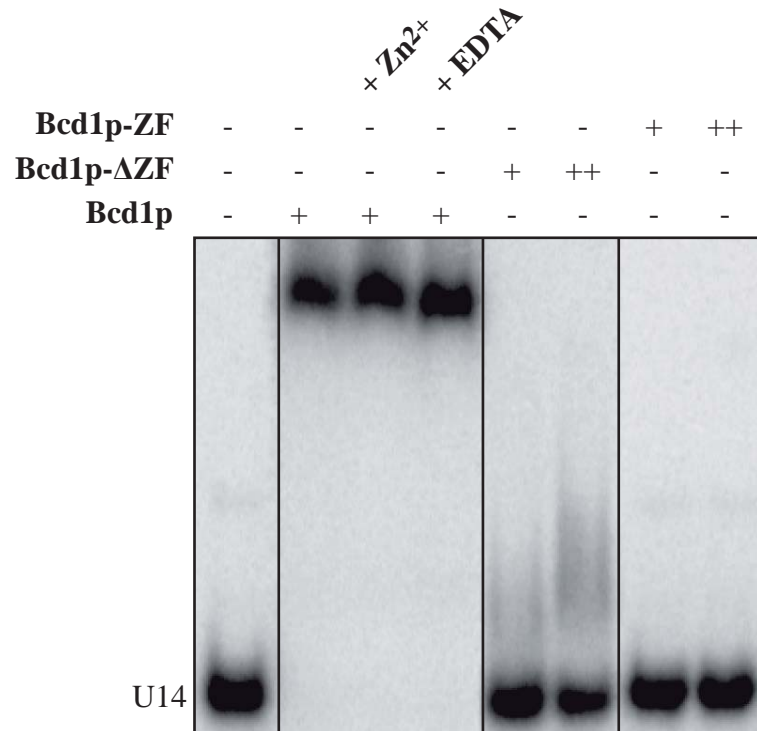


**Figure 88 : Bcd1p interagit avec les ARN en fonction de leur taille.** Tests d'interaction entre Bcd1p et différents ARN réalisés par l'approche de retard sur gel. L'affinité de Bcd1p pour chaque ARN a été testée en évaluant la capacité de chaque molécule d'ARN non radiomarquée à entrer en compétition pour la formation d'un complexe entre Bcd1p et l'ARN U14 radiomarqué au niveau de son extrémité 5'. La protéine Bcd1p a été incubée à une concentration de 1  $\mu$ M, en présence de 5 fmoles d'ARN U14 radiomarqué et des différents ARN non radiomarqués ajoutés en concentration croissante. L'ARN U14 (128 nt) a lui-même été utilisé en tant que compétiteur témoin. Le snR36 (182 nt) est un snoRNA à boîtes H/ACA. L'ARNU4/U6 (248 nt) mime le duplex formé par les U snRNA U4 et U6. L'ARN SelN (59 nt) est un motif retrouvé dans certains ARNm humains et permettant l'incorporation de sélénocystéines durant la traduction. Les ARNt utilisés (70 à 80 nt) correspondent à un mélange d'ARNt de levure *S. cerevisiae*. L'ARN RSV (266 nt) correspond à l'ARN viral du Sarcome de Rous.

à partir d'une concentration de 500 nM en ARN froid. Parmi les autres molécules testées, deux ARN, les ARNt de levure et le motif SECIS de l'ARNm codant la sélénoprotéine humaine SelN, ne présentent aucun effet de compétition. En revanche, le snoRNA à boîtes H/ACA snR36, le duplex de snRNA U4/U6 et même l'ARN génomique de virus du Sarcome de Rous, présentent un fort pouvoir compétiteur, avec des inhibitions totales observées respectivement pour des concentrations supérieures ou égales à 125 nM, 500 nM et 125 nM. Ce résultat indique donc que Bcd1p ne reconnaît pas spécifiquement les snoRNA à boîtes C/D. En revanche, la taille des molécules d'ARN pourrait constituer un critère déterminant pour la fixation de la protéine. En effet, le pouvoir compétiteur des ARN utilisés présente une bonne corrélation avec leur taille en nucléotides, puisque toutes les molécules inférieures à une centaine de nucléotides sont dépourvues de pouvoir compétiteur (**Fig. 88**) (+données non montrées). En dépit de cette absence de spécificité, il est donc probable que le ciblage spécifique des snoRNA à boîtes C/D repose, sur son interaction avec des partenaires protéiques, comme nous l'avons déjà démontré pour d'autres facteurs d'assemblage.

### 2.1.3. Le domaine à doigt de zinc de Bcd1p n'est pas un module d'interaction avec l'ARN

L'analyse de séquence de la protéine Bcd1p a révélé la présence, dans sa région N-terminale, d'un domaine à doigt de zinc, conservé dans la protéine humaine et similaire à celui retrouvé dans la protéine Hit1p/TRIP3 (Rothé *et al.*, Publication n°3). Bien que les domaines à doigt de zinc constituent généralement des modules d'interaction privilégiés avec l'ADN, certains sont dédiés à des interactions avec des partenaires protéiques, avec des lipides, avec des molécules d'ARN ou encore avec des hétéroduplex ADN/ARN (Hall, 2005; Matthews & Sunde, 2002). Nous avons donc cherché à déterminer si le domaine à doigt de zinc de Bcd1p est impliqué dans le mode d'interaction de la protéine avec l'ARN. Pour cela, R. Back a produit au laboratoire, deux fragments recombinants de la protéine Bcd1p : le fragment Bcd1p-ZF (1-41) qui correspond au domaine à doigt de zinc de la protéine, ainsi que le fragment Bcd1p-ΔZF (42-366), qui correspond à la protéine déléetée de son domaine à doigt de zinc (**Fig. 89A**). Nous avons ensuite reproduit les expériences de retard sur gel à partir du snoRNA U14 radiomarqué (**Fig. 89B**). L'interaction de la protéine Bcd1p entière a tout d'abord été testée en présence d'un excès d'ions  $Zn^{2+}$  ou en présence d'EDTA, un chélateur d'ions divalents.

**A****B**

**Figure 89 : Le domaine à doigt de zinc de Bcd1p est nécessaire mais non suffisant pour interagir avec l'ARN.** (A) Représentation schématique des fragments recombinants de la protéine Bcd1p utilisés pour les tests de retard sur gel. (B) Tests de retard sur gel entre le snoRNA U14 radiomarqué à son extrémité 5' et les protéines recombinantes Bcd1p (1  $\mu$ M), Bcd1p-ZF (1 ou 2  $\mu$ M) et Bcd1p- $\Delta$ ZF (1 ou 2  $\mu$ M). Les signes (+) et (++) font respectivement référence aux concentrations protéiques de 1 et 2  $\mu$ M. L'interaction de Bcd1p avec l'ARN a également été testée en présence de Zn<sup>2+</sup> (20  $\mu$ M) ou d'EDTA (2 mM).

Cependant, dans les deux cas, aucun effet n'a été constaté sur la capacité de la protéine à fixer l'ARN. L'ajout des protéines Bcd1p-ZF et Bcd1p- $\Delta$ ZF sur l'ARN a ensuite été testé, mais aucun des deux fragments recombinants n'a permis de reproduire le retard de migration observé pour la protéine entière. Seul le fragment Bcd1p- $\Delta$ ZF provoque l'apparition d'un faible *smear*, supposant la formation de contacts labiles avec le snoRNA U14. Ces résultats indiquent donc que le domaine à doigt de zinc ne constitue pas un module d'interaction avec l'ARN. En revanche, bien que non suffisant, ce domaine apparaît néanmoins nécessaire pour assurer la formation d'un complexe ribonucléoprotéique stable. Une possibilité est que sa présence au sein de la protéine soit requise pour assurer le repliement correct du polypeptide et l'acquisition d'une structure fonctionnelle.

## **2.2. Recherche des partenaires protéiques associés à Bcd1p**

D'après les données obtenues par l'équipe de T.R. Hugues, la protéine Bcd1p interagit de façon spécifique avec les snoRNA à boîtes C/D *in vivo*. En utilisant l'approche de retard sur gel, nous avons montré que ces associations pourraient reposer sur des contacts physiques entre la protéine et les ARN. *In vitro*, cette fonction est cependant apparue non spécifique, puisque Bcd1p reconnaît les ARN de taille supérieure à 80-100 nucléotides, indépendamment de leur nature et de leur séquence. Comme cela a été montré pour la plupart des facteurs d'assemblage, sa spécificité d'action dans le contexte cellulaire pourrait donc dépendre de son interaction avec d'autres protéines. Afin d'identifier les partenaires protéiques de Bcd1p, nous avons donc mis en place deux approches : une approche ciblée par double-hybride, ainsi qu'une approche systématique par co-immunosélection et analyse protéomique.

### **2.2.1. Rsa1p interagit avec Rrp9p, Pih1p et Rvb2p dans des tests de double-hybride**

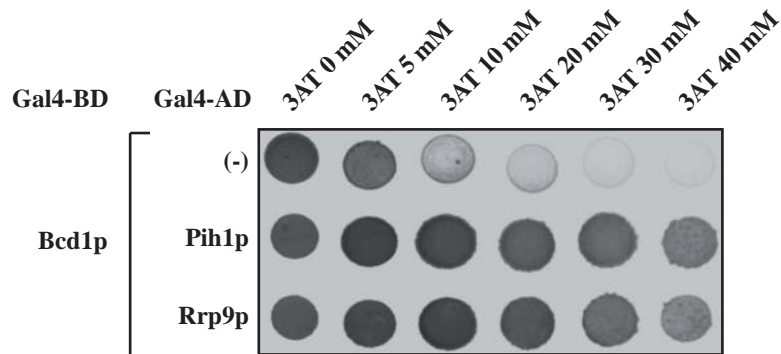
En collaboration avec l'équipe d'E. Bertrand, nous avons utilisé l'approche de double-hybride afin d'identifier, parmi les facteurs d'assemblage (Rsa1p, Pih1p, Rvb1p, Rvb2p) et les protéines des snoRNP (Snu13p, Nop1p, Nop56p, Nop58p, Rrp9p), les partenaires potentiels de Bcd1p. Ces analyses ont abouti à la détection de plusieurs interactions, conférant aux

cellules une forte capacité de résistance au 3-AT. Au laboratoire, nous avons en effet observé l'interaction de Bcd1p avec le facteur Pih1p et avec la protéine Rrp9p, spécifique de la snoRNP U3 (**Fig. 90A**). En plus, de ces deux interactions, l'équipe d'E. Bertrand a également observé l'association de Bcd1p avec l'hélicase Rvb2p, suggérant un lien probable entre Bcd1p et le complexe R2TP (E. Bertrand, Communication personnelle).

### 2.2.2. *In vivo*, Bcd1p est associée aux composants des pré-snoRNP

Afin de mieux définir l'environnement fonctionnel de la protéine Bcd1p, une perspective majeure était d'identifier ses partenaires *in vivo*. Cette démarche avait déjà été entreprise par l'équipe de T.R. Hugues en utilisant l'approche TAP-tag. Cependant à la suite des deux étapes de purification, la quantité de protéines co-immunosélectionnées s'était révélée trop faible pour permettre une analyse par spectrométrie de masse (Peng *et al.*, 2003). Cette situation apparaît relativement similaire à celle que nous avons rencontrée lors de l'étude des facteurs Rsa1p et Pih1p, et suggère l'abondance et/ou la stabilité restreinte des complexes contenant Bcd1p. Comme nous l'avons réalisé précédemment (cf. Résultats & Discussion, Chapitre 3) (Rothé *et al.*, Publication n°3), nous avons donc opté pour un protocole modifié, ne comptant que l'étape de purification par l'étiquette protA et l'éluion des complexes par clivage à la protéase TEV. Après fractionnement sur gel SDS-PAGE et découpage systématique des bandes de gel tous les 2 mm, les protéines co-immunosélectionnées avec Bcd1p ont été analysées par spectrométrie de masse MS/MS. Les analyses ont été réalisées par J.M. Saliou, dans l'équipe de Sarah Cianférani et d'Alain Van Dorsselaer du LSMBO à Strasbourg.

Les résultats obtenus à partir de la souche de levure TAP-*BCDI* ont été comparés avec ceux obtenus pour une souche de levure contrôle n'exprimant pas de protéine étiquetée. Les données obtenues révèlent un enrichissement significatif en protéine dans l'essai TAP-*BCDI*, puisque 359 peptides uniques sont détectés dans cette condition, contre 212 peptides dans l'expérience contrôle (données non montrées). De façon surprenante, seulement 8 peptides uniques correspondant à la protéine Bcd1p sont détectés dans l'essai TAP-*BCDI*. Outre la possibilité de biais liés aux protocoles de protéolyse et de détection par spectrométrie de masse, ceci semble donc confirmer la difficulté à immunosélectionner la protéine Bcd1p en quantité importante. En revanche, les résultats d'analyse indiquent la présence de nombreux

**A****B**

	Protéine	Numero d'accension	Masse moléculaire (Da)	Contrôle Peptides uniques	TAP-BCD1 Peptides uniques
	Bcd1p	P38772	42 316	-	8
	Rsa1p	Q08932	43 981	-	-
	Hit1p	P46973	18 358	-	-
Complexe R2TP	Pih1p	P38768	39 503	-	-
	Tah1p	P25638	12 493	-	-
	Rvb1p	Q03940	50 436	-	2
	Rvb2p	Q12464	51 595	-	2
snoRNP à boîtes C/D	Snu13p	P39990	13 551	-	4
	Nop1p	P15646	34 195	2	11
	Nop56p	Q12460	56 849	-	2
	Nop58p	A6ZPE5	57 176	-	9
snoRNP U3	Rrp9p	Q06506	65 054	-	-

**Figure 90 : Recherche des partenaires de Bcd1p *in cellulo*.** (A) Tests d'interaction par double-hybride entre Bcd1p et les protéines Pih1p et Rrp9p. Bcd1p est exprimée à partir du vecteur pGBKT7 (Gal4-BD), tandis que ces partenaires sont exprimés à partir du vecteur pACT2 (Gal4-AD). Les tests d'interaction ont été réalisés par croisement des souches Y190 (pACT2) et Y187 (pGBKT7), et étalement sur milieu sélectif sans histidine, en présence d'une gamme de concentration croissante en 3-aminotriazol. (B) Analyse par spectrométrie de masse de l'association entre Bcd1p, les facteurs d'assemblage et les composants des snoRNP C/D *in vivo*. Bcd1p a été purifiée à partir de l'étiquette protA du TAP-tag et les protéines co-immunosélectionnées ont été analysées de façon exhaustive par spectrométrie de masse, après fractionnement sur SDS-PAGE 12,5% (Rothé *et al.*, Publication n°3). La souche de levure BY4741, n'exprimant pas de protéine étiquetée, a été utilisée comme contrôle négatif. Le numéro d'accension à la base de donnée UniProtKB/Swiss-Prot est précisé pour chaque protéine.

contaminants, détectés dans l'essai TAP-Bcd1p et dans le contrôle, comme des synthétases d'acide gras, des protéines chaperonnes ou encore des protéines ribosomiques. Cet inconvénient est la contrepartie d'utiliser un protocole de purification à une seule étape et résulte vraisemblablement de la rétention aspécifique de ces protéines sur les billes de sépharose porteuses des IgG. Afin de simplifier l'analyse de ces données et d'éviter la prise en compte de faux-positifs, nous avons extrait de cette liste les protéines dont le lien avec les snoRNA à boîtes C/D a déjà été établi (**Fig. 90B**). De façon intéressante, l'essai TAP-Bcd1p révèle un enrichissement significatif pour les 4 protéines constitutives des snoRNP à boîtes C/D, puisque 4, 11, 2 et 9 peptides uniques sont respectivement détectés pour les protéines Snu13p, Nop1p, Nop56p et Nop58p, contre seulement deux peptides de la protéine Nop1p dans l'expérience contrôle. La protéine Rrp9p, qui interagit par double-hybride avec Bcd1p, n'est en revanche pas retrouvée parmi les protéines co-immunosélectionnées. Etant donné que cette dernière n'est représentée que dans une seule snoRNP, il est cependant possible que son enrichissement ne soit pas suffisant pour permettre son identification. Parmi, les facteurs d'assemblage, seules les hélicases Rvb1p et Rvb2p sont détectées. Bien que spécifique, leur association est cependant relativement faible puisque seulement 2 peptides uniques ont été identifiés pour chacune d'entre elle.

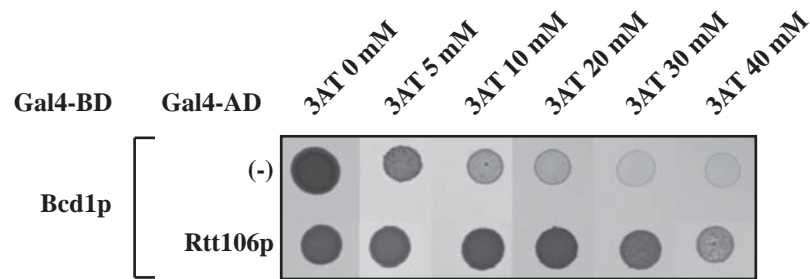
Ces résultats confirment donc les données obtenues par l'équipe de T.R. Hugues, puisque *in vivo*, Bcd1p est non seulement associée aux snoRNA à boîtes C/D (Peng *et al.*, 2003), mais aussi aux 4 protéines constitutives des snoRNP (**Fig. 90B**). Comme cela a été montré pour les autres facteurs d'assemblage, l'abondance relativement faible de ces protéines suggère que Bcd1p interagit avec une faible portion de snoRNP, correspondant vraisemblablement à des pré-snoRNP en cours de biogenèse. De plus, les résultats de co-immunosélection, couplés aux données de double-hybride, indiquent que Bcd1p pourrait interagir avec certains constituants du complexe R2TP. Compte tenu que l'approche de double-hybride n'a révélé aucune interaction entre les protéines constitutives des snoRNP à boîtes C/D et Bcd1p, il est envisageable que cette dernière soit recrutée sur les particules par l'intermédiaire de contacts avec le complexe R2TP. L'étude *in vitro* des interactions entre Bcd1p et Pih1p ou Rvb2p, constitue donc une perspective de travail intéressante. La mise en place d'approches de pull-down ou de co-expression chez *E. coli*, pourrait en effet nous permettre d'évaluer la capacité de Bcd1p à interagir physiquement avec ces deux facteurs et ainsi, de mieux comprendre comment cette protéine s'intègre dans la machinerie d'assemblage des snoRNP.

### 2.2.3. Rtt106p, un nouveau partenaire de Bcd1p en lien avec la régulation de la transcription

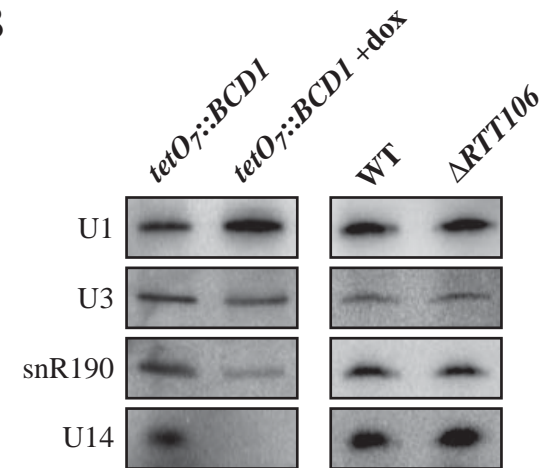
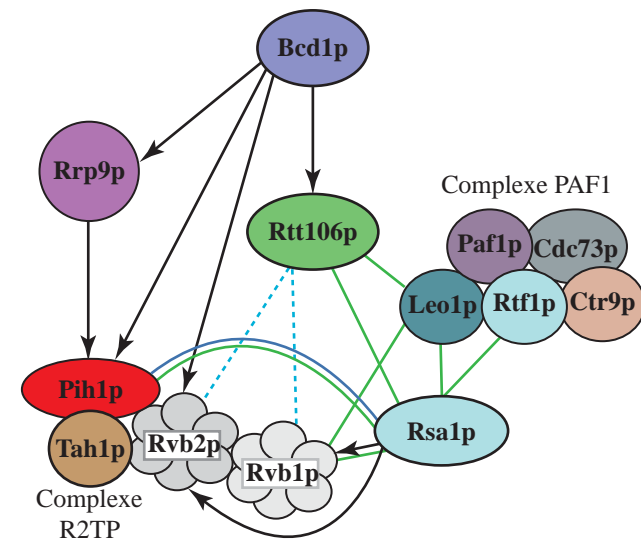
Sur la base de cribles à haut débit, des liens de différentes natures ont été établis entre Bcd1p et d'autres partenaires protéiques potentiels (*Saccharomyces Genome Database* ; (Cherry *et al.*, 1997)). L'un de ces candidats, la protéine Rtt106p, qui fait l'objet d'une co-localisation avec Bcd1p (Huh *et al.*, 2003), a tout particulièrement retenu notre attention. Cette dernière est une chaperonne d'histones impliquée dans la formation de l'hétérochromatine (Huang *et al.*, 2005; Huang *et al.*, 2007), mais également dans la répression de la transcription au niveau de gènes transcrits de façon active (Imbeault *et al.*, 2008). En coopérant avec d'autres chaperonnes, comme la protéine Asf1p ou les complexes HIR et CAF-1, elle permet en effet le dépôt des histones H3 et H4 sur l'ADN et l'assemblage des nucléosomes (Fillingham *et al.*, 2009; Huang *et al.*, 2005; Huang *et al.*, 2007; Lambert *et al.*, 2010), ainsi que le recrutement de complexes de remodelage de la chromatine comme SWI/SNF ou RSC (Ferreira *et al.*, 2011). Sa fonction repose en particulier sur sa capacité à interagir à la fois avec l'ADN et avec les histones H3 et H4 (Huang *et al.*, 2005; Liu *et al.*, 2010). Concernant son intervention au niveau des *loci* transcriptionnellement actifs, il a notamment été proposé que Rtt106p intervienne dans la régulation de la structure chromatinienne suite au passage de l'ARN polymérase II (Imbeault *et al.*, 2008). De façon intéressante, la délétion du gène non essentiel codant Rtt106p est épistatique avec celle du gène *RSA1* (cf. Résultats & Discussion, Chapitre 3, 2.2.4.2). De plus, des expériences d'immunoprécipitation de chromatine modifiée (mChIP, *modified Chromatin ImmunoPurification*), qui consistent à purifier des macro-complexes ADN-protéines et à en analyser le contenu protéique par spectrométrie de masse (Lambert *et al.*, 2009), ont permis de mettre en évidence l'association de Rtt106p et des hélicases Rvb1p et Rvb2p au sein des mêmes complexes (Lambert *et al.*, 2010).

Afin de vérifier si le lien entre Bcd1p et la chaperonne d'histone repose sur une interaction entre les deux partenaires, nous avons cloné l'ORF *RTT106* (*YNL206C*) et nous avons testé la capacité de la protéine à interagir par double-hybride avec Bcd1p (**Fig. 91A**). Les données obtenues indiquent qu'il existe une interaction relativement forte entre les deux facteurs, puisque les cellules résistent à des concentrations en 3-AT relativement élevées (30 à 40  $\mu$ M). Une telle résistance a déjà été observée pour d'autres interactions (Rsa1p-Snu13p, Pih1p-Nop58p, Pih1p-Tah1p) qui ont toutes été confirmées par des approches *in vitro* (cf. Résultats



**A**

**Figure 91 : Bcd1p interagit par double-hybride avec le facteur de remodelage des histones Rtt106p.** (A) Tests d'interaction par double-hybride entre Bcd1p et la protéine Rtt106p. Bcd1p et Rtt106p sont respectivement exprimées à partir des vecteurs pGBKT7 (Gal4-BD) et pACT2 (Gal4-AD). Les tests d'interaction ont été réalisés par croisement des souches Y190 (pACT2) et Y187 (pGBKT7), et étalement sur milieu sélectif sans histidine, en présence d'une gamme de concentration croissante en 3-aminotriazol. (B) Analyse du taux de quelques snoRNA à boîtes C/D dans la souche de levure  $\Delta RTT106$  par l'approche de Northern-blot. Le snRNA U1 a été utilisé comme contrôle interne. L'effet lié à la délétion du gène *RTT106* a été comparé à l'effet observé après déplétion de Bcd1p dans la souche *tetO7::BCD1* (R1158) cultivée pendant deux jours en milieu riche, en absence ou en présence de 50  $\mu\text{g}/\text{mL}$  de doxycycline (+dox). (C) Bilan des partenaires physiques et génétiques de Bcd1p et Rtt106p, en lien avec la biogenèse des snoRNP à boîtes C/D. Les flèches en trait plein symbolisent les interactions observées par double-hybride (la flèche est dirigée vers la protéine proie fusionnée au domaine Gal4-AD). Les traits bleu foncé représentent les interactions physiques observées *in vitro*. Les traits verts représentent les interactions génétiques observées au cours de ce travail (cf. Résultats & Discussion, Chapitre 3) (Rothé *et al.*, Publication n°3) ou par cribre SGA (*Synthetic Genetic Array*) (Imbeault *et al.*, 2008). Les traits en pointillés bleus indiquent que les protéines sont associées au sein des mêmes loci génomiques. Ces associations ont été observées par l'approche de mChIP (*modified Chomatin Immunopurification*) (Lambert *et al.*, 2010).

**B****C**

& Discussion, Chapitre 2 et 3). Outre leur association *in cellulo*, ce résultat suggère donc que Bcd1p et Rtt106p pourraient interagir de façon physique. Etant donné la possibilité de produire ces deux protéines sous forme recombinante (cf. paragraphe 2.1 et (Liu *et al.*, 2010)), une perspective directe de ce travail sera donc de vérifier par pull-down ou par co-expression chez *E. coli*, la possibilité de former un complexe stable entre les deux partenaires *in vitro*.

Dans le but de vérifier s'il existe un lien direct entre Rtt106p et la biogenèse des snoRNP à boîtes C/D, nous avons mesuré le taux de certains snoRNA dans une souche invalidée pour le gène *RTT106*. Cette dernière, qui présente un défaut de croissance à 30°C, a été cultivée jusqu'en phase stationnaire de croissance. Les ARN totaux ont été extraits et le taux de snoRNA U3, U14 et snR190 a été analysé par l'approche de Northern-blot (**Fig. 91B**). Une expérience parallèle a été réalisée à partir de cellules *tetO<sub>7</sub>-BCD1*, déplétées en protéines Bcd1p par traitement à la doxycycline. Comme cela a déjà été démontré (Hiley *et al.*, 2005; Peng *et al.*, 2003) et en comparaison avec le snRNA U1 utilisé comme contrôle interne, l'absence de Bcd1p conduit à une diminution marquée des différents snoRNA à boîtes C/D. En revanche, concernant la souche  $\Delta RTT106$ , aucun défaut d'accumulation de ces ARN n'a été observé.

Ce résultat indique donc que Rtt106p n'exerce pas d'influence majeure sur la biogenèse des snoRNP, en revanche il n'exclut pas la possibilité d'un couplage avec la régulation de la transcription impliquant Bcd1p et Rtt106p. En effet, cette éventualité doit être abordée avec d'autant plus d'intérêt, que Bcd1p et Rtt106p font tous deux l'objet d'une association *in vivo* avec les hélicases Rvb1p et Rvb2p, qui sont connues pour intervenir dans des événements de remodelage de la chromatine (cf. Introduction, Chapitre 3, 3.6.1). Un point important sera notamment de vérifier si ce réseau d'interaction est en lien direct avec l'assemblage des snoRNP à boîtes C/D (**Fig. 91C**), ou s'il implique une autre fonction associée à ces protéines. Dans le chapitre 3, nous avons présenté des résultats d'immunoprécipitation de la chromatine obtenus à partir des différents facteurs d'assemblage (cf. Résultats & Discussion, Chapitre 3, 2.2.2). Comme pour les autres facteurs, ces expériences n'ont pas abouti à la détection de Bcd1p au niveau des gènes de snoRNA (**Fig. 67**). Pour des raisons cinétiques, ces résultats ne nous avaient cependant pas permis de conclure quant à la nature co- ou post-transcriptionnel de l'assemblage des snoRNP étudiées. A la différence des facteurs d'assemblage, la protéine Rtt106p qui interagit directement avec l'ADN, est facilement détectée au niveau de la chromatine en utilisant l'approche de ChIP (Imbeault *et al.*, 2008; Lambert *et al.*, 2010). Une

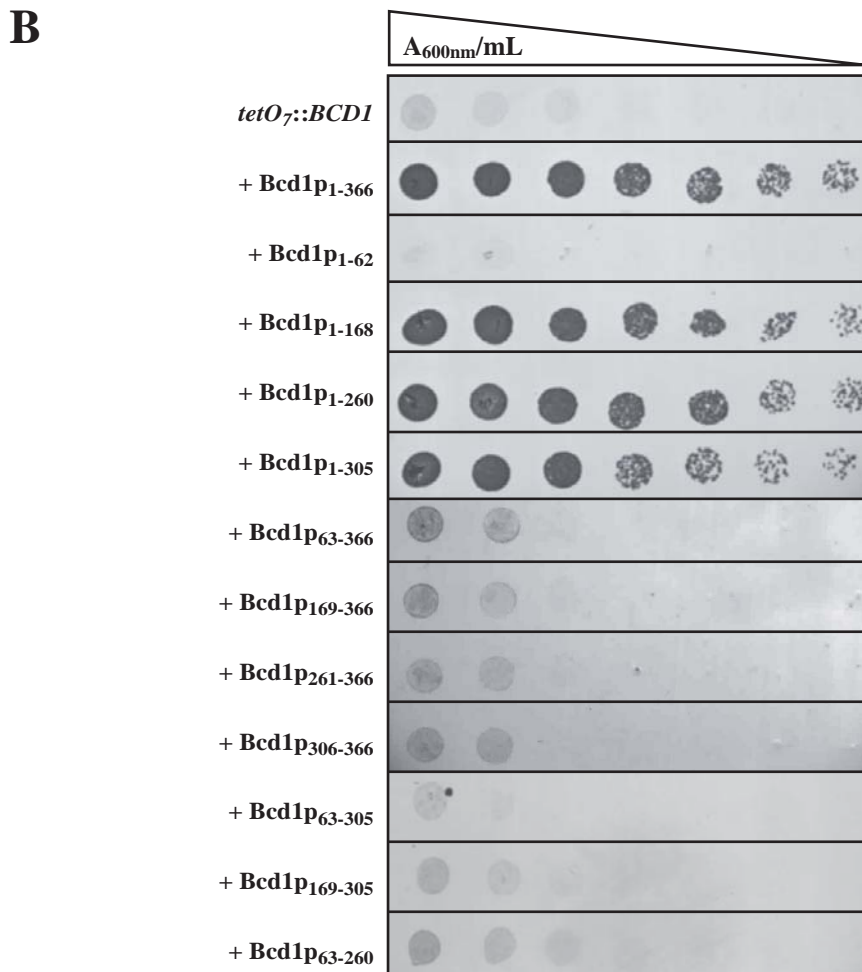
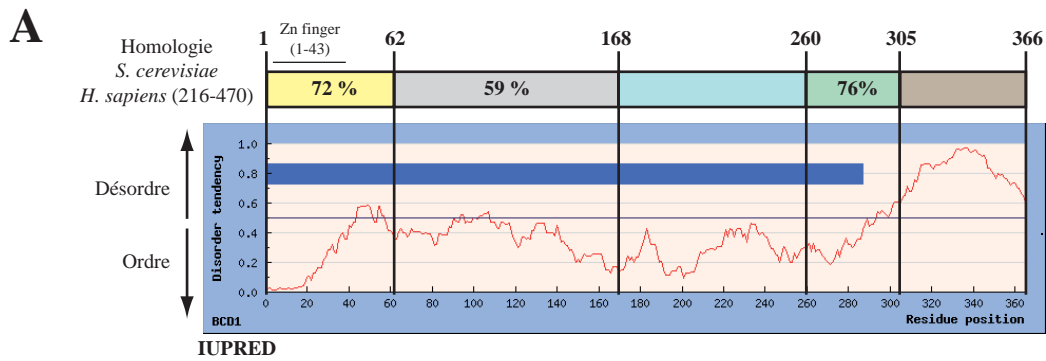
approche permettant de vérifier son implication dans la biogenèse des snoRNP, sera donc d'analyser sa présence sur les *loci* codant ou hébergeant des snoRNA à boîtes C/D. La validation d'une telle hypothèse pourrait alors nous permettre d'envisager d'autres approches expérimentales permettant d'étudier plus spécifiquement les liens entre la machinerie d'assemblage et la régulation de la transcription.

### 2.3. Cartographie fonctionnelle des domaines de Bcd1p

Afin de mieux appréhender la fonction de Bcd1p, nous avons entrepris de cloner différents fragments de la protéine et d'évaluer leur importance pour la croissance des cellules. Ce travail a été réalisé par D. Tiotiu, étudiant en thèse que j'ai encadré à son arrivée au laboratoire.

#### 2.3.1. La région N-terminale de Bcd1p, contenant un domaine à doigt de zinc, est requise pour la fonctionnalité de la protéine

Comme cela a déjà été réalisé pour les autres facteurs d'intérêt, des fragments de Bcd1p ont été définis sur la base de l'homologie de séquence avec la protéine BCD1 humaine et du profil de prédiction de structuration obtenu grâce au serveur IUPRED (**Fig. 92A**). La définition des bornes de ces fragments a également été affinée sur la base de prédiction de structure secondaire, afin d'éviter l'interruption d'éléments de structure importants pour le repliement de la protéine (données non montrées). L'alignement de séquence a révélé que seule la moitié C-terminale (acides aminés 216 à 470) de la protéine BCD1 humaine est homologue à la protéine Bcd1p de levure. Cette homologie est retrouvée principalement au niveau des fragments Bcd1p<sub>1-168</sub> et Bcd1p<sub>260-305</sub> de la protéine de levure et inclut notamment un domaine à doigt de zinc situé dans la portion N-terminale du polypeptide. En plus de la protéine entière, 11 fragments, combinant différentes bornes N- et C-terminales, ont été clonés dans le vecteur p413 et exprimés dans la souche de levure *tet<sub>07</sub>-BCD1* traitée à la doxycycline. La capacité de chaque fragment à compléter la déplétion de la protéine Bcd1p endogène a ainsi été évaluée par un test de croissance en goutte sur boîte de Pétri (**Fig. 92B**). Sur l'ensemble de ces conditions, seules les souches transformées par les fragments



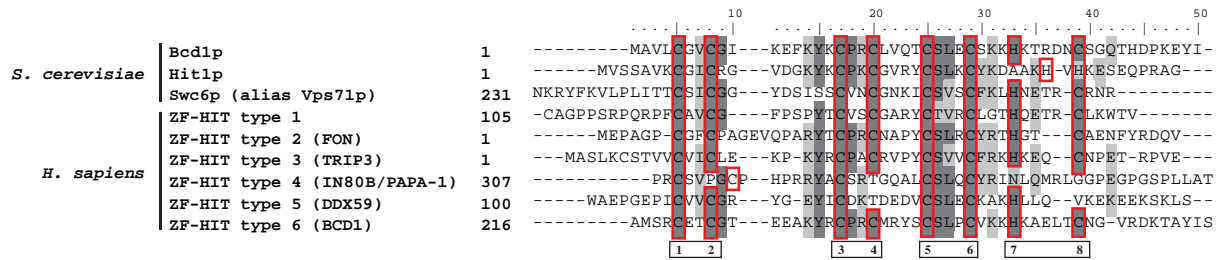
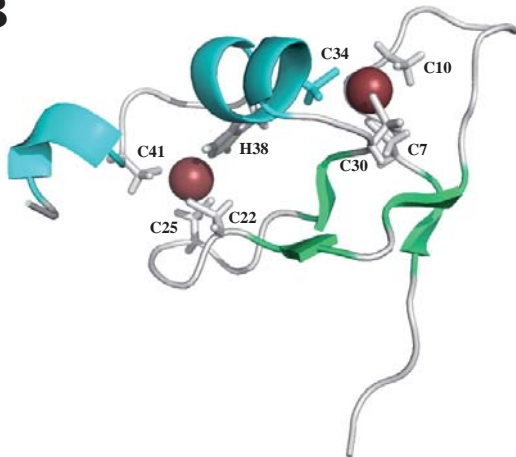
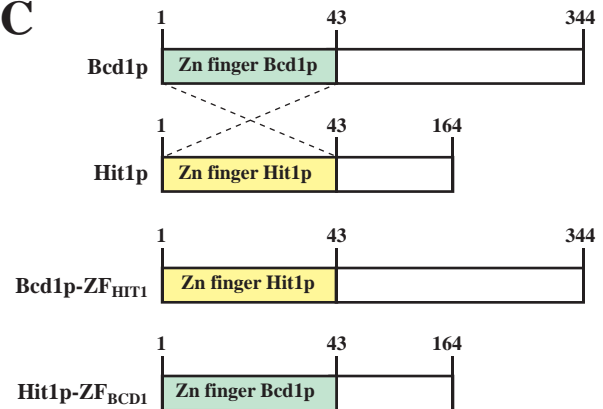
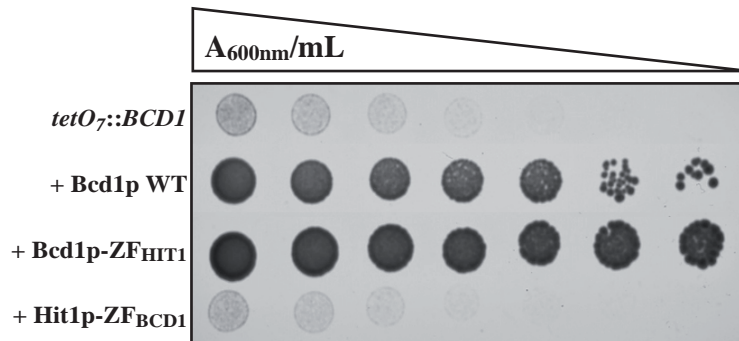
**Figure 92 : La portion N-terminale (acides aminés 1 à 168) contenant le domaine à doigt de zinc est le fragment minimal fonctionnel *in vivo* de la protéine Bcd1p. (A)** Représentation schématique de la protéine Bcd1p et de ses différents domaines. Les bornes des domaines ont été définies sur la base de l'homologie de séquence avec la région (216-470) de la protéine humaine BCD1 et du profil de prédiction de structuration IUPRED (Dosztányi *et al.*, 2005). Le % de similitude (i.e. % de résidus similaires et identiques) entre la protéine de levure et la protéine humaine est indiqué pour chaque fragment homologue. **(B)** Test de fonctionnalité des différents fragments de la protéine Bcd1p. Chaque fragment est exprimé à partir du vecteur d'expression p413 (*HIS3*) dans la souche de levure *tetO7::BCD1* (R1158), permettant de réprimer l'expression de Bcd1p en présence de doxycycline (Peng *et al.*, 2003). La croissance des cellules a été évaluée à 30°C, sur milieu sélectif (*HIS*<sup>-</sup>) et en présence de 50 µg/mL de doxycycline, par dilution en cascade (facteur : 1/2) à partir d'une suspension de cellules à 0,5 unités  $A_{600nm/mL}$ . Avant étalement, les cellules ont été cultivées à 30°C, pendant 2 jours, en milieu liquide sélectif et en présence de 50 µg/mL de doxycycline. La souche transformée par le vecteur p413 non recombinant a été utilisée comme témoin négatif de croissance (*tetO7::BCD1*).

Bcd1p<sub>1-168</sub>, Bcd1p<sub>1-260</sub> et Bcd1p<sub>260-305</sub> sont viables et présentent des capacités de croissances similaires à celles observées pour la souche exprimant la protéine entière. Ces résultats indiquent donc que le fragment N-terminal Bcd1p<sub>1-168</sub>, qui contient le domaine à doigt de zinc (acides aminés 1 à 43), constitue le domaine minimal fonctionnel de la protéine. Afin de déterminer le rôle de ce domaine il sera donc intéressant de tester sa capacité à interagir par double-hybride avec les différents partenaires protéiques de Bcd1p. De plus, sa production sous forme recombinante pourra être envisagée afin de réaliser de nouvelles expériences de retard sur gel et d'évaluer si la propriété d'interaction avec l'ARN est corrélée à la fonctionnalité de la protéine.

### 2.3.2. Le domaine à doigt de zinc de Hit1p est fonctionnel dans Bcd1p

Le domaine à doigt de zinc présent dans la protéine Bcd1p de levure et dans son homologue humaine appartient à une famille de domaines désignés sous le terme de « domaines à doigt de zinc de type HIT » (alias domaines ZF-HIT). Baptisés en référence à la protéine de levure Hit1p (Kawakami *et al.*, 1992), les domaines ZF-HIT contiennent généralement 8 acides aminés (C<sub>1</sub>C<sub>2</sub>-C<sub>3</sub>C<sub>4</sub>-C<sub>5</sub>C<sub>6</sub>-H<sub>7</sub>C/H<sub>8</sub>) potentiellement impliqués dans la coordination de deux atomes de zinc (**Fig. 93A**). La résolution de la structure en solution du domaine à doigt de zinc de la protéine humaine ZF-HIT type 2 (alias FON) a en effet révélé que les couples C<sub>1</sub>C<sub>2</sub> et C<sub>5</sub>C<sub>6</sub> d'une part, et les couples C<sub>3</sub>C<sub>4</sub> et H<sub>7</sub>C<sub>8</sub> d'autre part, s'associent respectivement deux à deux pour fixer deux ions Zn<sup>2+</sup> (He *et al.*, 2007) (**Fig. 93B**).

De façon intéressante, ce type de motif est donc retrouvé dans deux des facteurs impliqués dans la biogenèse des snoRNP à boîtes C/D, les protéines Bcd1p et Hit1p, ainsi que dans leurs homologues respectifs chez l'homme, BCD1 et TRIP3 (**Fig. 93A**). La caractérisation fonctionnelle de ces deux facteurs chez la levure a toutefois révélé que le domaine ZF-HIT ne présente pas la même importance pour ces deux protéines, puisque dans des conditions standard de croissance, ce domaine est nécessaire pour la fonctionnalité de Bcd1p (**Fig. 92B**), tandis qu'il n'est pas requis assurer la fonction de Hit1p (cf. Résultats & Discussion, Chapitre 3) (**Fig. 71**). En dépit de cette différence, la présence de ce motif dans ces deux protéines soulève cependant la question d'une redondance fonctionnelle. En effet, étant donné l'implication de Bcd1p et Hit1p dans le même processus cellulaire, une hypothèse

**A****B****C****D**

**Figure 93 : Le domaine à doigt de zinc de Hit1p est fonctionnel dans la protéine Bcd1p.** Alignement de séquence des domaines à doigt de zinc de type HIT, retrouvés chez la levure et chez l'homme. L'alignement a été réalisé par ClustalW Multiple Alignment à l'aide du logiciel BioEdit (Hall, T.A. 1999). Les acides aminés identiques dans plus de 50% des séquences alignées sont représentés sur fond gris foncé. Les acides aminés similaires dans plus de 50% des séquences alignées sont représentés sur fond gris clair. Les acides aminés (cystéines ou histidines) coordonnant les atomes de zinc sont encadrés en rouge et numérotés de 1 à 8. **(B)** Structure tridimensionnelle du domaine ZF-HIT (1-46) de la protéine humaine FON obtenue par RMN (He *et al.*, 2007; PDB : 1X4S). **(C)** Représentation schématique des protéines chimériques construites par échange des domaines à doigts de zinc entre les protéines Bcd1p et Hit1p. **(D)** Test de fonctionnalité des différentes protéines chimériques. Chaque protéine est exprimée à partir du vecteur d'expression p413 (*HIS3*) dans la souche de levure R1158-*tetO7::BCD1* (Peng *et al.*, 2003). La croissance des cellules a été évaluée à 30°C, sur milieu sélectif (*HIS*<sup>-</sup>) et en présence de 50 µg/mL de doxycycline, par dilution en cascade (facteur : 1/2) à partir d'une suspension de cellules à 0,5 unités *A*<sub>600nm</sub>/mL. Avant étalement, les cellules ont été cultivées à 30°C, pendant 2 jours, en milieu liquide sélectif et en présence de 50 µg/mL de doxycycline.

envisageable est que leurs domaines ZF-HIT puissent jouer un rôle similaire. Afin d'évaluer cette possibilité, nous avons donc entrepris d'échanger les domaines ZF-HIT entre les deux protéines (**Fig. 93C**) et de tester les conséquences d'une telle manipulation sur la croissance des cellules. Etant donné que le domaine ZF-HIT de Hit1p n'est pas nécessaire pour compléter le défaut de croissance d'une souche de levure  $\Delta HIT1$ , nous avons étudié l'impact des protéines chimères dans la souche *tetO7-BCD1*. Cette dernière a été transformée par le plasmide p413 non recombinant et par les vecteurs p413::*BCD1*, p413::*BCD1*<sub>ZF-HIT1</sub> et p413::*HIT1*<sub>ZF-BCD1</sub>, permettant respectivement l'expression de la protéine Bcd1p WT, de la protéine Bcd1p porteuse du domaine à doigt de zinc de Hit1p et de la protéine Hit1p porteuse du domaine à doigt de zinc de Bcd1p. Les différentes souches de levure ont été cultivées en présence de doxycycline et leur croissance a été évaluée par un test en goutte sur milieu solide (**Fig. 93D**). De façon intéressante, alors que la protéine Hit1p<sub>ZF-BCD1</sub> ne permet pas de compléter l'absence de la protéine Bcd1p endogène, la protéine Bcd1p<sub>ZF-HIT1</sub> s'est avérée parfaitement fonctionnelle *in vivo*. Ce résultat suggère donc que les deux protéines partagent non seulement un même domaine structural, mais également un même domaine fonctionnel. Afin de compléter ces observations, il serait donc intéressant de vérifier si ce domaine est réellement impliqué dans une fonction spécifique ou s'il exerce une influence exclusivement structurale. En effet, deux principales hypothèses sont désormais envisageables, puisque ces deux domaines ZF-HIT pourraient soit permettre d'établir des interactions avec les mêmes partenaires, soit contribuer au repliement de la protéine et permettre dans le cas de Bcd1p, l'acquisition d'une structure fonctionnelle.

La comparaison des domaines ZF-HIT de Bcd1p et Hit1p révèle une légère différence de configuration, puisque le domaine à doigt de zinc de Bcd1p présente une organisation de type (C<sub>1</sub>xxC<sub>2</sub>-C<sub>3</sub>xxC<sub>4</sub>-C<sub>5</sub>xxxC<sub>6</sub>-H<sub>7</sub>xxxxxC<sub>8</sub>), tandis que celui de Hit1p présente une organisation de type (C<sub>1</sub>xxC<sub>2</sub>-C<sub>3</sub>xxC<sub>4</sub>-C<sub>5</sub>xxxC<sub>6</sub>-H<sub>7</sub>xH<sub>8</sub>). Cette différence ne semble cependant pas influencer sur le rôle du domaine à doigt de zinc et suppose que ces deux domaines adoptent une structuration similaire. En termes de séquence, les domaines ZF-HIT de Bcd1p et Hit1p présentent une similitude de 36%, tandis que leur similitude respective avec le domaine à doigt de zinc de Swc6p, le troisième représentant de cette catégorie chez la levure, ne s'élève qu'à 24% et 10%. Il serait donc intéressant de construire un vecteur permettant l'expression d'une protéine chimère Bcd1p<sub>ZF-SWC6</sub>, afin de tester la capacité du domaine ZF-HIT de la protéine Swc6p à remplacer le domaine ZF-HIT de Bcd1p. Si cette propriété s'avère restreinte au domaine à doigt de zinc de Hit1p, il est en effet probable que les domaines ZF-HIT de

Bcd1p et de Hit1p partagent des déterminants communs et leur identification par des approches de mutagenèse pourra alors être envisagée.

Par ailleurs, la comparaison des différents membres de la famille des protéines contenant un domaine ZF-HIT révèle une remarquable convergence fonctionnelle et soulève une autre perspective. En effet, la protéine Swc6p (alias Vps71p) est un composant du complexe de remodelage de la chromatine SWR1 (alias SWR-C), qui contient les hélicases Rvb1p et Rvb2p (Morrison & Shen, 2009). De même, outre les protéines ZF-HIT type 3 (alias TRIP3) et ZF-HIT type 6 (alias BCD1), on retrouve également parmi les 6 protéines humaines à domaine ZF-HIT identifiées, des partenaires fonctionnels des hélicases TIP49 et TIP48. La protéine ZF-HIT type 1, qui pourrait être l'homologue de la protéine Swc6p, est en effet associée aux hélicases TIP49 et TIP48 au sein du complexe SCRAP, l'équivalent humain du complexe SWR1 (Morrison & Shen, 2009). La protéine ZF-HIT type 4 (alias INO80B/PAPA-1), appartient quant à elle au complexe de remodelage de la chromatine INO80, qui contient également ces deux protéines (Chen *et al.*, 2011; Jin *et al.*, 2005). L'hypothèse d'un lien fonctionnel entre les hélicases Rvb1p/TIP49 et Rvb2p/TIP48 et les protéines à domaine ZF-HIT est donc clairement posée. Chez la levure, une interaction par double-hybride a été identifiée entre Bcd1p et Rvb2p (E. Bertrand, Communication personnelle), tandis qu'une interaction directe a été mise en évidence chez l'homme entre BCD1 et le duplexe TIP49/TIP48 (McKeegan *et al.*, 2007; McKeegan *et al.*, 2009). Une perspective majeure sera donc de tester l'influence du domaine à doigt de zinc de Bcd1p sur la capacité de la protéine à interagir par double-hybride avec Rvb2p. De même, l'interaction de Hit1p avec les hélicases Rvb1p et Rvb2p, qui n'a pas encore fait l'objet d'une étude, devra également être testée par double-hybride. Enfin, outre l'éventualité d'une interaction physique entre les domaines ZF-HIT et les hélicases, une autre possibilité est que ces différentes entités jouent des rôles à part entière au sein des mêmes complexes. Etant donné que les complexes, contenant à la fois des protéines à domaine ZF-HIT et les hélicases, sont impliqués dans des phénomènes de régulation transcriptionnelles, il sera intéressant de vérifier si les protéines Bcd1p et Hit1p sont capables de se lier à l'ADN. Pour cela des expériences de retard sur gel à partir de protéines recombinantes entières ou de leurs domaines ZF-HIT devront être réalisées.



### 3. Conclusion & Perspectives

Les résultats présentés dans ce chapitre confirment donc l'intervention de la protéine Bcd1p dans la biogenèse des snoRNP à boîtes C/D et suggèrent que ce nouveau facteur pourrait faire partie intégrante de la machinerie d'assemblage. En effet, nos données et celles obtenues par l'équipe de T.R. Hugues (Peng *et al.*, 2003), indiquent que Bcd1p est associée *in vivo* avec des snoRNP en cours d'assemblage. Cette intervention pourrait notamment reposer sur sa capacité à s'associer physiquement avec l'ARN et à interagir avec les composants du complexe R2TP, Pih1p et Rvb2p. De plus, Bcd1p pourrait également intervenir de façon spécifique dans la biogenèse de la snoRNP U3, puisqu'une interaction avec la protéine Rrp9p a également été détectée par double-hybride.

De façon intéressante, des résultats obtenus chez l'homme par l'équipe de N.J. Watkins confirment les observations réalisées chez la levure et indiquent que la fonction de Bcd1p/BCD1 pourrait être conservée entre ces deux organismes (McKeegan *et al.*, 2007; McKeegan *et al.*, 2009). Sa déplétion par l'utilisation de siRNA, s'est en effet révélée délétère pour l'accumulation des snoRNA à boîtes C/D et des approches d'immunoprécipitation ont permis de mettre en évidence son association avec les précurseurs nucléoplasmiques des snoRNA U3 et U8 (McKeegan *et al.*, 2007). Cette association s'effectue vraisemblablement au sein de larges complexes d'assemblage, puisque BCD1 est co-purifiée avec la protéine NOP58, en compagnie des protéines NOP56, NUFIP, PIH1, TIP49 et TIP48 (McKeegan *et al.*, 2007). Sa production sous forme recombinante et la réalisation d'expériences de pull-down a également permis de mettre à jour son interaction avec plusieurs partenaires protéiques. La forme tronquée BCD1<sub>1-360</sub>, qui contient notamment le domaine ZF-HIT, interagit avec la Fibrillarine, avec NUFIP et avec le duplex TIP49-TIP48 en présence d'ATP (McKeegan *et al.*, 2007; McKeegan *et al.*, 2009). La présence du domaine C-terminal de BCD1 s'est avérée requise pour établir une interaction avec la protéine 15.5K. En revanche, aucune interaction n'a été observée avec la protéine PIH1.

De façon intéressante, les travaux de l'équipe de N.J. Watkins ont permis de mettre en évidence une nouvelle interaction entre BCD1 entière et la protéine TAF9 (McKeegan *et al.*, 2007). TAF9 est un facteur de transcription général, conservé chez les eucaryotes, qui est retrouvé dans les complexes TFIID et SAGA, et qui est impliqué chez la levure dans l'expression d'environ 60% des gènes (Grant *et al.*, 1998; Moqtaderi *et al.*, 1996). Ce facteur,

qui interagit également avec la Fibrillarine et avec les hélicases TIP49 et TIP48, pourrait être un composant des pré-snoRNP, puisqu'il est co-purifié avec NOP58 et qu'il s'est avéré associé *in vivo* aux précurseurs des snoRNA U3 et U8 (McKeegan *et al.*, 2007). Initialement, un lien entre ce facteur et les snoRNP à boîtes C/D avait été établi chez la levure grâce à la mise en évidence de son interaction par double-hybride avec la protéine Nop56p (Ito *et al.*, 2001). La présence de Taf9p n'a pas été observée parmi les protéines co-purifiées avec Bcd1p, Rsa1p, Hit1p ou Pih1p, lors de nos expériences d'analyses protéomiques, mais il serait toutefois intéressant de vérifier si son interaction avec Bcd1p est conservée chez la levure. En effet, si le rôle précis de Bcd1p n'a pas encore été élucidé, son étude a cependant révélé plusieurs liens avec les processus de régulation de la transcription. Nos travaux ont notamment permis de mettre en évidence une interaction par double-hybride entre Bcd1p et la chaperonne d'histones Rtt106p, qui a récemment été co-purifiée avec les hélicases Rvb1p et Rvb2p au sein de complexes chromatinien (Lambert *et al.*, 2010). De plus, nous avons également mis en évidence que la fonction de Bcd1p requiert la présence d'un domaine ZF-HIT dans sa région N-terminale. Même si sa fonction précise demeure inconnue, ce type de domaine est retrouvé dans plusieurs facteurs, associés aux protéines Rvb1p/TIP49 et Rvb2p/TIP48, au sein des complexes de remodelage de la chromatine SWR1 et INO80 (Huen *et al.*, 2010; Morrison & Shen, 2009). Leur lien avec les protéines à domaine ZF-HIT et avec la protéine Rtt106p, soulève donc de nouvelles perspectives et pose la question du couplage entre l'assemblage des RNP et le processus de transcription.

Comme nous l'avons abordé au cours du chapitre 3, l'approche de CHIP ne nous a pas permis de détecter les différents facteurs d'assemblage, y compris Bcd1p et les hélicases, au niveau des sites de transcription des snoRNA. A la différence du recrutement de ces facteurs sur les snoRNA à boîte C/D naissants, dont la sensibilité de détection peut être discutée (cf. Résultats & Discussion, Chapitre 3), leur association physique sur l'ADN, au sein de complexe de remodelage de la chromatine, devrait cependant être aisément observable. Il est donc peu probable que Rvb1p, Rvb2p et Bcd1p soient retrouvés au sein de tels complexes régulant directement la transcription des snoRNA. En revanche, il est possible que l'association de ces facteurs au sein des complexes d'assemblage des snoRNP repose sur des bases structurales similaires à celles retrouvées dans les complexes INO80 et SWR1. A la différence des protéines comme Swc6p, ZF-HIT 1 ou INO80B, qui accompagnent les hélicases dans des fonctions liées à la transcription, Bcd1p/BCD1 pourrait donc être une protéine à domaine ZF-

HIT, associée spécifiquement à Rvb1p/TIP49 et Rvb2p/TIP48, dans le processus de biogenèse des snoRNP à boîtes C/D.

Parmi les autres facteurs en lien avec le mécanisme de transcription, les protéines Taf9p et Rtt106p, demeurent toutefois des candidats intéressants pour agir en tant que médiateurs d'un couplage entre la transcription et l'assemblage des RNP. L'interaction de Taf9p avec certaines protéines, comme Nop56p et éventuellement Bcd1p, pourrait en effet permettre d'établir des contacts entre la machinerie de transcription et la machinerie d'assemblage. Sa présence au sein des pré-snoRNP U3 et U8 humaines constitue vraisemblablement la marque d'un tel couplage. Il serait donc intéressant d'intégrer ce facteur à nos futures études chez la levure, en évaluant par double-hybride sa capacité à interagir avec les différentes protéines d'intérêt, et en vérifiant, par des approches d'immunoprécipitation, son association avec les précurseurs du snoRNA U3. Concernant la protéine Rtt106p, l'approche de ChIP pourra également être mise à profit afin de vérifier sa présence sur les gènes de snoRNA. Etant donné son influence inhibitrice sur le processus de transcription et son absence d'impact sur le taux de snoRNA, une hypothèse envisageable est que Rtt106p soit recrutée spécifiquement sur les gènes de snoRNA, lorsque la biogenèse des particules est inhibée. Pour cela il sera donc intéressant de réaliser des expériences de ChIP à partir de cellules en phase stationnaire de croissance, période durant laquelle la transcription par l'ARN polymérase II est peu effective (cf. Résultats & Discussion, Chapitre 5).

Comme nous l'avons déjà évoqué, la poursuite de ce projet nécessitera également, d'étudier plus en détail les liens physiques s'établissant entre Bcd1p, Rtt106p et les hélicases Rvb1p et Rvb2p. Il sera notamment important de vérifier par des approches de pull-down ou de co-expression chez *E. coli*, la possibilité de former des complexes stables entre ces protéines *in vitro*. De plus, l'implication du domaine ZF-HIT de Bcd1p dans l'interaction avec les hélicases Rvb1p et Rvb2p pourra également être étudiée par double-hybride, tandis que sa capacité à fixer directement l'ADN devra faire l'objet d'une vérification par l'approche de retard sur gel. Comme cela a été réalisé pour la protéine Rsa1p, l'implication directe de Bcd1p dans le processus d'assemblage des snoRNP à boîtes C/D pourra être évaluée grâce au système de reconstitution de la snoRNP U14 en extrait cellulaire (cf. Résultats & Discussion, Chapitre 1). En effet, l'utilisation de la souche de levure *tetO7-BCD1* pourra nous permettre de tester l'assemblage de cette particule dans un extrait épuisé en protéine Bcd1p, mais aussi d'évaluer par complémentation, la contribution de différents fragments de la protéine. A plus

long terme, des formes variantes de Bcd1p pourront également être analysées, puisque D. Tiotiu a entrepris la mise en place, au laboratoire, d'une approche de mutagenèse aléatoire visant à identifier des acides aminés requis pour l'interaction de Bcd1p avec ses différents partenaires. Enfin, une souche de levure *BCDI-DAMP*, porteuse de la cassette de résistance à la nourseothricine, est en cours de préparation au laboratoire. La caractérisation de cette souche, par l'équipe de C. Saveanu, grâce à l'approche GIM (cf. Résultats & Discussion, Chapitre 3), constituera un élément important dans le but d'identifier l'environnement fonctionnel de Bcd1p et de progresser dans la compréhension de son rôle.

CONCLUSION  
&  
PERSPECTIVES



## CONCLUSION & PERSPECTIVES

Chez les eucaryotes, les RNP contenant des ARN non codants sont des acteurs clés de nombreuses fonctions essentielles et leur biogenèse constitue un processus fondamental, mobilisant une majeure partie de l'activité transcriptionnelle cellulaire et faisant appel à un grand nombre de facteurs spécialisés. L'étude de ces processus complexes, qui intègrent notamment la transcription et la maturation des ARN, la synthèse et la modification post-traductionnelle des protéines, le transport des différents constituants et leur assemblage sous forme d'entités fonctionnelles, constitue donc un domaine de recherche à part entière. Au cours de mon travail de thèse, nous nous sommes tout particulièrement intéressés aux RNP contenant des protéines de la famille L7Ae, qui interviennent dans de nombreux processus cellulaires comme la production des ribosomes, l'épissage, la réplication des télomères ou encore la synthèse des sélénoprotéines. En collaboration avec les équipes d'E. Bertrand et d'A. Krol, nous avons ainsi contribué à la mise à jour de facteurs protéiques impliqués spécifiquement dans l'assemblage de ces particules (Boulon *et al.*, 2008). Puis, en utilisant l'organisme modèle *S. cerevisiae*, nous avons entrepris la caractérisation de leur fonctionnement dans le cadre de la biogenèse des snoRNP à boîtes C/D.

### 1. Mise à jour d'une machinerie impliquée dans l'assemblage des RNP

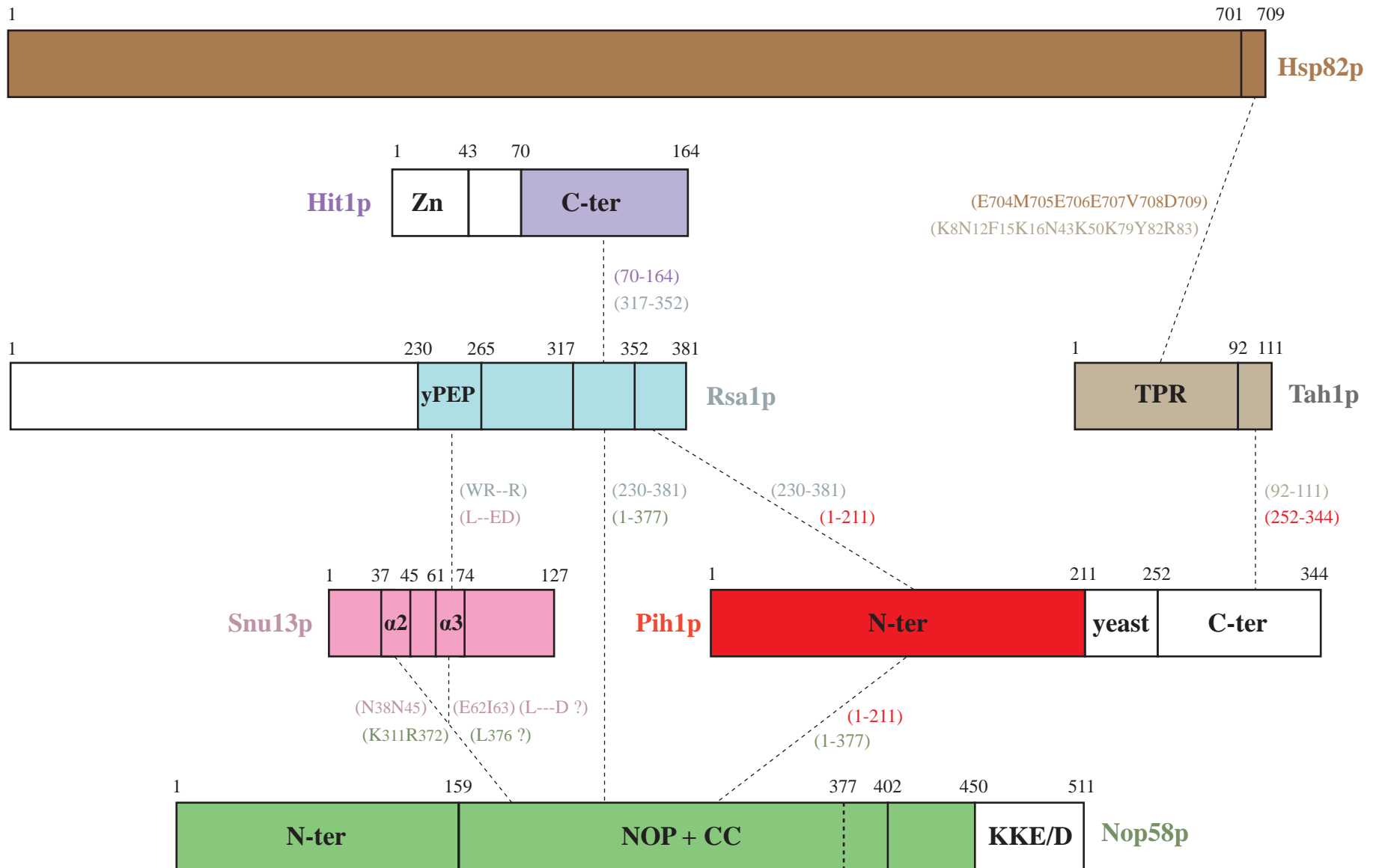
La machinerie moléculaire identifiée est composée de deux principales sous-unités fonctionnelles : (i) d'une part le facteur Rsa1p/NUFIP qui constitue un adaptateur spécialisé, capable d'interagir avec certaines protéines de la famille L7Ae et de faciliter le recrutement des autres protéines sur les pré-RNP, (ii) d'autre part, le complexe R2TP, composé des protéines Rvb1p/TIP49, Rvb2p/TIP48, Pih1p/PIH1 et Tah1p/SPAGH, dont le rôle majeur serait d'opérer des remodelages conformationnels conduisant à la formation des RNP matures. En plus de ces acteurs centraux, d'autres facteurs protéiques sont également apparus intimement liés au processus d'assemblage. C'est notamment le cas de la protéine Hit1p/TRIP3, un partenaire privilégié du facteur Rsa1p/NUFIP, qui est requise chez la levure, pour maintenir le taux de Rsa1p et pour potentialiser sa fonction. C'est également le cas de la chaperonne Hsp82p/HSP90, dont l'importance prévaut essentiellement dans le système

humain et qui contribue aux processus d'assemblage, en exerçant son activité sur plusieurs constituants protéiques des RNP. Enfin, la protéine Bcd1p/BCD1, dont le rôle précis reste encore inconnu, constitue également un facteur important, qui pourrait être une extension du complexe R2TP, spécifiquement associée aux hélicases Rvb1p/TIP49 et Rvb2p/TIP48 dans le cadre de l'assemblage des snoRNP à boîtes C/D.

## 2. Nouveaux éléments de réflexion sur l'assemblage des snoRNP C/D

En recoupant différentes approches de génétique de la levure, de biochimie et de biologie moléculaire, nous avons entrepris l'étude des liens physiques et fonctionnels s'établissant entre les différents facteurs d'assemblage et les constituants des snoRNP à boîtes C/D chez la levure. Ce travail a notamment permis d'établir un bilan relativement exhaustif des interactions impliquant ces différentes protéines, mais également de dresser une cartographie précise des domaines protéiques ou des acides aminés impliqués dans ces interactions (**Fig. 94**) (**Tableau 9**). L'exploitation de ces données *in vitro* ou dans le contexte cellulaire, nous a également permis d'obtenir des informations sur la contribution respective des principaux facteurs d'assemblage et sur la façon dont ils pourraient coopérer au cours de la formation des particules. Les résultats tirés de cette démarche constituent ainsi des éléments de réflexion nouveaux qui, en complément des données disponibles dans la littérature, nous permettent de mieux appréhender la dynamique d'assemblage des snoRNP à boîtes C/D (**Fig. 95**). Un élément clé de ce processus semble être la formation d'un complexe initial entre le pré-snoRNA, la protéine Snu13p/15.5K et le facteur d'assemblage Rsa1p/NUFIP (Rothé, Back *et al.*, Publication n°2) (**Fig. 95**). D'après les données actuellement disponibles, il semble que la protéine Hit1p/TRIP3, qui est fortement associée à la protéine Rsa1p/NUFIP *in vivo*, fasse également partie intégrante de ces pré-snoRNP (Rothé *et al.*, Publication n°3) (E. Bertrand, Communication personnelle). Un point restant à éclaircir est cependant de savoir si la protéine Snu13p/15.5K, qui interagit spécifiquement avec les motifs C/D des snoRNA, est recrutée sur l'ARN avant les deux facteurs d'assemblage ou bien si elle est initialement incluse dans un complexe purement protéique contenant Rsa1p/NUFIP et Hit1p/TRIP3, avant son interaction avec l'ARN ? Les résultats obtenus *in vitro* à partir de protéines recombinantes indiquent que les protéines de levure Snu13p, Rsa1p et Hit1p peuvent être co-exprimées chez *E. coli* et purifiées sous la forme d'un complexe trimérique stable, indiquant qu'une telle entité peut





**Figure 94 : Cartographie détaillée des interactions physiques entre les facteurs d'assemblage et les constituants des snoRNP à boîtes C/D.** Les domaines non colorés représentent les régions non essentielles à la fonctionnalité des protéines. Pour chaque interaction, les bornes des domaines impliqués ou les acides aminés impliqués sont précisés entre parenthèse, selon le code couleur attribué à chaque protéine. Les différentes protéines et leurs domaines sont représentés à l'échelle.

Facteurs d'assemblage		Interactions détectées		Effet Knock-out/Inhibition sur la stabilité des constituants des RNP					
		Protéines constitutives des RNP	Facteurs d'assemblage	C/D	H/ACA	U4	TERC	mRNA <sup>sel</sup>	Protéines affectée
Rsa1p NUFIP	Levure	Snu13p, Nop58p, Prp31p	Hit1p, Pih1p, Rvb1p, Rvb2p	↓↓	→	→			aucun effet
	Homme	15.5K, NHP2, SBP2, Fibrillarine, NOP56*, NOP58*, PRP31, U3-55K	PIH1, TIP49, TIP48, TRIP3	↓↓	(↓)	→	↓↓	non testé	non testé
Hit1p TRIP3	Levure	Prp31p	Rsa1p	↓↓	→	→			non testé
	Homme		NUFIP	non testé	non testé	non testé	non testé	non testé	non testé
Pih1p PIH1	Levure	Nop58p, Prp31p, Rrp9p	Tah1p, Rvb1/2p, Hsp90p, Rsa1p	↓↓	→	→			aucun effet**
	Homme	SBP2, Fibrillarine*, NOP56, NOP58	SPAGH, TIP49*, NUFIP	↓↓*	non testé	non testé	non testé	non testé	non testé
Tah1p SPAGH	Levure		Pih1p	→	→	→			aucun effet**
	Homme		PIH1, TIP48, HSP90	non testé	non testé	non testé	non testé	non testé	non testé
Rvb1p TIP49	Levure	Prp31p	Rsa1p, Pih1p, Rvb2p	↓↓**	↓↓**	non testé			aucun effet**
	Homme	15.5K*, Fibrillarine*, NOP56*, NOP58*	TIP48, PIH1*, BCD1*	↓↓*	non testé	non testé	non testé	non testé	non testé
Rvb2p TIP48	Levure	Prp31p	Rsa1p, Pih1p, Rvb1p	↓↓**	↓↓**	non testé			aucun effet**
	Homme	15.5K*, Fibrillarine*, NOP56*, NOP58*	TIP49, SPAGH, BCD1*	↓↓*	non testé	non testé	non testé	non testé	non testé
Hsp82p HSP90	Levure	non testé	Tah1p, Pih1p	→	→	non testé			aucun effet**
	Homme		SPAGH	↓↓	non testé	(↓)	↓↓	non testé	15.5K, SBP2, NHP2, NOP58
Bcd1p BCD1	Levure	Rrp9p	Pih1p, Rvb2p	→	→	non testé			aucun effet**
	Homme	15.5K	TIP48/TIP49*, NUFIP*	↓↓	non testé	(↓)	↓↓	non testé	15.5K, SBP2, NHP2, NOP58

**Tableau 9 : Bilan final des liens établis entre les facteurs d'assemblage et les composants des RNP de la famille L7Ae chez la levure *S. cerevisiae* et chez l'homme.**

Les noms de protéines mentionnés en bleu indiquent que l'interaction correspondante a été vérifiée par une approche *in vitro*. Les noms en italique indiquent que l'interaction correspondante requiert la présence d'ATP. Les noms qui apparaissent en vert correspondent à des interactions observées uniquement par double-hybride. Seules les interactions double-hybride observées à partir des protéines de pleine taille sont mentionnées. Les effets "Knock-out" correspondent à la caractérisation de souches de levures invalidées pour le gène d'intérêt, ou de cellules humaines traitées par des siRNA spécifiques du facteur étudié. L'effet "inhibition" correspond à la caractérisation de cellules dans lesquelles la chaperonne Hsp82p/HSP90 est inhibée par la geldanamycine. Le symbole ↓↓ indique un effet majeur sur le taux d'ARN, le symbole (↓) indique un effet modéré et/ou nécessitant des expériences complémentaires. Le symbole → indique l'absence d'effet négatif sur le taux d'ARN. Les cases grisées correspondent aux RNP de la famille L7Ae non conservées chez la levure. \*Résultats obtenus par l'équipe de N.J. Watkins (Watkins *et al.*, 2004; McKeegan *et al.*, 2007 et 2009). \*\*Résultats publiés dans Zhao *et al.*, 2008. Les autres données sont relatives à la publication Boulon *et al.*, 2008 et aux différents chapitres Résultats & Discussion de la thèse.

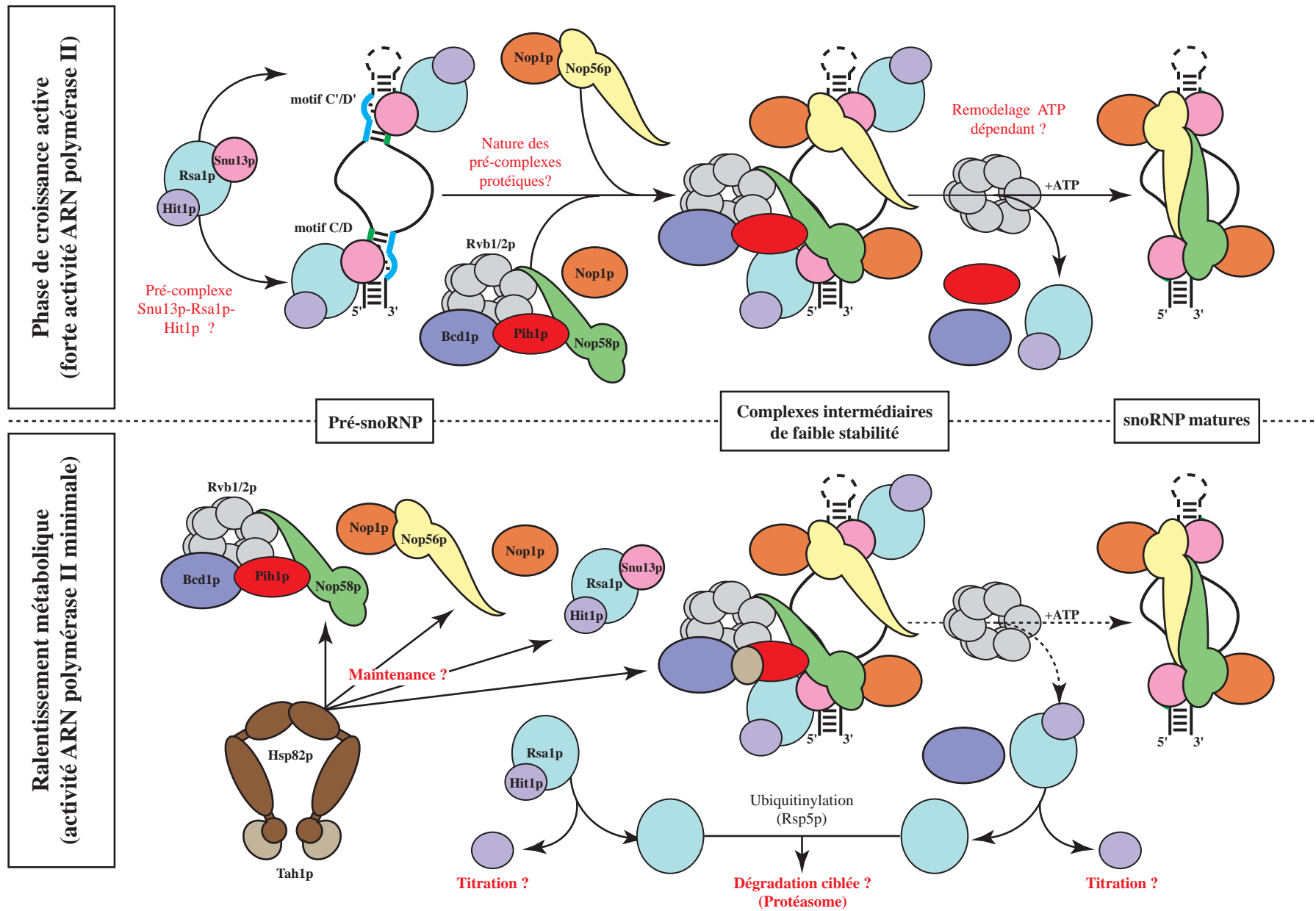


Figure 95 : Modèle proposé pour l'assemblage des snoRNP à boîtes C/D chez la levure *Saccharomyces cerevisiae*. Les contacts entre protéines schématisés dans ce modèle sont le reflet des interactions détectées *in vivo* et *in vitro*. Les annotations en rouge résument les principales questions et perspectives de notre travail.

exister indépendamment de l'ARN (Rothé *et al.*, Publication n°3 ). D'autre part, en utilisant l'approche d'immunoprécipitation SILAC (*Stable isotope labeling with amino acids in cell culture*), les groupes d'E. Bertrand et d'A.I. Lamond ont observé l'association des protéines humaines 15.5K, NUFIP et TRIP3 au sein de complexes cytoplasmiques, supposant la formation du complexe trimérique avant son import dans le noyau et son recrutement sur l'ARN (E. Bertrand, Communication personnelle).

Une fois formée, cette pré-snoRNP sert de socle pour la suite des événements d'assemblage. De nombreuses études ont en effet montré que la fixation de la protéine Snu13p/15.5K sur l'ARN constitue un événement essentiel pour le recrutement des autres protéines, qui sont stabilisées sur la RNP par le biais d'interactions bispécifiques avec l'ARN et avec la protéine de la famille L7Ae (Clery *et al.*, 2007b; Granneman *et al.*, 2002; Nottrott *et al.*, 2002; Schultz *et al.*, 2006a; Schultz *et al.*, 2006b; Watkins *et al.*, 2002). Nos résultats, ainsi que les données obtenues par l'équipe de N.J. Watkins, ont permis de mettre en évidence que la présence du facteur Rsa1p/NUFIP au sein des pré-snoRNP permet d'améliorer de façon considérable la nucléation des RNP (Boulon *et al.*, 2008; McKeegan *et al.*, 2007) (cf. Résultats & Discussion, Chapitre 1, 2.2.2). Chez la levure comme chez l'homme, le facteur d'assemblage présente en effet la capacité de faciliter spécifiquement l'interaction des protéines Nop56p/NOP56 et Nop58p/NOP58 avec le complexe U14-Snu13p/15.5K *in vitro*. Cette propriété pourrait notamment reposer sur une augmentation des surfaces de contacts entre la pré-snoRNP et les protéines Nop56p/NOP56 et Nop58p/NOP58, permettant de stabiliser la particule dans une conformation intermédiaire (**Fig. 95**). Par ailleurs, la comparaison entre la structure de la sRNP à boîtes C/D obtenue chez l'archée *S. solfataricus* (Lin *et al.*, 2011) et notre modèle de l'interface Snu13p/15.5K-Rsa1p/NUFIP (Rothé, Back *et al.*, Publication n°2), suggère que les protéines à domaine NOP et le facteur d'assemblage occupent des sites de fixation chevauchant au niveau de la protéine Snu13p/15.5K (cf. Résultats & Discussion, Chapitre 2). Cette observation, qui nécessitera une confirmation expérimentale, pourrait expliquer la différence de stabilité observée entre les snoRNP matures et les pré-snoRNP contenant les protéines NOP56 et NOP58 humaines (McKeegan *et al.*, 2009; Watkins *et al.*, 2004). En effet, notre hypothèse est que l'activité de remodelage portée par les hélicases Rvb1p/TIP49 et Rvb2p/TIP48 ait pour principal effet de déplacer le facteur Rsa1p/NUFIP de son site d'interaction, afin de permettre le positionnement correct des protéines à domaine NOP au sein des snoRNP matures (**Fig. 95**).

Récemment, des études structurales et fonctionnelles ont conduit à des avancées majeures concernant l'organisation des snoRNP à boîtes C/D (Lin *et al.*, 2011; Qu *et al.*, 2011) (cf. Introduction, Chapitre 2, 3.1.1.3). Ces travaux ont notamment abouti à la proposition d'un modèle d'assemblage pseudo-symétrique reposant sur la fixation d'un jeu de protéines Snu13p/15.5K-Nop58p/NOP58-Nop1p/Fibrillarine au niveau du motif C/D, et d'un jeu de protéines Snu13p/15.5K-Nop56p/NOP56-Nop1p/Fibrillarine au niveau du motif C'/D' (Qu *et al.*, 2011). La prise en compte de ces données sera donc essentielle dans le cadre de la poursuite de nos recherches. En effet, d'après les résultats obtenus par les groupes de N.J Watkins et E.S. Maxwell, le caractère sous-optimal des motifs C'/D', qui constituent généralement des sites de fixation de faible affinité pour la protéine Snu13p/15.5K, est compensé par la présence du motif C/D (Qu *et al.*, 2011). Comme cela est envisagé pour la particule C'/D de la snoRNP U3 (Knox *et al.*, 2011) (Travail de thèse N. Rolland), l'assemblage des particules C'/D' sous une forme stable semble donc constituer un événement secondaire, nécessitant la formation préalable de la particule C/D et l'établissement d'un dimère entre les domaines *Coiled-coil* des protéines Nop56p/NOP56 et Nop58p/NOP58 (Lin *et al.*, 2011; Qu *et al.*, 2011). Dans ce cadre, un repositionnement des protéines Nop56p/NOP56 et Nop58p/NOP58 induit par les hélicases pourrait conduire à l'acquisition d'une conformation de haute stabilité, en favorisant la formation de ce dimère et en contribuant au raccordement physique des particules C/D et C'/D' (**Fig. 95**).

Une question importante qui demeure posée, concerne les modalités de prise en charge et d'acheminement des protéines Nop1p/FIB, Nop56p/NOP56 et Nop58p/NOP58, jusqu'aux sites d'assemblage des particules (**Fig. 95**). Chez la levure, Nop58p fait l'objet d'une association spécifique avec la protéine Pih1p, qui pourrait se mettre en place de façon précoce et constituer la forme d'incorporation de la protéine au sein de la RNP ((Gonzales *et al.*, 2005) ; Rothé *et al.*, Publication n°3). A la différence de Nop58p, la protéine Nop56p n'est pas associée à Pih1p et son assemblage sur le snoRNA s'est révélé dépendant de la protéine Nop1p (Lafontaine & Tollervey, 2000). Des données récentes ont par ailleurs démontré que les protéines humaines NOP56 et Fibrillarine sont associées *in cellulo* indépendamment de leur présence au sein des snoRNP (Lechertier *et al.*, 2009). Par analogie avec le couple Nop58p/NOP58-Pih1p/PIH1, il est donc possible que la protéine Nop56p/NOP56 soit prise en charge et stabilisée par la 2'-O-méthylase, jusqu'à l'étape d'assemblage sur le motif C'/D'. L'assemblage de la protéine Nop1p/Fibrillarine au niveau du motif C/D, qui dans le cas du snoRNA U14 de levure, s'est révélé indépendant du facteur Rsa1p, pourrait en revanche

s'effectuer de façon relativement autonome (cf. Résultats & Discussion, Chapitre 1). En complément des structures du complexe FIB-NOP5 obtenues chez les archées (Aittaleb *et al.*, 2003; Lin *et al.*, 2011; Oruganti *et al.*, 2007; Xue *et al.*, 2010; Ye *et al.*, 2009), l'étude du complexe Pih1p<sub>1-250</sub>-Nop58p<sub>159-412</sub>, débutée au laboratoire par le groupe de X. Manival, constitue donc une perspective intéressante. L'acquisition de données structurales par radiocristallographie pourrait en effet nous permettre d'identifier les bases moléculaires de ces différences et ainsi de mieux comprendre les spécificités liées à l'assemblage de chacune de ces protéines.

### 3. Assemblage des snoRNP à boîtes C/D et régulations

Outre les aspects directement liés au mécanisme d'assemblage des snoRNP, nos travaux chez la levure, nous ont également conduits à mettre à jour un environnement régulateur potentiel (**Fig. 95**). La protéine Rsa1p, qui est un substrat privilégié de l'ubiquitine ligase E3 Rsp5p (Gupta *et al.*, 2007), s'est en effet révélée dépendante de la protéine Hit1p pour sa propre accumulation dans les cellules (Rothé *et al.*, Publication n°3). Etant donné son rôle central dans la formation des particules et son abondance limitée (Ghaemmaghami *et al.*, 2003), Rsa1p constitue une cible de choix pour la mise en place d'un processus de régulation. L'hypothèse d'un système de rétrocontrôle permettant d'ajuster le taux de production des snoRNP à celui de synthèse des ribosomes (cf. Résultats & Discussion, Chapitre 1, 3.2) devra donc être étudiée avec attention.

D'autre part, nous avons également montré que la chaperonne Hsp82p et son cofacteur Tah1p exercent une influence limitée sur le processus d'assemblage des snoRNP chez la levure (cf. Résultats & Discussion, Chapitre 4). En effet, la protéine Tah1p est détectée au sein des pré-snoRNP U3 uniquement durant la phase stationnaire de croissance, au cours de laquelle elle assure un maintien de la stabilité de Pih1p ((Zhao *et al.*, 2008); Rothé *et al.*, Publication n°3). De même, l'intervention de Hsp82p n'a pu être observée que dans des contextes défavorables, liés notamment à des défaillances du processus d'assemblage des snoRNP (Zhao *et al.*, 2008). Chez la levure, Tah1p et Hsp82p semblent donc exercer une fonction de maintenance visant à préserver l'intégrité des acteurs du processus de biogenèse en cas d'arrêt métabolique. Même si leur importance dans des conditions de croissance standard en laboratoire (milieu riche,

30°C, phase exponentielle de croissance) n'a pas clairement été établie, leur intervention dans le cadre de conditions environnementales défavorables doit être prise avec autant de considération. En effet, dans leur milieu naturel les cellules de levures sont soumises à de nombreuses variations qui les poussent à adopter, outre la phase végétative, des états métaboliques variés, comme des états de quiescence, des phases de sporulation ou encore des phases de croissance pseudohyphale (*i.e.* association des cellules sous forme de chaînes invasives leur permettant d'explorer le milieu à la recherche de nouvelles ressources nutritives). Dans ce cadre d'adaptation métabolique constante, les protéines Tah1p et Hsp82p pourraient donc jouer un rôle « tampon », afin de préserver la fonction d'assemblage et de permettre une réactivation rapide de la production des snoRNP lorsque les conditions redeviennent favorables.

#### 4. Spectre d'action de la machinerie d'assemblage chez la levure

Hormis les hélicases Rvb1p et Rvb2p qui sont requises pour l'accumulation des snoRNA à boîtes H/ACA (King *et al.*, 2001; Zhao *et al.*, 2008), les autres membres de la machinerie d'assemblage se sont avérés, chez la levure, spécifiques de l'assemblage des snoRNP à boîtes C/D (**Tableau 9**). En effet, l'absence de facteurs majeurs comme Rsa1p, Pih1p et Bcd1p, n'a aucun effet sur le taux des snoRNA à boîtes H/ACA (Peng *et al.*, 2003) (cf. Résultats & Discussion, Chapitre 3) et la protéine Rsa1p s'est avérée incapable d'interagir avec la protéine Nhp2p (cf. Résultats & Discussion, Chapitre 2). Ces données suggèrent donc que les facteurs Rsa1p, Pih1p et Bcd1p pourraient former, au sein de la machinerie d'assemblage, une composante spécifique des snoRNP à boîtes C/D. Son intervention est en effet guidée par la mise en place d'une interaction entre Rsa1p et Snu13p, qui semble constituer le socle du processus d'assemblage (Rothé, Back *et al.*, Publication n°2).

Etant donné la présence de Snu13p et d'une protéine à domaine NOP (Prp31p) au sein de la snRNP U4, il serait néanmoins intéressant d'établir si, par extension, la machinerie d'assemblage peut également assister la formation de cette particule. Des données préliminaires, obtenues par double-hybride indique une certaine similarité entre le profil d'interaction de Nop58p et de Prp31p, puisque cette dernière interagit également avec Rsa1p et avec Pih1p (Boulon *et al.*, 2008) (E. Bertand, Communication personnelle) (**Tableau 9**).

D'autre part, une interaction spécifique de Prp31p avec la protéine Hit1p a également été identifiée par l'équipe d'E. Bertrand, grâce à l'approche de double-hybride (E. Bertrand, Communication personnelle). De façon intéressante, ces informations suggèrent donc que la machinerie d'assemblage pourrait posséder une double spécificité, permettant de cibler la protéine Snu13p, mais également l'ensemble des protéines à domaine NOP (*i.e.* Nop56p, Nop58p, Prp31p). Néanmoins, en dépit de ces interactions, aucune diminution du taux de snRNA U4 n'a été observé dans les souches de levure  $\Delta RSA1$ ,  $\Delta HIT1$  et  $\Delta PIH1$  (Boulon *et al.*, 2008) (cf. Résultats & Discussion, Chapitre 3) (**Fig. 65**). Une explication de cette apparente stabilité pourrait notamment résider dans la présence d'un anneau de protéine Sm, formant une RNP stable dans la région 3' du snRNA U4 (cf. Introduction, Chapitre 2, 2.1.3) (**Fig. 31**). Afin d'évaluer l'influence réelle de la machinerie d'assemblage sur la biogenèse de cette particule, il serait donc intéressant de vérifier sa fonctionnalité en utilisant un système rapporteur d'épissage dans des souches de levure invalidées pour les gènes *RSA1*, *HIT1* et *PIH1*. De plus, il serait également nécessaire d'analyser l'association entre la protéine Prp31p et le snRNA U4 dans ces différents contextes génétiques, en utilisant les approches de co-immunoprécipitation et de Northern-blot.

## 5. Diversification fonctionnelle de la machinerie d'assemblage chez l'homme

La collaboration que nous avons menée avec le groupe d'E. Bertrand a constitué un atout important dans notre démarche puisque les résultats obtenus par nos deux équipes, respectivement chez la levure et chez l'homme, ont constamment été confrontés et comparés dans les deux systèmes d'étude. Un point majeur qui ressort de ce travail commun est donc la grande conservation de la machinerie d'assemblage puisque l'intégralité des facteurs mis à jour, ainsi qu'une majorité des interactions identifiées, sont conservés entre la levure et l'homme (Boulon *et al.*, 2008) (**Tableau 9**). Une étude récente, réalisée chez *A. thaliana* par l'équipe de M. Echeverria, a également confirmé la conservation de la protéine NUFIP et son implication dans la biogenèse des snoRNP à boîtes C/D chez les plantes (Rodor *et al.*, 2011). Ces différents constats témoignent donc du caractère essentiel et de l'origine ancienne de cette machinerie, qui pourrait avoir émergée de façon concomitante à l'apparition du *phylum* Eucaryote et à la multiplication des fonctions impliquant des RNP (**Tableau 1**). Au même



titre que le complexe SMN ou les facteurs d'assemblage des snoRNP H/ACA, la machinerie que nous avons caractérisée aurait ainsi pu contribuer à répondre à des problèmes d'ordre spatio-temporels et à des contraintes de cinétique et de spécificité, afin de préserver, en dépit de la complexification de l'environnement cellulaire, l'efficacité d'assemblage de certaines RNP essentielles.

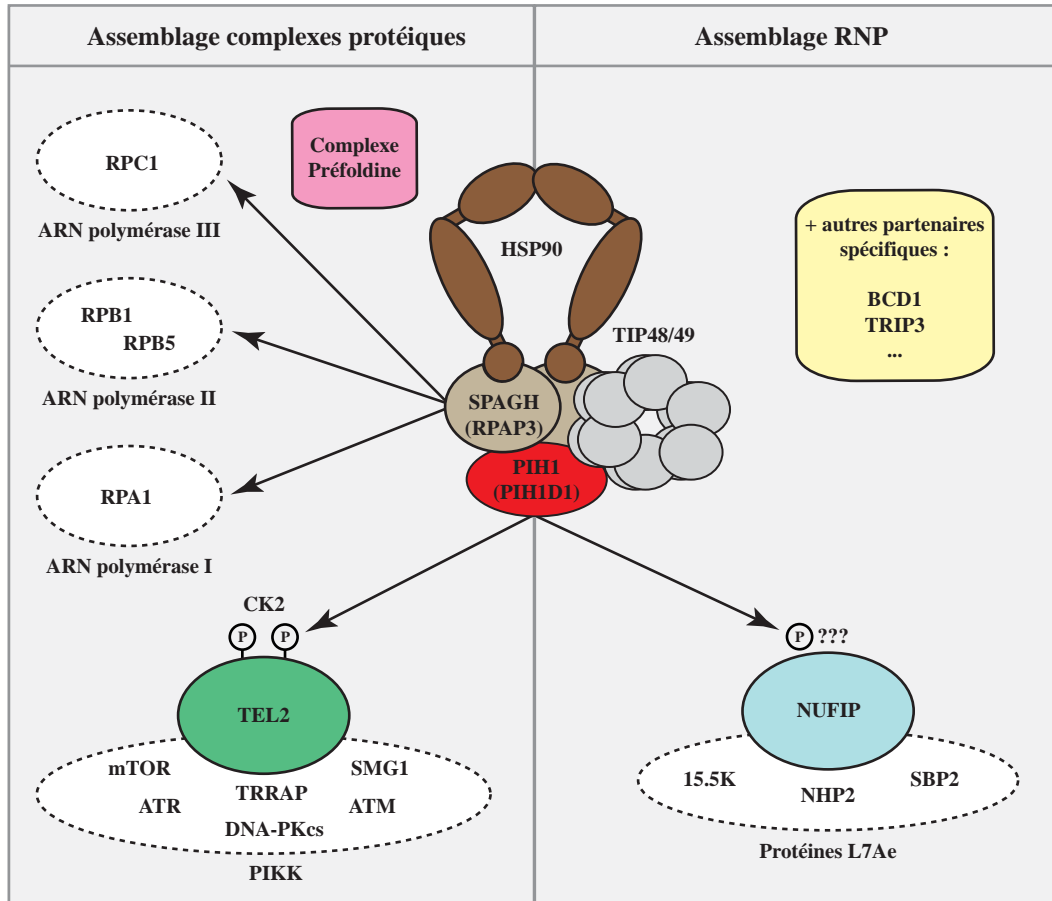
Parmi ces particules, les snoRNP, qui sont également retrouvées chez les archées sous l'appellation de sRNP, font figures de représentants ancestraux. En revanche d'autres RNP appartenant également à la famille L7Ae/L30 sont vraisemblablement apparues de façon secondaire (cf. Introduction, Chapitre 2, 1.4) (**Fig. 28**). C'est probablement le cas de la snRNP U4, retrouvée uniquement chez les eucaryotes, mais également de la composante H/ACA du complexe télomérase et des mRNP SECIS contenant la protéine SBP2, qui ont été identifiés sélectivement dans certains groupes d'eucaryotes. Sur la base de l'homologie structurale entre les protéines de la famille L7Ae/L30, l'émergence de ces nouvelles RNP s'est donc vraisemblablement accompagnée d'une diversification fonctionnelle de la machinerie d'assemblage. En effet, les études réalisées chez l'homme ont permis d'établir son implication dans l'assemblage de la snRNP U4, mais également dans la biogenèse des snoRNP à boîtes H/ACA et du complexe télomérase, ainsi que dans la formation des mRNP contenant la protéine SBP2 (Boulon *et al.*, 2008; Venteicher *et al.*, 2008) (cf. Résultats & Discussion, Chapitre 1). D'autre part, si l'implication de la chaperonne Hsp82p et de son cofacteur Tah1p s'est avérée restreinte à certaines conditions de stress chez la levure, les protéines HSP90 et SPAGH humaines constituent des acteurs à part entière du processus d'assemblage et sont requises pour la biogenèse des RNP contenant les protéines 15.5K, NHP2 et SBP2 (Boulon *et al.*, 2008).

Les différences constatées entre les attributs fonctionnels de la machinerie d'assemblage chez la levure et chez l'homme témoignent donc d'une réelle dynamique d'adaptation et d'un processus de co-évolution des RNP et de leur système d'assemblage. Par le biais de sa capacité à interagir à la fois avec les protéines 15.5K, NHP2 et SBP2, la protéine NUFIP constitue vraisemblablement un élément clé de ce phénomène (Boulon *et al.*, 2008). De part son rôle de protéine plateforme, cette dernière apparaît en effet comme un adaptateur, permettant de spécifier l'intervention de la machinerie d'assemblage dans la biogenèse des RNP de la famille L7Ae. En marge de cette fonction et de son lien avec NUFIP, le complexe R2TP pourrait constituer quant à lui un module d'assemblage au spectre d'action relativement

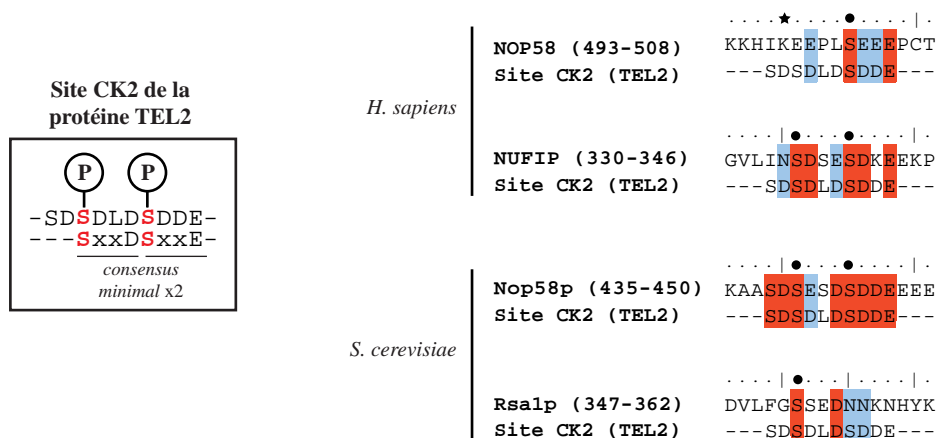
large. Plusieurs études récentes ont en effet mis en évidence son implication dans l'assemblage d'autres complexes macromoléculaires majeurs comme les ARN polymérases (Boulon *et al.*, 2010; Cloutier *et al.*, 2009; Jeronimo *et al.*, 2007) et les complexes contenant des kinases de la famille des PIKK (*Phosphatidylinositol 3-kinase-related kinase*) (Horejsi *et al.*, 2010; Izumi *et al.*, 2010) (**Fig. 96A**). Dans ce cadre fonctionnel, le complexe R2TP agit en coopération avec HSP90, mais il s'associe également avec une autre catégorie de protéines chaperonnes ubiquitaires, les préfoldines (Cloutier *et al.*, 2009; Jeronimo *et al.*, 2007). Le complexe R2TP/préfoldine ainsi formé renferme, en plus des membres du complexe R2TP, la protéine Monad (alias WDR92), les *prefoldins* PFDN2 et PFDN6, ainsi que les *prefoldins-like* UXT (alias ART-27), PDRG1 et URI (alias RMP ou C19orf2) (Cloutier *et al.*, 2009; Itsuki *et al.*, 2008; Jeronimo *et al.*, 2007).

De façon analogue au modèle établi pour l'assemblage des RNP de la famille L7Ae, l'implication du complexe R2TP/préfoldines dans la biogenèse des complexes PIKK semble également reposer sur une molécule adaptatrice, la protéine TEL2, capable d'interagir spécifiquement avec la protéine PIH1 et avec plusieurs membres de la famille des PIKK (*i.e.* mTOR, SMG1, ATM, ATR, DNA-PKcs et TRRAP) (Horejsi *et al.*, 2010; Takai *et al.*, 2007; Takai *et al.*, 2010) (**Fig. 96A**). Parmi ces différentes kinases, les taux des protéines mTOR et SMG1 se sont révélés particulièrement sensibles à un défaut d'association entre TEL2 et PIH1, soulignant le rôle clé du complexe R2TP/préfoldines dans la mise en place des fonctions de signalisation qui leurs sont associées (*i.e.* respectivement traduction/croissance cellulaire et *Non-sens Mediated Decay*) (Horejsi *et al.*, 2010). De plus, ce rôle spécifique du complexe R2TP/Préfoldine dans l'assemblage des complexes PIKK semble sujet à un phénomène de régulation, puisque la phosphorylation de deux sérines successives, par la caséine kinase CK2 dans la séquence 485-SDSDLSDDE-494 de la protéine TEL2, s'est avérée essentielle pour son interaction avec PIH1 (Horejsi *et al.*, 2010). Cette enzyme est une *seryl/threonyl kinase* ubiquitaire, impliquée dans de nombreuses fonctions essentielles, dont la régulation de l'activité chaperonne de HSP90 (Miyata, 2009). De façon intéressante, au cours de notre travail nous avons montré l'existence d'un lien génétique entre le gène *RSA1* et les ORF codant les sous-unités Cka2p et Ckb2p du complexe CK2 de levure (cf. Résultats & Discussion, Chapitre 3, 2.2.4) (**Tableau 6**). De plus, nos résultats indiquent également que ces deux sous-unités sont co-immunoséctionnées avec les protéines Rsa1p, Hit1p et Pih1p, suggérant une association possible entre la kinase et les facteurs d'assemblage (**Tableaux 6 et 7**) (Rothé *et al.*, Publication n°3).

**A**



**B**



**Figure 96 : Bilan des fonctions d'assemblage de complexes macromoléculaires impliquant le complexe R2TP et HSP90 chez l'homme.** (A) Schéma représentant les rôles de HSP90 et du complexe R2TP dans l'assemblage de complexes macromoléculaires (Hořejší *et al.*, 2010; Boulon *et al.*, 2010; Jeronimo *et al.*, 2007; Boulon *et al.*, 2008; Mc Keegan *et al.*, 2007). (Schéma adapté de Hořejší *et al.*, 2010). (B) Mise en évidence de sites potentiels de phosphorylation par la kinase CK2 au niveau des partenaires de la protéine Pih1p/PIH1. L'alignement a été réalisé par ClustalW Multiple Alignment à l'aide du logiciel BioEdit (Hall, T.A. 1999). Les acides aminés identiques sont représentés sur fond rouge. Les acides aminés similaires sont représentés sur fond bleu clair. Les sérines potentiellement ciblées par CK2 sont indiquées par une pastille noire. L'étoile noire fait référence à la sumoylation du résidu K497 de la protéine NOP58, qui s'est avérée dépendante de la phosphorylation du résidu S502 (Matic *et al.*, 2010).

Le dénominateur commun des substrats de la kinase CK2 est un motif consensus S/T-x-x-E/D, incluant généralement des charges négatives additionnelles dans son environnement (Meggio *et al.*, 1994). De façon intéressante, la protéine NOP58 humaine présente dans sa région C-terminale riche en résidus KKE/D, un motif 502-SEED-505 faisant également l'objet d'une phosphorylation par la kinase CK2 (Matic *et al.*, 2010) (cf. Introduction, Chapitre 3, 3.3). Etant donné que l'intervention de cette enzyme au niveau de la protéine TEL2 permet de créer une surface d'interaction favorable à la protéine PIH1, il est possible qu'une telle modification puisse également réguler l'interaction du facteur d'assemblage avec d'autres partenaires comme NOP58. Afin de vérifier cette éventualité nous avons donc recherché la présence de séquences homologues au motif SDSDLDSDE dans les protéines Rsa1p/NUFIP et Nop58p/NOP58, qui constituent les deux principaux partenaires de Pih1p/PIH1 dans le cadre de la biogenèse des snoRNP à boîtes C/D (**Fig. 96B**). Outre le motif déjà identifié dans la protéine NOP58 humaine, nous avons observé la présence d'un double site de phosphorylation potentiel dans la région C-terminal de la protéine Nop58p de levure. De même, un et deux sites de phosphorylation potentiels ont respectivement été détectés dans la région C-terminale des facteurs d'assemblage Rsa1p et NUFIP.

Concernant les protéines de levure, les résultats obtenus au cours de notre travail indiquent que le fragment C-terminal (230-381) de Rsa1p, qui contient ce motif, est suffisant pour établir une interaction stable avec Pih1p par co-expression chez *E. coli* (cf. Résultats Discussion, Chapitre 3, 2.2.3.2) (**Fig. 94**). En revanche, ce motif ne semble pas requis pour l'association de Nop58p avec Pih1p puisqu'une interaction entre le complexe <sup>His6</sup>Tah1p-Pih1p et la protéine Nop58p<sub>1-377</sub> obtenue par transcription/traduction dans un extrait bactérien, a été observée par pull-down (cf. Résultats & Discussion, Chapitre 3, 2.2.3.4) (**Fig. 94**). A défaut de constituer un déterminant essentiel de cette interaction, il est néanmoins envisageable qu'un tel motif puisse contribuer à la discrimination des protéines ciblées par Pih1p/PIH1 et par le complexe R2TP *in vivo*. A ce titre, une analyse protéomique réalisée par les groupes d'E. Bertrand et d'A.I. Lamond a révélé une différence majeure entre NOP56 et NOP58, puisque cette dernière constitue l'une des protéines nucléolaires les plus fortement sumoylées (Westman *et al.*, 2010) (cf. Introduction; Chapitre 3; 3.3). De façon remarquable, cette caractéristique pourrait avoir des conséquences importantes sur le mode d'assemblage différentiel de ces deux protéines, puisque la sumoylation des résidus K467 et K497 de NOP58 est requise pour son association optimale avec les snoRNA à boîtes C/D U3, U8, U13 et U14 *in vivo* (Westman *et al.*, 2010). Par ailleurs, une autre étude a également révélé que la

phosphorylation de la sérine en position 502 par la kinase CK2 constitue un pré-requis nécessaire à la sumoylation du résidu K497, suggérant l'existence d'un couplage entre ces événements de modification (Matic *et al.*, 2010). Ces différents résultats soulèvent donc la question d'un lien entre l'état de modification de NOP58, son association avec le complexe R2TP et son incorporation au sein des snoRNP. Outre le jeu d'interaction mis à jour entre les facteurs d'assemblage et les composants des RNP, il est probable que la modification post-traductionnelle de certaines protéines contribue à réguler l'efficacité et/ou à la spécificité du mécanisme d'assemblage. La prise en compte de ces phénomènes constitue donc une perspective nouvelle, qui par la suite, pourrait s'avérer déterminante dans notre façon d'aborder et d'appréhender la complexité des processus de biogenèse des RNP de la famille L7Ae.



REFERENCES  
BIBLIOGRAPHIQUES





- Abou Elela S, Ares M, Jr. (1998) Depletion of yeast RNase III blocks correct U2 3' end formation and results in polyadenylated but functional U2 snRNA. *EMBO J* **17**(13): 3738-3746
- Ach RA, Weiner AM (1987) The highly conserved U small nuclear RNA 3'-end formation signal is quite tolerant to mutation. *Mol Cell Biol* **7**(6): 2070-2079
- Achsel T, Brahm H, Kastner B, Bachi A, Wilm M, Luhrmann R (1999) A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J* **18**(20): 5789-5802
- Ahn SH, Kim M, Buratowski S (2004) Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol Cell* **13**(1): 67-76
- Aittaleb M, Rashid R, Chen Q, Palmer JR, Daniels CJ, Li H (2003) Structure and function of archaeal box C/D sRNP core proteins. *Nat Struct Biol* **10**(4): 256-263
- Ali MM, Roe SM, Vaughan CK, Meyer P, Panaretou B, Piper PW, Prodromou C, Pearl LH (2006) Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex. *Nature* **440**(7087): 1013-1017
- Allmang C, Carbon P, Krol A (2002) The SBP2 and 15.5 kD/Snu13p proteins share the same RNA binding domain: identification of SBP2 amino acids important to SECIS RNA binding. *RNA* **8**(10): 1308-1318
- Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D (1999) Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J* **18**(19): 5399-5410
- Allmang C, Mitchell P, Petfalski E, Tollervey D (2000) Degradation of ribosomal RNA precursors by the exosome. *Nucleic Acids Res* **28**(8): 1684-1691
- Amiri KA (1994) Fibrillar-like proteins occur in the domain Archaea. *J Bacteriol* **176**(7): 2124-2127
- Amrani N, Ganesan R, Kervestin S, Mangus DA, Ghosh S, Jacobson A (2004) A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* **432**(7013): 112-118
- Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M (2005) Nucleolar proteome dynamics. *Nature* **433**(7021): 77-83
- Andersen JS, Lyon CE, Fox AH, Leung AK, Lam YW, Steen H, Mann M, Lamond AI (2002) Directed proteomic analysis of the human nucleolus. *Curr Biol* **12**(1): 1-11
- Andrade LE, Tan EM, Chan EK (1993) Immunocytochemical analysis of the coiled body in the cell cycle and during cell proliferation. *Proc Natl Acad Sci U S A* **90**(5): 1947-1951
- Angrisani A, Turano M, Paparo L, Di Mauro C, Furia M (2011) A new human dyskerin isoform with cytoplasmic localization. *Biochim Biophys Acta* **1810**(12): 1361-1368
- Antal M, Boros E, Solymosy F, Kiss T (2002) Analysis of the structure of human telomerase RNA in vivo. *Nucleic Acids Res* **30**(4): 912-920
- Antal M, Mouglin A, Kis M, Boros E, Steger G, Jakab G, Solymosy F, Branlant C (2000) Molecular characterization at the RNA and gene levels of U3 snoRNA from a unicellular green alga, *Chlamydomonas reinhardtii*. *Nucleic Acids Res* **28**(15): 2959-2968
- Anthony JG, Weidenhammer EM, Woolford JL, Jr. (1997) The yeast Prp3 protein is a U4/U6 snRNP protein necessary for integrity of the U4/U6 snRNP and the U4/U6.U5 tri-snRNP. *RNA* **3**(10): 1143-1152
- Aravind L, Koonin EV (1999) Novel predicted RNA-binding domains associated with the translation machinery. *J Mol Evol* **48**(3): 291-302
- Arcus VL, Backbro K, Roos A, Daniel EL, Baker EN (2004) Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J Biol Chem* **279**(16): 16471-16478
- Armanios M, Chen JL, Chang YP, Brodsky RA, Hawkins A, Griffin CA, Eshleman JR, Cohen AR, Chakravarti A, Hamosh A, Greider CW (2005) Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. *Proc Natl Acad Sci U S A* **102**(44): 15960-15964
- Armistead J, Khatkar S, Meyer B, Mark BL, Patel N, Coghlan G, Lamont RE, Liu S, Wiechert J, Cattini PA, Koetter P, Wrogemann K, Greenberg CR, Entian KD, Zelinski T, Triggs-Raine B (2009) Mutation of a gene essential for ribosome biogenesis, EMG1, causes Bowen-Conradi syndrome. *Am J Hum Genet* **84**(6): 728-739
- Askjaer P, Bachi A, Wilm M, Bischoff FR, Weeks DL, Ogniewski V, Ohno M, Niehrs C, Kjems J, Mattaj JW, Fornerod M (1999) RanGTP-regulated interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase. *Mol Cell Biol* **19**(9): 6276-6285
- Ayadi L, Callebaut I, Saguez C, Villa T, Moron JP, Banroques J (1998) Functional and structural characterization of the prp3 binding domain of the yeast prp4 splicing factor. *J Mol Biol* **284**(3): 673-687
- Azzam R, Chen SL, Shou W, Mah AS, Alexandru G, Nasmyth K, Annan RS, Carr SA, Deshaies RJ (2004) Phosphorylation by cyclin B-Cdk underlies release of mitotic exit activator Cdc14 from the nucleolus. *Science* **305**(5683): 516-519
- Baccon J, Pellizzoni L, Rappsilber J, Mann M, Dreyfuss G (2002) Identification and characterization of Gemin7, a novel component of the survival of motor neuron complex. *J Biol Chem* **277**(35): 31957-31962
- Bachand F, Boisvert FM, Cote J, Richard S, Autexier C (2002) The product of the survival of motor neuron (SMN) gene is a human telomerase-associated protein. *Mol Biol Cell* **13**(9): 3192-3202

- Bachelier JP, Cavaille J (1997) Guiding ribose methylation of rRNA. *Trends Biochem Sci* **22**(7): 257-261
- Bachelier JP, Michot B, Nicoloso M, Balakin A, Ni J, Fournier MJ (1995) Antisense snoRNAs: a family of nucleolar RNAs with long complementarities to rRNA. *Trends Biochem Sci* **20**(7): 261-264
- Baillat D, Hakimi MA, Naar AM, Shilatifard A, Cooch N, Shiekhattar R (2005) Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**(2): 265-276
- Baker DL, Youssef OA, Chastkofsky MI, Dy DA, Terns RM, Terns MP (2005) RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP. *Genes Dev* **19**(10): 1238-1248
- Bakin A, Lane BG, Ofengand J (1994) Clustering of pseudouridine residues around the peptidyltransferase center of yeast cytoplasmic and mitochondrial ribosomes. *Biochemistry* **33**(45): 13475-13483
- Balakin AG, Smith L, Fournier MJ (1996) The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell* **86**(5): 823-834
- Ballarino M, Morlando M, Pagano F, Fatica A, Bozzoni I (2005) The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol* **25**(13): 5396-5403
- Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**(5481): 905-920
- Banroques J, Abelson JN (1989) PRP4: a protein of the yeast U4/U6 small nuclear ribonucleoprotein particle. *Mol Cell Biol* **9**(9): 3710-3719
- Bardoni B, Schenck A, Mandel JL (1999) A novel RNA-binding nuclear protein that interacts with the fragile X mental retardation (FMR1) protein. *Hum Mol Genet* **8**(13): 2557-2566
- Bardoni B, Willemsen R, Weiler I, Schenck A, Severijnen L, Hindelang C, Lalli E, Mandel J (2003) NUFIP1 (nuclear FMRP interacting protein 1) is a nucleocytoplasmic shuttling protein associated with active synaptosomes. *Experimental Cell Research* **289**(1): 95-107
- Baserga SJ, Yang XD, Steitz JA (1991) An intact Box C sequence in the U3 snRNA is required for binding of fibrillarin, the protein common to the major family of nucleolar snRNPs. *EMBO J* **10**(9): 2645-2651
- Bassler J, Grandi P, Gadal O, Lessmann T, Petfalski E, Tollervey D, Lechner J, Hurt E (2001) Identification of a 60S preribosomal particle that is closely linked to nuclear export. *Mol Cell* **8**(3): 517-529
- Battle DJ, Lau CK, Wan L, Deng H, Lotti F, Dreyfuss G (2006) The Gemin5 protein of the SMN complex identifies snRNAs. *Mol Cell* **23**(2): 273-279
- Baudin-Baillieu A, Fabret C, Liang XH, Piekna-Przybylska D, Fournier MJ, Rousset JP (2009) Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. *Nucleic Acids Res* **37**(22): 7665-7677
- Bauer A, Chauvet S, Huber O, Usseglio F, Rothbacher U, Aragnol D, Kemler R, Pradel J (2000) Pontin52 and reptin52 function as antagonistic regulators of beta-catenin signalling activity. *EMBO J* **19**(22): 6121-6130
- Bauer A, Huber O, Kemler R (1998) Pontin52, an interaction partner of beta-catenin, binds to the TATA box binding protein. *Proc Natl Acad Sci U S A* **95**(25): 14787-14792
- Bauer DW, Gall JG (1997) Coiled bodies without coilin. *Mol Biol Cell* **8**(1): 73-82
- Bax R, Vos HR, Raue HA, Vos JC (2006) *Saccharomyces cerevisiae* Sof1p associates with 35S Pre-rRNA independent from U3 snoRNA and Rrp5p. *Eukaryot Cell* **5**(3): 427-434
- Baxter-Roshek JL, Petrov AN, Dinman JD (2007) Optimization of ribosome structure and function by rRNA base modification. *PLoS One* **2**(1): e174
- Beaudoing E, Gautheret D (2001) Identification of alternate polyadenylation sites and analysis of their tissue distribution using EST data. *Genome Res* **11**(9): 1520-1526
- Behm-Ansmant I, Gatfield D, Rehwinkel J, Hilgers V, Izaurralde E (2007) A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. *EMBO J* **26**(6): 1591-1601
- Bell M, Schreiner S, Damianov A, Reddy R, Bindereif A (2002) p110, a novel human U6 snRNP protein and U4/U6 snRNP recycling factor. *EMBO J* **21**(11): 2724-2735
- Bellodi C, Kindle K, Bernassola F, Dinsdale D, Cossarizza A, Melino G, Heery D, Salomoni P (2006) Cytoplasmic function of mutant promyelocytic leukemia (PML) and PML-retinoic acid receptor-alpha. *J Biol Chem* **281**(20): 14465-14473
- Bellosta P, Hulf T, Balla Diop S, Usseglio F, Pradel J, Aragnol D, Gallant P (2005) Myc interacts genetically with Tip48/Reptin and Tip49/Pontin to control growth and proliferation during *Drosophila* development. *Proc Natl Acad Sci U S A* **102**(33): 11799-11804
- Beltrame M, Henry Y, Tollervey D (1994) Mutational analysis of an essential binding site for the U3 snoRNA in the 5' external transcribed spacer of yeast pre-rRNA. *Nucleic Acids Res* **22**(20): 4057-4065
- Beltrame M, Tollervey D (1992) Identification and functional analysis of two U3 binding sites on yeast pre-ribosomal RNA. *EMBO J* **11**(4): 1531-1542
- Beltrame M, Tollervey D (1995) Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis. *EMBO J* **14**(17): 4350-4356

- Ben-Shem A, Jenner L, Yusupova G, Yusupov M (2010) Crystal structure of the eukaryotic ribosome. *Science* **330**(6008): 1203-1209
- Berg OG, von Hippel PH (1985) Diffusion-controlled macromolecular interactions. *Annu Rev Biophys Biophys Chem* **14**: 131-160
- Bernardi R, Pandolfi PP (2003) Role of PML and the PML-nuclear body in the control of programmed cell death. *Oncogene* **22**(56): 9048-9057
- Bertrand E, Houser-Scott F, Kendall A, Singer RH, Engelke DR (1998) Nucleolar localization of early tRNA processing. *Genes Dev* **12**(16): 2463-2468
- Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F, Parker SC, Sabo PJ, Sandstrom R, Shafer A, Vetrie D, Weaver M, Wilcox S, Yu M, Collins FS, Dekker J, Lieb JD, Tullius TD, Crawford GE, Sunyaev S, Noble WS, Dunham I, Denoeud F, Reymond A, Kapranov P, Rozowsky J, Zheng D, Castelo R, Frankish A, Harrow J, Ghosh S, Sandelin A, Hofacker IL, Baertsch R, Keefe D, Dike S, Cheng J, Hirsch HA, Sekinger EA, Lagarde J, Abril JF, Shahab A, Flamm C, Fried C, Hackermuller J, Hertel J, Lindemeyer M, Missal K, Tanzer A, Washietl S, Korbel J, Emanuelsson O, Pedersen JS, Holroyd N, Taylor R, Swarbreck D, Matthews N, Dickson MC, Thomas DJ, Weirauch MT, Gilbert J, Drenkow J, Bell I, Zhao X, Srinivasan KG, Sung WK, Ooi HS, Chiu KP, Foissac S, Alioto T, Brent M, Pachter L, Tress ML, Valencia A, Choo SW, Choo CY, Ucla C, Manzano C, Wyss C, Cheung E, Clark TG, Brown JB, Ganesh M, Patel S, Tammana H, Chrast J, Henrichsen CN, Kai C, Kawai J, Nagalakshmi U, Wu J, Lian Z, Lian J, Newburger P, Zhang X, Bickel P, Mattick JS, Carninci P, Hayashizaki Y, Weissman S, Hubbard T, Myers RM, Rogers J, Stadler PF, Lowe TM, Wei CL, Ruan Y, Struhl K, Gerstein M, Antonarakis SE, Fu Y, Green ED, Karaoz U, Siepel A, Taylor J, Liefer LA, Wetterstrand KA, Good PJ, Feingold EA, Guyer MS, Cooper GM, Asimenos G, Dewey CN, Hou M, Nikolaev S, Montoya-Burgos JI, Loytynoja A, Whelan S, Pardi F, Massingham T, Huang H, Zhang NR, Holmes I, Mullikin JC, Ureta-Vidal A, Paten B, Seringhaus M, Church D, Rosenbloom K, Kent WJ, Stone EA, Batzoglu S, Goldman N, Hardison RC, Haussler D, Miller W, Sidow A, Trinklein ND, Zhang ZD, Barrera L, Stuart R, King DC, Ameur A, Enroth S, Bieda MC, Kim J, Bhing AA, Jiang N, Liu J, Yao F, Vega VB, Lee CW, Ng P, Yang A, Moqtaderi Z, Zhu Z, Xu X, Squazzo S, Oberley MJ, Inman D, Singer MA, Richmond TA, Munn KJ, Rada-Iglesias A, Wallerman O, Komorowski J, Fowler JC, Couttet P, Bruce AW, Dovey OM, Ellis PD, Langford CF, Nix DA, Euskirchen G, Hartman S, Urban AE, Kraus P, Van Calcar S, Heintzman N, Kim TH, Wang K, Qu C, Hon G, Luna R, Glass CK, Rosenfeld MG, Aldred SF, Cooper SJ, Halees A, Lin JM, Shulha HP, Xu M, Haidar JN, Yu Y, Iyer VR, Green RD, Wadelius C, Farnham PJ, Ren B, Harte RA, Hinrichs AS, Trumbower H, Clawson H, Hillman-Jackson J, Zweig AS, Smith K, Thakkapallayil A, Barber G, Kuhn RM, Karolchik D, Armengol L, Bird CP, de Bakker PI, Kern AD, Lopez-Bigas N, Martin JD, Stranger BE, Woodroffe A, Davydov E, Dimas A, Eyras E, Hallgrimsdottir IB, Huppert J, Zody MC, Abecasis GR, Estivill X, Bouffard GG, Guan X, Hansen NF, Idol JR, Maduro VV, Maskeri B, McDowell JC, Park M, Thomas PJ, Young AC, Blakesley RW, Muzny DM, Sodergren E, Wheeler DA, Worley KC, Jiang H, Weinstock GM, Gibbs RA, Graves T, Fulton R, Mardis ER, Wilson RK, Clamp M, Cuff J, Gnerre S, Jaffe DB, Chang JL, Lindblad-Toh K, Lander ES, Koriabine M, Nefedov M, Osoegawa K, Yoshinaga Y, Zhu B, de Jong PJ (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**(7146): 799-816
- Biswas S, Buhman G, Gagnon K, Mattos C, Brown BA, 2nd, Maxwell ES (2011) Comparative analysis of the 15.5kD box C/D snoRNP core protein in the primitive eukaryote *Giardia lamblia* reveals unique structural and functional features. *Biochemistry* **50**(14): 2907-2918
- Bjorn SP, Soltyk A, Beggs JD, Friesen JD (1989) PRP4 (RNA4) from *Saccharomyces cerevisiae*: its gene product is associated with the U4/U6 small nuclear ribonucleoprotein particle. *Mol Cell Biol* **9**(9): 3698-3709
- Bleichert F, Gagnon KT, Brown BA, 2nd, Maxwell ES, Leschziner AE, Unger VM, Baserga SJ (2009) A dimeric structure for archaeal box C/D small ribonucleoproteins. *Science* **325**(5946): 1384-1387
- Bleichert F, Granneman S, Osheim YN, Beyer AL, Baserga SJ (2006) The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proc Natl Acad Sci U S A* **103**(25): 9464-9469
- Bohnsack MT, Kos M, Tollervey D (2008) Quantitative analysis of snoRNA association with pre-ribosomes and release of snR30 by Rok1 helicase. *EMBO Rep* **9**(12): 1230-1236
- Bohnsack MT, Martin R, Granneman S, Ruprecht M, Schleiff E, Tollervey D (2009) Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. *Mol Cell* **36**(4): 583-592
- Boisvert FM, van Koningsbruggen S, Navascues J, Lamond AI (2007) The multifunctional nucleolus. *Nat Rev Mol Cell Biol* **8**(7): 574-585
- Bond CS, Fox AH (2009) Paraspeckles: nuclear bodies built on long noncoding RNA. *J Cell Biol* **186**(5): 637-644
- Borkovich KA, Farrelly FW, Finkelstein DB, Taulien J, Lindquist S (1989) hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. *Mol Cell Biol* **9**(9): 3919-3930
- Borovjagin AV, Gerbi SA (1999) U3 small nucleolar RNA is essential for cleavage at sites 1, 2 and 3 in pre-rRNA and determines which rRNA processing pathway is taken in *Xenopus* oocytes. *J Mol Biol* **286**(5): 1347-1363

- Borovjagin AV, Gerbi SA (2000) The spacing between functional Cis-elements of U3 snoRNA is critical for rRNA processing. *J Mol Biol* **300**(1): 57-74
- Bortolin ML, Ganot P, Kiss T (1999) Elements essential for accumulation and function of small nucleolar RNAs directing site-specific pseudouridylation of ribosomal RNAs. *EMBO J* **18**(2): 457-469
- Boudonck K, Dolan L, Shaw PJ (1999) The movement of coiled bodies visualized in living plant cells by the green fluorescent protein. *Mol Biol Cell* **10**(7): 2297-2307
- Boulon S (2005) Trafic et biogenèse des petits ARNs nucléolaires à boîtes C/D. *Thèse*
- Boulon S, Marmier-Gourrier N, Pradet-Balade B, Wurth L, Verheggen C, Jady BE, Rothe B, Pescia C, Robert MC, Kiss T, Bardoni B, Krol A, Branlant C, Allmang C, Bertrand E, Charpentier B (2008) The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J Cell Biol* **180**(3): 579-595
- Boulon S, Pradet-Balade B, Verheggen C, Molle D, Boireau S, Georgieva M, Azzag K, Robert MC, Ahmad Y, Neel H, Lamond AI, Bertrand E (2010) HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol Cell* **39**(6): 912-924
- Boulon S, Verheggen C, Jady BE, Girard C, Pescia C, Paul C, Ospina JK, Kiss T, Matera AG, Bordonne R, Bertrand E (2004) PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. *Mol Cell* **16**(5): 777-787
- Bousquet-Antonelli C, Henry Y, G'Elugne J P, Caizergues-Ferrer M, Kiss T (1997) A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs. *EMBO J* **16**(15): 4770-4776
- Box JA, Bunch JT, Tang W, Baumann P (2008) Spliceosomal cleavage generates the 3' end of telomerase RNA. *Nature* **456**(7224): 910-914
- Bradatsch B, Katahira J, Kowalinski E, Bange G, Yao W, Sekimoto T, Baumgartel V, Boese G, Bassler J, Wild K, Peters R, Yoneda Y, Sinning I, Hurt E (2007) Arx1 functions as an unorthodox nuclear export receptor for the 60S preribosomal subunit. *Mol Cell* **27**(5): 767-779
- Brahms H, Meheus L, de Brabandere V, Fischer U, Luhrmann R (2001) Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA* **7**(11): 1531-1542
- Brand RC, Klootwijk J, Van Steenberg T, De Kok AJ, Planta RJ (1977) Secondary methylation of yeast ribosomal precursor RNA. *Eur J Biochem* **75**(1): 311-318
- Branlant C, Krol A, Ebel JP, Lazar E, Haendler B, Jacob M (1982) U2 RNA shares a structural domain with U1, U4, and U5 RNAs. *EMBO J* **1**(10): 1259-1265
- Branlant C, Krol A, Machatt MA, Pouyet J, Ebel JP, Edwards K, Kossel H (1981) Primary and secondary structures of Escherichia coli MRE 600 23S ribosomal RNA. Comparison with models of secondary structure for maize chloroplast 23S rRNA and for large portions of mouse and human 16S mitochondrial rRNAs. *Nucleic Acids Res* **9**(17): 4303-4324
- Brimacombe R, Mitchell P, Osswald M, Stade K, Bochkariov D (1993) Clustering of modified nucleotides at the functional center of bacterial ribosomal RNA. *FASEB J* **7**(1): 161-167
- Brule F, Gregoire A, Segault V, Mougou A, Branlant C (1995) Secondary structure conservation of the U3 small nucleolar RNA introns in Saccharomyces. *C R Acad Sci III* **318**(12): 1197-1206
- Brule F, Venema J, Segault V, Tollervey D, Branlant C (1996) The yeast Hansenula wingei U3 snoRNA gene contains an intron and its coding sequence co-evolved with the 5' ETS region of the pre-ribosomal RNA. *RNA* **2**(2): 183-197
- Buhler D, Raker V, Luhrmann R, Fischer U (1999) Essential role for the tudor domain of SMN in spliceosomal U snRNP assembly: implications for spinal muscular atrophy. *Hum Mol Genet* **8**(13): 2351-2357
- Buratowski S (2003) The CTD code. *Nat Struct Biol* **10**(9): 679-680
- 
- Cabart P, Chew HK, Murphy S (2004) BRCA1 cooperates with NUFIP and P-TEFb to activate transcription by RNA polymerase II. *Oncogene* **23**(31): 5316-5329
- Cabello-Villegas J, Nikonowicz EP (2005) Solution structure of psi32-modified anticodon stem-loop of Escherichia coli tRNA<sup>Phe</sup>. *Nucleic Acids Res* **33**(22): 6961-6971
- Caffarelli E, Fatica A, Prislei S, De Gregorio E, Frapapan P, Bozzoni I (1996) Processing of the intron-encoded U16 and U18 snoRNAs: the conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA. *EMBO J* **15**(5): 1121-1131
- Cahill NM, Friend K, Speckmann W, Li ZH, Terns RM, Terns MP, Steitz JA (2002) Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snoRNP. *EMBO J* **21**(14): 3816-3828
- Cai Y, Jin J, Florens L, Swanson SK, Kusch T, Li B, Workman JL, Washburn MP, Conaway RC, Conaway JW (2005) The mammalian YL1 protein is a shared subunit of the TRRAP/TIP60 histone acetyltransferase and SRCAP complexes. *J Biol Chem* **280**(14): 13665-13670
- Caplan AJ, Mandal AK, Theodoraki MA (2007) Molecular chaperones and protein kinase quality control. *Trends Cell Biol* **17**(2): 87-92
- Carbone R, Pearson M, Minucci S, Pelicci PG (2002) PML NBs associate with the hMre11 complex and p53 at sites of irradiation induced DNA damage. *Oncogene* **21**(11): 1633-1640
- Carissimi C, Baccon J, Straccia M, Chiarella P, Maiolica A, Sawyer A, Rappsilber J, Pellizzoni L (2005) Unrip is a

- component of SMN complexes active in snRNP assembly. *FEBS Lett* **579**(11): 2348-2354
- Carissimi C, Saieva L, Baccon J, Chiarella P, Maiolica A, Sawyer A, Rappsilber J, Pellizzoni L (2006) Gemin8 is a novel component of the survival motor neuron complex and functions in small nuclear ribonucleoprotein assembly. *J Biol Chem* **281**(12): 8126-8134
- Carmo-Fonseca M, Ferreira J, Lamond AI (1993) Assembly of snRNP-containing coiled bodies is regulated in interphase and mitosis--evidence that the coiled body is a kinetic nuclear structure. *J Cell Biol* **120**(4): 841-852
- Carmo-Fonseca M, Mendes-Soares L, Campos I (2000) To be or not to be in the nucleolus. *Nat Cell Biol* **2**(6): E107-112
- Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C, Kodzius R, Shimokawa K, Bajic VB, Brenner SE, Batalov S, Forrest AR, Zavolan M, Davis MJ, Wilming LG, Aidinis V, Allen JE, Ambesi-Impiombato A, Apweiler R, Aturaliya RN, Bailey TL, Bansal M, Baxter L, Beisel KW, Bersano T, Bono H, Chalk AM, Chiu KP, Choudhary V, Christoffels A, Clutterbuck DR, Crowe ML, Dalla E, Dalrymple BP, de Bono B, Della Gatta G, di Bernardo D, Down T, Engstrom P, Fagiolini M, Faulkner G, Fletcher CF, Fukushima T, Furuno M, Futaki S, Gariboldi M, Georgii-Hemming P, Gingeras TR, Gojobori T, Green RE, Gustincich S, Harbers M, Hayashi Y, Hensch TK, Hirokawa N, Hill D, Huminiecki L, Iacono M, Ikeo K, Iwama A, Ishikawa T, Jakt M, Kanapin A, Katoh M, Kawasawa Y, Kelso J, Kitamura H, Kitano H, Kollias G, Krishnan SP, Kruger A, Kummerfeld SK, Kurochkin IV, Lareau LF, Lazarevic D, Lipovich L, Liu J, Liuni S, McWilliam S, Madan Babu M, Madera M, Marchionni L, Matsuda H, Matsuzawa S, Miki H, Mignone F, Miyake S, Morris K, Mottagui-Tabar S, Mulder N, Nakano N, Nakauchi H, Ng P, Nilsson R, Nishiguchi S, Nishikawa S, Nori F, Ohara O, Okazaki Y, Orlando V, Pang KC, Pavan WJ, Pavesi G, Pesole G, Petrovsky N, Piazza S, Reed J, Reid JF, Ring BZ, Ringwald M, Rost B, Ruan Y, Salzberg SL, Sandelin A, Schneider C, Schonbach C, Sekiguchi K, Semple CA, Seno S, Sessa L, Sheng Y, Shibata Y, Shimada H, Shimada K, Silva D, Sinclair B, Sperling S, Stupka E, Sugiura K, Sultana R, Takenaka Y, Taki K, Tammoya K, Tan SL, Tang S, Taylor MS, Tegner J, Teichmann SA, Ueda HR, van Nimwegen E, Verardo R, Wei CL, Yagi K, Yamanishi H, Zabarovsky E, Zhu S, Zimmer A, Hide W, Bult C, Grimmond SM, Teasdale RD, Liu ET, Brusci V, Quackenbush J, Wahlestedt C, Mattick JS, Hume DA, Kai C, Sasaki D, Tomaru Y, Fukuda S, Kanamori-Katayama M, Suzuki M, Aoki J, Arakawa T, Iida J, Imamura K, Itoh M, Kato T, Kawaji H, Kawagashira N, Kawashima T, Kojima M, Kondo S, Konno H, Nakano K, Ninomiya N, Nishio T, Okada M, Plessy C, Shibata K, Shiraki T, Suzuki S, Tagami M, Waki K, Watahiki A, Okamura-Oho Y, Suzuki H, Kawai J, Hayashizaki Y (2005) The transcriptional landscape of the mammalian genome. *Science* **309**(5740): 1559-1563
- Carroll KL, Pradhan DA, Granek JA, Clarke ND, Corden JL (2004) Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. *Mol Cell Biol* **24**(14): 6241-6252
- Cartegni L, Chew SL, Krainer AR (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* **3**(4): 285-298
- Carter AP, Clemons WM, Brodersen DE, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **407**(6802): 340-348
- Cavaillé J, Bachellerie JP (1996) Processing of fibrillarin-associated snoRNAs from pre-mRNA introns: an exonucleolytic process exclusively directed by the common stem-box terminal structure. *Biochimie* **78**(6): 443-456
- Cavaillé J, Buiting K, Kiefmann M, Lalande M, Brannan CI, Horsthemke B, Bachellerie JP, Brosius J, Huttenhofer A (2000) Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc Natl Acad Sci U S A* **97**(26): 14311-14316
- Cavaillé J, Nicoloso M, Bachellerie JP (1996) Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides. *Nature* **383**(6602): 732-735
- Cavaillé J, Seitz H, Paulsen M, Ferguson-Smith AC, Bachellerie JP (2002) Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. *Hum Mol Genet* **11**(13): 1527-1538
- Cech TR (2009) Evolution of biological catalysis: ribozyme to RNP enzyme. *Cold Spring Harb Symp Quant Biol* **74**: 11-16
- Chakarova CF, Hims MM, Bolz H, Abu-Safieh L, Patel RJ, Papaioannou MG, Inglehearn CF, Keen TJ, Willis C, Moore AT, Rosenberg T, Webster AR, Bird AC, Gal A, Hunt D, Vithana EN, Bhattacharya SS (2002) Mutations in HPRP3, a third member of pre-mRNA splicing factor genes, implicated in autosomal dominant retinitis pigmentosa. *Hum Mol Genet* **11**(1): 87-92
- Chanfreau G, Buckle M, Jacquier A (2000) Recognition of a conserved class of RNA tetraloops by *Saccharomyces cerevisiae* RNase III. *Proc Natl Acad Sci U S A* **97**(7): 3142-3147
- Chanfreau G, Elela SA, Ares M, Jr., Guthrie C (1997) Alternative 3'-end processing of U5 snRNA by RNase III. *Genes Dev* **11**(20): 2741-2751
- Chanfreau G, Legrain P, Jacquier A (1998a) Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. *J Mol Biol* **284**(4): 975-988
- Chanfreau G, Rotondo G, Legrain P, Jacquier A (1998b) Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. *EMBO J* **17**(13): 3726-3737
- Chang DD, Clayton DA (1989) Mouse RNAase MRP RNA is encoded by a nuclear gene and contains a decamer sequence complementary to a conserved region of mitochondrial RNA substrate. *Cell* **56**(1): 131-139

- Chao JA, Prasad GS, White SA, Stout CD, Williamson JR (2003) Inherent protein structural flexibility at the RNA-binding interface of L30e. *J Mol Biol* **326**(4): 999-1004
- Chao JA, Williamson JR (2004) Joint X-ray and NMR refinement of the yeast L30e-mRNA complex. *Structure* **12**(7): 1165-1176
- Chapon C, Cech TR, Zaug AJ (1997) Polyadenylation of telomerase RNA in budding yeast. *RNA* **3**(11): 1337-1351
- Charette M, Gray MW (2000) Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* **49**(5): 341-351
- Chari A, Fischer U (2010) Cellular strategies for the assembly of molecular machines. *Trends Biochem Sci* **35**(12): 676-683
- Chari A, Paknia E, Fischer U (2009) The role of RNP biogenesis in spinal muscular atrophy. *Curr Opin Cell Biol* **21**(3): 387-393
- Charpentier B, Fourmann JB, Branlant C (2007) Reconstitution of archaeal H/ACA sRNPs and test of their activity. *Methods Enzymol* **425**: 389-405
- Charpentier B, Muller S, Branlant C (2005) Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic Acids Res* **33**(10): 3133-3144
- Charron C, Manival X, Clery A, Senty-Segault V, Charpentier B, Marmier-Gourrier N, Branlant C, Aubry A (2004) The archaeal sRNA binding protein L7Ae has a 3D structure very similar to that of its eukaryal counterpart while having a broader RNA-binding specificity. *J Mol Biol* **342**(3): 757-773
- Charroux B, Pellizzoni L, Perkinson RA, Shevchenko A, Mann M, Dreyfuss G (1999) Gemin3: A novel DEAD box protein that interacts with SMN, the spinal muscular atrophy gene product, and is a component of gems. *J Cell Biol* **147**(6): 1181-1194
- Charroux B, Pellizzoni L, Perkinson RA, Yong J, Shevchenko A, Mann M, Dreyfuss G (2000) Gemin4. A novel component of the SMN complex that is found in both gems and nucleoli. *J Cell Biol* **148**(6): 1177-1186
- Chavatte L, Brown BA, Driscoll DM (2005) Ribosomal protein L30 is a component of the UGA-selenocysteine recoding machinery in eukaryotes. *Nat Struct Mol Biol* **12**(5): 408-416
- Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR, Lao KQ, Livak KJ, Guegler KJ (2005a) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* **33**(20): e179
- Chen D, Dundr M, Wang C, Leung A, Lamond A, Misteli T, Huang S (2005b) Condensed mitotic chromatin is accessible to transcription factors and chromatin structural proteins. *J Cell Biol* **168**(1): 41-54
- Chen L, Cai Y, Jin J, Florens L, Swanson SK, Washburn MP, Conaway JW, Conaway RC (2011) Subunit organization of the human INO80 chromatin remodeling complex: an evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodeling. *J Biol Chem* **286**(13): 11283-11289
- Chen LL, Carmichael GG (2009) Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol Cell* **35**(4): 467-478
- Chen LL, DeCerbo JN, Carmichael GG (2008) Alu element-mediated gene silencing. *EMBO J* **27**(12): 1694-1705
- Chen YW, Bycroft M, Wong KB (2003) Crystal structure of ribosomal protein L30e from the extreme thermophile *Thermococcus celer*: thermal stability and RNA binding. *Biochemistry* **42**(10): 2857-2865
- Cherry JM, Ball C, Weng S, Juvik G, Schmidt R, Adler C, Dunn B, Dwight S, Riles L, Mortimer RK, Botstein D (1997) Genetic and physical maps of *Saccharomyces cerevisiae*. *Nature* **387**(6632 Suppl): 67-73
- Cheung KL, Huen J, Houry WA, Ortega J (2010) Comparison of the multiple oligomeric structures observed for the Rvb1 and Rvb2 proteins. *Biochem Cell Biol* **88**(1): 77-88
- Cho EJ, Takagi T, Moore CR, Buratowski S (1997) mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev* **11**(24): 3319-3326
- Cho IM, Lai LB, Susanti D, Mukhopadhyay B, Gopalan V (2010) Ribosomal protein L7Ae is a subunit of archaeal RNase P. *Proc Natl Acad Sci U S A* **107**(33): 14573-14578
- Choesmel V, Fribourg S, Aguisa-Toure AH, Pinaud N, Legrand P, Gazda HT, Gleizes PE (2008) Mutation of ribosomal protein RPS24 in Diamond-Blackfan anemia results in a ribosome biogenesis disorder. *Hum Mol Genet* **17**(9): 1253-1263
- Cioce M, Lamond AI (2005) Cajal bodies: a long history of discovery. *Annu Rev Cell Dev Biol* **21**: 105-131
- Clamp M, Fry B, Kamal M, Xie X, Cuff J, Lin MF, Kellis M, Lindblad-Toh K, Lander ES (2007) Distinguishing protein-coding and noncoding genes in the human genome. *Proc Natl Acad Sci U S A* **104**(49): 19428-19433
- Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB (2009) An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* **33**(6): 717-726
- Clery A, Bourguignon-Igel V, Allmang C, Krol A, Branlant C (2007a) An improved definition of the RNA-binding specificity of SECIS-binding protein 2, an essential component of the selenocysteine incorporation machinery. *Nucleic Acids Res* **35**(6): 1868-1884
- Clery A, Senty-Segault V, Leclerc F, Raue HA, Branlant C (2007b) Analysis of sequence and structural features that identify the B/C motif of U3 small nucleolar RNA as the recognition site for the Snu13p-Rrp9p protein pair. *Mol Cell Biol* **27**(4): 1191-1206

- Clouet d'Orval B, Bortolin ML, Gaspin C, Bachellerie JP (2001) Box C/D RNA guides for the ribose methylation of archaeal tRNAs. The tRNA<sup>Trp</sup> intron guides the formation of two ribose-methylated nucleosides in the mature tRNA<sup>Trp</sup>. *Nucleic Acids Res* **29**(22): 4518-4529
- Cloutier P, Al-Khoury R, Lavalée-Adam M, Faubert D, Jiang H, Poitras C, Bouchard A, Forget D, Blanchette M, Coulombe B (2009) High-resolution mapping of the protein interaction network for the human transcription machinery and affinity purification of RNA polymerase II-associated complexes. *Methods* **48**(4): 381-386
- Cojocaru V, Nottrott S, Klement R, Jovin TM (2005) The snRNP 15.5K protein folds its cognate K-turn RNA: a combined theoretical and biochemical study. *RNA* **11**(2): 197-209
- Colley A, Beggs JD, Tollervey D, Lafontaine DL (2000) Dhr1p, a putative DEAH-box RNA helicase, is associated with the box C+D snoRNP U3. *Mol Cell Biol* **20**(19): 7238-7246
- Collins LJ, Penny D (2009) The RNA infrastructure: dark matter of the eukaryotic cell? *Trends Genet* **25**(3): 120-128
- Collins SR, Miller KM, Maas NL, Roguev A, Fillingham J, Chu CS, Schuldiner M, Gebbia M, Recht J, Shales M, Ding H, Xu H, Han J, Ingvarsdottir K, Cheng B, Andrews B, Boone C, Berger SL, Hieter P, Zhang Z, Brown GW, Ingles CJ, Emili A, Allis CD, Toczyski DP, Weissman JS, Greenblatt JF, Krogan NJ (2007) Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* **446**(7137): 806-810
- Colot HV, Stutz F, Rosbash M (1996) The yeast splicing factor Mud13p is a commitment complex component and corresponds to CBP20, the small subunit of the nuclear cap-binding complex. *Genes Dev* **10**(13): 1699-1708
- Conlan LA, McNeese CJ, Heierhorst J (2004) Proteasome-dependent dispersal of PML nuclear bodies in response to alkylating DNA damage. *Oncogene* **23**(1): 307-310
- Conrad NK, Wilson SM, Steinmetz EJ, Patturajan M, Brow DA, Swanson MS, Corden JL (2000) A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. *Genetics* **154**(2): 557-571
- Copeland PR, Fletcher JE, Carlson BA, Hatfield DL, Driscoll DM (2000) A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J* **19**(2): 306-314
- Costello JL, Stead JA, Feigenbutz M, Jones RM, Mitchell P (2011) The C-terminal region of the exosome-associated protein Rrp47 is specifically required for box C/D small nucleolar RNA 3'-maturation. *J Biol Chem* **286**(6): 4535-4543
- Cowling VH (2009) Regulation of mRNA cap methylation. *Biochem J* **425**(2): 295-302
- Cremer T, Cremer C (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* **2**(4): 292-301
- Cremer T, Cremer M, Dietzel S, Muller S, Solovei I, Fakan S (2006) Chromosome territories--a functional nuclear landscape. *Curr Opin Cell Biol* **18**(3): 307-316
- Cullen BR (2003) Nuclear RNA export. *J Cell Sci* **116**(Pt 4): 587-597
- D**'Andrea LD, Regan L (2003) TPR proteins: the versatile helix. *Trends Biochem Sci* **28**(12): 655-662
- D'Angelo MA, Hetzer MW (2008) Structure, dynamics and function of nuclear pore complexes. *Trends Cell Biol* **18**(10): 456-466
- Danckwardt S, Hentze MW, Kulozik AE (2008) 3' end mRNA processing: molecular mechanisms and implications for health and disease. *EMBO J* **27**(3): 482-498
- Darzacq X, Jady BE, Verheggen C, Kiss AM, Bertrand E, Kiss T (2002) Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J* **21**(11): 2746-2756
- Darzacq X, Kiss T (2000) Processing of intron-encoded box C/D small nucleolar RNAs lacking a 5',3'-terminal stem structure. *Mol Cell Biol* **20**(13): 4522-4531
- Darzacq X, Kittur N, Roy S, Shav-Tal Y, Singer RH, Meier UT (2006) Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *J Cell Biol* **173**(2): 207-218
- Davey MJ, Indiani C, O'Donnell M (2003) Reconstitution of the Mcm2-7p heterohexamer, subunit arrangement, and ATP site architecture. *J Biol Chem* **278**(7): 4491-4499
- Davis DR (1995) Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Res* **23**(24): 5020-5026
- Dawkins R (1976) *The Selfish Gene*.
- De Jonge P, Klootwijk J, Planta RJ (1977) Terminal nucleotide sequences of 17-S ribosomal RNA and its immediate precursor 18-S RNA in yeast. *Eur J Biochem* **72**(2): 361-369
- de la Cruz J, Lacombe T, Deloche O, Linder P, Kressler D (2004) The putative RNA helicase Dbp6p functionally interacts with Rpl3p, Nop8p and the novel trans-acting Factor Rsa3p during biogenesis of 60S ribosomal subunits in *Saccharomyces cerevisiae*. *Genetics* **166**(4): 1687-1699
- De Rubeis S, Bagni C (2010) Fragile X mental retardation protein control of neuronal mRNA metabolism: Insights into mRNA stability. *Mol Cell Neurosci* **43**(1): 43-50
- Decatur WA, Fournier MJ (2003) RNA-guided nucleotide modification of ribosomal and other RNAs. *J Biol Chem* **278**(2): 695-698
- Decatur WA, Schnare MN (2008) Different mechanisms for pseudouridine formation in yeast 5S and 5.8S rRNAs. *Mol Cell Biol* **28**(10): 3089-3100

- Decourty L, Saveanu C, Zemam K, Hantraye F, Frachon E, Rousselle JC, Fromont-Racine M, Jacquier A (2008) Linking functionally related genes by sensitive and quantitative characterization of genetic interaction profiles. *Proc Natl Acad Sci U S A* **105**(15): 5821-5826
- Deery EC, Vithana EN, Newbold RJ, Gallon VA, Bhattacharya SS, Warren MJ, Hunt DM, Wilkie SE (2002) Disease mechanism for retinitis pigmentosa (RP11) caused by mutations in the splicing factor gene PRPF31. *Hum Mol Genet* **11**(25): 3209-3219
- DeLabre ML, Kessl J, Karamanou S, Trumppower BL (2002) RPL29 codes for a non-essential protein of the 60S ribosomal subunit in *Saccharomyces cerevisiae* and exhibits synthetic lethality with mutations in genes for proteins required for subunit coupling. *Biochim Biophys Acta* **1574**(3): 255-261
- Dellaire G, Bazett-Jones DP (2004) PML nuclear bodies: dynamic sensors of DNA damage and cellular stress. *Bioessays* **26**(9): 963-977
- Dellaire G, Farrall R, Bickmore WA (2003) The Nuclear Protein Database (NPD): sub-nuclear localisation and functional annotation of the nuclear proteome. *Nucleic Acids Res* **31**(1): 328-330
- Deng L, Starostina NG, Liu ZJ, Rose JP, Terns RM, Terns MP, Wang BC (2004) Structure determination of fibrillar from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Biochem Biophys Res Commun* **315**(3): 726-732
- Dez C, Froment C, Noaillac-Depeyre J, Monsarrat B, Caizergues-Ferrer M, Henry Y (2004) Npa1p, a component of very early pre-60S ribosomal particles, associates with a subset of small nucleolar RNPs required for peptidyl transferase center modification. *Mol Cell Biol* **24**(14): 6324-6337
- Dez C, Henras A, Faucon B, Lafontaine D, Caizergues-Ferrer M, Henry Y (2001) Stable expression in yeast of the mature form of human telomerase RNA depends on its association with the box H/ACA small nucleolar RNP proteins Cbf5p, Nhp2p and Nop10p. *Nucleic Acids Res* **29**(3): 598-603
- Dez C, Noaillac-Depeyre J, Caizergues-Ferrer M, Henry Y (2002) Naf1p, an essential nucleoplasmic factor specifically required for accumulation of box H/ACA small nucleolar RNPs. *Mol Cell Biol* **22**(20): 7053-7065
- Dheur S, Voile TA, Voisin-Hakil F, Minet M, Schmitter JM, Lacroute F, Wyers F, Minvielle-Sebastia L (2003) Pti1p and Ref2p found in association with the mRNA 3' end formation complex direct snoRNA maturation. *EMBO J* **22**(11): 2831-2840
- Dick FA, Eisinger DP, Trumppower BL (1997a) Exchangeability of Qsr1p, a large ribosomal subunit protein required for subunit joining, suggests a novel translational regulatory mechanism. *FEBS Lett* **419**(1): 1-3
- Dick FA, Karamanou S, Trumppower BL (1997b) QSR1, an essential yeast gene with a genetic relationship to a subunit of the mitochondrial cytochrome bc1 complex, codes for a 60 S ribosomal subunit protein. *J Biol Chem* **272**(20): 13372-13379
- Diebold ML, Fribourg S, Koch M, Metzger T, Romier C (2011) Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. *J Struct Biol* **175**(2): 178-188
- Dobbyn HC, O'Keefe RT (2004) Analysis of Snu13p mutations reveals differential interactions with the U4 snRNA and U3 snoRNA. *RNA* **10**(2): 308-320
- Dominski Z, Yang XC, Marzluff WF (2005a) The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. *Cell* **123**(1): 37-48
- Dominski Z, Yang XC, Purdy M, Wagner EJ, Marzluff WF (2005b) A CPSF-73 homologue is required for cell cycle progression but not cell growth and interacts with a protein having features of CPSF-100. *Mol Cell Biol* **25**(4): 1489-1500
- Donalies UE, Stahl U (2001) Phase-specific gene expression in *Saccharomyces cerevisiae*, using maltose as carbon source under oxygen-limiting conditions. *Curr Genet* **39**(3): 150-155
- Donmez G, Hartmuth K, Luhrmann R (2004) Modified nucleotides at the 5' end of human U2 snRNA are required for spliceosomal E-complex formation. *RNA* **10**(12): 1925-1933
- Donovan J, Copeland PR (2009) Evolutionary history of selenocysteine incorporation from the perspective of SECIS binding proteins. *BMC Evol Biol* **9**: 229
- Dragon F, Gallagher JE, Compagnone-Post PA, Mitchell BM, Porwancher KA, Wehner KA, Wormsley S, Settlage RE, Shabanowitz J, Osheim Y, Beyer AL, Hunt DF, Baserga SJ (2002) A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* **417**(6892): 967-970
- Dragon F, Pogacic V, Filipowicz W (2000) In vitro assembly of human H/ACA small nucleolar RNPs reveals unique features of U17 and telomerase RNAs. *Mol Cell Biol* **20**(9): 3037-3048
- Duan J, Li L, Lu J, Wang W, Ye K (2009) Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. *Mol Cell* **34**(4): 427-439
- Dutca LM, Gallagher JE, Baserga SJ (2011) The initial U3 snoRNA:pre-rRNA base pairing interaction required for pre-18S rRNA folding revealed by in vivo chemical probing. *Nucleic Acids Res*
- 
- Eckert K, Saliou JM, Monlezun L, Vigouroux A, Atmane N, Caillat C, Quevillon-Cheruel S, Madiona K, Nicaise M, Lazereg S, Van Dorsselaer A, Sanglier-Cianferani S, Meyer P, Morera S (2010) The Pih1-Tah1 cochaperone complex inhibits Hsp90 molecular chaperone ATPase activity. *J Biol Chem* **285**(41): 31304-31312



- Eckner R, Ellmeier W, Birnstiel ML (1991) Mature mRNA 3' end formation stimulates RNA export from the nucleus. *EMBO J* **10**(11): 3513-3522
- Ederly P, Marcaillou C, Sahbatou M, Labalme A, Chastang J, Touraine R, Tubacher E, Senni F, Bober MB, Nampoothiri S, Jouk PS, Steichen E, Berland S, Toutain A, Wise CA, Sanlaville D, Rousseau F, Clerget-Darpoux F, Leutenegger AL (2011) Association of TALS developmental disorder with defect in minor splicing component U4atac snRNA. *Science* **332**(6026): 240-243
- Egecioglu DE, Henras AK, Chanfreau GF (2006) Contributions of Trf4p- and Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. *RNA* **12**(1): 26-32
- Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S (2007) Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* **318**(5857): 1777-1779
- Eisinger DP, Dick FA, Trumpower BL (1997) Qsr1p, a 60S ribosomal subunit protein, is required for joining of 40S and 60S subunits. *Mol Cell Biol* **17**(9): 5136-5145
- Ellis RJ (2001) Macromolecular crowding: obvious but underappreciated. *Trends Biochem Sci* **26**(10): 597-604
- Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, Pfeffer S, Rajewsky N, Meister G (2008) A human snoRNA with microRNA-like functions. *Mol Cell* **32**(4): 519-528
- Eng FJ, Warner JR (1991) Structural basis for the regulation of splicing of a yeast messenger RNA. *Cell* **65**(5): 797-804
- Eppens NA, Faber AW, Rondaij M, Jahangir RS, van Hemert S, Vos JC, Venema J, Raue HA (2002) Deletions in the S1 domain of Rrp5p cause processing at a novel site in ITS1 of yeast pre-rRNA that depends on Rex4p. *Nucleic Acids Res* **30**(19): 4222-4231
- Epstein P, Reddy R, Busch H (1984) Multiple states of U3 RNA in Novikoff hepatoma nucleoli. *Biochemistry* **23**(23): 5421-5425
- Erkman JA, Kutay U (2004) Nuclear export of mRNA: from the site of transcription to the cytoplasm. *Exp Cell Res* **296**(1): 12-20
- Etard C, Gradl D, Kunz M, Eilers M, Wedlich D (2005) Pontin and Reptin regulate cell proliferation in early *Xenopus* embryos in collaboration with c-Myc and Miz-1. *Mech Dev* **122**(4): 545-556
- Falb M, Amata I, Gabel F, Simon B, Carlomagno T (2010) Structure of the K-turn U4 RNA: a combined NMR and SANS study. *Nucleic Acids Res* **38**(18): 6274-6285
- Fatica A, Cronshaw AD, Dlakic M, Tollervey D (2002a) Ssf1p prevents premature processing of an early pre-60S ribosomal particle. *Mol Cell* **9**(2): 341-351
- Fatica A, Dlakic M, Tollervey D (2002b) Naf1 p is a box H/ACA snoRNP assembly factor. *RNA* **8**(12): 1502-1514
- Fatica A, Galardi S, Altieri F, Bozzoni I (2000a) Fibrillarin binds directly and specifically to U16 box C/D snoRNA. *RNA* **6**(1): 88-95
- Fatica A, Morlando M, Bozzoni I (2000b) Yeast snoRNA accumulation relies on a cleavage-dependent/polyadenylation-independent 3'-processing apparatus. *EMBO J* **19**(22): 6218-6229
- Fatica A, Oeffinger M, Dlakic M, Tollervey D (2003) Nob1p is required for cleavage of the 3' end of 18S rRNA. *Mol Cell Biol* **23**(5): 1798-1807
- Fayet-Lebaron E, Atzorn V, Henry Y, Kiss T (2009) 18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences. *EMBO J* **28**(9): 1260-1270
- Feder M, Pas J, Wyrwicz LS, Bujnicki JM (2003) Molecular phylogenetics of the RrmJ/fibrillarin superfamily of ribose 2'-O-methyltransferases. *Gene* **302**(1-2): 129-138
- Fedorov R, Nevskaya N, Khairullina A, Tishchenko S, Mikhailov A, Garber M, Nikonov S (1999) Structure of ribosomal protein L30 from *Thermus thermophilus* at 1.9 Å resolution: conformational flexibility of the molecule. *Acta Crystallogr D Biol Crystallogr* **55**(Pt 11): 1827-1833
- Felts SJ, Karnitz LM, Toft DO (2007) Functioning of the Hsp90 machine in chaperoning checkpoint kinase I (Chk1) and the progesterone receptor (PR). *Cell Stress Chaperones* **12**(4): 353-363
- Feng Y, Lee N, Fearon ER (2003) TIP49 regulates beta-catenin-mediated neoplastic transformation and T-cell factor target gene induction via effects on chromatin remodeling. *Cancer Res* **63**(24): 8726-8734
- Fernandez CF, Pannone BK, Chen X, Fuchs G, Wolin SL (2004) An Lsm2-Lsm7 complex in *Saccharomyces cerevisiae* associates with the small nucleolar RNA snR5. *Mol Biol Cell* **15**(6): 2842-2852
- Ferreira ME, Flaherty K, Prochasson P (2011) The *Saccharomyces cerevisiae* histone chaperone Rtt106 mediates the cell cycle recruitment of SWI/SNF and RSC to the HIR-dependent histone genes. *PLoS One* **6**(6): e21113
- Filipowicz W, Pelczar P, Pogacic V, Dragon F (1999) Structure and biogenesis of small nucleolar RNAs acting as guides for ribosomal RNA modification. *Acta Biochim Pol* **46**(2): 377-389
- Fillingham J, Kainth P, Lambert JP, van Bakel H, Tsui K, Pena-Castillo L, Nislow C, Figeys D, Hughes TR, Greenblatt J, Andrews BJ (2009) Two-color cell array
- F**aber AW, Van Dijk M, Raue HA, Vos JC (2002) Ngl2p is a Ccr4p-like RNA nuclease essential for the final step in 3'-end processing of 5.8S rRNA in *Saccharomyces cerevisiae*. *RNA* **8**(9): 1095-1101
- Faber AW, Vos JC, Vos HR, Ghazal G, Elela SA, Raue HA (2004) The RNA catabolic enzymes Rex4p, Rnt1p, and Dbr1p show genetic interaction with trans-acting factors involved in processing of ITS1 in *Saccharomyces cerevisiae* pre-rRNA. *RNA* **10**(12): 1946-1956

- screen reveals interdependent roles for histone chaperones and a chromatin boundary regulator in histone gene repression. *Mol Cell* **35**(3): 340-351
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**(6669): 806-811
- Fischer U, Liu Q, Dreyfuss G (1997) The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* **90**(6): 1023-1029
- Fischer U, Sumpter V, Sekine M, Satoh T, Luhrmann R (1993) Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap. *EMBO J* **12**(2): 573-583
- Flaherty SM, Fortes P, Izaurralde E, Mattaj IW, Gilmartin GM (1997) Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proc Natl Acad Sci U S A* **94**(22): 11893-11898
- Fletcher JE, Copeland PR, Driscoll DM, Krol A (2001) The selenocysteine incorporation machinery: interactions between the SECIS RNA and the SECIS-binding protein SBP2. *RNA* **7**(10): 1442-1453
- Fong YW, Zhou Q (2001) Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414**(6866): 929-933
- Fourmann JB (2009) Analyse structurale et fonctionnelle des particules sRNP à boîtes H/ACA, catalysant la formation ciblée des pseudouridines dans les ARN d'archaea. *Thèse*
- Fournier R, Brule F, Segault V, Mougin A, Branlant C (1998) U3 snoRNA genes with and without intron in the *Kluyveromyces* genus: yeasts can accommodate great variations of the U3 snoRNA 3'-terminal domain. *RNA* **4**(3): 285-302
- Fox AH, Lam YW, Leung AK, Lyon CE, Andersen J, Mann M, Lamond AI (2002) Paraspeckles: a novel nuclear domain. *Curr Biol* **12**(1): 13-25
- Francastel C, Schubeler D, Martin DI, Groudine M (2000) Nuclear compartmentalization and gene activity. *Nat Rev Mol Cell Biol* **1**(2): 137-143
- Franke J, Gehlen J, Ehrenhofer-Murray AE (2008) Hypermethylation of yeast telomerase RNA by the snRNA and snoRNA methyltransferase Tgs1. *J Cell Sci* **121**(Pt 21): 3553-3560
- Frankel A, Clarke S (1999) RNase treatment of yeast and mammalian cell extracts affects in vitro substrate methylation by type I protein arginine N-methyltransferases. *Biochem Biophys Res Commun* **259**(2): 391-400
- Fraser P, Bickmore W (2007) Nuclear organization of the genome and the potential for gene regulation. *Nature* **447**(7143): 413-417
- Freed EF, Bleichert F, Dutca LM, Baserga SJ (2010) When ribosomes go bad: diseases of ribosome biogenesis. *Mol Biosyst* **6**(3): 481-493
- Frey MR, Matera AG (1995) Coiled bodies contain U7 small nuclear RNA and associate with specific DNA sequences in interphase human cells. *Proc Natl Acad Sci U S A* **92**(13): 5915-5919
- Frey MR, Matera AG (2001) RNA-mediated interaction of Cajal bodies and U2 snRNA genes. *J Cell Biol* **154**(3): 499-509
- Friesen WJ, Massenet S, Paushkin S, Wyce A, Dreyfuss G (2001a) SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. *Mol Cell* **7**(5): 1111-1117
- Friesen WJ, Paushkin S, Wyce A, Massenet S, Pesiridis GS, Van Duyne G, Rappsilber J, Mann M, Dreyfuss G (2001b) The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol Cell Biol* **21**(24): 8289-8300
- Friesen WJ, Wyce A, Paushkin S, Abel L, Rappsilber J, Mann M, Dreyfuss G (2002) A novel WD repeat protein component of the methylosome binds Sm proteins. *J Biol Chem* **277**(10): 8243-8247
- Fromont-Racine M, Rain JC, Legrain P (1997) Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet* **16**(3): 277-282
- Frydman J, Hohfeld J (1997) Chaperones get in touch: the Hip-Hop connection. *Trends Biochem Sci* **22**(3): 87-92
- Fu D, Collins K (2003) Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs. *Mol Cell* **11**(5): 1361-1372
- Fuchs M, Gerber J, Drapkin R, Sif S, Ikura T, Ogryzko V, Lane WS, Nakatani Y, Livingston DM (2001) The p400 complex is an essential E1A transformation target. *Cell* **106**(3): 297-307
- 
- Gadal O, Strauss D, Kessler J, Trumpower B, Tollervey D, Hurt E (2001) Nuclear export of 60S ribosomal subunits depends on Xpo1p and requires a nuclear export sequence-containing factor, Nmd3p, that associates with the large subunit protein Rpl10p. *Mol Cell Biol* **21**(10): 3405-3415
- Gagnon KT, Zhang X, Qu G, Biswas S, Suryadi J, Brown BA, 2nd, Maxwell ES (2010) Signature amino acids enable the archaeal L7Ae box C/D RNP core protein to recognize and bind the K-loop RNA motif. *RNA* **16**(1): 79-90
- Galardi S, Fatica A, Bachi A, Scaloni A, Presutti C, Bozzoni I (2002) Purified box C/D snoRNPs are able to reproduce site-specific 2'-O-methylation of target RNA in vitro. *Mol Cell Biol* **22**(19): 6663-6668
- Gall JG, Bellini M, Wu Z, Murphy C (1999) Assembly of the nuclear transcription and processing machinery: Cajal bodies (coiled bodies) and transcriptosomes. *Mol Biol Cell* **10**(12): 4385-4402

- Gallagher JE, Dunbar DA, Granneman S, Mitchell BM, Osheim Y, Beyer AL, Baserga SJ (2004) RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components. *Genes Dev* **18**(20): 2506-2517
- Galtier N, Lobry JR (1997) Relationships between genomic G+C content, RNA secondary structures, and optimal growth temperature in prokaryotes. *J Mol Evol* **44**(6): 632-636
- Ganapathi KA, Austin KM, Lee CS, Dias A, Malsch MM, Reed R, Shimamura A (2007) The human Shwachman-Diamond syndrome protein, SBDS, associates with ribosomal RNA. *Blood* **110**(5): 1458-1465
- Ganot P, Bortolin ML, Kiss T (1997a) Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* **89**(5): 799-809
- Ganot P, Caizergues-Ferrer M, Kiss T (1997b) The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev* **11**(7): 941-956
- Ganot P, Jady BE, Bortolin ML, Darzacq X, Kiss T (1999) Nucleolar factors direct the 2'-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. *Mol Cell Biol* **19**(10): 6906-6917
- Gao L, Frey MR, Matera AG (1997) Human genes encoding U3 snRNA associate with coiled bodies in interphase cells and are clustered on chromosome 17p11.2 in a complex inverted repeat structure. *Nucleic Acids Res* **25**(23): 4740-4747
- Gaspin C, Cavaille J, Erauso G, Bachellerie JP (2000) Archaeal homologs of eukaryotic methylation guide small nucleolar RNAs: lessons from the *Pyrococcus* genomes. *J Mol Biol* **297**(4): 895-906
- Gatbonton T, Imbesi M, Nelson M, Akey JM, Ruderfer DM, Kruglyak L, Simon JA, Bedalov A (2006) Telomere length as a quantitative trait: genome-wide survey and genetic mapping of telomere length-control genes in yeast. *PLoS Genet* **2**(3): e35
- Gautier T, Berges T, Tollervey D, Hurt E (1997) Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. *Mol Cell Biol* **17**(12): 7088-7098
- Gazda HT, Sheen MR, Vlachos A, Choessel V, O'Donohue MF, Schneider H, Darras N, Hasman C, Sieff CA, Newburger PE, Ball SE, Niewiadomska E, Matysiak M, Zaucha JM, Glader B, Niemeyer C, Meerpohl JJ, Atsidaftos E, Lipton JM, Gleizes PE, Beggs AH (2008) Ribosomal protein L5 and L11 mutations are associated with cleft palate and abnormal thumbs in Diamond-Blackfan anemia patients. *Am J Hum Genet* **83**(6): 769-780
- Geerlings TH, Faber AW, Bister MD, Vos JC, Raue HA (2003) Rio2p, an evolutionarily conserved, low abundant protein kinase essential for processing of 20 S Pre-rRNA in *Saccharomyces cerevisiae*. *J Biol Chem* **278**(25): 22537-22545
- Geerlings TH, Vos JC, Raue HA (2000) The final step in the formation of 25S rRNA in *Saccharomyces cerevisiae* is performed by 5'→3' exonucleases. *RNA* **6**(12): 1698-1703
- Gelperin D, Horton L, Beckman J, Hensold J, Lemmon SK (2001) Bms1p, a novel GTP-binding protein, and the related Tsr1p are required for distinct steps of 40S ribosome biogenesis in yeast. *RNA* **7**(9): 1268-1283
- Gerbi SA, Borovjagin AV, Lange TS (2003a) The nucleolus: a site of ribonucleoprotein maturation. *Curr Opin Cell Biol* **15**(3): 318-325
- Gerbi SA, Borovjagin AV, Odreman FE, Lange TS (2003b) U4 snRNA nucleolar localization requires the NHPX/15.5-kD protein binding site but not Sm protein or U6 snRNA association. *J Cell Biol* **162**(5): 821-832
- Gerbi SA, Lange TS (2002) All small nuclear RNAs (snRNAs) of the [U4/U6.U5] Tri-snRNP localize to nucleoli; Identification of the nucleolar localization element of U6 snRNA. *Mol Biol Cell* **13**(9): 3123-3137
- Gerczei T, Correll CC (2004) Imp3p and Imp4p mediate formation of essential U3-precursor rRNA (pre-rRNA) duplexes, possibly to recruit the small subunit processome to the pre-rRNA. *Proc Natl Acad Sci U S A* **101**(43): 15301-15306
- Gerczei T, Shah BN, Manzo AJ, Walter NG, Correll CC (2009) RNA chaperones stimulate formation and yield of the U3 snoRNA-Pre-rRNA duplexes needed for eukaryotic ribosome biogenesis. *J Mol Biol* **390**(5): 991-1006
- Gerlich D, Beaudouin J, Kalbfuss B, Daigle N, Eils R, Ellenberg J (2003) Global chromosome positions are transmitted through mitosis in mammalian cells. *Cell* **112**(6): 751-764
- Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS (2003) Global analysis of protein expression in yeast. *Nature* **425**(6959): 737-741
- Gilbert W (1986) Origin of life: The RNA world. *Nature* **319**(618)
- Gill T, Cai T, Aulds J, Wierzbicki S, Schmitt ME (2004) RNase MRP cleaves the CLB2 mRNA to promote cell cycle progression: novel method of mRNA degradation. *Mol Cell Biol* **24**(3): 945-953
- Ginisty H, Sicard H, Roger B, Bouvet P (1999) Structure and functions of nucleolin. *J Cell Sci* **112** ( Pt 6): 761-772
- Giorgi C, Fatica A, Nagel R, Bozzoni I (2001) Release of U18 snoRNA from its host intron requires interaction of Nop1p with the Rnt1p endonuclease. *EMBO J* **20**(23): 6856-6865
- Girard C, Verheggen C, Neel H, Cammas A, Vagner S, Soret J, Bertrand E, Bordonne R (2008) Characterization of a short isoform of human Tgs1 hypermethylase associating with small nucleolar ribonucleoprotein core proteins and produced by limited proteolytic processing. *J Biol Chem* **283**(4): 2060-2069

- Girard JP, Lehtonen H, Caizergues-Ferrer M, Amalric F, Tollervey D, Lapeyre B (1992) GAR1 is an essential small nucleolar RNP protein required for pre-rRNA processing in yeast. *EMBO J* **11**(2): 673-682
- Godin KS, Walbott H, Leulliot N, van Tilbeurgh H, Varani G (2009) The box H/ACA snoRNP assembly factor Shq1p is a chaperone protein homologous to Hsp90 cochaperones that binds to the Cbf5p enzyme. *J Mol Biol* **390**(2): 231-244
- Gohshi T, Shimada M, Kawahire S, Imai N, Ichimura T, Omata S, Horigome T (1999) Molecular cloning of mouse p47, a second group mammalian RuvB DNA helicase-like protein: homology with those from human and *Saccharomyces cerevisiae*. *J Biochem* **125**(5): 939-946
- Goldfeder MB, Oliveira CC (2008) Cwc24p, a novel *Saccharomyces cerevisiae* nuclear ring finger protein, affects pre-snoRNA U3 splicing. *J Biol Chem* **283**(5): 2644-2653
- Golembe TJ, Yong J, Dreyfuss G (2005) Specific sequence features, recognized by the SMN complex, identify snRNAs and determine their fate as snRNPs. *Mol Cell Biol* **25**(24): 10989-11004
- Gondor A, Ohlsson R (2009) Chromosome crosstalk in three dimensions. *Nature* **461**(7261): 212-217
- Gonzales FA, Zanchin NI, Luz JS, Oliveira CC (2005) Characterization of *Saccharomyces cerevisiae* Nop17p, a novel Nop58p-interacting protein that is involved in Pre-rRNA processing. *J Mol Biol* **346**(2): 437-455
- Gonzalez-Santos JM, Wang A, Jones J, Ushida C, Liu J, Hu J (2002) Central region of the human splicing factor Hprp3p interacts with Hprp4p. *J Biol Chem* **277**(26): 23764-23772
- Gonzalez S, Pisano DG, Serrano M (2008) Mechanistic principles of chromatin remodeling guided by siRNAs and miRNAs. *Cell Cycle* **7**(16): 2601-2608
- Goodstadt L, Ponting CP (2006) Phylogenetic reconstruction of orthology, paralogy, and conserved synteny for dog and human. *PLoS Comput Biol* **2**(9): e133
- Goody TA, Melcher SE, Norman DG, Lilley DM (2004) The kink-turn motif in RNA is dimorphic, and metal ion-dependent. *RNA* **10**(2): 254-264
- Gorisch SM, Wachsmuth M, Ittrich C, Bacher CP, Rippe K, Lichter P (2004) Nuclear body movement is determined by chromatin accessibility and dynamics. *Proc Natl Acad Sci U S A* **101**(36): 13221-13226
- Gottschalk A, Neubauer G, Banroques J, Mann M, Luhrmann R, Fabrizio P (1999) Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6.U5] tri-snoRNP. *EMBO J* **18**(16): 4535-4548
- Grainger RM, Maizels N (1980) Dictyostelium ribosomal RNA is processed during transcription. *Cell* **20**(3): 619-623
- Granato DC, Gonzales FA, Luz JS, Cassiola F, Machado-Santelli GM, Oliveira CC (2005) Nop53p, an essential nucleolar protein that interacts with Nop17p and Nip7p, is required for pre-rRNA processing in *Saccharomyces cerevisiae*. *FEBS J* **272**(17): 4450-4463
- Grandi P, Rybin V, Bassler J, Petfalski E, Strauss D, Marzioch M, Schafer T, Kuster B, Tschochner H, Tollervey D, Gavin AC, Hurt E (2002) 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. *Mol Cell* **10**(1): 105-115
- Granneman S, Gallagher JE, Vogelzangs J, Horstman W, van Venrooij WJ, Baserga SJ, Pruijn GJ (2003) The human Imp3 and Imp4 proteins form a ternary complex with hMpp10, which only interacts with the U3 snoRNA in 60-80S ribonucleoprotein complexes. *Nucleic Acids Res* **31**(7): 1877-1887
- Granneman S, Kudla G, Petfalski E, Tollervey D (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc Natl Acad Sci U S A* **106**(24): 9613-9618
- Granneman S, Nandineni MR, Baserga SJ (2005) The putative NTPase Fap7 mediates cytoplasmic 20S pre-rRNA processing through a direct interaction with Rps14. *Mol Cell Biol* **25**(23): 10352-10364
- Granneman S, Pruijn GJ, Horstman W, van Venrooij WJ, Luhrmann R, Watkins NJ (2002) The hU3-55K protein requires 15.5K binding to the box B/C motif as well as flanking RNA elements for its association with the U3 small nucleolar RNA in Vitro. *J Biol Chem* **277**(50): 48490-48500
- Granneman S, Vogelzangs J, Luhrmann R, van Venrooij WJ, Pruijn GJ, Watkins NJ (2004) Role of pre-rRNA base pairing and 80S complex formation in subnucleolar localization of the U3 snoRNP. *Mol Cell Biol* **24**(19): 8600-8610
- Grant PA, Schieltz D, Pray-Grant MG, Steger DJ, Reese JC, Yates JR, 3rd, Workman JL (1998) A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* **94**(1): 45-53
- Gribun A, Cheung KL, Huen J, Ortega J, Houry WA (2008) Yeast Rvb1 and Rvb2 are ATP-dependent DNA helicases that form a heterohexameric complex. *J Mol Biol* **376**(5): 1320-1333
- Grimmler M, Bauer L, Nousiainen M, Korner R, Meister G, Fischer U (2005a) Phosphorylation regulates the activity of the SMN complex during assembly of spliceosomal U snRNPs. *EMBO Rep* **6**(1): 70-76
- Grimmler M, Otter S, Peter C, Muller F, Chari A, Fischer U (2005b) Unrip, a factor implicated in cap-independent translation, associates with the cytosolic SMN complex and influences its intracellular localization. *Hum Mol Genet* **14**(20): 3099-3111
- Grobelyny JV, Godwin AK, Broccoli D (2000) ALT-associated PML bodies are present in viable cells and are enriched in cells in the G(2)/M phase of the cell cycle. *J Cell Sci* **113 Pt 24**: 4577-4585
- Gross JD, Moerke NJ, von der Haar T, Lugovskoy AA, Sachs AB, McCarthy JE, Wagner G (2003) Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* **115**(6): 739-750

- Grozdanov PN, Fernandez-Fuentes N, Fiser A, Meier UT (2009a) Pathogenic NAP57 mutations decrease ribonucleoprotein assembly in dyskeratosis congenita. *Hum Mol Genet* **18**(23): 4546-4551
- Grozdanov PN, Roy S, Kittur N, Meier UT (2009b) SHQ1 is required prior to NAF1 for assembly of H/ACA small nucleolar and telomerase RNPs. *RNA* **15**(6): 1188-1197
- Gruenbaum Y, Margalit A, Goldman RD, Shumaker DK, Wilson KL (2005) The nuclear lamina comes of age. *Nat Rev Mol Cell Biol* **6**(1): 21-31
- Grund SE, Fischer T, Cabal GG, Antunez O, Perez-Ortin JE, Hurt E (2008) The inner nuclear membrane protein Src1 associates with subtelomeric genes and alters their regulated gene expression. *J Cell Biol* **182**(5): 897-910
- Grzechnik P, Kufel J (2008) Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast. *Mol Cell* **32**(2): 247-258
- Gubitz AK, Feng W, Dreyfuss G (2004) The SMN complex. *Exp Cell Res* **296**(1): 51-56
- Gubitz AK, Mourelatos Z, Abel L, Rappsilber J, Mann M, Dreyfuss G (2002) Gemin5, a novel WD repeat protein component of the SMN complex that binds Sm proteins. *J Biol Chem* **277**(7): 5631-5636
- Gudipati RK, Villa T, Boulay J, Libri D (2008) Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. *Nat Struct Mol Biol* **15**(8): 786-794
- Guo A, Salomoni P, Luo J, Shih A, Zhong S, Gu W, Pandolfi PP (2000) The function of PML in p53-dependent apoptosis. *Nat Cell Biol* **2**(10): 730-736
- Gupta R, Kus B, Fladd C, Wasmuth J, Tonikian R, Sidhu S, Krogan NJ, Parkinson J, Rotin D (2007) Ubiquitination screen using protein microarrays for comprehensive identification of Rsp5 substrates in yeast. *Mol Syst Biol* **3**: 116
- Gurha P, Joardar A, Chaurasia P, Gupta R (2007) Differential roles of archaeal box H/ACA proteins in guide RNA-dependent and independent pseudouridine formation. *RNA Biol* **4**(2): 101-109
- Gutell RR, Larsen N, Woese CR (1994) Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. *Microbiol Rev* **58**(1): 10-26
- 
- Haeusler RA, Engelke DR (2006) Spatial organization of transcription by RNA polymerase III. *Nucleic Acids Res* **34**(17): 4826-4836
- Hage AE, Tollervey D (2004) A surfeit of factors: why is ribosome assembly so much more complicated in eukaryotes than bacteria? *RNA Biol* **1**(1): 10-15
- Hahn S (2004) Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* **11**(5): 394-403
- Hall TM (2005) Multiple modes of RNA recognition by zinc finger proteins. *Curr Opin Struct Biol* **15**(3): 367-373
- Hamma T, Ferre-D'Amare AR (2004) Structure of protein L7Ae bound to a K-turn derived from an archaeal box H/ACA sRNA at 1.8 Å resolution. *Structure* **12**(5): 893-903
- Hamma T, Reichow SL, Varani G, Ferre-D'Amare AR (2005) The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat Struct Mol Biol* **12**(12): 1101-1107
- Hanson PI, Whiteheart SW (2005) AAA+ proteins: have engine, will work. *Nat Rev Mol Cell Biol* **6**(7): 519-529
- Hardin JW, Reyes FE, Batey RT (2009) Analysis of a critical interaction within the archaeal box C/D small ribonucleoprotein complex. *J Biol Chem* **284**(22): 15317-15324
- Harnpicharnchai P, Jakovljevic J, Horsey E, Miles T, Roman J, Rout M, Meagher D, Imai B, Guo Y, Brame CJ, Shabanowitz J, Hunt DF, Woolford JL, Jr. (2001) Composition and functional characterization of yeast 66S ribosome assembly intermediates. *Mol Cell* **8**(3): 505-515
- Hartshorne T, Agabian N (1994) A common core structure for U3 small nucleolar RNAs. *Nucleic Acids Res* **22**(16): 3354-3364
- He F, Umehara T, Tsuda K, Inoue M, Kigawa T, Matsuda T, Yabuki T, Aoki M, Seki E, Terada T, Shirouzu M, Tanaka A, Sugano S, Muto Y, Yokoyama S (2007) Solution structure of the zinc finger HIT domain in protein FON. *Protein Sci* **16**(8): 1577-1587
- He H, Liyanarachchi S, Akagi K, Nagy R, Li J, Dietrich RC, Li W, Sebastian N, Wen B, Xin B, Singh J, Yan P, Alder H, Haan E, Wiczorek D, Albrecht B, Puffenberger E, Wang H, Westman JA, Padgett RA, Symer DE, de la Chapelle A (2011) Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD I. *Science* **332**(6026): 238-240
- Heard E, Bickmore W (2007) The ins and outs of gene regulation and chromosome territory organisation. *Curr Opin Cell Biol* **19**(3): 311-316
- Hebert MD, Shpargel KB, Ospina JK, Tucker KE, Matera AG (2002) Coilin methylation regulates nuclear body formation. *Dev Cell* **3**(3): 329-337
- Hebert MD, Szymczyk PW, Shpargel KB, Matera AG (2001) Coilin forms the bridge between Cajal bodies and SMN, the spinal muscular atrophy protein. *Genes Dev* **15**(20): 2720-2729
- Hedges J, West M, Johnson AW (2005) Release of the export adapter, Nmd3p, from the 60S ribosomal subunit requires Rpl10p and the cytoplasmic GTPase Lsg1p. *EMBO J* **24**(3): 567-579
- Heiss NS, Knight SW, Vulliamy TJ, Klauk SM, Wiemann S, Mason PJ, Poustka A, Dokal I (1998) X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet* **19**(1): 32-38

- Hengartner CJ, Myer VE, Liao SM, Wilson CJ, Koh SS, Young RA (1998) Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Mol Cell* **2**(1): 43-53
- Henras A, Dez C, Noaillac-Depeyre J, Henry Y, Caizergues-Ferrer M (2001) Accumulation of H/ACA snoRNPs depends on the integrity of the conserved central domain of the RNA-binding protein Nhp2p. *Nucleic Acids Res* **29**(13): 2733-2746
- Henras A, Henry Y, Bousquet-Antonelli C, Noaillac-Depeyre J, Gelugne JP, Caizergues-Ferrer M (1998) Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *EMBO J* **17**(23): 7078-7090
- Henras AK, Bertrand E, Chanfreau G (2004) A cotranscriptional model for 3'-end processing of the *Saccharomyces cerevisiae* pre-ribosomal RNA precursor. *RNA* **10**(10): 1572-1585
- Henras AK, Soudet J, Gerus M, Lebaron S, Caizergues-Ferrer M, Mougin A, Henry Y (2008) The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci* **65**(15): 2334-2359
- Henry Y, Wood H, Morrissey JP, Petfalski E, Kearsey S, Tollervey D (1994) The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. *EMBO J* **13**(10): 2452-2463
- Hermann H, Fabrizio P, Raker VA, Foulaki K, Hornig H, Brahm H, Luhrmann R (1995) snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J* **14**(9): 2076-2088
- Hernandez N (1985) Formation of the 3' end of U1 snRNA is directed by a conserved sequence located downstream of the coding region. *EMBO J* **4**(7): 1827-1837
- Hessling M, Richter K, Buchner J (2009) Dissection of the ATP-induced conformational cycle of the molecular chaperone Hsp90. *Nat Struct Mol Biol* **16**(3): 287-293
- Hetzer MW, Wente SR (2009) Border control at the nucleus: biogenesis and organization of the nuclear membrane and pore complexes. *Dev Cell* **17**(5): 606-616
- Hiley SL, Babak T, Hughes TR (2005) Global analysis of yeast RNA processing identifies new targets of RNase III and uncovers a link between tRNA 5' end processing and tRNA splicing. *Nucleic Acids Res* **33**(9): 3048-3056
- Hilleren P, Parker R (2001) Defects in the mRNA export factors Rat7p, Gle1p, Mex67p, and Rat8p cause hyperadenylation during 3'-end formation of nascent transcripts. *RNA* **7**(5): 753-764
- Hinsby AM, Kiemer L, Karlberg EO, Lage K, Fausboll A, Juncker AS, Andersen JS, Mann M, Brunak S (2006) A wiring of the human nucleolus. *Mol Cell* **22**(2): 285-295
- Hirose T, Ideue T, Nagai M, Hagiwara M, Shu MD, Steitz JA (2006) A spliceosomal intron binding protein, IBP160, links position-dependent assembly of intron-encoded box C/D snoRNP to pre-mRNA splicing. *Mol Cell* **23**(5): 673-684
- Hirose T, Shu MD, Steitz JA (2003) Splicing-dependent and -independent modes of assembly for intron-encoded box C/D snoRNPs in mammalian cells. *Mol Cell* **12**(1): 113-123
- Hirose T, Steitz JA (2001) Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. *Proc Natl Acad Sci U S A* **98**(23): 12914-12919
- Hirose Y, Ohkuma Y (2007) Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression. *J Biochem* **141**(5): 601-608
- Hishida T, Han YW, Fujimoto S, Iwasaki H, Shinagawa H (2004) Direct evidence that a conserved arginine in RuvB AAA+ ATPase acts as an allosteric effector for the ATPase activity of the adjacent subunit in a hexamer. *Proc Natl Acad Sci U S A* **101**(26): 9573-9577
- Ho CK, Shuman S (1999) Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol Cell* **3**(3): 405-411
- Ho JH, Kallstrom G, Johnson AW (2000) Nascent 60S ribosomal subunits enter the free pool bound by Nmd3p. *RNA* **6**(11): 1625-1634
- Hoareau-Aveilla C, Bonoli M, Caizergues-Ferrer M, Henry Y (2006) hNaf1 is required for accumulation of human box H/ACA snoRNPs, scaRNPs, and telomerase. *RNA* **12**(5): 832-840
- Hobor F, Pergoli R, Kubicek K, Hrossova D, Bacikova V, Zimmermann M, Pasulka J, Hofr C, Vanacova S, Stefl R (2011) Recognition of transcription termination signal by the nuclear polyadenylated RNA-binding (NAB) 3 protein. *J Biol Chem* **286**(5): 3645-3657
- Hocine S, Singer RH, Grunwald D (2010) RNA processing and export. *Cold Spring Harb Perspect Biol* **2**(12): a000752
- Hogg JR, Collins K (2008) Structured non-coding RNAs and the RNP Renaissance. *Curr Opin Chem Biol* **12**(6): 684-689
- Holzmann K, Gerner C, Korosec T, Polt A, Grimm R, Sauermaun G (1998) Identification and characterization of the ubiquitously occurring nuclear matrix protein NMP 238. *Biochem Biophys Res Commun* **252**(1): 39-45
- Hong B, Brockenbrough JS, Wu P, Aris JP (1997) Nop2p is required for pre-rRNA processing and 60S ribosome subunit synthesis in yeast. *Mol Cell Biol* **17**(1): 378-388
- Horejsi Z, Takai H, Adelman CA, Collis SJ, Flynn H, Maslen S, Skehel JM, de Lange T, Boulton SJ (2010) CK2 phospho-dependent binding of R2TP complex to TEL2 is essential for mTOR and SMG1 stability. *Mol Cell* **39**(6): 839-850
- Horowitz DS, Kobayashi R, Krainer AR (1997) A new cyclophilin and the human homologues of yeast Prp3 and

- Prp4 form a complex associated with U4/U6 snRNPs. *RNA* **3**(12): 1374-1387
- Horowitz DS, Lee EJ, Mabon SA, Misteli T (2002) A cyclophilin functions in pre-mRNA splicing. *EMBO J* **21**(3): 470-480
- Hosoda N, Kim YK, Lejeune F, Maquat LE (2005) CBP80 promotes interaction of Upf1 with Upf2 during nonsense-mediated mRNA decay in mammalian cells. *Nat Struct Mol Biol* **12**(10): 893-901
- Houseley J, LaCava J, Tollervey D (2006) RNA-quality control by the exosome. *Nat Rev Mol Cell Biol* **7**(7): 529-539
- Houseley J, Tollervey D (2006) Yeast Trf5p is a nuclear poly(A) polymerase. *EMBO Rep* **7**(2): 205-211
- Hsu CL, Stevens A (1993) Yeast cells lacking 5'→3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure. *Mol Cell Biol* **13**(8): 4826-4835
- Huang GM, Jarmolowski A, Struck JC, Fournier MJ (1992) Accumulation of U14 small nuclear RNA in *Saccharomyces cerevisiae* requires box C, box D, and a 5', 3' terminal stem. *Mol Cell Biol* **12**(10): 4456-4463
- Huang S (2002) Building an efficient factory: where is pre-rRNA synthesized in the nucleolus? *J Cell Biol* **157**(5): 739-741
- Huang S, Zhou H, Katzmann D, Hochstrasser M, Atanasova E, Zhang Z (2005) Rtt106p is a histone chaperone involved in heterochromatin-mediated silencing. *Proc Natl Acad Sci U S A* **102**(38): 13410-13415
- Huang S, Zhou H, Tarara J, Zhang Z (2007) A novel role for histone chaperones CAF-1 and Rtt106p in heterochromatin silencing. *EMBO J* **26**(9): 2274-2283
- Huber J, Cronshagen U, Kadokura M, Marshallsay C, Wada T, Sekine M, Luhrmann R (1998) Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *EMBO J* **17**(14): 4114-4126
- Huen J, Kakahara Y, Ugwu F, Cheung KL, Ortega J, Houry WA (2010) Rvb1-Rvb2: essential ATP-dependent helicases for critical complexes. *Biochem Cell Biol* **88**(1): 29-40
- Hughes JM (1996) Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit RNA. *J Mol Biol* **259**(4): 645-654
- Hughes JM, Ares M, Jr. (1991) Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA. *EMBO J* **10**(13): 4231-4239
- Hughes JM, Konings DA, Cesareni G (1987) The yeast homologue of U3 snRNA. *EMBO J* **6**(7): 2145-2155
- Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK (2003) Global analysis of protein localization in budding yeast. *Nature* **425**(6959): 686-691
- Hung NJ, Lo KY, Patel SS, Helmke K, Johnson AW (2008) Arx1 is a nuclear export receptor for the 60S ribosomal subunit in yeast. *Mol Biol Cell* **19**(2): 735-744
- Huranova M, Hnilicova J, Fleischer B, Cvackova Z, Stanek D (2009) A mutation linked to retinitis pigmentosa in HPRP31 causes protein instability and impairs its interactions with spliceosomal snRNPs. *Hum Mol Genet* **18**(11): 2014-2023
- Hwang J, Sato H, Tang Y, Matsuda D, Maquat LE (2010) UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps. *Mol Cell* **39**(3): 396-409
- Iborra FJ, Pombo A, Jackson DA, Cook PR (1996) Active RNA polymerases are localized within discrete transcription "factories" in human nuclei. *J Cell Sci* **109** ( Pt 6): 1427-1436
- Ikura T, Ogryzko VV, Grigoriev M, Groisman R, Wang J, Horikoshi M, Scully R, Qin J, Nakatani Y (2000) Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. *Cell* **102**(4): 463-473
- Imbeault D, Gamar L, Rufiange A, Paquet E, Nourani A (2008) The Rtt106 histone chaperone is functionally linked to transcription elongation and is involved in the regulation of spurious transcription from cryptic promoters in yeast. *J Biol Chem* **283**(41): 27350-27354
- Inada M, Guthrie C (2004) Identification of Lhp1p-associated RNAs by microarray analysis in *Saccharomyces cerevisiae* reveals association with coding and noncoding RNAs. *Proc Natl Acad Sci U S A* **101**(2): 434-439
- Ingelfinger D, Arndt-Jovin DJ, Luhrmann R, Achsel T (2002) The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrn1 in distinct cytoplasmic foci. *RNA* **8**(12): 1489-1501
- Isaac C, Yang Y, Meier UT (1998) Nopp140 functions as a molecular link between the nucleolus and the coiled bodies. *J Cell Biol* **142**(2): 319-329
- Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A* **98**(8): 4569-4574
- Itsuki Y, Saeki M, Nakahara H, Egusa H, Irie Y, Terao Y, Kawabata S, Yatani H, Kamisaki Y (2008) Molecular cloning of novel Monad binding protein containing tetratricopeptide repeat domains. *FEBS Lett* **582**(16): 2365-2370
- Izaurrealde E, Lewis J, Gamberi C, Jarmolowski A, McGuigan C, Mattaj IW (1995) A cap-binding protein complex mediating U snRNA export. *Nature* **376**(6542): 709-712

- Izquierdo JM, Valcarcel J (2006) A simple principle to explain the evolution of pre-mRNA splicing. *Genes Dev* **20**(13): 1679-1684
- Izumi N, Yamashita A, Iwamatsu A, Kurata R, Nakamura H, Saari B, Hirano H, Anderson P, Ohno S (2010) AAA+ proteins RUVBL1 and RUVBL2 coordinate PIKK activity and function in nonsense-mediated mRNA decay. *Sci Signal* **3**(116): ra27
- 
- J**acobson MR, Pederson T (1998) Localization of signal recognition particle RNA in the nucleolus of mammalian cells. *Proc Natl Acad Sci U S A* **95**(14): 7981-7986
- Jady BE, Bertrand E, Kiss T (2004) Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *J Cell Biol* **164**(5): 647-652
- Jady BE, Darzacq X, Tucker KE, Matera AG, Bertrand E, Kiss T (2003) Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *EMBO J* **22**(8): 1878-1888
- Jady BE, Kiss T (2001) A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO J* **20**(3): 541-551
- Jady BE, Richard P, Bertrand E, Kiss T (2006) Cell cycle-dependent recruitment of telomerase RNA and Cajal bodies to human telomeres. *Mol Biol Cell* **17**(2): 944-954
- Jansen RP, Hurt EC, Kern H, Lehtonen H, Carmo-Fonseca M, Lapeyre B, Tollervey D (1991) Evolutionary conservation of the human nucleolar protein fibrillarin and its functional expression in yeast. *J Cell Biol* **113**(4): 715-729
- Jarmolowski A, Zagorski J, Li HV, Fournier MJ (1990) Identification of essential elements in U14 RNA of *Saccharomyces cerevisiae*. *EMBO J* **9**(13): 4503-4509
- Jeffares DC, Poole AM, Penny D (1998) Relics from the RNA world. *J Mol Evol* **46**(1): 18-36
- Jeppesen C, Stebbins-Boaz B, Gerbi SA (1988) Nucleotide sequence determination and secondary structure of *Xenopus* U3 snRNA. *Nucleic Acids Res* **16**(5): 2127-2148
- Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, Therien C, Bergeron D, Bourassa S, Greenblatt J, Chabot B, Poirier GG, Hughes TR, Blanchette M, Price DH, Coulombe B (2007) Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell* **27**(2): 262-274
- Jha S, Dutta A (2009) RVB1/RVB2: running rings around molecular biology. *Mol Cell* **34**(5): 521-533
- Jha S, Shibata E, Dutta A (2008) Human Rvb1/Tip49 is required for the histone acetyltransferase activity of Tip60/NuA4 and for the downregulation of phosphorylation on H2AX after DNA damage. *Mol Cell Biol* **28**(8): 2690-2700
- Jiang W, Middleton K, Yoon HJ, Fouquet C, Carbon J (1993) An essential yeast protein, CBF5p, binds in vitro to centromeres and microtubules. *Mol Cell Biol* **13**(8): 4884-4893
- Jun J, Cai Y, Yao T, Gottschalk AJ, Florens L, Swanson SK, Gutierrez JL, Coleman MK, Workman JL, Mushegian A, Washburn MP, Conaway RC, Conaway JW (2005) A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex. *J Biol Chem* **280**(50): 41207-41212
- Joardar A, Gurha P, Skariah G, Gupta R (2008) Box C/D RNA-guided 2'-O methylations and the intron of tRNATrp are not essential for the viability of *Haloflex volcanii*. *J Bacteriol* **190**(21): 7308-7313
- Johnson A, O'Donnell M (2003) Ordered ATP hydrolysis in the gamma complex clamp loader AAA+ machine. *J Biol Chem* **278**(16): 14406-14413
- Jones KW, Gorzynski K, Hales CM, Fischer U, Badbanchi F, Terns RM, Terns MP (2001) Direct interaction of the spinal muscular atrophy disease protein SMN with the small nucleolar RNA-associated protein fibrillarin. *J Biol Chem* **276**(42): 38645-38651
- Jonsson ZO, Dhar SK, Narlikar GJ, Auty R, Wagle N, Pellman D, Pratt RE, Kingston R, Dutta A (2001) Rvb1p and Rvb2p are essential components of a chromatin remodeling complex that regulates transcription of over 5% of yeast genes. *J Biol Chem* **276**(19): 16279-16288
- Jonsson ZO, Jha S, Wohlschlegel JA, Dutta A (2004) Rvb1p/Rvb2p recruit Arp5p and assemble a functional Ino80 chromatin remodeling complex. *Mol Cell* **16**(3): 465-477
- Jorgensen R (1990) Altered gene expression in plants due to trans interactions between homologous genes. *Trends Biotechnol* **8**(12): 340-344
- Joyce GF (2002) The antiquity of RNA-based evolution. *Nature* **418**(6894): 214-221
- 
- K**ambach C, Walke S, Young R, Avis JM, de la Fortelle E, Raker VA, Luhrmann R, Li J, Nagai K (1999) Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* **96**(3): 375-387
- Kameoka S, Duque P, Konarska MM (2004) p54(nrb) associates with the 5' splice site within large transcription/splicing complexes. *EMBO J* **23**(8): 1782-1791
- Kanemaki M, Kurokawa Y, Matsu-ura T, Makino Y, Masani A, Okazaki K, Morishita T, Tamura TA (1999) TIP49b, a new RuvB-like DNA helicase, is included in a complex together with another RuvB-like DNA helicase, TIP49a. *J Biol Chem* **274**(32): 22437-22444
- Kanemaki M, Makino Y, Yoshida T, Kishimoto T, Koga A, Yamamoto K, Yamamoto M, Moncollin V, Egly JM, Muramatsu M, Tamura T (1997) Molecular cloning of a rat 49-kDa TBP-interacting protein (TIP49) that is highly



- homologous to the bacterial RuvB. *Biochem Biophys Res Commun* **235**(1): 64-68
- Karata K, Inagawa T, Wilkinson AJ, Tatsuta T, Ogura T (1999) Dissecting the role of a conserved motif (the second region of homology) in the AAA family of ATPases. Site-directed mutagenesis of the ATP-dependent protease FtsH. *J Biol Chem* **274**(37): 26225-26232
- Karbstein K, Jonas S, Doudna JA (2005) An essential GTPase promotes assembly of preribosomal RNA processing complexes. *Mol Cell* **20**(4): 633-643
- Karijolich J, Stephenson D, Yu YT (2007) Biochemical purification of box H/ACA RNPs involved in pseudouridylation. *Methods Enzymol* **425**: 241-262
- Kass S, Tyc K, Steitz JA, Sollner-Webb B (1990) The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. *Cell* **60**(6): 897-908
- Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap CC, Suzuki M, Kawai J, Suzuki H, Carninci P, Hayashizaki Y, Wells C, Frith M, Ravasi T, Pang KC, Hallinan J, Mattick J, Hume DA, Lipovich L, Batalov S, Engstrom PG, Mizuno Y, Faghihi MA, Sandelin A, Chalk AM, Mottagui-Tabar S, Liang Z, Lenhard B, Wahlestedt C (2005) Antisense transcription in the mammalian transcriptome. *Science* **309**(5740): 1564-1566
- Kato N, Harada F (1984) Nucleotide sequence of nuclear 5.4 S RNA of mouse cells. *Biochim Biophys Acta* **782**(2): 127-131
- Kawai G, Yamamoto Y, Kamimura T, Masegi T, Sekine M, Hata T, Iimori T, Watanabe T, Miyazawa T, Yokoyama S (1992) Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2'-hydroxyl group. *Biochemistry* **31**(4): 1040-1046
- Kawakami K, Shafer BK, Garfinkel DJ, Strathern JN, Nakamura Y (1992) Ty element-induced temperature-sensitive mutations of *Saccharomyces cerevisiae*. *Genetics* **131**(4): 821-832
- Kenna M, Stevens A, McCammon M, Douglas MG (1993) An essential yeast gene with homology to the exonuclease-encoding XRN1/KEM1 gene also encodes a protein with exoribonuclease activity. *Mol Cell Biol* **13**(1): 341-350
- Kenniston JA, Baker TA, Fernandez JM, Sauer RT (2003) Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* **114**(4): 511-520
- Kessler SH, Sachs AB (1998) RNA recognition motif 2 of yeast Pab1p is required for its functional interaction with eukaryotic translation initiation factor 4G. *Mol Cell Biol* **18**(1): 51-57
- Khanna M, Wu H, Johansson C, Caizergues-Ferrer M, Feigon J (2006) Structural study of the H/ACA snoRNP components Nop10p and the 3' hairpin of U65 snoRNA. *RNA* **12**(1): 40-52
- Kim JH, Kim B, Cai L, Choi HJ, Ohgi KA, Tran C, Chen C, Chung CH, Huber O, Rose DW, Sawyers CL, Rosenfeld MG, Baek SH (2005) Transcriptional regulation of a metastasis suppressor gene by Tip60 and beta-catenin complexes. *Nature* **434**(7035): 921-926
- Kim JW, Kim HC, Kim GM, Yang JM, Boeke JD, Nam K (2000) Human RNA lariat debranching enzyme cDNA complements the phenotypes of *Saccharomyces cerevisiae* *dbl1* and *Schizosaccharomyces pombe* *dbl1* mutants. *Nucleic Acids Res* **28**(18): 3666-3673
- Kim M, Krogan NJ, Vasiljeva L, Rando OJ, Nedeia E, Greenblatt JF, Buratowski S (2004) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**(7016): 517-522
- Kim VN, Han J, Siomi MC (2009a) Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* **10**(2): 126-139
- Kim YS, Alarcon SV, Lee S, Lee MJ, Giaccone G, Neckers L, Trepel JB (2009b) Update on Hsp90 inhibitors in clinical trial. *Curr Top Med Chem* **9**(15): 1479-1492
- King TH, Decatur WA, Bertrand E, Maxwell ES, Fournier MJ (2001) A well-connected and conserved nucleoplasmic helicase is required for production of box C/D and H/ACA snoRNAs and localization of snoRNP proteins. *Mol Cell Biol* **21**(22): 7731-7746
- King TH, Liu B, McCully RR, Fournier MJ (2003) Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol Cell* **11**(2): 425-435
- Kiss-Laszlo Z, Henry Y, Bachellerie JP, Caizergues-Ferrer M, Kiss T (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* **85**(7): 1077-1088
- Kiss-Laszlo Z, Henry Y, Kiss T (1998) Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *EMBO J* **17**(3): 797-807
- Kiss AM, Jady BE, Bertrand E, Kiss T (2004) Human box H/ACA pseudouridylation guide RNA machinery. *Mol Cell Biol* **24**(13): 5797-5807
- Kiss AM, Jady BE, Darzacq X, Verheggen C, Bertrand E, Kiss T (2002) A Cajal body-specific pseudouridylation guide RNA is composed of two box H/ACA snoRNA-like domains. *Nucleic Acids Res* **30**(21): 4643-4649
- Kiss T, Fayet-Lebaron E, Jady BE (2010) Box H/ACA small ribonucleoproteins. *Mol Cell* **37**(5): 597-606
- Kiss T, Filipowicz W (1993) Small nucleolar RNAs encoded by introns of the human cell cycle regulatory gene *RCC1*. *EMBO J* **12**(7): 2913-2920
- Kiss T, Filipowicz W (1995) Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. *Genes Dev* **9**(11): 1411-1424
- Kitao S, Segref A, Kast J, Wilm M, Mattaj IW, Ohno M (2008) A compartmentalized phosphorylation/

- dephosphorylation system that regulates U snRNA export from the nucleus. *Mol Cell Biol* **28**(1): 487-497
- Kittur N, Darzacq X, Roy S, Singer RH, Meier UT (2006) Dynamic association and localization of human H/ACA RNP proteins. *RNA* **12**(12): 2057-2062
- Klein DJ, Schmeing TM, Moore PB, Steitz TA (2001) The kink-turn: a new RNA secondary structure motif. *EMBO J* **20**(15): 4214-4221
- Knox AA, McKeegan KS, Debieux CM, Traynor A, Richardson H, Watkins NJ (2011) A weak C' box renders U3 snoRNA levels dependent on hU3-55K binding. *Mol Cell Biol* **31**(12): 2404-2412
- Kobor MS, Venkatasubrahmanyam S, Meneghini MD, Gin JW, Jennings JL, Link AJ, Madhani HD, Rine J (2004) A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol* **2**(5): E131
- Komarnitsky P, Cho EJ, Buratowski S (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev* **14**(19): 2452-2460
- Koo BK, Park CJ, Fernandez CF, Chim N, Ding Y, Chanfreau G, Feigon J (2011) Structure of H/ACA RNP Protein Nhp2p Reveals Cis/Trans Isomerization of a Conserved Proline at the RNA and Nop10 Binding Interface. *J Mol Biol*
- Koonin EV (1996) Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases. *Nucleic Acids Res* **24**(12): 2411-2415
- Koonin EV, Bork P, Sander C (1994) A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins and a ribosome modification enzyme? *Nucleic Acids Res* **22**(11): 2166-2167
- Kos M, Tollervey D (2005) The Putative RNA Helicase Dbp4p Is Required for Release of the U14 snoRNA from Preribosomes in *Saccharomyces cerevisiae*. *Mol Cell* **20**(1): 53-64
- Kostova FV, Williams VC, Heemskerk J, Iannaccone S, Didonato C, Swoboda K, Maria BL (2007) Spinal muscular atrophy: classification, diagnosis, management, pathogenesis, and future research directions. *J Child Neurol* **22**(8): 926-945
- Kraft C, Deplazes A, Sohrmann M, Peter M (2008) Mature ribosomes are selectively degraded upon starvation by an autophagy pathway requiring the Ubp3p/Bre5p ubiquitin protease. *Nat Cell Biol* **10**(5): 602-610
- Kramer A, Keller W (1985) Purification of a protein required for the splicing of pre-mRNA and its separation from the lariat debranching enzyme. *EMBO J* **4**(13A): 3571-3581
- Kressler D, Doere M, Rojo M, Linder P (1999) Synthetic lethality with conditional dbp6 alleles identifies rsa1p, a nucleoplasmic protein involved in the assembly of 60S ribosomal subunits. *Mol Cell Biol* **19**(12): 8633-8645
- Kressler D, Hurt E, Bassler J (2010) Driving ribosome assembly. *Biochim Biophys Acta* **1803**(6): 673-683
- Kressler D, Roser D, Pertschy B, Hurt E (2008) The AAA ATPase Rix7 powers progression of ribosome biogenesis by stripping Nsa1 from pre-60S particles. *J Cell Biol* **181**(6): 935-944
- Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng MA, Dean K, Ryan OW, Golshani A, Johnston M, Greenblatt JF, Shilatifard A (2003a) The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell* **11**(3): 721-729
- Krogan NJ, Keogh MC, Datta N, Sawa C, Ryan OW, Ding H, Haw RA, Pootoolal J, Tong A, Canadien V, Richards DP, Wu X, Emili A, Hughes TR, Buratowski S, Greenblatt JF (2003b) A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol Cell* **12**(6): 1565-1576
- Krogan NJ, Peng WT, Cagney G, Robinson MD, Haw R, Zhong G, Guo X, Zhang X, Canadien V, Richards DP, Beattie BK, Lalev A, Zhang W, Davierwala AP, Mnaimneh S, Starostine A, Tikuisis AP, Grigull J, Datta N, Bray JE, Hughes TR, Emili A, Greenblatt JF (2004) High-definition macromolecular composition of yeast RNA-processing complexes. *Mol Cell* **13**(2): 225-239
- Kubicek K, Pasulka J, Cerna H, Lohr F, Stefl R (2011) (1)H, (13)C, and (15)N resonance assignments for the CTD-interacting domain of Nrd1 bound to Ser5-phosphorylated CTD of RNA polymerase II. *Biomol NMR Assign*
- Kucera NJ, Hodsdon ME, Wolin SL (2011) An intrinsically disordered C terminus allows the La protein to assist the biogenesis of diverse noncoding RNA precursors. *Proc Natl Acad Sci U S A* **108**(4): 1308-1313
- Kufel J, Allmang C, Chanfreau G, Petfalski E, Lafontaine DL, Tollervey D (2000) Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. *Mol Cell Biol* **20**(15): 5415-5424
- Kufel J, Allmang C, Verdone L, Beggs J, Tollervey D (2003) A complex pathway for 3' processing of the yeast U3 snoRNA. *Nucleic Acids Res* **31**(23): 6788-6797
- Kufel J, Allmang C, Verdone L, Beggs JD, Tollervey D (2002) Lsm proteins are required for normal processing of pre-tRNAs and their efficient association with La-homologous protein Lhp1p. *Mol Cell Biol* **22**(14): 5248-5256
- Kuhn JF, Tran EJ, Maxwell ES (2002) Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein. *Nucleic Acids Res* **30**(4): 931-941
- Kuras L, Struhl K (1999) Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme. *Nature* **399**(6736): 609-613

- Kwapisz M, Beckouet F, Thuriaux P (2008) Early evolution of eukaryotic DNA-dependent RNA polymerases. *Trends Genet* **24**(5): 211-215
- 
- L**aCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121**(5): 713-724
- Lafontaine D, Delcour J, Glasser AL, Desgres J, Vandenhaute J (1994) The DIM1 gene responsible for the conserved m6(2)Am6(2)A dimethylation in the 3'-terminal loop of 18 S rRNA is essential in yeast. *J Mol Biol* **241**(3): 492-497
- Lafontaine DL, Bousquet-Antonelli C, Henry Y, Caizergues-Ferrer M, Tollervey D (1998) The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev* **12**(4): 527-537
- Lafontaine DL, Tollervey D (1998) Birth of the snoRNPs: the evolution of the modification-guide snoRNAs. *Trends Biochem Sci* **23**(10): 383-388
- Lafontaine DL, Tollervey D (1999) Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. *RNA* **5**(3): 455-467
- Lafontaine DL, Tollervey D (2000) Synthesis and assembly of the box C+D small nucleolar RNPs. *Mol Cell Biol* **20**(8): 2650-2659
- Lamanna AC, Karbstein K (2009) Nob1 binds the single-stranded cleavage site D at the 3'-end of 18S rRNA with its PIN domain. *Proc Natl Acad Sci U S A* **106**(34): 14259-14264
- Lambert JP, Fillingham J, Siahbazi M, Greenblatt J, Baetz K, Figeys D (2010) Defining the budding yeast chromatin-associated interactome. *Mol Syst Biol* **6**: 448
- Lambert JP, Mitchell L, Rudner A, Baetz K, Figeys D (2009) A novel proteomics approach for the discovery of chromatin-associated protein networks. *Mol Cell Proteomics* **8**(4): 870-882
- Lamond AI, Spector DL (2003) Nuclear speckles: a model for nuclear organelles. *Nat Rev Mol Cell Biol* **4**(8): 605-612
- Lanctot C, Cheutin T, Cremer M, Cavalli G, Cremer T (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* **8**(2): 104-115
- Lane BG, Ofengand J, Gray MW (1995) Pseudouridine and O<sup>2</sup>-methylated nucleosides. Significance of their selective occurrence in rRNA domains that function in ribosome-catalyzed synthesis of the peptide bonds in proteins. *Biochimie* **77**(1-2): 7-15
- Lange TS, Ezrokhi M, Amaldi F, Gerbi SA (1999) Box H and box ACA are nucleolar localization elements of U17 small nucleolar RNA. *Mol Biol Cell* **10**(11): 3877-3890
- Lange TS, Ezrokhi M, Borovjagin AV, Rivera-Leon R, North MT, Gerbi SA (1998) Nucleolar localization elements of *Xenopus laevis* U3 small nucleolar RNA. *Mol Biol Cell* **9**(10): 2973-2985
- Lange TS, Gerbi SA (2000) Transient nucleolar localization of U6 small nuclear RNA in *Xenopus laevis* oocytes. *Mol Biol Cell* **11**(7): 2419-2428
- Lapeyre B, Purushothaman SK (2004) Spb1p-directed formation of Gm2922 in the ribosome catalytic center occurs at a late processing stage. *Mol Cell* **16**(4): 663-669
- Larimer FW, Hsu CL, Maupin MK, Stevens A (1992) Characterization of the XRN1 gene encoding a 5'→3' exoribonuclease: sequence data and analysis of disparate protein and mRNA levels of gene-disrupted yeast cells. *Gene* **120**(1): 51-57
- Le Hir H, Nott A, Moore MJ (2003) How introns influence and enhance eukaryotic gene expression. *Trends Biochem Sci* **28**(4): 215-220
- Lebaron S, Froment C, Fromont-Racine M, Rain JC, Monsarrat B, Caizergues-Ferrer M, Henry Y (2005) The splicing ATPase prp43p is a component of multiple preribosomal particles. *Mol Cell Biol* **25**(21): 9269-9282
- Lechertier T, Grob A, Hernandez-Verdun D, Roussel P (2009) Fibrillarin and Nop56 interact before being co-assembled in box C/D snoRNPs. *Exp Cell Res* **315**(6): 928-942
- Lecoq K, Belloc I, Desgranges C, Daignan-Fornier B (2001) Role of adenosine kinase in *Saccharomyces cerevisiae*: identification of the ADO1 gene and study of the mutant phenotypes. *Yeast* **18**(4): 335-342
- Lee CY, Lee A, Chanfreau G (2003) The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5'-end processing of *S. cerevisiae* box C/D snoRNAs. *RNA* **9**(11): 1362-1370
- Lee SJ, Baserga SJ (1999) Imp3p and Imp4p, two specific components of the U3 small nucleolar ribonucleoprotein that are essential for pre-18S rRNA processing. *Mol Cell Biol* **19**(8): 5441-5452
- Lee Y, Nazar RN (1997) Ribosomal 5 S rRNA maturation in *Saccharomyces cerevisiae*. *J Biol Chem* **272**(24): 15206-15212
- Lefebvre S, Burglen L, Reboullet S, Clermont O, Bulet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M, et al. (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **80**(1): 155-165
- Leger-Silvestre I, Trumtel S, Noaillac-Depeyre J, Gas N (1999) Functional compartmentalization of the nucleus in the budding yeast *Saccharomyces cerevisiae*. *Chromosoma* **108**(2): 103-113
- Lejeune F, Ishigaki Y, Li X, Maquat LE (2002) The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. *EMBO J* **21**(13): 3536-3545
- Lempiainen H, Shore D (2009) Growth control and ribosome biogenesis. *Curr Opin Cell Biol* **21**(6): 855-863

- Lesage G, Shapiro J, Specht CA, Sdicu AM, Menard P, Hussein S, Tong AH, Boone C, Bussey H (2005) An interactional network of genes involved in chitin synthesis in *Saccharomyces cerevisiae*. *BMC Genet* **6**: 8
- Lestrade L, Weber MJ (2006) snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res* **34**(Database issue): D158-162
- Leulliot N, Godin KS, Hoareau-Aveilla C, Quevillon-Cheruel S, Varani G, Henry Y, Van Tilbeurgh H (2007) The box H/ACA RNP assembly factor Naf1p contains a domain homologous to Gar1p mediating its interaction with Cbf5p. *J Mol Biol* **371**(5): 1338-1353
- Leung AK, Andersen JS, Mann M, Lamond AI (2003) Bioinformatic analysis of the nucleolus. *Biochem J* **376**(Pt 3): 553-569
- Leung AK, Lamond AI (2002) In vivo analysis of NHPX reveals a novel nucleolar localization pathway involving a transient accumulation in splicing speckles. *J Cell Biol* **157**(4): 615-629
- Leung AK, Nagai K, Li J (2011) Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature* **473**(7348): 536-539
- Lewis JD, Izaurralde E, Jarmolowski A, McGuigan C, Mattaj JW (1996) A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev* **10**(13): 1683-1698
- Li D, Fournier MJ (1992) U14 function in *Saccharomyces cerevisiae* can be provided by large deletion variants of yeast U14 and hybrid mouse-yeast U14 RNAs. *EMBO J* **11**(2): 683-689
- Li H, White SA (1997) RNA aptamers for yeast ribosomal protein L32 have a conserved purine-rich internal loop. *RNA* **3**(3): 245-254
- Li HD, Zagorski J, Fournier MJ (1990) Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in *Saccharomyces cerevisiae*. *Mol Cell Biol* **10**(3): 1145-1152
- Li J, Richter K, Buchner J (2011a) Mixed Hsp90-cochaperone complexes are important for the progression of the reaction cycle. *Nat Struct Mol Biol* **18**(1): 61-66
- Li L, Ye K (2006) Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* **443**(7109): 302-307
- Li S, Duan J, Li D, Ma S, Ye K (2011b) Structure of the Shq1-Cbf5-Nop10-Gar1 complex and implications for H/ACA RNP biogenesis and dyskeratosis congenita. *EMBO J* **30**(24): 5010-5020
- Li S, Duan J, Li D, Yang B, Dong M, Ye K (2011c) Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase. *Genes Dev* **25**(22): 2409-2421
- Liang B, Xue S, Terns RM, Terns MP, Li H (2007a) Substrate RNA positioning in the archaeal H/ACA ribonucleoprotein complex. *Nat Struct Mol Biol* **14**(12): 1173-1180
- Liang B, Zhou J, Kahen E, Terns RM, Terns MP, Li H (2009a) Structure of a functional ribonucleoprotein pseudouridine synthase bound to a substrate RNA. *Nat Struct Mol Biol* **16**(7): 740-746
- Liang XH, Fournier MJ (2006) The helicase Has1p is required for snoRNA release from pre-rRNA. *Mol Cell Biol* **26**(20): 7437-7450
- Liang XH, Liu Q, Fournier MJ (2007b) rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Mol Cell* **28**(6): 965-977
- Liang XH, Liu Q, Fournier MJ (2009b) Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA* **15**(9): 1716-1728
- Liang XH, Liu Q, King TH, Fournier MJ (2010) Strong dependence between functional domains in a dual-function snoRNA infers coupling of rRNA processing and modification events. *Nucleic Acids Res* **38**(10): 3376-3387
- Liautard JP, Sri-Widada J, Brunel C, Jeanteur P (1982) Structural organization of ribonucleoproteins containing small nuclear RNAs from HeLa cells. Proteins interact closely with a similar structural domain of U1, U2, U4 and U5 small nuclear RNAs. *J Mol Biol* **162**(3): 623-643
- Licalosi DD, Geiger G, Minet M, Schroeder S, Cilli K, McNeil JB, Bentley DL (2002) Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. *Mol Cell* **9**(5): 1101-1111
- Lim CR, Kimata Y, Ohdate H, Kokubo T, Kikuchi N, Horigome T, Kohno K (2000) The *Saccharomyces cerevisiae* RuvB-like protein, Tih2p, is required for cell cycle progression and RNA polymerase II-directed transcription. *J Biol Chem* **275**(29): 22409-22417
- Lin J, Jin R, Zhang B, Chen H, Bai YX, Yang PX, Han SW, Xie YH, Huang PT, Huang C, Huang JJ (2008a) Nucleolar localization of TERT is unrelated to telomerase function in human cells. *J Cell Sci* **121**(Pt 13): 2169-2176
- Lin J, Lai S, Jia R, Xu A, Zhang L, Lu J, Ye K (2011) Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. *Nature* **469**(7331): 559-563
- Lin S, Coutinho-Mansfield G, Wang D, Pandit S, Fu XD (2008b) The splicing factor SC35 has an active role in transcriptional elongation. *Nat Struct Mol Biol* **15**(8): 819-826
- Lin Y, Kielkopf CL (2008) X-ray structures of U2 snRNA-branchpoint duplexes containing conserved pseudouridines. *Biochemistry* **47**(20): 5503-5514
- Lincoln TA, Joyce GF (2009) Self-sustained replication of an RNA enzyme. *Science* **323**(5918): 1229-1232
- Lindahl L, Archer RH, Zengel JM (1994) Alternate pathways for processing in the internal transcribed spacer 1 in pre-rRNA of *Saccharomyces cerevisiae*. *Nucleic Acids Res* **22**(24): 5399-5407

- Ling J, Morley SJ, Pain VM, Marzluff WF, Gallie DR (2002) The histone 3'-terminal stem-loop-binding protein enhances translation through a functional and physical interaction with eukaryotic initiation factor 4G (eIF4G) and eIF3. *Mol Cell Biol* **22**(22): 7853-7867
- Linger RJ, Kruk PA (2010) BRCA1 16 years later: risk-associated BRCA1 mutations and their functional implications. *FEBS J* **277**(15): 3086-3096
- Liu C, Apodaca J, Davis LE, Rao H (2007a) Proteasome inhibition in wild-type yeast *Saccharomyces cerevisiae* cells. *Biotechniques* **42**(2): 158, 160, 162
- Liu J, Hebert MD, Ye Y, Templeton DJ, Kung H, Matera AG (2000) Cell cycle-dependent localization of the CDK2-cyclin E complex in Cajal (coiled) bodies. *J Cell Sci* **113** (Pt 9): 1543-1552
- Liu PC, Thiele DJ (2001) Novel stress-responsive genes EMG1 and NOP14 encode conserved, interacting proteins required for 40S ribosome biogenesis. *Mol Biol Cell* **12**(11): 3644-3657
- Liu Q, Dreyfuss G (1996) A novel nuclear structure containing the survival of motor neurons protein. *EMBO J* **15**(14): 3555-3565
- Liu Q, Fischer U, Wang F, Dreyfuss G (1997) The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell* **90**(6): 1013-1021
- Liu S, Li P, Dybkov O, Nottrott S, Hartmuth K, Luhrmann R, Carlomagno T, Wahl MC (2007b) Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP. *Science* **316**(5821): 115-120
- Liu S, Rauhut R, Vornlocher HP, Luhrmann R (2006) The network of protein-protein interactions within the human U4/U6.U5 tri-snRNP. *RNA* **12**(7): 1418-1430
- Liu Y, Huang H, Zhou BO, Wang SS, Hu Y, Li X, Liu J, Zang J, Niu L, Wu J, Zhou JQ, Teng M, Shi Y (2010) Structural analysis of Rtt106p reveals a DNA binding role required for heterochromatin silencing. *J Biol Chem* **285**(6): 4251-4262
- Lorson CL, Hahnen E, Androphy EJ, Wirth B (1999) A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc Natl Acad Sci U S A* **96**(11): 6307-6311
- Lowe TM, Eddy SR (1999) A computational screen for methylation guide snoRNAs in yeast. *Science* **283**(5405): 1168-1171
- Lu Q, Wierzbicki S, Krasilnikov AS, Schmitt ME (2011) Comparison of mitochondrial and nucleolar RNase MRP reveals identical RNA components with distinct enzymatic activities and protein components. *RNA* **16**(3): 529-537
- Lubben B, Marshallsay C, Rottmann N, Luhrmann R (1993) Isolation of U3 snoRNP from CHO cells: a novel 55 kDa protein binds to the central part of U3 snoRNA. *Nucleic Acids Res* **21**(23): 5377-5385
- Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**(6648): 251-260
- Lukowiak AA, Granneman S, Mattox SA, Speckmann WA, Jones K, Pluk H, Venrooij WJ, Terns RM, Terns MP (2000) Interaction of the U3-55k protein with U3 snoRNA is mediated by the box B/C motif of U3 and the WD repeats of U3-55k. *Nucleic Acids Res* **28**(18): 3462-3471
- Lunde BM, Horner M, Meinhart A (2011) Structural insights into cis element recognition of non-polyadenylated RNAs by the Nab3-RRM. *Nucleic Acids Res* **39**(1): 337-346
- Luo J, Duggan DJ, Chen Y, Sauvageot J, Ewing CM, Bittner ML, Trent JM, Isaacs WB (2001) Human prostate cancer and benign prostatic hyperplasia: molecular dissection by gene expression profiling. *Cancer Res* **61**(12): 4683-4688
- Lustig AJ, Lin RJ, Abelson J (1986) The yeast RNA gene products are essential for mRNA splicing in vitro. *Cell* **47**(6): 953-963
- Lygerou Z, Allmann C, Tollervey D, Seraphin B (1996) Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP in vitro. *Science* **272**(5259): 268-270
- 
- Ma T, Van Tine BA, Wei Y, Garrett MD, Nelson D, Adams PD, Wang J, Qin J, Chow LT, Harper JW (2000) Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes Dev* **14**(18): 2298-2313
- Ma X, Yang C, Alexandrov A, Grayhack EJ, Behm-Ansmant I, Yu YT (2005a) Pseudouridylation of yeast U2 snRNA is catalyzed by either an RNA-guided or RNA-independent mechanism. *EMBO J* **24**(13): 2403-2413
- Ma X, Zhao X, Yu YT (2003) Pseudouridylation (Psi) of U2 snRNA in *S. cerevisiae* is catalyzed by an RNA-independent mechanism. *EMBO J* **22**(8): 1889-1897
- Ma Y, Dostie J, Dreyfuss G, Van Duyne GD (2005b) The Gemin6-Gemin7 heterodimer from the survival of motor neurons complex has an Sm protein-like structure. *Structure* **13**(6): 883-892
- Mager WH, Planta RJ, Ballesta JG, Lee JC, Mizuta K, Suzuki K, Warner JR, Woolford J (1997) A new nomenclature for the cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae*. *Nucleic Acids Res* **25**(24): 4872-4875
- Majewski F, Spranger J (1976) [A new (brachymelic) type of primordial dwarfism (author's transl)]. *Monatsschr Kinderheilkd* **124**(6): 499-503
- Majewski F, Stoeckenius M, Kemperdick H (1982) Studies of microcephalic primordial dwarfism III: an intrauterine dwarf with platyspondyly and anomalies of pelvis and clavicles--osteodysplastic primordial dwarfism type III. *Am J Med Genet* **12**(1): 37-42
- Makarova OV, Makarov EM, Liu S, Vornlocher HP, Luhrmann R (2002) Protein 61K, encoded by a gene

- (PRPF31) linked to autosomal dominant retinitis pigmentosa, is required for U4/U6\*U5 tri-snRNP formation and pre-mRNA splicing. *EMBO J* **21**(5): 1148-1157
- Makino Y, Kanemaki M, Kurokawa Y, Koji T, Tamura T (1999) A rat RuvB-like protein, TIP49a, is a germ cell-enriched novel DNA helicase. *J Biol Chem* **274**(22): 15329-15335
- Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL, Tong L (2006) Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* **444**(7121): 953-956
- Manival X, Charron C, Fourmann JB, Godard F, Charpentier B, Branlant C (2006) Crystal structure determination and site-directed mutagenesis of the Pyrococcus abyssi aCBF5-aNOP10 complex reveal crucial roles of the C-terminal domains of both proteins in H/ACA sRNP activity. *Nucleic Acids Res* **34**(3): 826-839
- Mao H, White SA, Williamson JR (1999) A novel loop-loop recognition motif in the yeast ribosomal protein L30 autoregulatory RNA complex. *Nat Struct Biol* **6**(12): 1139-1147
- Mao H, Williamson JR (1999) Local folding coupled to RNA binding in the yeast ribosomal protein L30. *J Mol Biol* **292**(2): 345-359
- Mao X, Schwer B, Shuman S (1995) Yeast mRNA cap methyltransferase is a 50-kilodalton protein encoded by an essential gene. *Mol Cell Biol* **15**(8): 4167-4174
- Marcello A, Ferrari A, Pellegrini V, Pegoraro G, Lusic M, Beltram F, Giacca M (2003) Recruitment of human cyclin T1 to nuclear bodies through direct interaction with the PML protein. *EMBO J* **22**(9): 2156-2166
- Marmier-Gourrier N (2003) Etudes sur le mécanisme d'action du petit ARN nucléolaire U3 de la levure *Saccharomyces cerevisiae* et sur la protéine Snu13p/15,5 kD associée à cet ARN et son homologue d'archaea L7Ae. *Thèse*
- Marmier-Gourrier N, Clery A, Schlotter F, Senty-Segault V, Branlant C (2011) A second base pair interaction between U3 small nucleolar RNA and the 5'-ETS region is required for early cleavage of the yeast pre-ribosomal RNA. *Nucleic Acids Res* **39**(22): 9731-9745
- Marmier-Gourrier N, Clery A, Senty-Segault V, Charpentier B, Schlotter F, Leclerc F, Fournier R, Branlant C (2003) A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA. *RNA* **9**(7): 821-838
- Marrone A, Walne A, Dokal I (2005) Dyskeratosis congenita: telomerase, telomeres and anticipation. *Curr Opin Genet Dev* **15**(3): 249-257
- Marrone A, Walne A, Tamary H, Masunari Y, Kirwan M, Beswick R, Vulliamy T, Dokal I (2007) Telomerase reverse-transcriptase homozygous mutations in autosomal recessive dyskeratosis congenita and Hoyeraal-Hreidarsson syndrome. *Blood* **110**(13): 4198-4205
- Martin AN, Li Y (2007) RNase MRP RNA and human genetic diseases. *Cell Res* **17**(3): 219-226
- Martin F, Schaller A, Eglite S, Schumperli D, Muller B (1997) The gene for histone RNA hairpin binding protein is located on human chromosome 4 and encodes a novel type of RNA binding protein. *EMBO J* **16**(4): 769-778
- Marzluff WF, Wagner EJ, Duronio RJ (2008) Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet* **9**(11): 843-854
- Massenet S, Motorin Y, Lafontaine DL, Hurt EC, Grosjean H, Branlant C (1999) Pseudouridine mapping in the *Saccharomyces cerevisiae* spliceosomal U small nuclear RNAs (snRNAs) reveals that pseudouridine synthase pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA. *Mol Cell Biol* **19**(3): 2142-2154
- Massenet S, Mougin A, Branlant C (1998) Posttranscriptional modifications in the U small nuclear RNAs. *The Modification and Editing of RNA Grosjean H and Benne R (eds) ASM Press, Washington, DC: pp. 201-228*
- Massenet S, Pellizzoni L, Paushkin S, Mattaj IW, Dreyfuss G (2002) The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway. *Mol Cell Biol* **22**(18): 6533-6541
- Matera AG, Frey MR (1998) Coiled bodies and gems: Janus or gemini? *Am J Hum Genet* **63**(2): 317-321
- Matera AG, Terns RM, Terns MP (2007) Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol* **8**(3): 209-220
- Matias PM, Gorynia S, Donner P, Carrondo MA (2006) Crystal structure of the human AAA+ protein RuvBL1. *J Biol Chem* **281**(50): 38918-38929
- Matic I, Schimmel J, Hendriks IA, van Santen MA, van de Rijke F, van Dam H, Gnad F, Mann M, Vertegaal AC (2010) Site-specific identification of SUMO-2 targets in cells reveals an inverted SUMOylation motif and a hydrophobic cluster SUMOylation motif. *Mol Cell* **39**(4): 641-652
- Matsumura S, Ikawa Y, Inoue T (2003) Biochemical characterization of the kink-turn RNA motif. *Nucleic Acids Res* **31**(19): 5544-5551
- Matthews JM, Sunde M (2002) Zinc fingers--folds for many occasions. *IUBMB Life* **54**(6): 351-355
- Maul GG, Negorev D, Bell P, Ishov AM (2000) Review: properties and assembly mechanisms of ND10, PML bodies, or PODs. *J Struct Biol* **129**(2-3): 278-287
- Maurel (2003) *La Naissance de la Vie : de l'Evolution Prébiotique à l'Evolution Biologique*. Edition Dunod
- Mayer C, Grummt I (2006) Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene* **25**(48): 6384-6391

- Mayer C, Zhao J, Yuan X, Grummt I (2004) mTOR-dependent activation of the transcription factor TIF-IA links rRNA synthesis to nutrient availability. *Genes Dev* **18**(4): 423-434
- Mayer MP (2010) Gymnastics of molecular chaperones. *Mol Cell* **39**(3): 321-331
- McKeegan KS, Debieux CM, Boulon S, Bertrand E, Watkins NJ (2007) A dynamic scaffold of pre-snoRNP factors facilitates human box C/D snoRNP assembly. *Mol Cell Biol* **27**(19): 6782-6793
- McKeegan KS, Debieux CM, Watkins NJ (2009) Evidence that the AAA+ proteins TIP48 and TIP49 bridge interactions between 15.5K and the related NOP56 and NOP58 proteins during box C/D snoRNP biogenesis. *Mol Cell Biol* **29**(18): 4971-4981
- McKie AB, McHale JC, Keen TJ, Tarttelin EE, Goliath R, van Lith-Verhoeven JJ, Greenberg J, Ramesar RS, Hoyng CB, Cremers FP, Mackey DA, Bhattacharya SS, Bird AC, Markham AF, Inglehearn CF (2001) Mutations in the pre-mRNA splicing factor gene PRPC8 in autosomal dominant retinitis pigmentosa (RP13). *Hum Mol Genet* **10**(15): 1555-1562
- Medenbach J, Schreiner S, Liu S, Luhrmann R, Bindereif A (2004) Human U4/U6 snRNP recycling factor p110: mutational analysis reveals the function of the tetratricopeptide repeat domain in recycling. *Mol Cell Biol* **24**(17): 7392-7401
- Meggio F, Marin O, Pinna LA (1994) Substrate specificity of protein kinase CK2. *Cell Mol Biol Res* **40**(5-6): 401-409
- Meister G, Buhler D, Pillai R, Lottspeich F, Fischer U (2001a) A multiprotein complex mediates the ATP-dependent assembly of spliceosomal U snRNPs. *Nat Cell Biol* **3**(11): 945-949
- Meister G, Eggert C, Buhler D, Brahms H, Kambach C, Fischer U (2001b) Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. *Curr Biol* **11**(24): 1990-1994
- Meister G, Eggert C, Fischer U (2002) SMN-mediated assembly of RNPs: a complex story. *Trends Cell Biol* **12**(10): 472-478
- Meister G, Fischer U (2002) Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. *EMBO J* **21**(21): 5853-5863
- Meister G, Hannus S, Plottner O, Baars T, Hartmann E, Fakan S, Lagerbauer B, Fischer U (2001c) SMNrp is an essential pre-mRNA splicing factor required for the formation of the mature spliceosome. *EMBO J* **20**(9): 2304-2314
- Melcak I, Cermanova S, Jirsova K, Koberna K, Malinsky J, Raska I (2000) Nuclear pre-mRNA compartmentalization: trafficking of released transcripts to splicing factor reservoirs. *Mol Biol Cell* **11**(2): 497-510
- Menne TF, Goyenechea B, Sanchez-Puig N, Wong CC, Tonkin LM, Ancliff PJ, Brost RL, Costanzo M, Boone C, Warren AJ (2007) The Shwachman-Bodian-Diamond syndrome protein mediates translational activation of ribosomes in yeast. *Nat Genet* **39**(4): 486-495
- Mereau A, Fournier R, Gregoire A, Mougou A, Fabrizio P, Luhrmann R, Branlant C (1997) An in vivo and in vitro structure-function analysis of the *Saccharomyces cerevisiae* U3A snoRNP: protein-RNA contacts and base-pair interaction with the pre-ribosomal RNA. *J Mol Biol* **273**(3): 552-571
- Meyer P, Prodromou C, Hu B, Vaughan C, Roe SM, Panaretou B, Piper PW, Pearl LH (2003) Structural and functional analysis of the middle segment of hsp90: implications for ATP hydrolysis and client protein and cochaperone interactions. *Mol Cell* **11**(3): 647-658
- Meyer P, Prodromou C, Liao C, Hu B, Roe SM, Vaughan CK, Vlasic I, Panaretou B, Piper PW, Pearl LH (2004) Structural basis for recruitment of the ATPase activator Aha1 to the Hsp90 chaperone machinery. *EMBO J* **23**(6): 1402-1410
- Meyer S, Urbanke C, Wahle E (2002) Equilibrium studies on the association of the nuclear poly(A) binding protein with poly(A) of different lengths. *Biochemistry* **41**(19): 6082-6089
- Miller OL, Jr., Beatty BR (1969) Visualization of nucleolar genes. *Science* **164**(882): 955-957
- Millevoi S, Vagner S (2009) Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res* **38**(9): 2757-2774
- Minvielle-Sebastia L, Preker PJ, Wiederkehr T, Strahm Y, Keller W (1997) The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. *Proc Natl Acad Sci U S A* **94**(15): 7897-7902
- Misteli T, Caceres JF, Spector DL (1997) The dynamics of a pre-mRNA splicing factor in living cells. *Nature* **387**(6632): 523-527
- Mitchell JR, Cheng J, Collins K (1999a) A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Mol Cell Biol* **19**(1): 567-576
- Mitchell JR, Wood E, Collins K (1999b) A telomerase component is defective in the human disease dyskeratosis congenita. *Nature* **402**(6761): 551-555
- Mitchell P, Petfalski E, Tollervey D (1996) The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism. *Genes Dev* **10**(4): 502-513
- Miyata Y (2009) Protein kinase CK2 in health and disease: CK2: the kinase controlling the Hsp90 chaperone machinery. *Cell Mol Life Sci* **66**(11-12): 1840-1849
- Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C (2004) ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**(5656): 343-348
- Menne TF, Goyenechea B, Sanchez-Puig N, Wong CC, Tonkin LM, Ancliff PJ, Brost RL, Costanzo M, Boone C,

- Moazed D (2009) Small RNAs in transcriptional gene silencing and genome defence. *Nature* **457**(7228): 413-420
- Montange RK, Batey RT (2006) Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature* **441**(7097): 1172-1175
- Moore MJ, Proudfoot NJ (2009) Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136**(4): 688-700
- Moore PB, Steitz TA (2002) The involvement of RNA in ribosome function. *Nature* **418**(6894): 229-235
- Moore T, Zhang Y, Fenley MO, Li H (2004) Molecular basis of box C/D RNA-protein interactions; cocrystal structure of archaeal L7Ae and a box C/D RNA. *Structure* **12**(5): 807-818
- Moqtaderi Z, Yale JD, Struhl K, Buratowski S (1996) Yeast homologues of higher eukaryotic TFIID subunits. *Proc Natl Acad Sci U S A* **93**(25): 14654-14658
- Mordes D, Yuan L, Xu L, Kawada M, Molday RS, Wu JY (2007) Identification of photoreceptor genes affected by PRPF31 mutations associated with autosomal dominant retinitis pigmentosa. *Neurobiol Dis* **26**(2): 291-300
- Morlando M, Ballarino M, Greco P, Caffarelli E, Dichtl B, Bozzoni I (2004) Coupling between snoRNP assembly and 3' processing controls box C/D snoRNA biosynthesis in yeast. *EMBO J* **23**(12): 2392-2401
- Morlando M, Greco P, Dichtl B, Fatica A, Keller W, Bozzoni I (2002) Functional analysis of yeast snoRNA and snRNA 3'-end formation mediated by uncoupling of cleavage and polyadenylation. *Mol Cell Biol* **22**(5): 1379-1389
- Morris DP, Greenleaf AL (2000) The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* **275**(51): 39935-39943
- Morris GE (2008) The Cajal body. *Biochim Biophys Acta* **1783**(11): 2108-2115
- Morrison AJ, Shen X (2009) Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. *Nat Rev Mol Cell Biol* **10**(6): 373-384
- Morrissey JP, Tollervey D (1993) Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol Cell Biol* **13**(4): 2469-2477
- Mouaikel J, Narayanan U, Verheggen C, Matera AG, Bertrand E, Tazi J, Bordonne R (2003) Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO Rep* **4**(6): 616-622
- Mouaikel J, Verheggen C, Bertrand E, Tazi J, Bordonne R (2002) Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol Cell* **9**(4): 891-901
- Mougey EB, O'Reilly M, Osheim Y, Miller OL, Jr., Beyer A, Sollner-Webb B (1993) The terminal balls characteristic of eukaryotic rRNA transcription units in chromatin spreads are rRNA processing complexes. *Genes Dev* **7**(8): 1609-1619
- Mougin A, Gottschalk A, Fabrizio P, Luhrmann R, Branlant C (2002) Direct probing of RNA structure and RNA-protein interactions in purified HeLa cell's and yeast spliceosomal U4/U6.U5 tri-snRNP particles. *J Mol Biol* **317**(5): 631-649
- Muhrad D, Decker CJ, Parker R (1995) Turnover mechanisms of the stable yeast PGK1 mRNA. *Mol Cell Biol* **15**(4): 2145-2156
- Muller S, Fourmann JB, Loegler C, Charpentier B, Branlant C (2007) Identification of determinants in the protein partners aCBF5 and aNOP10 necessary for the tRNA:Psi55-synthase and RNA-guided RNA:Psi-synthase activities. *Nucleic Acids Res* **35**(16): 5610-5624
- Muller S, Leclerc F, Behm-Ansmant I, Fourmann JB, Charpentier B, Branlant C (2008) Combined in silico and experimental identification of the Pyrococcus abyssi H/ACA sRNAs and their target sites in ribosomal RNAs. *Nucleic Acids Res* **36**(8): 2459-2475
- Munro TP, Magee RJ, Kidd GJ, Carson JH, Barbarese E, Smith LM, Smith R (1999) Mutational analysis of a heterogeneous nuclear ribonucleoprotein A2 response element for RNA trafficking. *J Biol Chem* **274**(48): 34389-34395
- Muratani M, Gerlich D, Janicki SM, Gebhard M, Eils R, Spector DL (2002) Metabolic-energy-dependent movement of PML bodies within the mammalian cell nucleus. *Nat Cell Biol* **4**(2): 106-110
- Myslinski E, Segault V, Branlant C (1990) An intron in the genes for U3 small nucleolar RNAs of the yeast *Saccharomyces cerevisiae*. *Science* **247**(4947): 1213-1216
- Myslyuk I, Doniger T, Horesh Y, Hury A, Hoffer R, Ziporen Y, Michaeli S, Unger R (2008) Psiscan: a computational approach to identify H/ACA-like and AGA-like non-coding RNA in trypanosomatid genomes. *BMC Bioinformatics* **9**: 471
- 
- Nakamura T, Prestayko AW, Busch H (1968) Studies on nucleolar 4 to 6 S ribonucleic acid of Novikoff hepatoma cells. *J Biol Chem* **243**(7): 1368-1375
- Narayanan A, Lukowiak A, Jady BE, Dragon F, Kiss T, Terns RM, Terns MP (1999a) Nucleolar localization signals of box H/ACA small nucleolar RNAs. *EMBO J* **18**(18): 5120-5130
- Narayanan A, Speckmann W, Terns R, Terns MP (1999b) Role of the box C/D motif in localization of small nucleolar RNAs to coiled bodies and nucleoli. *Mol Biol Cell* **10**(7): 2131-2147
- Narayanan U, Ospina JK, Frey MR, Hebert MD, Matera AG (2002) SMN, the spinal muscular atrophy protein, forms a pre-import snRNP complex with snurportin1 and importin beta. *Hum Mol Genet* **11**(15): 1785-1795



- Narla A, Ebert BL (2010) Ribosomopathies: human disorders of ribosome dysfunction. *Blood* **115**(16): 3196-3205
- Nedeia E, He X, Kim M, Pootoolal J, Zhong G, Canadien V, Hughes T, Buratowski S, Moore CL, Greenblatt J (2003) Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3'-ends. *J Biol Chem* **278**(35): 33000-33010
- Neefs JM, Van de Peer Y, De Rijk P, Chapelle S, De Wachter R (1993) Compilation of small ribosomal subunit RNA structures. *Nucleic Acids Res* **21**(13): 3025-3049
- Nesic D, Tanackovic G, Kramer A (2004) A role for Cajal bodies in the final steps of U2 snRNP biogenesis. *J Cell Sci* **117**(Pt 19): 4423-4433
- Neuwald AF, Aravind L, Spouge JL, Koonin EV (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res* **9**(1): 27-43
- Newby MI, Greenbaum NL (2001) A conserved pseudouridine modification in eukaryotic U2 snRNA induces a change in branch-site architecture. *RNA* **7**(6): 833-845
- Newman DR, Kuhn JF, Shanab GM, Maxwell ES (2000) Box C/D snoRNA-associated proteins: two pairs of evolutionarily ancient proteins and possible links to replication and transcription. *RNA* **6**(6): 861-879
- Ng HH, Robert F, Young RA, Struhl K (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **11**(3): 709-719
- Ni J, Tien AL, Fournier MJ (1997) Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* **89**(4): 565-573
- Nissan TA, Bassler J, Petfalski E, Tollervey D, Hurt E (2002) 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. *EMBO J* **21**(20): 5539-5547
- Nissen P, Hansen J, Ban N, Moore PB, Steitz TA (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**(5481): 920-930
- Noble CG, Hollingworth D, Martin SR, Ennis-Adeniran V, Smerdon SJ, Kelly G, Taylor IA, Ramos A (2005) Key features of the interaction between Pcf11 CID and RNA polymerase II CTD. *Nat Struct Mol Biol* **12**(2): 144-151
- Noble M, Lewis SA, Cowan NJ (1989) The microtubule binding domain of microtubule-associated protein MAP1B contains a repeated sequence motif unrelated to that of MAP2 and tau. *J Cell Biol* **109**(6 Pt 2): 3367-3376
- Nolivos S, Carpousis AJ, Clouet-d'Orval B (2005) The K-loop, a general feature of the Pyrococcus C/D guide RNAs, is an RNA structural motif related to the K-turn. *Nucleic Acids Res* **33**(20): 6507-6514
- Noller HF (2004) The driving force for molecular evolution of translation. *RNA* **10**(12): 1833-1837
- Nottrott S, Hartmuth K, Fabrizio P, Urlaub H, Vidovic I, Ficner R, Luhrmann R (1999) Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5' stem-loop of U4 snRNA. *EMBO J* **18**(21): 6119-6133
- Nottrott S, Urlaub H, Luhrmann R (2002) Hierarchical, clustered protein interactions with U4/U6 snRNA: a biochemical role for U4/U6 proteins. *EMBO J* **21**(20): 5527-5538
- Novotny I, Blazikova M, Stanek D, Herman P, Malinsky J (2010) In vivo kinetics of U4/U6.U5 tri-snRNP formation in Cajal bodies. *Mol Biol Cell* **22**(4): 513-523
- 
- O'Keefe RT, Mayeda A, Sadowski CL, Krainer AR, Spector DL (1994) Disruption of pre-mRNA splicing in vivo results in reorganization of splicing factors. *J Cell Biol* **124**(3): 249-260
- Oakes M, Aris JP, Brockenbrough JS, Wai H, Vu L, Nomura M (1998) Mutational analysis of the structure and localization of the nucleolus in the yeast *Saccharomyces cerevisiae*. *J Cell Biol* **143**(1): 23-34
- Ochs RL, Lischwe MA, Spohn WH, Busch H (1985) Fibrillarin: a new protein of the nucleolus identified by autoimmune sera. *Biol Cell* **54**(2): 123-133
- Ohdate H, Lim CR, Kokubo T, Matsubara K, Kimata Y, Kohno K (2003) Impairment of the DNA binding activity of the TATA-binding protein renders the transcriptional function of Rvb2p/Tih2p, the yeast RuvB-like protein, essential for cell growth. *J Biol Chem* **278**(17): 14647-14656
- Ohi MD, Link AJ, Ren L, Jennings JL, McDonald WH, Gould KL (2002) Proteomics analysis reveals stable multiprotein complexes in both fission and budding yeasts containing Myb-related Cdc5p/Cef1p, novel pre-mRNA splicing factors, and snRNAs. *Mol Cell Biol* **22**(7): 2011-2024
- Ohno M, Sakamoto H, Shimura Y (1987) Preferential excision of the 5' proximal intron from mRNA precursors with two introns as mediated by the cap structure. *Proc Natl Acad Sci U S A* **84**(15): 5187-5191
- Ohno M, Segref A, Bachi A, Wilm M, Mattaj IW (2000) PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* **101**(2): 187-198
- Oliveira CC, Gonzales FA, Zanchin NI (2002) Temperature-sensitive mutants of the exosome subunit Rrp43p show a deficiency in mRNA degradation and no longer interact with the exosome. *Nucleic Acids Res* **30**(19): 4186-4198
- Oliver SG, McLaughlin CS (1977) The regulation of RNA synthesis in yeast. I: Starvation experiments. *Mol Gen Genet* **154**(2): 145-153

- Omer AD, Lowe TM, Russell AG, Ebhardt H, Eddy SR, Dennis PP (2000) Homologs of small nucleolar RNAs in Archaea. *Science* **288**(5465): 517-522
- Omer AD, Ziesche S, Ebhardt H, Dennis PP (2002) In vitro reconstitution and activity of a C/D box methylation guide ribonucleoprotein complex. *Proc Natl Acad Sci U S A* **99**(8): 5289-5294
- Ono M, Scott MS, Yamada K, Avolio F, Barton GJ, Lamond AI (2010) Identification of human miRNA precursors that resemble box C/D snoRNAs. *Nucleic Acids Res*
- Ooi SL, Samarsky DA, Fournier MJ, Boeke JD (1998) Intronic snoRNA biosynthesis in *Saccharomyces cerevisiae* depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. *RNA* **4**(9): 1096-1110
- Orlando V (2000) Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem Sci* **25**(3): 99-104
- Oruganti S, Zhang Y, Li H (2005) Structural comparison of yeast snoRNP and spliceosomal protein Snu13p with its homologs. *Biochem Biophys Res Commun* **333**(2): 550-554
- Oruganti S, Zhang Y, Li H, Robinson H, Terns MP, Terns RM, Yang W (2007) Alternative conformations of the archaeal Nop56/58-fibrillarin complex imply flexibility in box C/D RNPs. *J Mol Biol* **371**(5): 1141-1150
- Osenberg S, Dominissini D, Rechavi G, Eisenberg E (2009) Widespread cleavage of A-to-I hyperediting substrates. *RNA* **15**(9): 1632-1639
- Osheim YN, French SL, Keck KM, Champion EA, Spasov K, Dragon F, Baserga SJ, Beyer AL (2004) Pre-18S ribosomal RNA is structurally compacted into the SSU processome prior to being cleaved from nascent transcripts in *Saccharomyces cerevisiae*. *Mol Cell* **16**(6): 943-954
- Owen N, Doe CL, Mellor J, Davies KE (2000) Characterization of the *Schizosaccharomyces pombe* orthologue of the human survival motor neuron (SMN) protein. *Hum Mol Genet* **9**(5): 675-684
- 
- P**alacios I, Hetzer M, Adam SA, Mattaj IW (1997) Nuclear import of U snRNPs requires importin beta. *EMBO J* **16**(22): 6783-6792
- Pandey NB, Marzluff WF (1987) The stem-loop structure at the 3' end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability. *Mol Cell Biol* **7**(12): 4557-4559
- Pannone BK, Kim SD, Noe DA, Wolin SL (2001) Multiple functional interactions between components of the Lsm2-Lsm8 complex, U6 snRNA, and the yeast La protein. *Genetics* **158**(1): 187-196
- Parada LA, McQueen PG, Misteli T (2004a) Tissue-specific spatial organization of genomes. *Genome Biol* **5**(7): R44
- Parada LA, Sotiriou S, Misteli T (2004b) Spatial genome organization. *Exp Cell Res* **296**(1): 64-70
- Parker KA, Steitz JA (1987) Structural analysis of the human U3 ribonucleoprotein particle reveal a conserved sequence available for base pairing with pre-rRNA. *Mol Cell Biol* **7**(8): 2899-2913
- Patel AA, Steitz JA (2003) Splicing double: insights from the second spliceosome. *Nat Rev Mol Cell Biol* **4**(12): 960-970
- Patel SB, Bellini M (2008) The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res* **36**(20): 6482-6493
- Paushkin S, Charroux B, Abel L, Perkinson RA, Pellizzoni L, Dreyfuss G (2000) The survival motor neuron protein of *Schizosaccharomyces pombe*. Conservation of survival motor neuron interaction domains in divergent organisms. *J Biol Chem* **275**(31): 23841-23846
- Pearl LH, Prodromou C (2006) Structure and mechanism of the Hsp90 molecular chaperone machinery. *Annu Rev Biochem* **75**: 271-294
- Peculis BA (1997) The sequence of the 5' end of the U8 small nucleolar RNA is critical for 5.8S and 28S rRNA maturation. *Mol Cell Biol* **17**(7): 3702-3713
- Peculis BA, Steitz JA (1993) Disruption of U8 nucleolar snRNA inhibits 5.8S and 28S rRNA processing in the *Xenopus* oocyte. *Cell* **73**(6): 1233-1245
- Pellizzoni L, Baccon J, Charroux B, Dreyfuss G (2001) The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1. *Curr Biol* **11**(14): 1079-1088
- Pellizzoni L, Baccon J, Rappsilber J, Mann M, Dreyfuss G (2002a) Purification of native survival of motor neurons complexes and identification of Gemin6 as a novel component. *J Biol Chem* **277**(9): 7540-7545
- Pellizzoni L, Kataoka N, Charroux B, Dreyfuss G (1998) A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell* **95**(5): 615-624
- Pellizzoni L, Yong J, Dreyfuss G (2002b) Essential role for the SMN complex in the specificity of snRNP assembly. *Science* **298**(5599): 1775-1779
- Peng WT, Robinson MD, Mnaimneh S, Krogan NJ, Cagney G, Morris Q, Davierwala AP, Grigull J, Yang X, Zhang W, Mitsakakis N, Ryan OW, Datta N, Jovic V, Pal C, Canadien V, Richards D, Beattie B, Wu LF, Altschuler SJ, Roweis S, Frey BJ, Emili A, Greenblatt JF, Hughes TR (2003) A panoramic view of yeast noncoding RNA processing. *Cell* **113**(7): 919-933
- Perez-Fernandez J, Roman A, De Las Rivas J, Bustelo XR, Dosil M (2007) The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. *Mol Cell Biol* **27**(15): 5414-5429
- Pertschy B, Saveanu C, Zisser G, Lebreton A, Tengg M, Jacquier A, Liebming E, Nobis B, Kappel L, van der Klei

- I, Hogenauer G, Fromont-Racine M, Bergler H (2007) Cytoplasmic recycling of 60S preribosomal factors depends on the AAA protein Drg1. *Mol Cell Biol* **27**(19): 6581-6592
- Pertschy B, Schneider C, Gnadig M, Schafer T, Tollervey D, Hurt E (2009) RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J Biol Chem* **284**(50): 35079-35091
- Perumal K, Reddy R (2002) The 3' end formation in small RNAs. *Gene Expr* **10**(1-2): 59-78
- Petes TD (1979) Yeast ribosomal DNA genes are located on chromosome XII. *Proc Natl Acad Sci U S A* **76**(1): 410-414
- Petfalski E, Dandekar T, Henry Y, Tollervey D (1998) Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol Cell Biol* **18**(3): 1181-1189
- Phipps KR, Charette JM, Baserga SJ (2011) The SSU Processome in Ribosome Biogenesis - Progress and Prospects. *WIREs RNA* **2**(1): 1-21
- Piazzon N, Rage F, Schlotter F, Moine H, Branlant C, Massenet S (2008) In vitro and in cellulo evidences for association of the survival of motor neuron complex with the fragile X mental retardation protein. *J Biol Chem* **283**(9): 5598-5610
- Piekna-Przybylska D, Decatur WA, Fournier MJ (2007) New bioinformatic tools for analysis of nucleotide modifications in eukaryotic rRNA. *RNA* **13**(3): 305-312
- Piekna-Przybylska D, Decatur WA, Fournier MJ (2008a) The 3D rRNA modification maps database: with interactive tools for ribosome analysis. *Nucleic Acids Res* **36**(Database issue): D178-183
- Piekna-Przybylska D, Przybylski P, Baudin-Baillieu A, Rousset JP, Fournier MJ (2008b) Ribosome performance is enhanced by a rich cluster of pseudouridines in the A-site finger region of the large subunit. *J Biol Chem* **283**(38): 26026-26036
- Pillai RS, Grimm M, Meister G, Will CL, Luhrmann R, Fischer U, Schumperli D (2003) Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev* **17**(18): 2321-2333
- Pillai RS, Will CL, Luhrmann R, Schumperli D, Muller B (2001) Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO J* **20**(19): 5470-5479
- Piper PW, Bellatin JA, Lockheart A (1983) Altered maturation of sequences at the 3' terminus of 5S gene transcripts in a *Saccharomyces cerevisiae* mutant that lacks a RNA processing endonuclease. *EMBO J* **2**(3): 353-359
- Platani M, Goldberg I, Lamond AI, Swedlow JR (2002) Cajal body dynamics and association with chromatin are ATP-dependent. *Nat Cell Biol* **4**(7): 502-508
- Platani M, Goldberg I, Swedlow JR, Lamond AI (2000) In vivo analysis of Cajal body movement, separation, and joining in live human cells. *J Cell Biol* **151**(7): 1561-1574
- Pluk H, Soffner J, Luhrmann R, van Venrooij WJ (1998) cDNA cloning and characterization of the human U3 small nucleolar ribonucleoprotein complex-associated 55-kilodalton protein. *Mol Cell Biol* **18**(1): 488-498
- Pomeranz Krummel DA, Oubridge C, Leung AK, Li J, Nagai K (2009) Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Nature* **458**(7237): 475-480
- Poole AM, Jeffares DC, Penny D (1998) The path from the RNA world. *J Mol Evol* **46**(1): 1-17
- Pradet-Balade B, Girard C, Boulon S, Paul C, Azzag K, Bordonne R, Bertrand E, Verheggen C (2011) CRM1 controls the composition of nucleoplasmic pre-snoRNA complexes to licence them for nucleolar transport. *EMBO J* **30**(11): 2205-2218
- Prasanth KV, Prasanth SG, Xuan Z, Hearn S, Freier SM, Bennett CF, Zhang MQ, Spector DL (2005) Regulating gene expression through RNA nuclear retention. *Cell* **123**(2): 249-263
- Prestayko AW, Tonato M, Busch H (1970) Low molecular weight RNA associated with 28 s nucleolar RNA. *J Mol Biol* **47**(3): 505-515
- Prestayko AW, Tonato M, Lewis BC, Busch H (1971) Heterogeneity of nucleolar U3 ribonucleic acid of the Novikoff hepatoma. *J Biol Chem* **246**(1): 182-187
- Prodromou C, Roe SM, O'Brien R, Ladbury JE, Piper PW, Pearl LH (1997) Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. *Cell* **90**(1): 65-75
- Pu WT, Krapivinsky GB, Krapivinsky L, Clapham DE (1999) pICln inhibits snRNP biogenesis by binding core spliceosomal proteins. *Mol Cell Biol* **19**(6): 4113-4120
- Puri T, Wendler P, Sigala B, Saibil H, Tsaneva IR (2007) Dodecameric structure and ATPase activity of the human TIP48/TIP49 complex. *J Mol Biol* **366**(1): 179-192
- Qiu H, Eifert J, Wacheul L, Thiry M, Berger AC, Jakovljevic J, Woolford JL, Jr., Corbett AH, Lafontaine DL, Terns RM, Terns MP (2008) Identification of genes that function in the biogenesis and localization of small nucleolar RNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol* **28**(11): 3686-3699
- Qiu XB, Lin YL, Thome KC, Pian P, Schlegel BP, Weremowicz S, Parvin JD, Dutta A (1998) An eukaryotic RuvB-like protein (RUVBL1) essential for growth. *J Biol Chem* **273**(43): 27786-27793
- Qu G, van Nues RW, Watkins NJ, Maxwell ES (2011) The spatial-functional coupling of box C/D and C'/D' RNPs is an evolutionarily conserved feature of the eukaryotic box C/D snoRNP nucleotide modification complex. *Mol Cell Biol* **31**(2): 365-374

- Qu LH, Henras A, Lu YJ, Zhou H, Zhou WX, Zhu YQ, Zhao J, Henry Y, Caizergues-Ferrer M, Bachellerie JP (1999) Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. *Mol Cell Biol* **19**(2): 1144-1158
- Quignon F, De Bels F, Koken M, Feunteun J, Ameisen JC, de The H (1998) PML induces a novel caspase-independent death process. *Nat Genet* **20**(3): 259-265
- 
- R**abl J, Leibundgut M, Ataide SF, Haag A, Ban N (2011) Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* **331**(6018): 730-736
- Raker VA, Hartmuth K, Kastner B, Luhrmann R (1999) Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Mol Cell Biol* **19**(10): 6554-6565
- Rashid R, Aittaleb M, Chen Q, Spiegel K, Demeler B, Li H (2003) Functional requirement for symmetric assembly of archaeal box C/D small ribonucleoprotein particles. *J Mol Biol* **333**(2): 295-306
- Rashid R, Liang B, Baker DL, Youssef OA, He Y, Phipps K, Terns RM, Terns MP, Li H (2006) Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol Cell* **21**(2): 249-260
- Raska I, Andrade LE, Ochs RL, Chan EK, Chang CM, Roos G, Tan EM (1991) Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies. *Exp Cell Res* **195**(1): 27-37
- Rasmussen TP, Culbertson MR (1998) The putative nucleic acid helicase Sen1p is required for formation and stability of termini and for maximal rates of synthesis and levels of accumulation of small nucleolar RNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol* **18**(12): 6885-6896
- Razga F, Spackova N, Reblova K, Koca J, Leontis NB, Sponer J (2004) Ribosomal RNA kink-turn motif--a flexible molecular hinge. *J Biomol Struct Dyn* **22**(2): 183-194
- Reddy R, Busch H (1983) Small nuclear RNAs and RNA processing. *Prog Nucleic Acid Res Mol Biol* **30**: 127-162
- Reddy R, Henning D, Busch H (1980) Substitutions, insertions, and deletions in two highly conserved U3 RNA species. *J Biol Chem* **255**(14): 7029-7033
- Reddy R, Henning D, Chirala S, Rothblum L, Wright D, Busch H (1985) Isolation and characterization of three rat U3 RNA pseudogenes colinear with U3 RNA. *J Biol Chem* **260**(9): 5715-5719
- Regad T, Chelbi-Alix MK (2001) Role and fate of PML nuclear bodies in response to interferon and viral infections. *Oncogene* **20**(49): 7274-7286
- Reichow SL, Varani G (2008) Nop10 is a conserved H/ACA snoRNP molecular adaptor. *Biochemistry* **47**(23): 6148-6156
- Reidt U, Reuter K, Achsel T, Ingelfinger D, Luhrmann R, Ficner R (2000) Crystal structure of the human U4/U6 small nuclear ribonucleoprotein particle-specific SnuCyp-20, a nuclear cyclophilin. *J Biol Chem* **275**(11): 7439-7442
- Reidt U, Wahl MC, Fasshauer D, Horowitz DS, Luhrmann R, Ficner R (2003) Crystal structure of a complex between human spliceosomal cyclophilin H and a U4/U6 snRNP-60K peptide. *J Mol Biol* **331**(1): 45-56
- Reilly TH, Schmitt ME (1995) The yeast, *Saccharomyces cerevisiae*, RNase P/MRP ribonucleoprotein endoribonuclease family. *Mol Biol Rep* **22**(2-3): 87-93
- Renvoise B, Colasse S, Burlet P, Viollet L, Meier UT, Lefebvre S (2009) The loss of the snoRNP chaperone Nopp140 from Cajal bodies of patient fibroblasts correlates with the severity of spinal muscular atrophy. *Hum Mol Genet* **18**(7): 1181-1189
- Richard P, Darzacq X, Bertrand E, Jady BE, Verheggen C, Kiss T (2003) A common sequence motif determines the Cajal body-specific localization of box H/ACA scaRNAs. *EMBO J* **22**(16): 4283-4293
- Richard P, Kiss AM, Darzacq X, Kiss T (2006) Cotranscriptional recognition of human intronic box H/ACA snoRNAs occurs in a splicing-independent manner. *Mol Cell Biol* **26**(7): 2540-2549
- Rife JP, Moore PB (1998) The structure of a methylated tetraloop in 16S ribosomal RNA. *Structure* **6**(6): 747-756
- Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**(10): 1030-1032
- Rio Frio T, Wade NM, Ransijn A, Berson EL, Beckmann JS, Rivolta C (2008) Premature termination codons in PRPF31 cause retinitis pigmentosa via haploinsufficiency due to nonsense-mediated mRNA decay. *J Clin Invest* **118**(4): 1519-1531
- Robledo S, Idol RA, Crimmins DL, Ladenson JH, Mason PJ, Bessler M (2008) The role of human ribosomal proteins in the maturation of rRNA and ribosome production. *RNA* **14**(9): 1918-1929
- Rodley CD, Bertels F, Jones B, O'Sullivan JM (2009) Global identification of yeast chromosome interactions using Genome conformation capture. *Fungal Genet Biol* **46**(11): 879-886
- Rodor J, Jobet E, Bizarro J, Vignols F, Carles C, Suzuki T, Nakamura K, Echeverria M (2011) AtNUFIP, an essential protein for plant development, reveals the impact of snoRNA gene organisation on the assembly of snoRNPs and rRNA methylation in *Arabidopsis thaliana*. *Plant J* **65**(5): 807-819

- Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A (2004) Identification of mammalian microRNA host genes and transcription units. *Genome Res* **14**(10A): 1902-1910
- Rodriguez CR, Cho EJ, Keogh MC, Moore CL, Greenleaf AL, Buratowski S (2000) Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Mol Cell Biol* **20**(1): 104-112
- Roe SM, Ali MM, Meyer P, Vaughan CK, Panaretou B, Piper PW, Prodromou C, Pearl LH (2004) The Mechanism of Hsp90 regulation by the protein kinase-specific cochaperone p50(cdc37). *Cell* **116**(1): 87-98
- Roovers M, Hale C, Tricot C, Terns MP, Terns RM, Grosjean H, Droogmans L (2006) Formation of the conserved pseudouridine at position 55 in archaeal tRNA. *Nucleic Acids Res* **34**(15): 4293-4301
- Rosen EM, Fan S, Ma Y (2006) BRCA1 regulation of transcription. *Cancer Lett* **236**(2): 175-185
- Rozhdestvensky TS, Tang TH, Tchirkova IV, Brosius J, Bachellerie JP, Huttenhofer A (2003) Binding of L7Ae protein to the K-turn of archaeal snoRNAs: a shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea. *Nucleic Acids Res* **31**(3): 869-877
- Ruggero D, Pandolfi PP (2003) Does the ribosome translate cancer? *Nat Rev Cancer* **3**(3): 179-192
- Ruggero D, Wang ZG, Pandolfi PP (2000) The puzzling multiple lives of PML and its role in the genesis of cancer. *Bioessays* **22**(9): 827-835
- Russell AG, Schnare MN, Gray MW (2004) Pseudouridine-guide RNAs and other Cbf5p-associated RNAs in *Euglena gracilis*. *RNA* **10**(7): 1034-1046
- Rutherford SL (2003) Between genotype and phenotype: protein chaperones and evolvability. *Nat Rev Genet* **4**(4): 263-274
- 
- Saha S, Murthy S, Rangarajan PN (2006) Identification and characterization of a virus-inducible non-coding RNA in mouse brain. *J Gen Virol* **87**(Pt 7): 1991-1995
- Sahoo T, del Gaudio D, German JR, Shinawi M, Peters SU, Person RE, Garnica A, Cheung SW, Beaudet AL (2008) Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. *Nat Genet* **40**(6): 719-721
- Saitoh N, Spahr CS, Patterson SD, Bubulya P, Neuwald AF, Spector DL (2004) Proteomic analysis of interchromatin granule clusters. *Mol Biol Cell* **15**(8): 3876-3890
- Salzer U, Kubicek M, Prohaska R (1999) Isolation, molecular characterization, and tissue-specific expression of ECP-51 and ECP-54 (TIP49), two homologous, interacting erythroid cytosolic proteins. *Biochim Biophys Acta* **1446**(3): 365-370
- Samarsky DA, Fournier MJ (1998) Functional mapping of the U3 small nucleolar RNA from the yeast *Saccharomyces cerevisiae*. *Mol Cell Biol* **18**(6): 3431-3444
- Samarsky DA, Fournier MJ (1999) A comprehensive database for the small nucleolar RNAs from *Saccharomyces cerevisiae*. *Nucleic Acids Res* **27**(1): 161-164
- Samarsky DA, Fournier MJ, Singer RH, Bertrand E (1998) The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. *EMBO J* **17**(13): 3747-3757
- Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* **419**(6905): 407-411
- Sasaki T, Toh EA, Kikuchi Y (2000) Yeast Krr1p physically and functionally interacts with a novel essential Kri1p, and both proteins are required for 40S ribosome biogenesis in the nucleolus. *Mol Cell Biol* **20**(21): 7971-7979
- Sauer RT, Bolon DN, Burton BM, Burton RE, Flynn JM, Grant RA, Hersch GL, Joshi SA, Kenniston JA, Levchenko I, Neher SB, Oakes ES, Siddiqui SM, Wah DA, Baker TA (2004) Sculpting the proteome with AAA(+) proteases and disassembly machines. *Cell* **119**(1): 9-18
- Saveanu C, Bienvenu D, Namane A, Gleizes PE, Gas N, Jacquier A, Fromont-Racine M (2001) Nog2p, a putative GTPase associated with pre-60S subunits and required for late 60S maturation steps. *EMBO J* **20**(22): 6475-6484
- Schafer T, Strauss D, Petfalski E, Tollervey D, Hurt E (2003) The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. *EMBO J* **22**(6): 1370-1380
- Schaffert N, Hossbach M, Heintzmann R, Achsel T, Luhrmann R (2004) RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies. *EMBO J* **23**(15): 3000-3009
- Schattner P, Decatur WA, Davis CA, Ares M, Jr., Fournier MJ, Lowe TM (2004) Genome-wide searching for pseudouridylation guide snoRNAs: analysis of the *Saccharomyces cerevisiae* genome. *Nucleic Acids Res* **32**(14): 4281-4296
- Scherl A, Coute Y, Deon C, Calle A, Kindbeiter K, Sanchez JC, Greco A, Hochstrasser D, Diaz JJ (2002) Functional proteomic analysis of human nucleolus. *Mol Biol Cell* **13**(11): 4100-4109
- Scheuffler C, Brinker A, Bourenkov G, Pegoraro S, Moroder L, Bartunik H, Hartl FU, Moarefi I (2000) Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell* **101**(2): 199-210
- Schimmang T, Tollervey D, Kern H, Frank R, Hurt EC (1989) A yeast nucleolar protein related to mammalian fibrillarin is associated with small nucleolar RNA and is essential for viability. *EMBO J* **8**(13): 4015-4024
- Schmeing TM, Voorhees RM, Kelley AC, Gao YG, Murphy FVt, Weir JR, Ramakrishnan V (2009) The crystal structure

- of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* **326**(5953): 688-694
- Schneider M, Hsiao HH, Will CL, Giet R, Urlaub H, Luhrmann R (2010) Human PRP4 kinase is required for stable tri-snRNP association during spliceosomal B complex formation. *Nat Struct Mol Biol* **17**(2): 216-221
- Schubert HL, Blumenthal RM, Cheng X (2003) Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem Sci* **28**(6): 329-335
- Schul W, Adelaar B, van Driel R, de Jong L (1999a) Coiled bodies are predisposed to a spatial association with genes that contain snoRNA sequences in their introns. *J Cell Biochem* **75**(3): 393-403
- Schul W, Groenhout B, Koberna K, Takagaki Y, Jenny A, Manders EM, Raska I, van Driel R, de Jong L (1996) The RNA 3' cleavage factors CstF 64 kDa and CPSF 100 kDa are concentrated in nuclear domains closely associated with coiled bodies and newly synthesized RNA. *EMBO J* **15**(11): 2883-2892
- Schul W, van Der Kraan I, Matera AG, van Driel R, de Jong L (1999b) Nuclear domains enriched in RNA 3'-processing factors associate with coiled bodies and histone genes in a cell cycle-dependent manner. *Mol Biol Cell* **10**(11): 3815-3824
- Schul W, van Driel R, de Jong L (1998) Coiled bodies and U2 snRNA genes adjacent to coiled bodies are enriched in factors required for snRNA transcription. *Mol Biol Cell* **9**(5): 1025-1036
- Schultz A, Nottrott S, Hartmuth K, Luhrmann R (2006a) RNA structural requirements for the association of the spliceosomal hPrp31 protein with the U4 and U4atac small nuclear ribonucleoproteins. *J Biol Chem* **281**(38): 28278-28286
- Schultz A, Nottrott S, Watkins NJ, Luhrmann R (2006b) Protein-protein and protein-RNA contacts both contribute to the 15.5K-mediated assembly of the U4/U6 snRNP and the box C/D snoRNPs. *Mol Cell Biol* **26**(13): 5146-5154
- Scofield DG, Lynch M (2008) Evolutionary diversification of the Sm family of RNA-associated proteins. *Mol Biol Evol* **25**(11): 2255-2267
- Scott MS, Avolio F, Ono M, Lamond AI, Barton GJ (2009) Human miRNA precursors with box H/ACA snoRNA features. *PLoS Comput Biol* **5**(9): e1000507
- Segault V, Mougou A, Gregoire A, Banroques J, Branlant C (1992) An experimental study of *Saccharomyces cerevisiae* U3 snRNA conformation in solution. *Nucleic Acids Res* **20**(13): 3443-3451
- Segref A, Mattaj IW, Ohno M (2001) The evolutionarily conserved region of the U snRNA export mediator PHAX is a novel RNA-binding domain that is essential for U snRNA export. *RNA* **7**(3): 351-360
- Seipelt RL, Zheng B, Asuru A, Rymond BC (1999) U1 snRNA is cleaved by RNase III and processed through an Sm site-dependent pathway. *Nucleic Acids Res* **27**(2): 587-595
- Selmer M, Dunham CM, Murphy FVt, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**(5795): 1935-1942
- Seraphin B (1995) Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. *EMBO J* **14**(9): 2089-2098
- Seto AG, Zaug AJ, Sobel SG, Wolin SL, Cech TR (1999) *Saccharomyces cerevisiae* telomerase is an Sm small nuclear ribonucleoprotein particle. *Nature* **401**(6749): 177-180
- Shafaatian R, Payton MA, Reid JD (1996) PWP2, a member of the WD-repeat family of proteins, is an essential *Saccharomyces cerevisiae* gene involved in cell separation. *Mol Gen Genet* **252**(1-2): 101-114
- Shan J, Munro TP, Barbarese E, Carson JH, Smith R (2003) A molecular mechanism for mRNA trafficking in neuronal dendrites. *J Neurosci* **23**(26): 8859-8866
- Sharma K, Tollervey D (1999) Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing. *Mol Cell Biol* **19**(9): 6012-6019
- Sheldon KE, Mauger DM, Arndt KM (2005) A Requirement for the *Saccharomyces cerevisiae* Paf1 complex in snoRNA 3' end formation. *Mol Cell* **20**(2): 225-236
- Shen X, Mizuguchi G, Hamiche A, Wu C (2000) A chromatin remodelling complex involved in transcription and DNA processing. *Nature* **406**(6795): 541-544
- Shi Y, Di Giammartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates JR, 3rd, Frank J, Manley JL (2009) Molecular architecture of the human pre-mRNA 3' processing complex. *Mol Cell* **33**(3): 365-376
- Shibagaki Y, Itoh N, Yamada H, Nagata S, Mizumoto K (1992) mRNA capping enzyme. Isolation and characterization of the gene encoding mRNA guanylyltransferase subunit from *Saccharomyces cerevisiae*. *J Biol Chem* **267**(14): 9521-9528
- Shimba S, Buckley B, Reddy R, Kiss T, Filipowicz W (1992) Cap structure of U3 small nucleolar RNA in animal and plant cells is different. gamma-Monomethyl phosphate cap structure in plant RNA. *J Biol Chem* **267**(19): 13772-13777
- Shou W, Seol JH, Shevchenko A, Baskerville C, Moazed D, Chen ZW, Jang J, Charbonneau H, Deshaies RJ (1999) Exit from mitosis is triggered by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex. *Cell* **97**(2): 233-244
- Sigaudy S, Toutain A, Moncla A, Fredouille C, Bourliere B, Ayme S, Philip N (1998) Microcephalic osteodysplastic primordial dwarfism Taybi-Linder type: report of four cases and review of the literature. *Am J Med Genet* **80**(1): 16-24
- Sims RJ, 3rd, Rojas LA, Beck D, Bonasio R, Schuller R, Drury WJ, 3rd, Eick D, Reinberg D (2011) The C-terminal

- domain of RNA polymerase II is modified by site-specific methylation. *Science* **332**(6025): 99-103
- Singh R, Reddy R (1989) Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proc Natl Acad Sci U S A* **86**(21): 8280-8283
- Singh R, Valcarcel J (2005) Building specificity with nonspecific RNA-binding proteins. *Nat Struct Mol Biol* **12**(8): 645-653
- Singh SK, Gurha P, Tran EJ, Maxwell ES, Gupta R (2004) Sequential 2'-O-methylation of archaeal pre-tRNA<sup>Trp</sup> nucleotides is guided by the intron-encoded but trans-acting box C/D ribonucleoprotein of pre-tRNA. *J Biol Chem* **279**(46): 47661-47671
- Siomi H, Siomi MC (2009) On the road to reading the RNA-interference code. *Nature* **457**(7228): 396-404
- Sjoblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, Mandelker D, Leary RJ, Ptak J, Silliman N, Szabo S, Buckhaults P, Farrell C, Meeh P, Markowitz SD, Willis J, Dawson D, Willson JK, Gazdar AF, Hartigan J, Wu L, Liu C, Parmigiani G, Park BH, Bachman KE, Papadopoulos N, Vogelstein B, Kinzler KW, Velculescu VE (2006) The consensus coding sequences of human breast and colorectal cancers. *Science* **314**(5797): 268-274
- Sleeman JE, Lamond AI (1999a) Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. *Curr Biol* **9**(19): 1065-1074
- Sleeman JE, Lamond AI (1999b) Nuclear organization of pre-mRNA splicing factors. *Curr Opin Cell Biol* **11**(3): 372-377
- Smith JE, Cooperman BS, Mitchell P (1992) Methylation sites in *Escherichia coli* ribosomal RNA: localization and identification of four new sites of methylation in 23S rRNA. *Biochemistry* **31**(44): 10825-10834
- Smith KP, Lawrence JB (2000) Interactions of U2 gene loci and their nuclear transcripts with Cajal (coiled) bodies: evidence for PreU2 within Cajal bodies. *Mol Biol Cell* **11**(9): 2987-2998
- Smith MJ, Kulkarni S, Pawson T (2004) FF domains of CA150 bind transcription and splicing factors through multiple weak interactions. *Mol Cell Biol* **24**(21): 9274-9285
- Speckmann W, Narayanan A, Terns R, Terns MP (1999) Nuclear retention elements of U3 small nucleolar RNA. *Mol Cell Biol* **19**(12): 8412-8421
- Speckmann WA, Terns RM, Terns MP (2000) The box C/D motif directs snoRNA 5'-cap hypermethylation. *Nucleic Acids Res* **28**(22): 4467-4473
- Spector DL (2006) SnapShot: Cellular bodies. *Cell* **127**(5): 1071
- Sperger JM, Chen X, Draper JS, Antosiewicz JE, Chon CH, Jones SB, Brooks JD, Andrews PW, Brown PO, Thomson JA (2003) Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. *Proc Natl Acad Sci U S A* **100**(23): 13350-13355
- Stanek D, Pridalova-Hnilicova J, Novotny I, Huranova M, Blazikova M, Wen X, Sapra AK, Neugebauer KM (2008) Spliceosomal small nuclear ribonucleoprotein particles repeatedly cycle through Cajal bodies. *Mol Biol Cell* **19**(6): 2534-2543
- Stanek D, Rader SD, Klingauf M, Neugebauer KM (2003) Targeting of U4/U6 small nuclear RNP assembly factor SART3/p110 to Cajal bodies. *J Cell Biol* **160**(4): 505-516
- Stark H, Dube P, Luhrmann R, Kastner B (2001) Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature* **409**(6819): 539-542
- Stefano JE (1984) Purified lupus antigen La recognizes an oligouridylylate stretch common to the 3' termini of RNA polymerase III transcripts. *Cell* **36**(1): 145-154
- Steinmetz EJ, Brow DA (1996) Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. *Mol Cell Biol* **16**(12): 6993-7003
- Steinmetz EJ, Brow DA (1998) Control of pre-mRNA accumulation by the essential yeast protein Nrd1 requires high-affinity transcript binding and a domain implicated in RNA polymerase II association. *Proc Natl Acad Sci U S A* **95**(12): 6699-6704
- Steinmetz EJ, Brow DA (2003) Ssu72 protein mediates both poly(A)-coupled and poly(A)-independent termination of RNA polymerase II transcription. *Mol Cell Biol* **23**(18): 6339-6349
- Steinmetz EJ, Conrad NK, Brow DA, Corden JL (2001) RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature* **413**(6853): 327-331
- Steinmetz EJ, Ng SB, Cloute JP, Brow DA (2006) cis- and trans-Acting determinants of transcription termination by yeast RNA polymerase II. *Mol Cell Biol* **26**(7): 2688-2696
- Stevens A, Hsu CL, Isham KR, Larimer FW (1991) Fragments of the internal transcribed spacer 1 of pre-rRNA accumulate in *Saccharomyces cerevisiae* lacking 5'----3' exoribonuclease 1. *J Bacteriol* **173**(21): 7024-7028
- Stevens SW, Abelson J (1999) Purification of the yeast U4/U6.U5 small nuclear ribonucleoprotein particle and identification of its proteins. *Proc Natl Acad Sci U S A* **96**(13): 7226-7231
- Stewart M (2007) Molecular mechanism of the nuclear protein import cycle. *Nat Rev Mol Cell Biol* **8**(3): 195-208
- Stotz A, Linder P (1990) The ADE2 gene from *Saccharomyces cerevisiae*: sequence and new vectors. *Gene* **95**(1): 91-98

- Stroke IL, Weiner AM (1985) Genes and pseudogenes for rat U3A and U3B small nuclear RNA. *J Mol Biol* **184**(2): 183-193
- Strunk BS, Karbstein K (2009) Powering through ribosome assembly. *RNA* **15**(12): 2083-2104
- Suryadi J, Tran EJ, Maxwell ES, Brown BA, 2nd (2005) The crystal structure of the Methanocaldococcus jannaschii multifunctional L7Ae RNA-binding protein reveals an induced-fit interaction with the box C/D RNAs. *Biochemistry* **44**(28): 9657-9672
- Sutherland H, Bickmore WA (2009) Transcription factories: gene expression in unions? *Nat Rev Genet* **10**(7): 457-466
- Sylvester JE, Whiteman DA, Podolsky R, Pozsgay JM, Respass J, Schmickel RD (1986) The human ribosomal RNA genes: structure and organization of the complete repeating unit. *Hum Genet* **73**(3): 193-198
- Szewczak LB, DeGregorio SJ, Strobel SA, Steitz JA (2002) Exclusive interaction of the 15.5 kD protein with the terminal box C/D motif of a methylation guide snoRNP. *Chem Biol* **9**(10): 1095-1107
- Szewczak LB, Gabrielsen JS, Degregorio SJ, Strobel SA, Steitz JA (2005) Molecular basis for RNA kink-turn recognition by the h15.5K small RNP protein. *RNA* **11**(9): 1407-1419
- 
- Taddei A, Hediger F, Neumann FR, Bauer C, Gasser SM (2004) Separation of silencing from perinuclear anchoring functions in yeast Ku80, Sir4 and Esc1 proteins. *EMBO J* **23**(6): 1301-1312
- Taddei A, Schober H, Gasser SM (2010) The budding yeast nucleus. *Cold Spring Harb Perspect Biol* **2**(8): a000612
- Taft RJ, Glazov EA, Lassmann T, Hayashizaki Y, Carninci P, Mattick JS (2009) Small RNAs derived from snoRNAs. *RNA* **15**(7): 1233-1240
- Takai H, Wang RC, Takai KK, Yang H, de Lange T (2007) Tel2 regulates the stability of PI3K-related protein kinases. *Cell* **131**(7): 1248-1259
- Takai H, Xie Y, de Lange T, Pavletich NP (2010) Tel2 structure and function in the Hsp90-dependent maturation of mTOR and ATR complexes. *Genes Dev* **24**(18): 2019-2030
- Tanackovic G, Ransijn A, Thibault P, Abou Elela S, Klinck R, Berson EL, Chabot B, Rivolta C (2011) PRPF mutations are associated with generalized defects in spliceosome formation and pre-mRNA splicing in patients with retinitis pigmentosa. *Hum Mol Genet* **20**(11): 2116-2130
- Tang TH, Bachelier JP, Rozhdestvensky T, Bortolin ML, Huber H, Drungowski M, Elge T, Brosius J, Huttenhofer A (2002) Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc Natl Acad Sci U S A* **99**(11): 7536-7541
- Tapia H, Morano KA (2010) Hsp90 nuclear accumulation in quiescence is linked to chaperone function and spore development in yeast. *Mol Biol Cell* **21**(1): 63-72
- Tarun SZ, Jr., Sachs AB (1996) Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J* **15**(24): 7168-7177
- Te J, Jia L, Rogers J, Miller A, Hartson SD (2007) Novel subunits of the mammalian Hsp90 signal transduction chaperone. *J Proteome Res* **6**(5): 1963-1973
- Teigelkamp S, Achsel T, Mundt C, Gotthel SF, Cronshagen U, Lane WS, Marahiel M, Luhrmann R (1998) The 20kD protein of human [U4/U6.U5] tri-snRNPs is a novel cyclophilin that forms a complex with the U4/U6-specific 60kD and 90kD proteins. *RNA* **4**(2): 127-141
- Theimer CA, Jady BE, Chim N, Richard P, Breece KE, Kiss T, Feigon J (2007) Structural and functional characterization of human telomerase RNA processing and cajal body localization signals. *Mol Cell* **27**(6): 869-881
- Thiry M (1995) The interchromatin granules. *Histol Histopathol* **10**(4): 1035-1045
- Thiry M, Lafontaine DL (2005) Birth of a nucleolus: the evolution of nucleolar compartments. *Trends Cell Biol* **15**(4): 194-199
- Thomson E, Tollervey D (2010) The final step in 5.8S rRNA processing is cytoplasmic in *Saccharomyces cerevisiae*. *Mol Cell Biol* **30**(4): 976-984
- Tian H, Chen HJ, Cross TH, Edenberg HJ (2005) Alternative splicing and promoter use in the human GABRA2 gene. *Brain Res Mol Brain Res* **137**(1-2): 174-183
- Tiedge H (2006) K-turn motifs in spatial RNA coding. *RNA Biol* **3**(4): 133-139
- Tollervey D, Lehtonen H, Carmo-Fonseca M, Hurt EC (1991) The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre-rRNA processing in yeast. *EMBO J* **10**(3): 573-583
- Tollervey D, Lehtonen H, Jansen R, Kern H, Hurt EC (1993) Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell* **72**(3): 443-457
- Tong AH, Lesage G, Bader GD, Ding H, Xu H, Xin X, Young J, Berriz GF, Brost RL, Chang M, Chen Y, Cheng X, Chua G, Friesen H, Goldberg DS, Haynes J, Humphries C, He G, Hussein S, Ke L, Krogan N, Li Z, Levinson JN, Lu H, Menard P, Munyana C, Parsons AB, Ryan O, Tonikian R, Roberts T, Sdicu AM, Shapiro J, Sheikh B, Suter B, Wong SL, Zhang LV, Zhu H, Burd CG, Munro S, Sander C, Rine J, Greenblatt J, Peter M, Bretscher A, Bell G, Roth FP, Brown GW, Andrews B, Bussey H, Boone C (2004) Global mapping of the yeast genetic interaction network. *Science* **303**(5659): 808-813
- Torreira E, Jha S, Lopez-Blanco JR, Arias-Palomo E, Chacon P, Canas C, Ayora S, Dutta A, Llorca O (2008) Architecture of the pontin/reptin complex, essential in the assembly of several macromolecular complexes. *Structure* **16**(10): 1511-1520



- Toyota CG, Davis MD, Cosman AM, Hebert MD (2010) Coilin phosphorylation mediates interaction with SMN and SmB'. *Chromosoma* **119**(2): 205-215
- Trahan C, Dragon F (2009) Dyskeratosis congenita mutations in the H/ACA domain of human telomerase RNA affect its assembly into a pre-RNP. *RNA* **15**(2): 235-243
- Trahan C, Martel C, Dragon F (2010) Effects of dyskeratosis congenita mutations in dyskerin, NHP2 and NOP10 on assembly of H/ACA pre-RNPs. *Hum Mol Genet* **19**(5): 825-836
- Tran E, Brown J, Maxwell ES (2004) Evolutionary origins of the RNA-guided nucleotide-modification complexes: from the primitive translation apparatus? *Trends Biochem Sci* **29**(7): 343-350
- Tran E, Zhang X, Lackey L, Maxwell ES (2005) Conserved spacing between the box C/D and C'/D' RNPs of the archaeal box C/D sRNP complex is required for efficient 2'-O-methylation of target RNAs. *RNA* **11**(3): 285-293
- Tran EJ, Zhang X, Maxwell ES (2003) Efficient RNA 2'-O-methylation requires juxtaposed and symmetrically assembled archaeal box C/D and C'/D' RNPs. *EMBO J* **22**(15): 3930-3940
- Trapman J, Planta RJ (1976) Maturation of ribosomes in yeast. I Kinetic analysis by labelling of high molecular weight rRNA species. *Biochim Biophys Acta* **442**(3): 265-274
- Trapman J, Retel J, Planta RJ (1975) Ribosomal precursor particles from yeast. *Exp Cell Res* **90**(1): 95-104
- Trepel J, Mollapour M, Giaccone G, Neckers L (2010) Targeting the dynamic HSP90 complex in cancer. *Nat Rev Cancer* **10**(8): 537-549
- Trumtel S, Leger-Silvestre I, Gleizes PE, Teulieres F, Gas N (2000) Assembly and functional organization of the nucleolus: ultrastructural analysis of *Saccharomyces cerevisiae* mutants. *Mol Biol Cell* **11**(6): 2175-2189
- Tsaneva IR, Muller B, West SC (1993) RuvA and RuvB proteins of *Escherichia coli* exhibit DNA helicase activity in vitro. *Proc Natl Acad Sci U S A* **90**(4): 1315-1319
- Tschochner H, Hurt E (2003) Pre-ribosomes on the road from the nucleolus to the cytoplasm. *Trends Cell Biol* **13**(5): 255-263
- Tsakamoto T, Shibagaki Y, Imajoh-Ohmi S, Murakoshi T, Suzuki M, Nakamura A, Gotoh H, Mizumoto K (1997) Isolation and characterization of the yeast mRNA capping enzyme beta subunit gene encoding RNA 5'-triphosphatase, which is essential for cell viability. *Biochem Biophys Res Commun* **239**(1): 116-122
- Tsakamoto T, Shibagaki Y, Niikura Y, Mizumoto K (1998) Cloning and characterization of three human cDNAs encoding mRNA (guanine-7-)-methyltransferase, an mRNA cap methylase. *Biochem Biophys Res Commun* **251**(1): 27-34
- Tucker KE, Berciano MT, Jacobs EY, LePage DF, Shpargel KB, Rossire JJ, Chan EK, Lafarga M, Conlon RA, Matera AG (2001) Residual Cajal bodies in coilin knockout mice fail to recruit Sm snRNPs and SMN, the spinal muscular atrophy gene product. *J Cell Biol* **154**(2): 293-307
- Turner AJ, Knox AA, Prieto JL, McStay B, Watkins NJ (2009) A novel small-subunit processome assembly intermediate that contains the U3 snoRNP, nucleolin, RRP5, and DBP4. *Mol Cell Biol* **29**(11): 3007-3017
- Turner B, Melcher SE, Wilson TJ, Norman DG, Lilley DM (2005) Induced fit of RNA on binding the L7Ae protein to the kink-turn motif. *RNA* **11**(8): 1192-1200
- Tycowski KT, Shu MD, Kukoyi A, Steitz JA (2009) A conserved WD40 protein binds the Cajal body localization signal of scaRNP particles. *Mol Cell* **34**(1): 47-57
- Tycowski KT, Shu MD, Steitz JA (1993) A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. *Genes Dev* **7**(7A): 1176-1190
- Tycowski KT, Shu MD, Steitz JA (1994) Requirement for intron-encoded U22 small nucleolar RNA in 18S ribosomal RNA maturation. *Science* **266**(5190): 1558-1561
- Tycowski KT, Smith CM, Shu MD, Steitz JA (1996) A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. *Proc Natl Acad Sci U S A* **93**(25): 14480-14485
- Tycowski KT, You ZH, Graham PJ, Steitz JA (1998) Modification of U6 spliceosomal RNA is guided by other small RNAs. *Mol Cell* **2**(5): 629-638
- 
- Udem SA, Warner JR (1972) Ribosomal RNA synthesis in *Saccharomyces cerevisiae*. *J Mol Biol* **65**(2): 227-242
- Ulbrich C, Diepholz M, Bassler J, Kressler D, Pertschy B, Galani K, Bottcher B, Hurt E (2009) Mechanochemical removal of ribosome biogenesis factors from nascent 60S ribosomal subunits. *Cell* **138**(5): 911-922
- Urlaub H, Raker VA, Kostka S, Luhrmann R (2001) Sm protein-Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure. *EMBO J* **20**(1-2): 187-196
- Ursic D, Himmel KL, Gurley KA, Webb F, Culbertson MR (1997) The yeast SEN1 gene is required for the processing of diverse RNA classes. *Nucleic Acids Res* **25**(23): 4778-4785
- 
- Van Hoof A, Lennertz P, Parker R (2000a) Three conserved members of the RNase D family have unique and overlapping functions in the processing of 5S, 5.8S, U4, U5, RNase MRP and RNase P RNAs in yeast. *EMBO J* **19**(6): 1357-1365
- van Hoof A, Lennertz P, Parker R (2000b) Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol Cell Biol* **20**(2): 441-452

- van Nues RW, Beggs JD (2001) Functional contacts with a range of splicing proteins suggest a central role for Brr2p in the dynamic control of the order of events in spliceosomes of *Saccharomyces cerevisiae*. *Genetics* **157**(4): 1451-1467
- Vanacova S, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith G, Keller W (2005) A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol* **3**(6): e189
- Vanrobays E, Gleizes PE, Bousquet-Antonelli C, Noaillac-Depeyre J, Caizergues-Ferrer M, Gelugne JP (2001) Processing of 20S pre-rRNA to 18S ribosomal RNA in yeast requires Rrp10p, an essential non-ribosomal cytoplasmic protein. *EMBO J* **20**(15): 4204-4213
- Vasiljeva L, Buratowski S (2006) Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. *Mol Cell* **21**(2): 239-248
- Vasiljeva L, Kim M, Mutschler H, Buratowski S, Meinhart A (2008) The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* **15**(8): 795-804
- Veldman GM, Klootwijk J, de Jonge P, Leer RJ, Planta RJ (1980) The transcription termination site of the ribosomal RNA operon in yeast. *Nucleic Acids Res* **8**(22): 5179-5192
- Veldman GM, Klootwijk J, de Regt VC, Planta RJ, Branlant C, Krol A, Ebel JP (1981) The primary and secondary structure of yeast 26S rRNA. *Nucleic Acids Res* **9**(24): 6935-6952
- Venema J, Vos HR, Faber AW, van Venrooij WJ, Raue HA (2000) Yeast Rrp9p is an evolutionarily conserved U3 snoRNP protein essential for early pre-rRNA processing cleavages and requires box C for its association. *RNA* **6**(11): 1660-1671
- Venteicher AS, Meng Z, Mason PJ, Veenstra TD, Artandi SE (2008) Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. *Cell* **132**(6): 945-957
- Verdecia MA, Bowman ME, Lu KP, Hunter T, Noel JP (2000) Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat Struct Biol* **7**(8): 639-643
- Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D (2004) RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**(5658): 672-676
- Verheggen C, Lafontaine DL, Samarsky D, Mouaikel J, Blanchard JM, Bordonne R, Bertrand E (2002) Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. *EMBO J* **21**(11): 2736-2745
- Verheggen C, Mouaikel J, Thiry M, Blanchard JM, Tollervey D, Bordonne R, Lafontaine DL, Bertrand E (2001) Box C/D small nucleolar RNA trafficking involves small nucleolar RNP proteins, nucleolar factors and a novel nuclear domain. *EMBO J* **20**(19): 5480-5490
- Vidovic I, Nottrott S, Hartmuth K, Luhrmann R, Ficner R (2000) Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol Cell* **6**(6): 1331-1342
- Vilardell J, Warner JR (1994) Regulation of splicing at an intermediate step in the formation of the spliceosome. *Genes Dev* **8**(2): 211-220
- Villa T, Ceradini F, Bozzoni I (2000) Identification of a novel element required for processing of intron-encoded box C/D small nucleolar RNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol* **20**(4): 1311-1320
- Villa T, Ceradini F, Presutti C, Bozzoni I (1998) Processing of the intron-encoded U18 small nucleolar RNA in the yeast *Saccharomyces cerevisiae* relies on both exo- and endonucleolytic activities. *Mol Cell Biol* **18**(6): 3376-3383
- Vincenti S, De Chiara V, Bozzoni I, Presutti C (2007) The position of yeast snoRNA-coding regions within host introns is essential for their biosynthesis and for efficient splicing of the host pre-mRNA. *RNA* **13**(1): 138-150
- Visa N, Izaurralde E, Ferreira J, Daneholt B, Mattaj IW (1996) A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. *J Cell Biol* **133**(1): 5-14
- Vitali P, Basyuk E, Le Meur E, Bertrand E, Muscatelli F, Cavaille J, Huttenhofer A (2005) ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs. *J Cell Biol* **169**(5): 745-753
- Vithana EN, Abu-Safieh L, Allen MJ, Carey A, Papaioannou M, Chakarova C, Al-Magthteh M, Ebenezer ND, Willis C, Moore AT, Bird AC, Hunt DM, Bhattacharya SS (2001) A human homolog of yeast pre-mRNA splicing gene, PRP31, underlies autosomal dominant retinitis pigmentosa on chromosome 19q13.4 (RP11). *Mol Cell* **8**(2): 375-381
- von der Haar T, Gross JD, Wagner G, McCarthy JE (2004) The mRNA cap-binding protein eIF4E in post-transcriptional gene expression. *Nat Struct Mol Biol* **11**(6): 503-511
- Vulliamy T, Beswick R, Kirwan M, Marrone A, Digweed M, Walne A, Dokal I (2008) Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. *Proc Natl Acad Sci U S A* **105**(23): 8073-8078
- Vulliamy T, Marrone A, Goldman F, Dearlove A, Bessler M, Mason PJ, Dokal I (2001) The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature* **413**(6854): 432-435
- Vulliamy TJ, Dokal I (2008) Dyskeratosis congenita: the diverse clinical presentation of mutations in the telomerase complex. *Biochimie* **90**(1): 122-130

---

Wahl MC, Will CL, Luhrmann R (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**(4): 701-718

- Walbott H, Machado-Pinilla R, Liger D, Blaud M, Rety S, Grozdanov PN, Godin K, van Tilbeurgh H, Varani G, Meier UT, Leulliot N (2011) The H/ACA RNP assembly factor SHQ1 functions as an RNA mimic. *Genes Dev* **25**(22): 2398-2408
- Waldron C, Jund R, Lacroute F (1977) Evidence for a high proportion of inactive ribosomes in slow-growing yeast cells. *Biochem J* **168**(3): 409-415
- Walne AJ, Dokal I (2009) Advances in the understanding of dyskeratosis congenita. *Br J Haematol* **145**(2): 164-172
- Walne AJ, Vulliamy T, Marrone A, Beswick R, Kirwan M, Masunari Y, Al-Qurashi FH, Aljurf M, Dokal I (2007) Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10. *Hum Mol Genet* **16**(13): 1619-1629
- Walther TN, Wittop Koning TH, Schumperli D, Muller B (1998) A 5'-3' exonuclease activity involved in forming the 3' products of histone pre-mRNA processing in vitro. *RNA* **4**(9): 1034-1046
- Wang C, Meier UT (2004) Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *EMBO J* **23**(8): 1857-1867
- Wang C, Query CC, Meier UT (2002) Immunopurified small nucleolar ribonucleoprotein particles pseudouridylate rRNA independently of their association with phosphorylated Nopp140. *Mol Cell Biol* **22**(24): 8457-8466
- Wang H, Boisvert D, Kim KK, Kim R, Kim SH (2000) Crystal structure of a fibrillarin homologue from *Methanococcus jannaschii*, a hyperthermophile, at 1.6 Å resolution. *EMBO J* **19**(3): 317-323
- Wang ZF, Whitfield ML, Ingledue TC, 3rd, Dominski Z, Marzluff WF (1996) The protein that binds the 3' end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing. *Genes Dev* **10**(23): 3028-3040
- Warner JR (1999) The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci* **24**(11): 437-440
- Watkins NJ, Dickmanns A, Luhrmann R (2002) Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Mol Cell Biol* **22**(23): 8342-8352
- Watkins NJ, Gottschalk A, Neubauer G, Kastner B, Fabrizio P, Mann M, Luhrmann R (1998) Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA* **4**(12): 1549-1568
- Watkins NJ, Lemm I, Ingelfinger D, Schneider C, Hossbach M, Urlaub H, Luhrmann R (2004) Assembly and maturation of the U3 snoRNP in the nucleoplasm in a large dynamic multiprotein complex. *Mol Cell* **16**(5): 789-798
- Watkins NJ, Lemm I, Luhrmann R (2007) Involvement of nuclear import and export factors in U8 box C/D snoRNP biogenesis. *Mol Cell Biol* **27**(20): 7018-7027
- Watkins NJ, Leverette RD, Xia L, Andrews MT, Maxwell ES (1996) Elements essential for processing intronic U14 snoRNA are located at the termini of the mature snoRNA sequence and include conserved nucleotide boxes C and D. *RNA* **2**(2): 118-133
- Watkins NJ, Segault V, Charpentier B, Nottrott S, Fabrizio P, Bachi A, Wilm M, Rosbash M, Branlant C, Luhrmann R (2000) A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP. *Cell* **103**(3): 457-466
- Wegierski T, Billy E, Nasr F, Filipowicz W (2001) Bms1p, a G-domain-containing protein, associates with Rcl1p and is required for 18S rRNA biogenesis in yeast. *RNA* **7**(9): 1254-1267
- Weidenhammer EM, Ruiz-Noriega M, Woolford JL, Jr. (1997) Prp31p promotes the association of the U4/U6 x U5 tri-snRNP with pre-spliceosomes to form spliceosomes in *Saccharomyces cerevisiae*. *Mol Cell Biol* **17**(7): 3580-3588
- Weidenhammer EM, Singh M, Ruiz-Noriega M, Woolford JL, Jr. (1996) The PRP31 gene encodes a novel protein required for pre-mRNA splicing in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **24**(6): 1164-1170
- Weinberg RA, Penman S (1968) Small molecular weight monodisperse nuclear RNA. *J Mol Biol* **38**(3): 289-304
- Weinstein LB, Steitz JA (1999) Guided tours: from precursor snoRNA to functional snoRNP. *Curr Opin Cell Biol* **11**(3): 378-384
- West S, Gromak N, Proudfoot NJ (2004) Human 5' → 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* **432**(7016): 522-525
- Westman BJ, Lamond AI (2011) A role for SUMOylation in snoRNP biogenesis revealed by quantitative proteomics. *Nucleus* **2**(1): 30-37
- Westman BJ, Verheggen C, Hutten S, Lam YW, Bertrand E, Lamond AI (2010) A proteomic screen for nucleolar SUMO targets shows SUMOylation modulates the function of Nop5/Nop58. *Mol Cell* **39**(4): 618-631
- White SA, Hoeger M, Schweppe JJ, Shillingford A, Shipilov V, Zarutskie J (2004) Internal loop mutations in the ribosomal protein L30 binding site of the yeast L30 RNA transcript. *RNA* **10**(3): 369-377
- Whittom AA, Xu H, Hebert MD (2008) Coilin levels and modifications influence artificial reporter splicing. *Cell Mol Life Sci* **65**(7-8): 1256-1271
- Wiederhold K, Passmore LA (2010) Cytoplasmic deadenylation: regulation of mRNA fate. *Biochem Soc Trans* **38**(6): 1531-1536
- Will CL, Luhrmann R (2001) Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol* **13**(3): 290-301

- Wilmes GM, Bergkessel M, Bandyopadhyay S, Shales M, Braberg H, Cagney G, Collins SR, Whitworth GB, Kress TL, Weissman JS, Ideker T, Guthrie C, Krogan NJ (2008) A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. *Mol Cell* **32**(5): 735-746
- Wimberly BT, Brodersen DE, Clemons WM, Jr., Morgan-Warren RJ, Carter AP, Vornrhein C, Hartsch T, Ramakrishnan V (2000) Structure of the 30S ribosomal subunit. *Nature* **407**(6802): 327-339
- Winkler AA, Bobok A, Zonneveld BJ, Steensma HY, Hooykaas PJ (1998) The lysine-rich C-terminal repeats of the centromere-binding factor 5 (Cbf5) of *Kluyveromyces lactis* are not essential for function. *Yeast* **14**(1): 37-48
- Winkler C, Eggert C, Gradl D, Meister G, Giegerich M, Wedlich D, Lagerbauer B, Fischer U (2005) Reduced U snRNP assembly causes motor axon degeneration in an animal model for spinal muscular atrophy. *Genes Dev* **19**(19): 2320-2330
- Wise JA, Weiner AM (1980) Dictyostelium small nuclear RNA D2 is homologous to rat nucleolar RNA U3 and is encoded by a dispersed multigene family. *Cell* **22**(1 Pt 1): 109-118
- Wolin SL, Cedervall T (2002) The La protein. *Annu Rev Biochem* **71**: 375-403
- Wood MA, McMahon SB, Cole MD (2000) An ATPase/helicase complex is an essential cofactor for oncogenic transformation by c-Myc. *Mol Cell* **5**(2): 321-330
- Wormsley S, Samarsky DA, Fournier MJ, Baserga SJ (2001) An unexpected, conserved element of the U3 snoRNA is required for Mpp10p association. *RNA* **7**(6): 904-919
- Wu G, Xiao M, Yang C, Yu YT (2011) U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *EMBO J* **30**(1): 79-89
- Wu H, Min J, Zeng H, Loppnau P, Weigelt J, Sundstrom M, Arrowsmith CH, Edwards AM, Bochkarev A, Plotnikov AN (non publié) Structure rRNA 2'-O-methyltransferase fibrillar.
- Wu H, Yang PK, Butcher SE, Kang S, Chanfreau G, Feigon J (2001) A novel family of RNA tetraloop structure forms the recognition site for *Saccharomyces cerevisiae* RNase III. *EMBO J* **20**(24): 7240-7249
- Wu P, Brockenbrough JS, Metcalfe AC, Chen S, Aris JP (1998) Nop5p is a small nucleolar ribonucleoprotein component required for pre-18 S rRNA processing in yeast. *J Biol Chem* **273**(26): 16453-16463
- Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Regnault B, Devaux F, Namane A, Seraphin B, Libri D, Jacquier A (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**(5): 725-737
- Xia L, Watkins NJ, Maxwell ES (1997) Identification of specific nucleotide sequences and structural elements required for intronic U14 snoRNA processing. *RNA* **3**(1): 17-26
- Xiao M, Yang C, Schattner P, Yu YT (2009) Functionality and substrate specificity of human box H/ACA guide RNAs. *RNA* **15**(1): 176-186
- Xu C, Henry PA, Setya A, Henry MF (2003) In vivo analysis of nucleolar proteins modified by the yeast arginine methyltransferase Hmt1/Rmt1p. *RNA* **9**(6): 746-759
- Xue D, Rubinson DA, Pannone BK, Yoo CJ, Wolin SL (2000a) U snRNP assembly in yeast involves the La protein. *EMBO J* **19**(7): 1650-1660
- Xue S, Wang R, Yang F, Terns RM, Terns MP, Zhang X, Maxwell ES, Li H (2010) Structural basis for substrate placement by an archaeal box C/D ribonucleoprotein particle. *Mol Cell* **39**(6): 939-949
- Xue Y, Bai X, Lee I, Kallstrom G, Ho J, Brown J, Stevens A, Johnson AW (2000b) *Saccharomyces cerevisiae* RAI1 (YGL246c) is homologous to human DOM3Z and encodes a protein that binds the nuclear exoribonuclease Rat1p. *Mol Cell Biol* **20**(11): 4006-4015
- 
- Yamada-Okabe T, Doi R, Shimmi O, Arisawa M, Yamada-Okabe H (1998) Isolation and characterization of a human cDNA for mRNA 5'-capping enzyme. *Nucleic Acids Res* **26**(7): 1700-1706
- Yang PK, Hoareau C, Froment C, Monsarrat B, Henry Y, Chanfreau G (2005) Cotranscriptional recruitment of the pseudouridyltransferase Cbf5p and of the RNA binding protein Naf1p during H/ACA snoRNP assembly. *Mol Cell Biol* **25**(8): 3295-3304
- Yang PK, Rotondo G, Porras T, Legrain P, Chanfreau G (2002) The Shq1p.Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis. *J Biol Chem* **277**(47): 45235-45242
- Yang Y, Isaac C, Wang C, Dragon F, Pogacic V, Meier UT (2000) Conserved composition of mammalian box H/ACA and box C/D small nucleolar ribonucleoprotein particles and their interaction with the common factor Nopp140. *Mol Biol Cell* **11**(2): 567-577
- Yao W, Roser D, Kohler A, Bradatsch B, Bassler J, Hurt E (2007) Nuclear export of ribosomal 60S subunits by the general mRNA export receptor Mex67-Mtr2. *Mol Cell* **26**(1): 51-62
- Ye K, Jia R, Lin J, Ju M, Peng J, Xu A, Zhang L (2009) Structural organization of box C/D RNA-guided RNA methyltransferase. *Proc Natl Acad Sci U S A* **106**(33): 13808-13813
- Ye X, Wei Y, Nalepa G, Harper JW (2003) The cyclin E/Cdk2 substrate p220(NPAT) is required for S-phase entry, histone gene expression, and Cajal body maintenance in human somatic cells. *Mol Cell Biol* **23**(23): 8586-8600

- Yeager TR, Neumann AA, Englezou A, Huschtscha LI, Noble JR, Reddel RR (1999) Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Res* **59**(17): 4175-4179
- Yong J, Golembe TJ, Battle DJ, Pellizzoni L, Dreyfuss G (2004) snRNAs contain specific SMN-binding domains that are essential for snRNP assembly. *Mol Cell Biol* **24**(7): 2747-2756
- Yong J, Kasim M, Bachorik JL, Wan L, Dreyfuss G (2010) Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. *Mol Cell* **38**(4): 551-562
- Yong J, Pellizzoni L, Dreyfuss G (2002) Sequence-specific interaction of U1 snRNA with the SMN complex. *EMBO J* **21**(5): 1188-1196
- Young PJ, Day PM, Zhou J, Androphy EJ, Morris GE, Lorson CL (2002) A direct interaction between the survival motor neuron protein and p53 and its relationship to spinal muscular atrophy. *J Biol Chem* **277**(4): 2852-2859
- Young PJ, Le TT, Dunckley M, Nguyen TM, Burghes AH, Morris GE (2001) Nuclear gems and Cajal (coiled) bodies in fetal tissues: nucleolar distribution of the spinal muscular atrophy protein, SMN. *Exp Cell Res* **265**(2): 252-261
- Yu YT, Shu MD, Narayanan A, Terns RM, Terns MP, Steitz JA (2001) Internal modification of U2 small nuclear (sn)RNA occurs in nucleoli of *Xenopus* oocytes. *J Cell Biol* **152**(6): 1279-1288
- Yu YT, Shu MD, Steitz JA (1998) Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J* **17**(19): 5783-5795
- Yuryev A, Patturajan M, Litingtung Y, Joshi RV, Gentile C, Gebara M, Corden JL (1996) The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc Natl Acad Sci U S A* **93**(14): 6975-6980
- Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, Cate JH, Noller HF (2001) Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**(5518): 883-896
- Zagorski J, Tollervey D, Fournier MJ (1988) Characterization of an SNR gene locus in *Saccharomyces cerevisiae* that specifies both dispensable and essential small nuclear RNAs. *Mol Cell Biol* **8**(8): 3282-3290
- Zhang J, Harnpicharnchai P, Jakovljevic J, Tang L, Guo Y, Oeffinger M, Rout MP, Hiley SL, Hughes T, Woolford JL, Jr. (2007) Assembly factors Rpf2 and Rrs1 recruit 5S rRNA and ribosomal proteins rpL5 and rpL11 into nascent ribosomes. *Genes Dev* **21**(20): 2580-2592
- Zhao R, Bodnar MS, Spector DL (2009) Nuclear neighborhoods and gene expression. *Curr Opin Genet Dev* **19**(2): 172-179
- Zhao R, Davey M, Hsu YC, Kaplanek P, Tong A, Parsons AB, Krogan N, Cagney G, Mai D, Greenblatt J, Boone C, Emili A, Houry WA (2005) Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the hsp90 chaperone. *Cell* **120**(5): 715-727
- Zhao R, Houry WA (2005) Hsp90: a chaperone for protein folding and gene regulation. *Biochem Cell Biol* **83**(6): 703-710
- Zhao R, Houry WA (2007) Molecular interaction network of the Hsp90 chaperone system. *Adv Exp Med Biol* **594**: 27-36
- Zhao R, Kakihara Y, Gribun A, Huen J, Yang G, Khanna M, Costanzo M, Brost RL, Boone C, Hughes TR, Yip CM, Houry WA (2008) Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation. *J Cell Biol* **180**(3): 563-578
- Zhao X, Yu YT (2004) Pseudouridines in and near the branch site recognition region of U2 snRNA are required for snRNP biogenesis and pre-mRNA splicing in *Xenopus* oocytes. *RNA* **10**(4): 681-690
- Zhou J, Liang B, Li H (2011) Structural and functional evidence of high specificity of Cbf5 for ACA trinucleotide. *RNA* **17**(2): 244-250
- Zhu Y, Tomlinson RL, Lukowiak AA, Terns RM, Terns MP (2004) Telomerase RNA accumulates in Cajal bodies in human cancer cells. *Mol Biol Cell* **15**(1): 81-90
- Zieve G, Penman S (1976) Small RNA species of the HeLa cell: metabolism and subcellular localization. *Cell* **8**(1): 19-31





## **Étude des processus de biogenèse des petites particules ribonucléoprotéiques nucléolaires à boîtes C/D (snoRNP C/D) chez la levure *Saccharomyces cerevisiae*. Caractérisation fonctionnelle et structurale d'une machinerie dédiée à l'assemblage de ces RNP.**

Les protéines de la famille L7Ae sont les constituants de nombreuses RNP essentielles. Chez les vertébrés, les particules snoRNP C/D et H/ACA sont impliquées dans la biogenèse des ribosomes, la UsnRNP U4 dans l'épissage des pré-ARNm, le complexe télomérase dans la réplication des télomères, et les mRNP SECIS dans la traduction des sélénoprotéines. Comme c'est le cas pour la majorité des RNP eucaryotes, leur assemblage, sous forme d'entités fonctionnelles, ne constitue pas un processus autonome et requiert l'intervention de facteurs spécialisés. En basant notre étude sur l'assemblage des snoRNP C/D, dans l'organisme modèle *Saccharomyces cerevisiae*, et en utilisant des approches de biologie moléculaire, de biochimie et de génétique, nous avons entrepris de caractériser ces événements. Nos travaux ont contribué à identifier un ensemble de protéines, agissant de façon coordonnée au sein d'une machinerie conservée entre la levure et l'homme. Cette dernière est composée de deux principales sous-unités : (i) Rsa1p/NUFIP, une protéine plate-forme, qui interagit avec certaines protéines de la famille L7Ae et facilite l'assemblage des RNP, (ii) le complexe R2TP (Rvb1p/TIP49, Rvb2p/TIP48, Pih1p/PIH1, Tah1p/SPAGH), qui pourrait opérer des remodelages conformationnels nécessaires à la formation des RNP matures. En plus de ces acteurs centraux, d'autres facteurs sont apparus intimement liés à ce mécanisme. La protéine Hit1p/TRIP3, interagit notamment avec Rsa1p/NUFIP et s'est avéré requise pour assurer sa stabilité chez la levure. La chaperonne HSP90, dont le rôle est prédominant chez l'homme, exerce son activité sur certains constituants des RNP. Enfin, la protéine Bcd1p/BCD1 pourrait être associée à cette machinerie dans le cadre spécifique de l'assemblage des snoRNP C/D.

**Mots clés :** *Saccharomyces cerevisiae*, snoRNP, snoRNA à boîtes C/D, snoRNP U3, complexes ARN-protéines, analyse structure-fonction, assemblage de macro-complexes, protéines de la famille L7Ae, Snu13p, Nop56p, Nop58p, Rsa1p, Hit1p, Pih1p, Tah1p, Bcd1p, HSP90, NUFIP, complexe R2TP.

---

### **Study of the biogenesis process of box C/D small nucleolar ribonucleoparticles (C/D snoRNPs) in the yeast *Saccharomyces cerevisiae*. Functional and structural characterization of a machinery dedicated to assembly of these RNPs.**

The L7Ae family proteins are essential components of many RNPs. In vertebrates, C/D and H/ACA snoRNPs are involved in ribosome biogenesis, the U4 snRNP in pre-mRNA splicing, the telomerase complex in telomeres replication, and mRNP SECIS in selenoproteins translation. Like most eukaryotic RNPs, assembly in functional entities is not an autonomous process and requires the intervention of specialized factors. Basing our study on the assembly of C/D snoRNP in the model organism *Saccharomyces cerevisiae*, and using approaches of molecular biology, biochemistry and genetics, we undertook to decipher these mechanisms. Our work has helped to identify a set of proteins, acting in a coordinated manner within a machinery conserved between yeast and human. This machinery consists of two major subunits: (i) Rsa1p/NUFIP, a platform protein that interacts with some proteins of the L7Ae family and facilitates the RNPs assembly, (ii) the R2TP complex (Rvb1p/TIP49, Rvb2p/TIP48, Pih1p/PIH1, Tah1p/SPAGH), which could induce conformational remodeling necessary for the formation of mature RNPs. In addition to these key players, other factors appeared closely linked to this mechanism. The Hit1p/TRIP3 protein interacts with Rsa1p/NUFIP and is required to ensure its stability in yeast. HSP90 chaperone, whose role is predominant in human, operates on some components of the RNPs. Finally, the Bcd1p/BCD1 protein is associated specifically with this machinery during C/D snoRNPs assembly.

**Key words :** *Saccharomyces cerevisiae*, snoRNP, box C/D snoRNA, U3 snoRNP, RNA-proteins complexes, structure-function analysis, macro-complexes assembly, L7Ae family proteins, Snu13p, Nop56p, Nop58p, Rsa1p, Hit1p, Pih1p, Tah1p, Bcd1p, HSP90, NUFIP, R2TP complex.