



HAL
open science

Identification des résidus pseudouridines dans les ARN ribosomiques "d'Archaea" et les petits ARN nucléaires des spliceosomes majeurs et mineurs : recherche et caractérisation de pseudouridine-synthases chez "Saccharomyces cerevisiae"

Séverine Massenet

► **To cite this version:**

Séverine Massenet. Identification des résidus pseudouridines dans les ARN ribosomiques "d'Archaea" et les petits ARN nucléaires des spliceosomes majeurs et mineurs : recherche et caractérisation de pseudouridine-synthases chez "Saccharomyces cerevisiae". Biochimie, Biologie Moléculaire. Université Henri Poincaré - Nancy 1, 1999. Français. NNT : 1999NAN10311 . tel-01747696

HAL Id: tel-01747696

<https://hal.univ-lorraine.fr/tel-01747696>

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

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.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Faculté des Sciences
UFR Sciences et Techniques Biologiques
Ecole Doctorale Biologie et Santé
Laboratoire de Maturation des ARN et Enzymologie Moléculaire

Thèse

Présentée pour l'obtention du titre de

Docteur de l'Université Henri Poincaré, Nancy I
Biologie Structurale Moléculaire et Cellulaire

Par Séverine MASSENET

Identification des résidus pseudouridines dans les ARN ribosomiques d'*Archaea* et les petits ARN nucléaires des spliceosomes majeurs et mineurs, recherche et caractérisation de pseudouridine-synthases chez *Saccharomyces cerevisiae*.

Soutenance publique prévue le 23 décembre 1999 devant la commission d'examen:

Membres du Jury:

Président:

M. B. Decaris Professeur, Université Henri Poincaré, Nancy

Rapporteurs:

Mme J. Banroques Chargée de Recherche CNRS, Gif-sur-Yvette

M. R. Giegé Directeur de Recherche CNRS, Strasbourg

Examineur:

M. H. Grosjean Directeur de Recherche CNRS, Gif-sur-Yvette

Directeur de thèse:

Mme C. Branlant Directeur de Recherche CNRS, Nancy

S. C. D. - U.H.F. NANCY I
BIBLIOTHÈQUE DES SCIENCES
Rue de Jardin Botanique
54600 VILLERS-LES-NANCY

A mes parents

Je tiens à exprimer ma gratitude à Monsieur le Professeur Guy Branlant pour m'avoir accueillie dans son laboratoire et permis de réaliser ce travail.

J'ai à cœur d'exprimer ma reconnaissance à Madame Christiane Branlant qui a dirigé ce travail et qui m'a permis de découvrir le monde passionnant des ARN. Par ses conseils continus, sa rigueur et son enthousiasme, elle a su me guider sans relâche jusque dans la rédaction de ce mémoire. Je lui suis très reconnaissante de la formation scientifique qu'elle m'a donnée, ainsi que de la possibilité qu'elle m'a offerte de présenter mon travail dans de nombreux congrès nationaux et internationaux.

Je suis très sensible à l'honneur que me font Madame Josette Banroques et Monsieur Richard Giegé, d'accepter d'être rapporteurs de ce mémoire. Je remercie aussi vivement Monsieur Henri Grosjean et Monsieur Bernard Decaris d'avoir accepté de juger cette thèse.

Je tiens à remercier spécialement Henri Grosjean pour sa collaboration qui a été essentielle au bon déroulement de mon travail, pour m'avoir permis de travailler à plusieurs reprises dans son laboratoire et pour les nombreux échanges scientifiques et amicaux que nous avons eus.

Je remercie également Yuri Motorin qui a beaucoup contribué à l'avancement de ce travail. Ses conseils, lors de la rédaction de ce manuscrit, m'ont également été très précieux.

Mes remerciements vont également à Florence et à Hubert qui ont rendu, par leur gentillesse, mes séjours à Gif-sur-Yvette très agréables.

Mes pensées vont tout particulièrement "aux filles du box 3" pour leur aide tout au long de ce travail, leur gaieté et la chaleureuse ambiance qu'elles ont su créer: Annie, Sandrine, Véronique et Françoise. Mes remerciements vont également à Isabelle, la petite dernière, qui a apporté un regard neuf sur le sujet et qui m'a beaucoup soutenue aussi bien scientifiquement que moralement.

Je remercie sincèrement Véronique S. et Nathalie G. pour leur bonne humeur, leur encouragement et leur précieux concours lors de la rédaction de ce manuscrit.

Je remercie également Agnès et Annie pour leurs conseils avisés sur la partie épissage et Nicolas Brosse pour son aide concernant la réaction chimique de l'aniline.

Mes remerciements vont également à tous les membres du laboratoire pour l'ambiance agréable de travail qu'ils ont su créer et plus particulièrement à Jacky qui m'a gentiment évité les allers et retours à la bibliothèque lors de la rédaction de ce manuscrit.

Je remercie tout spécialement Valérie (et Alain), Sandrine et Vincent pour leur amitié et leur aide. Leur soutien au laboratoire et nos moments d'évasion le temps d'un week-end ou d'une semaine m'ont été très précieux. Mes pensées vont particulièrement à Vincent pour le chemin que nous avons parcouru ensemble

10/10/2023 10:10:10
10/10/2023 10:10:10
10/10/2023 10:10:10
10/10/2023 10:10:10

depuis le DEA, pour sa compréhension et son soutien moral dans les moments difficiles.

Je tiens sincèrement à remercier ma sœur et sa petite famille pour leur présence à mes côtés et les moments de détente qu'ils m'ont apportés tout au long de ces années. Mes remerciements vont également à tous mes amis qui ont su me supporter sans relâche et particulièrement à Annabelle qui, bien qu'éloignée géographiquement, m'a apporté un soutien continu.

Enfin, mes pensées vont tout spécialement à mes parents que je remercie sincèrement d'avoir depuis toujours été à mes côtés, d'avoir toujours cru en moi et de m'avoir soutenue sans condition dans mes projets. Sans leur présence, ce travail n'aurait jamais vu le jour.

ABREVIATIONS ET ANGLICISME

ADN	: acide désoxyribonucléique
ADNc	: acide désoxyribonucléique complémentaire
ARN	: acide ribonucléique
ARNm	: acide ribonucléique messenger
ARNr	: acide ribonucléique ribosomique
ARNt	: acide ribonucléique de transfert
ATP	: adénosine 5' triphosphate
°C	: degré celsius
Ci	: curie
CMCT	: 1-cyclohexyl-3-(2-morpholinoéthyl) carbodiimide métha-p-toluène
cpm	: coups par minute
Da, KDa	: dalton, kilodalton
db	: double brin
ddNTP	: didésoxyribonucléotide triphosphate
DNase	: désoxyribonucléase
DMS	: diméthylsulfate
dNTP	: désoxyribonucléotide triphosphate
DTT	: dithiothréitol
EDTA	: éthylène diamine tétraacétate
h	: heure
intron U2-dépendant	: intron épissé par les spliceosomes U2-dépendants
intron U12-dépendant	: intron épissé par les spliceosomes U12-dépendants
Kb	: kilobase
min	: minute
nt	: nucléotide
pb	: paire de base
PEG	: polyéthylène glycol
pGAL	: promoteur galactose dépendant
pré-ARNm	: précurseur de l'ARN messenger
pré-ARNr	: précurseur de l'ARN ribosomique
pré-ARNt	: précurseur de l'ARN de transfert
qsp	: quantité suffisante pour
R	: purine
RNasine	: inhibiteur des ribonucléases
RT	: reverse transcriptase
S	: Svedberg
sb	: simple brin
structure 2D	: structure secondaire
structure 3D	: structure tertiaire
RNase	: ribonucléase
U	: unité enzymatique
Ψ	: pseudouridine
Y	: pyrimidine

Anglicisme: certains termes n'ont pas de traduction française, ou leur traduction est d'usage peu courant. Ces termes ont été gardés en anglais.

BMV	: brome mosaic virus
ETS	: external transcribed sequence

hnRNP : particule ribonucléoprotéique nucléaire hétérogène (Heterologous Nuclear Ribonucleoprotein Particules)

ORF : phase ouverte de lecture (Open Reading Frame)

PCR : amplification en chaîne par une polymérase (polymerase chain reaction)

snoRNA : petit ARN nucléolaire (Small nucleolar RNA)

snRNA : petit ARN nucléaire (Small nuclear RNA)

snoRNP : particule ribonucléoprotéique contenant un ou des petits ARN nucléolaires (Small nucleolar ribonucleoprotein particule)

snRNP : particule ribonucléoprotéique contenant un ou des petits ARN nucléaires (Small nuclear Ribonucleoprotein Particule)

spliceosome : complexe nucléaire ribonucléoprotéique dans lequel ont lieu les réactions d'épissage

spliceosome U2-dépendant : spliceosome nécessitant le snRNA U2 afin de catalyser la réaction d'épissage

spliceosome U12-dépendant: spliceosome nécessitant le snRNA U12 afin de catalyser la réaction d'épissage

TYMV : turnip yellow mosaic virus

Sommaire

Avant-propos

Introduction

Partie 1: L'épissage des introns des ARN pré-messagers nucléaires	3
A/ Généralités: l'épissage, plusieurs mécanismes pour une même fonction.....	4
B/ Les ARN impliqués dans l'épissage spliceosomal.....	6
1. Les ARN pré-messagers: introns et bordures exoniques.....	6
1. 1. La séquence en 5'.....	7
1. 2. Les boîtes de branchement.....	8
1. 3. Le site d'épissage en 3'.....	10
2. Les UsnRNA	11
2. 1. Un motif commun aux UsnRNA: le site Sm	13
2. 2. Les motifs spécifiques à chaque UsnRNA	13
2. 2. 1. Les snRNA U1 et U11	13
a) L'interaction entre le snRNA U1 ou U11 et le site 5' d'épissage	13
b) Les régions d'interaction des protéines spécifiques au snRNA U1	15
2. 2. 2. Les snRNA U2 et U12	16
a) L'interaction entre le snRNA U2 ou U12 et la boîte de branchement	16
b) La structure primaire et secondaire des snRNA U2 et U12	17
c) Les régions d'interaction des protéines spécifiques du snRNA U2	18
2. 2. 3. Le snRNA U5	19
a) Les différents variants du snRNA U5	19
b) L'interaction du snRNA U5 avec les deux exons	19
c) Interaction du snRNA U5 avec les protéines de la tri-snRNP [U4/U6.U5]	20
2. 2. 4. Les snRNA U4 et U6 et les snRNA U4atac et U6atac	21
a) Les séquences primaires et secondaires des snRNA U4, U4atac, U6 et U6atac	21
b) Les duplex d'ARN U4/U6 et U4atac/U6atac	21
c) L'interaction des snRNA U6 et U6atac avec le site 5' d'épissage	23
d) L'interaction du duplex d'ARN U4/U6 avec les protéines qui lui sont associées	24
C/ Les UsnRNP	24
1. Biogenèse des UsnRNP.....	25
2. Les protéines communes.....	26
2. 1. Les protéines Sm.....	26
2. 1. 1. Nature et conservation phylogénétique	26
2. 1. 2. Les protéines Sm sont essentielles pour la constitution des spliceosomes actifs	27
2. 1. 3. La formation des pré-particules	27
2. 1. 4. Reconnaissance entre protéines Sm et site Sm	28
2. 1. 5. Les protéines "Sm like"	29
2. 2. Le complexe SMN/SIP1	29
3. Maturation des extrémités 3' des UsnRNA et formation de la pré-particule	30
4. Les protéines spécifiques à chaque UsnRNP.....	31
4. 1. Les domaines fonctionnels des protéines spécifiques à chaque UsnRNP.....	31
4. 2. La snRNP U1.....	33
4. 3. La snRNP U2.....	34
4. 4. La snRNP U6.....	35
4. 5. La snRNP U4/U6.....	35

4. 6. La snRNP U5.....	35
4. 7. La tri-snRNP [U4/U6.U5]	37
5. Les protéines des UsnRNP des spliceosomes U12-dépendants.....	38
D/ Les complexes spliceosomaux et les interactions ARN-ARN mises en jeu lors de l'assemblage et de l'action du spliceosome.....	38
1. Le complexe H ou hnRNP	39
1. 1. La composition en protéines	39
1. 2. Reconnaissance des exons et des introns.....	40
2. Le complexe d'engagement.....	41
3. Les pré-spliceosomes (complexe A chez les mammifères ou B chez <i>S. cerevisiae</i>).....	43
4. Le complexe B (nomenclature vertébrés).....	44
4. 1. Les interactions U2/U6 et U12/U6atac	45
4. 2. Les transitions U4/U6 \Rightarrow U2/U6 et U4atac/U6atac \Rightarrow U2/U6atac.....	47
4. 3. Les transitions U1/site 5' \Rightarrow U6/site 5' et U11/site 5' \Rightarrow U6atac/site 5'.....	48
4. 4. La formation du premier complexe actif est un mécanisme concerté, faisant intervenir à la fois les UsnRNA et les protéines des UsnRNP	48
5. Le complexe C: le spliceosome (nomenclature vertébrés).....	49
6. Les interactions ARN-ARN au sein des complexes spliceosomaux actifs.....	52
7. Les régulations de l'épissage et rôles des protéines SR	53
E/ Le point de vue phylogénétique.....	55
1. Un ancêtre commun aux introns des ARN pré-messagers nucléaires et aux introns de groupe II?.....	55
2. Modèle proposé pour la génération de deux systèmes distincts d'épissage des ARN pré-messagers nucléaires	57

Partie 2: Les résidus pseudouridines portés par les ARN: localisation, biosynthèse et rôles.....

A/ Généralités.....	60
B/ Localisation des résidus Ψ et des autres nucléotides modifiés dans les ARN: un premier pas vers la connaissance de leur fonction.....	61
1. Localisation dans les ARNt.....	61
2. Localisation dans les ARNr.....	62
3. Localisation dans les UsnRNA	63
C/ Les Ψ-synthases	64
1. Identification des Ψ -synthases	64
2. Mécanisme d'action des Ψ -synthases.....	65
3. Spécificité de reconnaissance des Ψ -synthases.....	67
3. 1. La spécificité de reconnaissance des ARN par les Ψ -synthases ne nécessitant pas de guides de modification.....	68
3. 1. 1. Les différentes classes d'enzymes	68
3. 1. 2. Reconnaissance d'une séquence ou d'une structure?	69
a) Cas des ARNr: Ψ -synthases d' <i>E. coli</i>	69
b) Cas des ARNt et UsnRNA: Ψ -synthases procaryotes et eucaryotes. Comparaison avec les autres enzymes de modification des ARNt	70
3. 2. Les snoRNA guides nucléolaires des Ψ -synthases modifiant les ARNr eucaryotes.....	72
3. 2. 1. Les snoRNP à boîtes H/ACA	73
a) Structure et mode d'action des snoRNA H/ACA	73
b) Les protéines des snoRNP H/ACA	74
3. 2. 2. Les snoRNP à boîtes C/D	76
a) Structure et mode d'action des snoRNA C/D	76
b) Les protéines des snoRNP C/D	76
3. 2. 3. Y a-t-il des ARN guides pour la modification d'autres ARN que les ARNr eucaryotes	77

3. 2. 4. Certains snoRNAs portent eux-mêmes des modifications post-transcriptionnelles	77
D/ Relations entre maturation par des RNases et modification de nucléotides dans les ARN. Localisation cellulaire des enzymes de modification eucaryotes.....	78
1. Cas des ARNr.....	78
2. Cas des ARNt.....	79
2. 1. ARNt procaryotes et ARNt cytoplasmiques eucaryotes.....	79
2. 2. ARNt mitochondriaux eucaryotes.....	80
3. Cas des UsnRNA.....	80
E/ Rôles des résidus pseudouridines dans les ARN.....	81
1. Rôles des résidus Ψ dans les ARNt.....	82
1. 1 Rôles structuraux des résidus Ψ	82
1. 2 Rôles probables des résidus Ψ dans le bras anticodon.....	83
1. 2. 1. Les résidus Ψ en positions 32, 38 et 39	83
1. 2. 2. Pas de résidu Ψ localisé en position 37	84
1. 2. 3. Rôles des résidus Ψ localisés dans l'anticodon (positions 34, 35 et 36)	84
a) Rôles dans l'interaction ARNt/aminoacyl-ARNt synthase	84
b) Rôles dans l'interaction codon/anticodon	85
1. 3. Rôles probables du résidu Ψ en position 55.....	86
2. Rôles possibles des résidus Ψ dans les ARNr.....	87
2. 1. Rôles des résidus Ψ localisés dans l'ARN de la petite sous-unité ribosomique	87
2. 2. Rôles des résidus Ψ localisés dans le grand ARN de la grande sous-unité ribosomique	87
3. Rôles putatifs des résidus Ψ dans les UsnRNA.....	89

Revue: Posttranscriptional modifications in the U small nuclear RNAs.

S. Massenet, A. Mouglin and C. Branlant

In H. Grosjean and R. Benne (eds), 1998. The modification and editing of RNA, ASM press, Washington DC, p201-227.

Résultats/Discussion

Partie 1: Localisation des résidus Ψ dans la grande sous-unité ribosomique d'une *Archaea* hyperthermophile: *S. acidocaldarius*90

A/ Introduction et Résultats91

Publication n°1:

The first determination of pseudouridine residues in 23S ribosomal RNA from hyperthermophilic *Archaea Sulfolobus acidocaldarius*.

S. Massenet, I. Ansmant, Y. Motorin and C. Branlant.

FEBS letters, sous presse.

B/ Discussion et Perspectives.....94

1. Des snoRNAs guidés chez les *Archaea*?94

2. La méthylation sur le ribose, une modification post-transcriptionnelle sélectionnée par les organismes se développant à température élevée?.....95

Partie 2: Localisation des résidus Ψ dans les UsnRNA de la levure *S. cerevisiae* et dans les UsnRNA des spliceosomes U12-dépendants de cellules HeLa. Identification chez *S. cerevisiae* d'une pseudouridine-synthase présentant une double spécificité de substrats.....97

A/ Introduction et Résultats98
1. Identification des résidus Ψ portés par les UsnRNA de *S. cerevisiae* et de la première Ψ-synthase agissant sur les UsnRNA98
1. 1. Localisation des résidus Ψ dans les UsnRNA de *S. cerevisiae*98
1. 2. Pus1p: une Ψ-synthase à double spécificité de substrats.....100

Publication n°2:

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.

Massenet S., Motorin Y., Lafontaine D.L.J., Hurt E.C., Grosjean H. and Branlant C.
Mol. Cell. Biol. 19, 2142-2154, 1999.

1. 3. Recherche des résidus 2'-O-méthylés portés par le snRNA U5 de *S. cerevisiae*.....103
1. 3. 1. intérêt de l'étude et problèmes techniques rencontrés 103
1. 3. 2. La boucle terminale I du snRNA U5 de *S. cerevisiae* ne contient vraisemblablement pas de résidus 2'-O-méthylés. Le résidu Ψ₉₉ est présent à la fois dans le snRNA U5S et le snRNA U5L 104
a) Développement de la méthode d'analyse 104
b) Examen des résultats obtenus 104
2. Recherche de résidus Ψ dans le snoRNA U3 de *S. cerevisiae*106
3. Identification des résidus Ψ dans les UsnRNA des spliceosomes U12-dépendants de cellules HeLa107

Publication n°3:

A limited number of pseudouridine residues in the human atax splicing UsnRNAs as compared to human major spliceosomal UsnRNAs.

Massenet S. and C. Branlant.
RNA, sous presse.

B/ Discussion et Perspectives109
1. Certains résidus Ψ sont très conservés dans les UsnRNA des spliceosomes U2-dépendants et sont présents dans les UsnRNA des spliceosomes U12-dépendants109
2. A quelles étapes de la biogenèse des UsnRNP sont formées les modifications post-transcriptionnelles des UsnRNA?112

Partie 3: Etude de la reconnaissance de l'ARN par la Ψ synthase Pus1p de *S. cerevisiae*. Comparaison des enzymes Pus1p de levure et de souris.....117

A/ Recherche d'un motif ARN reconnu par l'enzyme Pus1p de *S. cerevisiae*.....118
1. Présentation du problème.....118

1. 1. L'enzyme Pus1p de <i>S. cerevisiae</i> est responsable de la formation de plusieurs résidus Ψ dans les ARNt	118
1. 2. La formation des résidus Ψ par Pus1p dans les ARNt ne dépend pas de la structure 3D des ARNt. Pus1p peut agir sur des petits fragments d'ARNt	119
1. 3. Pus1p peut agir <i>in vitro</i> sur des ARN présentant un repliement de type ARNt.....	119
1. 4. L'enzyme Pus1p reconnaît les ARNt de manière large.....	120
1. 5. Quel est le requis en structure et en séquence pour, qu'après s'être lié à Pus1p, un ARN soit modifié par cette Ψ-synthase?.....	120

Publication n°4:

Recognition of RNA substrates by yeast pseudouridine synthase I (Pus1p) *in vitro*.
 Motorin Y., Massenet S., Becker H.F., Raué A., Branlant C. and Grosjean H.
 Soumis à RNA.

2. Discussion	123
2. 1. Pus1p modifie <i>in vitro</i> des ARN qu'elle ne modifie pas <i>in vivo</i>	123
2. 2. Motif de reconnaissance de Pus1p sur l'ARN	124
2. 2. 1. La structure 3D des ARNt n'est pas nécessaire pour la reconnaissance par Pus1p	124
2. 2. 2. Un motif structural spécifique est reconnu par Pus1p	125
B/ Recherche des motifs du snRNA U2 reconnus par la Ψ synthase Pus1p de <i>S. cerevisiae</i>.	126
1. Introduction.....	126
2. Méthodes	127
2. 1. Etude de la structure secondaire de l'extrémité 3'-terminale d'un transcrit, produit <i>in vitro</i> , correspondant au snRNA U2 de <i>S. cerevisiae</i> ,.....	127
2. 2. Production des variants du snRNA U2 de <i>S. cerevisiae</i>	127
2. 2. 1. Délétion des différentes tige-boucles du snRNA U2 par la méthode PCR	127
2. 2. 2. Génération de mutations ponctuelles dans le snRNA U2 de <i>S. cerevisiae</i>	127
2. 2. 3. Modification <i>in vitro</i> des variants obtenus par Pus1p et analyse des résidus Ψ formés	128
3. Résultats.....	128
3. 1. Analyse de la structure du snRNA U2 de <i>S. cerevisiae</i> , produit <i>in vitro</i> , dans les conditions où les modifications par Pus1p sont réalisées.....	128
3. 2. Les 139 nts en 5' du snRNA U2 de <i>S. cerevisiae</i> sont substrats de Pus1p	129
3. 3. Influence des structures tige-boucles localisées à proximité de la position 44 sur la reconnaissance et la catalyse par Pus1p	130
3. 4. Influence de l'identité des nucléotides autour de la position 44 sur la reconnaissance et la catalyse par Pus1p.....	130
4. Discussion et Perspectives	132
C/ Etude comparative des Ψ synthases Pus1p de <i>S. cerevisiae</i> et de souris	135
1. Introduction.....	135
2. Méthodes	136
2. 1. Identification des résidus Ψ portés par le snRNA U2 de cellules HeLa	136
2. 2. Production des snRNA U2 humain et de <i>S. cerevisiae in vitro</i>	136
2. 3. Modification des snRNA U2 humain et de <i>S. cerevisiae</i> produits par transcription <i>in vitro</i> , par les Ψ-synthases mPus1p et ScPus1p purifiées	137
2. 4. Formation <i>in vitro</i> de particules ribonucléoprotéiques contenant le snRNA U2 humain et modification du snRNA U2 au sein de ces particules, par les Ψ-synthases mPus1p et ScPus1p purifiées.....	138
2. 5. Etude de la structure du snRNA U2 humain produit <i>in vitro</i>	138
3. Résultats.....	139

3. 1. <i>In vitro</i> , la Ψ -synthase mPus1p catalyse la formation du résidu Ψ_{44} dans le snRNA U2 de <i>S. cerevisiae</i> produit <i>in vitro</i>	139
3. 2. Contrairement à la Ψ -synthase ScPus1p, la Ψ synthase mPus1p ne catalyse pas <i>in vitro</i> la formation de résidus Ψ dans un snRNA U2 humain produit <i>in vitro</i>	139
3. 2. 1. Localisation des résidus Ψ dans le snRNA U2 de cellules HeLa	140
3. 2. 2. Analyse des résidus Ψ formés <i>in vitro</i> par les enzymes ScPus1p et mPus1p dans un snRNA U2 humain, produit <i>in vitro</i> , libre ou associé aux protéines Sm	141
4. Discussion et Perspectives	143
4. 1. La Ψ -synthase scPus1p purifiée modifie le snRNA U2 humain produit par transcription <i>in vitro</i>	143
4. 2. La présence des protéines Sm limite le taux de modification, par la Ψ -synthase ScPus1p, du snRNA U2 humain produit <i>in vitro</i>	144
4. 3. La Ψ -synthase mPus1p modifie le snRNA U2 de <i>S. cerevisiae</i> produit <i>in vitro</i>	145
4. 4. Le snRNA U2 humain renferme 14 résidus Ψ	146

Partie 4: Recherche des substrats de quatre Ψ synthases potentielles de *S. cerevisiae*

A/ Introduction et Résultats	148
1. Approche utilisée pour abolir l'activité des 4 gènes codant des Ψ -synthases potentielles	148
2. Caractérisation des souches sélectionnées après mutation.....	149
3. Analyse des résidus Ψ présents dans les UsnRNA des souches portant les délétions.....	151
4. Analyse des résidus Ψ présents aux positions 13, 32 et 35 des ARNt cytoplasmiques, des souches portant une délétion dans l'une des 4 ORF étudiées	152
5. Recherche de la présence du résidu Ψ localisé en position 2819 dans l'ARNr 21S mitochondrial, dans les différentes souches mutantes.....	152

Publication n°5:

The YLR165c ORF is implicated in formation of the Ψ residue at position 2819 in the *Saccharomyces cerevisiae* 21S mitochondrial large ribosomal RNA.

Ansmant I., Massenet S., Motorin Y., Grosjean H. and Branlant C.

Soumis à Nucleic. Acids Res.

B/ Discussion et Perspectives	154
1. Etudes envisagées sur la Ψ -synthase Pus5p.....	154
2. La formation du résidu Ψ_{2819} n'est pas essentielle, mais elle pourrait apporter un avantage sélectif.....	154
3. Comment expliquer la présence d'un faible nombre de Ψ -synthases caractérisées ou potentielles chez <i>S. cerevisiae</i> ?.....	155
4. D'autres substrats ARN possibles pour les différentes Ψ -synthases déjà caractérisées	157

Conclusion

Bibliographie

Avant-propos

La série d'événements qui conduit à un ARN mature et fonctionnel, à partir du transcrit primaire, est appelée la maturation. Il s'agit en fait souvent d'une cascade de processus cellulaires mettant en jeu un grand nombre de molécules ou complexes de macromolécules à activité enzymatique et de cofacteurs.

Deux types de maturation peuvent être distingués:

⇒ Le premier type fait intervenir la formation et/ou la cassure de liaisons phosphodiester dans l'ARN. On peut regrouper, dans ce type de maturation, les processus d'addition ou de délétion de nucléotides internes, l'épissage, les polyadénylations et polyuridylation, l'addition d'une structure coiffe en 5'...

⇒ Le deuxième type de maturation fait intervenir des modifications chimiques des bases et/ou des riboses sans modifier l'intégrité de la molécule d'ARN.

Une des thématiques abordée dans l'équipe du Dr C. Branlant concerne une des étapes de maturation des ARN pré-messagers eucaryotes, l'épissage, qui est réalisé dans une macrostructure nucléoprotéique, appelée spliceosome. Un des points les plus étudiés, dans le laboratoire, est l'analyse des petits ARN nucléaires ou UsnRNA intervenant dans ce processus. Ces petits ARN subissent eux-mêmes des étapes de maturation, dont l'une consiste en la formation de nucléotides modifiés. Peu de choses étaient connues sur ce sujet, quand j'ai débuté mon travail. En effet, les données disponibles sur la localisation des modifications post-transcriptionnelles portées par les UsnRNA provenaient du séquençage de ces ARN, il y a une vingtaine d'années, notamment par les équipes de R. Reddy et de C. Branlant, et elles portaient essentiellement sur des UsnRNA d'eucaryotes multi-cellulaires. De même, aucune enzyme de modification, spécifique des UsnRNA, n'était connue.

Dans ce contexte, j'ai relancé les études sur ce sujet au laboratoire et mon travail de thèse a débuté avec, comme premier objectif, une comparaison entre espèces de la localisation des résidus pseudouridine ou Ψ dans les UsnRNA. J'ai d'abord entrepris l'identification des résidus Ψ portés par les UsnRNA de *S. cerevisiae*. Les connaissances apportées par les approches génétiques sur le spliceosome de cet organisme étaient susceptibles de contribuer à la définition du rôle joué par les nucléotides modifiés dans le processus d'épissage. Nous partions de l'hypothèse que les modifications importantes étaient conservées au cours de l'évolution. Pendant que je réalisais ce travail, un second type de spliceosomes, coexistant avec les spliceosomes majeurs déjà bien étudiés, a été détecté et caractérisé par plusieurs équipes, chez les métazoaires. Ce deuxième type de spliceosomes contenait des UsnRNA qui lui étaient spécifiques. J'ai alors entrepris d'identifier les résidus Ψ portés par ces UsnRNA, afin de comparer la distribution de ce nucléotide modifié dans les deux types de UsnRNA.

Le choix de la levure *S. cerevisiae* pour nos premiers travaux nous a permis d'aborder un deuxième objectif, à savoir la caractérisation des Ψ -synthases, responsables de la formation des résidus Ψ dans les UsnRNA. Quelques Ψ -synthases de levure avaient déjà été caractérisées par l'équipe d'H. Grosjean, j'ai étudié leur action sur les UsnRNA et une de ces enzymes s'est avérée agir sur l'un deux. Nous avons également débuté la recherche et la caractérisation de nouvelles Ψ -synthases de *S. cerevisiae* et nous avons étudié leur action sur différents types d'ARN.

Les résultats obtenus sont présentés sous forme de publications déjà publiées, soumises ou à soumettre. Chaque publication est précédée d'une introduction, présentant l'intérêt de l'étude effectuée, et est suivie d'une discussion des résultats obtenus avec la présentation des perspectives. Une partie des résultats, pour laquelle des expériences complémentaires sont encore nécessaires afin de rédiger un article, est présentée en français. Pour cette partie, les méthodes utilisées sont également décrites. Pour les autres parties, elles sont résumées dans les publications.

La description des résultats est précédée d'une partie bibliographique. La connaissance du mécanisme d'épissage et du rôle des différents UsnRNA, dans ce mécanisme sont indispensables pour essayer de comprendre quel peut être le rôle des nucléotides modifiés portés par ces UsnRNA. La première partie de l'introduction est donc consacrée à la présentation des connaissances actuelles sur le mécanisme de base de l'épissage des ARN pré-messagers nucléaires. La nature des UsnRNA, la mise en place des interactions ARN/ARN, ainsi que les interactions ARN/protéines mises en jeu sont présentées dans cette première partie. La connaissance actuelle des différents rôles joués par les nucléotides modifiés dans les ARNt, ARN pour lesquels cette étude est la plus avancée, est susceptible d'apporter des informations sur les rôles possibles de ces mêmes nucléotides modifiés dans les UsnRNA. C'est ce que nous présenterons dans une deuxième partie de l'introduction. Nous présenterons également dans cette deuxième partie les connaissances actuelles sur la nature et la spécificité des différentes Ψ -synthases modifiant les ARNr et les ARNt eucaryotes et procaryotes et les connaissances limitées sur la formation des résidus Ψ dans les UsnRNA. Toutes les données disponibles sur l'identité, la localisation et la synthèse des nucléotides modifiés portés par les UsnRNA sont présentées sous la forme d'une revue que nous avons rédigée pour le livre "Modification and Editing of RNA" publié par l'American Society of Microbiology. Dans cette revue, nous faisons, à partir de l'ensemble des données disponibles, des propositions sur les rôles potentiels des nucléotides modifiés portés par les UsnRNA.

S.C.D. - U.N.P. NANCY 1
BIBLIOTHÈQUE DES SCIENCES
Rue du Jardin Botanique
54600 VILLERS-LES-NANCY

Introduction

*PARTIE 1: L'ÉPISSAGE DES INTRONS DES ARN PRÉ-MESSAGERS
NUCLEAIRES*

A/ GENERALITES: L'ÉPISSAGE, PLUSIEURS MÉCANISMES POUR UNE MÊME FONCTION

L'organisation morcelée des génomes existe à des degrés divers dans quasiment tous les organismes.

Les gènes nucléaires de protéines, chez les eucaryotes, peuvent contenir jusqu'à une cinquantaine d'introns, dont la taille peut varier d'une dizaine à une dizaine de milliers de nucléotides. Les génomes d'organismes hautement différenciés renferment beaucoup plus d'introns que ceux d'eucaryotes unicellulaires. En effet, chez *S. cerevisiae*, moins de 5% des gènes contiennent un intron (pour revue, Rymond et Rosbash, 1992) alors que, chez les mammifères, un gène ne contenant pas d'intron est une exception. Des introns ont été mis en évidence dans presque toutes les catégories de gènes eucaryotes. En effet, en plus des introns de gènes de protéines, il existe des introns dans les gènes d'ARNt (Goodman *et al.*, 1977; Valenzuela *et al.*, 1978), d'ARNr nucléaires (Wild et Gall, 1979), d'ARNr mitochondriaux (Bos *et al.*, 1978) et chloroplastiques (Rochaix et Malnoe, 1978), ainsi que dans les gènes de certains UsnRNA: les snRNA U6 (Tani et Ohshima, 1989), U3 (Myslinski *et al.*, 1990) et U2 (Takahashi *et al.*, 1993) d'organismes unicellulaires.

Chez les procaryotes, la présence d'introns est rare: quelques gènes d'*Archaea* (Kaine *et al.*, 1983; Daniels *et al.*, 1985) et de bactéries (Kuhnel *et al.*, 1990; Xu *et al.*, 1990a; Ferat *et al.*, 1994) possèdent des introns. Des introns ont également été découverts dans des bactériophages (Chu *et al.*, 1986; Ehrenman *et al.*, 1986; Goodrich-Blair *et al.*, 1990)

On peut classer les introns en trois grandes catégories en fonction de leur mécanisme d'excision:

⇒ Les introns éliminés par voie enzymatique classique. Il s'agit essentiellement d'introns de gènes codant des ARN de transfert (Peebles *et al.*, 1983).

⇒ Les introns auto-épissables, introns de groupes I et II. Les introns de groupe I sont très répandus, ils sont trouvés dans les gènes d'ARNm, d'ARNt et d'ARNr d'organites eucaryotes, dans les génomes nucléaires d'eucaryotes (mais pas de vertébrés) et dans les génomes de bactériophages et de bactéries (Michel et Westhof, 1990; Reinhold-Hurek et Shub, 1992). Les introns de groupe II sont trouvés fréquemment dans les génomes des organelles chez les eucaryotes (pour revue, Michel *et al.*, 1989), mais également dans les cyanobactéries et les protéobactéries (Ferat et Michel, 1993).

Dans les deux cas (introns de groupes I ou II), le mécanisme d'excision repose sur deux étapes de transestérification (Fig. 1). Pour les introns de groupe I, le premier clivage au site 5' d'épissage implique une molécule extérieure, une guanosine, qui se lie à un site généré par le repliement de la séquence intronique (Cech, 1985). C'est l'hydroxyle en 3' de cette guanosine

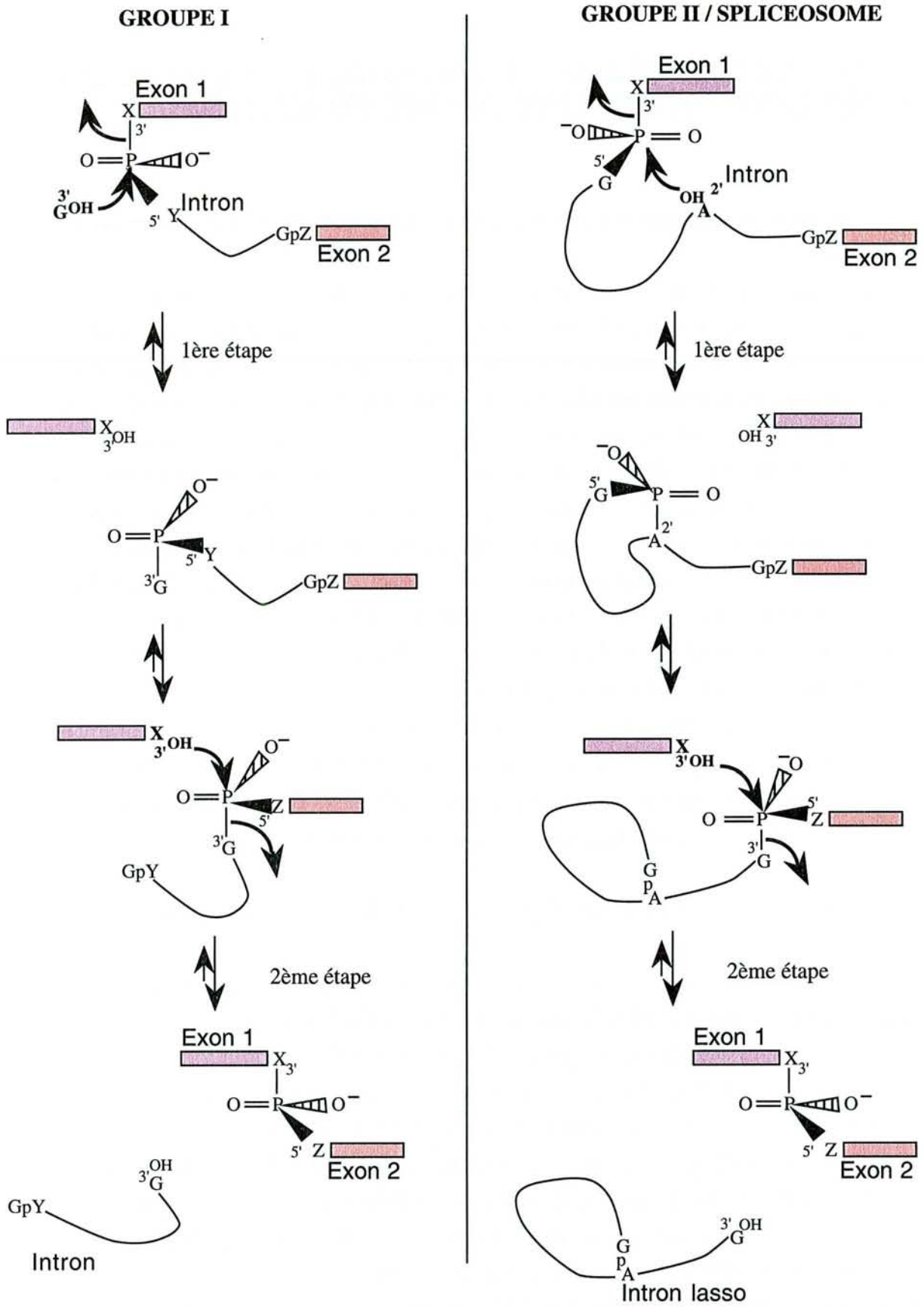


Figure 1: Les étapes de la réaction d'épissage des introns de groupe I (à gauche), des introns de groupe II et des ARN pré-messagers nucléaires (à droite).
 D'après le schéma de la revue Moore *et al.*, 1993.

qui réalise la première attaque nucléophile. Dans le cas des introns de groupe II, la première attaque nucléophile, conduisant au clivage au site 5' d'épissage, est réalisée par l'hydroxyle en 2' d'une adénosine interne de l'intron, activée par la structure formée par l'intron. Il en résulte un intermédiaire en lasso (Peebles *et al.*, 1986; van der Veen *et al.*, 1986). La deuxième étape de la réaction est similaire pour les 2 types d'introns. Les introns de groupe I, comme les introns de groupe II, possèdent donc toutes les informations nécessaires à leur excision.

Ces introns sont caractérisés par de courtes séquences conservées, notamment au niveau des sites 3' et 5' d'épissage, et par la capacité de l'ARN, codé par ces introns, à se replier selon une structure secondaire et tertiaire complexe caractéristique de chaque groupe. Cette structure est nécessaire à la création du site catalytique (Michel *et al.*, 1982; Michel et Jacquier, 1987; Burke *et al.*, 1987; Michel et Westhof, 1990; Cate *et al.*, 1996). Cependant beaucoup de ces introns nécessitent la présence de protéines pour augmenter leur efficacité catalytique *in vivo* (pour revue, Michel et Ferat, 1995).

⇒ Les introns contenus dans les ARN pré-messagers nucléaires. L'épissage des introns des ARN pré-messagers implique les 2 mêmes étapes de transestérification que l'épissage des introns de groupe II (Fig. 1), à savoir une attaque nucléophile du phosphate en 5' de l'intron par l'hydroxyle en 2' d'un résidu adénosine de l'intron situé en amont du site d'épissage en 3'. Cette première étape libère le premier exon et un intermédiaire en forme de lasso. Elle est suivie d'une seconde attaque nucléophile du phosphate en 3' de l'intron par l'hydroxyle en 3' du premier exon. Une liaison phosphodiester est ainsi générée entre les 2 exons et l'intron est libéré (Ruskin *et al.*, 1984; Padgett *et al.*, 1984; Konarska *et al.*, 1985).

Ces deux attaques nucléophiles sont catalysées par un large complexe ribonucléoprotéique, le spliceosome (Grabowski *et al.*, 1985; Brody et Abelson, 1985), qui s'assemble autour de l'intron et des bordures exoniques. Le spliceosome renferme cinq petits ARN nucléaires (snRNA U1, U2, U4, U5 et U6), qui sont des composants essentiels de cette machinerie (Grabowski et Sharp, 1986), et un très grand nombre de protéines. Les UsnRNA sont assemblés sous la forme de particules ribonucléoprotéiques.

Depuis longtemps, il était considéré que tous les introns des ARN pré-messagers eucaryotes étaient épissés par un type unique de spliceosome. Ce n'est que très récemment, qu'une deuxième machinerie d'épissage nécessaire à l'épissage d'introns très peu fréquents, a été mise en évidence (pour revue, Wu et Krainer, 1999). Ce deuxième type de spliceosome renferme les snRNA U11, U12, U4atac, U5 et U6atac et il est dénommé, comme les introns qu'il épisse, par les termes de "mineur" (en référence à sa faible abondance dans les cellules), ou ATAC (en référence aux bordures introniques de certains introns) ou encore "U12-dépendant". Pour la suite de cette introduction, nous emploierons la nomenclature "U12-dépendant", par opposition aux autres spliceosomes que nous appellerons "U2-dépendants".

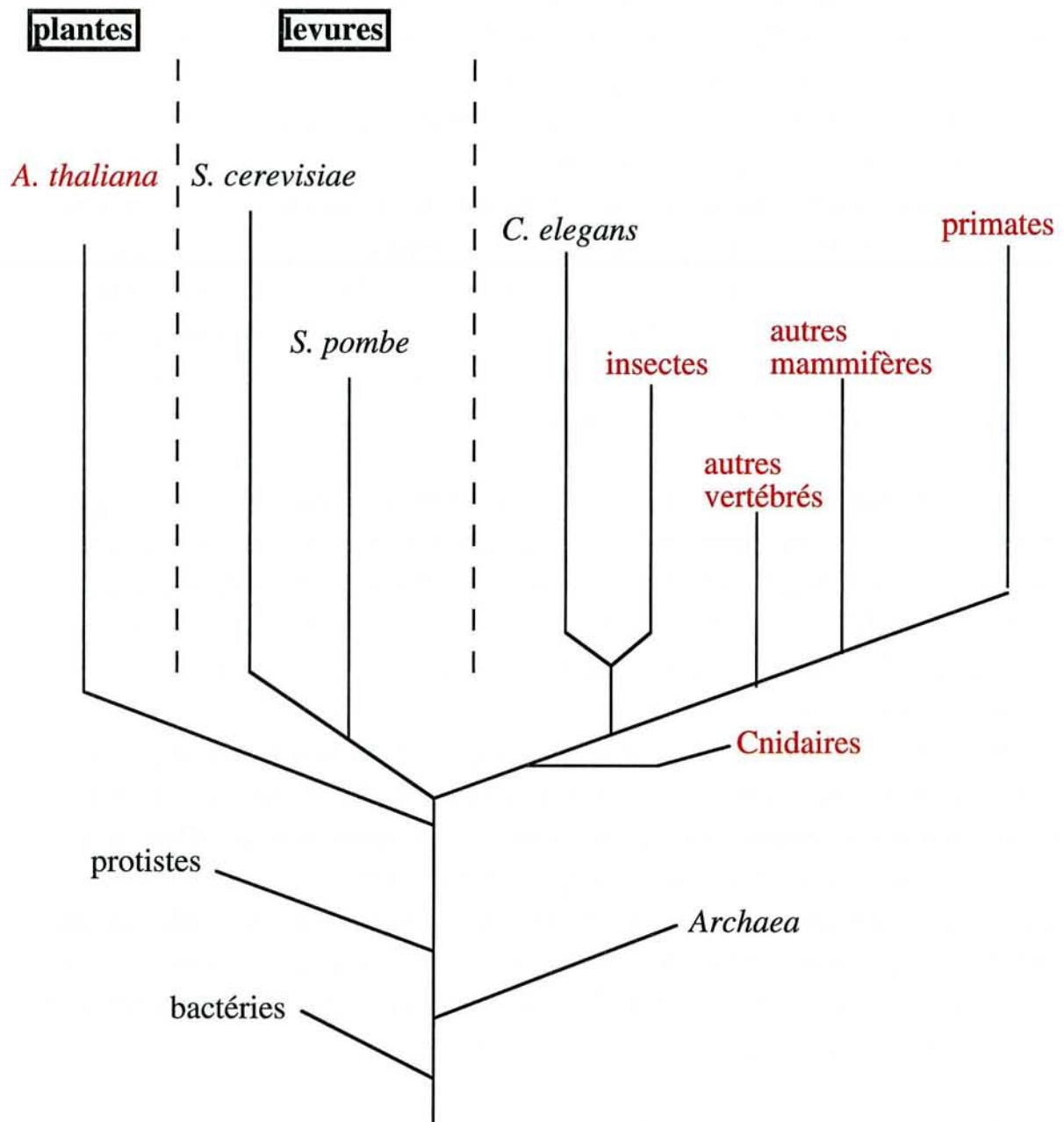


Figure 2: Espèces dans lesquelles des introns épissés par des spliceosomes U12-dépendants ont été détectés.

D'après une figure de Burge *et al.*, 1998.

Les espèces ou groupes d'espèces dans lesquels des introns U12-dépendants ont été mis en évidence sont indiqués en rouge.

Des introns éliminés par les spliceosomes U12-dépendants ont été trouvés chez les plantes, dans de multiples classes de vertébrés (les poissons, les amphibiens, les oiseaux et les mammifères) et chez les insectes (Fig. 2) (Burge *et al.*, 1998). D'après la connaissance de la séquence complète du génome de *S. cerevisiae*, aucun intron non conventionnel ne semble présent chez cette levure. Bien que la séquence complète des génomes de *C. elegans* et *S. pombe* ne soit pas encore disponible, ces espèces ne semblent pas non plus contenir d'introns inhabituels dans leurs gènes nucléaires.

Le premier chapitre de cette introduction sera consacré au mécanisme de base de l'épissage des ARN pré-messagers nucléaires. L'épissage par les deux types de spliceosomes sera abordé. Nous ferons d'abord une présentation des séquences introniques et exoniques, puis des UsnRNA et UsnRNP mis en jeu dans le processus. Ensuite, nous verrons comment ces UsnRNP s'assemblent autour de l'ARN pré-messager pour former un spliceosome et quelles sont les protéines qui interviennent dans ce processus. Enfin, nous présenterons l'épissage des ARN pré-messagers nucléaires sous un angle évolutif.

B/ LES ARN IMPLIQUES DANS L'ÉPISSAGE SPLICEOSOMAL

1. LES ARN PRE-MESSAGERS: INTRONS ET BORDURES EXONIQUES

Très tôt, l'alignement des séquences introniques a permis de mettre en évidence la conservation des séquences aux bordures exon/intron en 5' et en 3', ainsi que d'une séquence interne de l'intron, la boîte de branchement. Les nucléotides les plus conservés sont les dinucléotides GT et AG, localisés respectivement aux extrémités 5' et 3' de l'intron (Mount, 1982) (Fig. 3). L'idée que tous les introns nucléaires de protéines répondaient au même consensus a été acceptée pendant 15 ans. Plus récemment, l'explosion des données de séquence sur les génomes et les ADNc a conduit à reconsidérer ce postulat. La situation s'est avérée être beaucoup plus complexe qu'on ne l'imaginait. Dans un premier temps, une nouvelle classe d'introns a été observée. Ceux-ci ont été repérés parce qu'ils présentaient des dinucléotides AT-AC à leurs extrémités (Jackson, 1991; Hall et Padgett, 1994). Ces introns présentaient en plus des séquences différentes en 5' et au niveau de la boîte de branchement. Ils ont été appelés introns AT-AC de classe I (Fig. 3). Ils sont présents dans les génomes de métazoaires à une fréquence d'un intron AT-AC I pour 5000 à 10000 introns (Sharp et Burge, 1997). Très récemment, on s'est aperçu que quelques introns épissés par les spliceosomes U2-

Introns épissés par les spliceosomes U2-dépendants d'eucaryotes supérieurs

site 5' d'épissage	boîte de branchement	séquence polypyrimidine	site 3' d'épissage	nombre d'introns identifiés
AG GT RAGT	YNYTRA*Y	(Y) _n	Y AG GT	la majorité des introns
- introns AT.AC de classe II				
AG A TRAGT	YNYTRA*Y	(Y) _n	Y AC GT	7
- introns à bordures GC.AG et AT.AA				
AG G CRAGT	YNYTRA*Y	(Y) _n	Y AG GT	45
AG A TRAGT	YNYTRA*Y	(Y) _n	Y AA GT	1

Introns épissés par les spliceosomes U12-dépendants d'eucaryotes supérieurs

- introns AT.AC de classe I				
A TATCCTY	TCCTTRA*Y		Y AC	20
- introns à bordure GT.AG, AT.AA et AT.AG				
G TATCCTY	TCCTTRA*Y		Y AG	104
A TATCCTY	TCCTTRA*Y		Y AA	2
A TATCCTY	TCCTTRA*Y		Y AG	3

Introns de *S. cerevisiae*

AG G TATGT	TACTAA*C		C AG G	tous les introns
-------------------	----------	--	---------------	------------------

Figure 3: Séquences introniques essentielles à la réaction d'épissage des introns éliminés par les spliceosomes U2- et U12-dépendants des eucaryotes supérieurs. Comparaison avec les séquences fonctionnelles des introns de *S. cerevisiae*.

D'après Jackson, 1991; Hall et Padgett, 1994; Sharpe et Burge, 1997; Wu et Krainer, 1997; Dietrich *et al.*, 1997; Burge *et al.*, 1998 et la revue, Wu et Krainer, 1999.

Les dinucléotides GT et AG sont indiqués en vert, les nucléotides terminaux de nature différente par rapport au consensus sont indiqués en rouge.

L'astérisque désigne le point de branchement.

Le nombre d'introns de chaque type, identifiés à ce jour, est indiqué.

R, Y et N désignent, respectivement, une purine, une pyrimidine ou tous les nucléotides.

dépendants présentait également les dinucléotides AT-AC à leurs bordures. Ils ont été appelés introns AT-AC de classe II (Fig. 3). De plus, des résultats récents montrent que de nombreux introns épissés par les spliceosomes U12-dépendants portent en fait les dinucléotides GT-AG à leurs extrémités (Dietrich *et al.*, 1997; Sharp et Burge, 1997; Wu et Krainer, 1997; Burge *et al.*, 1998). Les introns épissés par les spliceosomes U12-dépendants, que nous appellerons introns U12-dépendants, sont souvent trouvés dans des familles de gènes bien définies: gènes de protéines impliquées dans la réparation de l'ADN, gènes de facteurs de transcription, gènes codant des protéines, qui sont des constituants de canaux ioniques calcium et sodium dépendants, et gènes de protéines de la matrice du cartilage (Tableau 1) (Dietrich *et al.*, 1997; Sharp et Burge, 1997; Burge *et al.*, 1998; pour revue, Wu et Krainer, 1999). Ces introns sont également souvent conservés phylogénétiquement (Tableau 1) (pour revue, Wu et Krainer, 1999). Peu de gènes contiennent des introns U12-dépendants. En général, dans un gène donné, un intron U12-dépendant est encadré par des introns U2-dépendants. Seuls quelques gènes contiennent deux introns U12-dépendants (Tableau 1) et certaines familles de gènes contiennent à la fois des introns U12-dépendants et des introns U2-dépendants aux bordures non canoniques (pour revue, Wu et Krainer, 1999).

Il s'avère donc que l'identité des dinucléotides terminaux des différents introns varie et qu'elle ne définit pas le type de spliceosome qui catalyse l'élimination de l'intron. Les déterminants de ce choix reposent essentiellement sur les capacités d'interaction de l'intron avec les UsnRNA.

Dans un premier temps, nous allons décrire les séquences conservées des introns reconnues par ces UsnRNA.

1.1. La séquence en 5'

Chez les eucaryotes supérieurs, la conservation de séquence au site 5' des introns majoritaires épissés par les spliceosomes U2-dépendants s'étend sur 8 nucléotides (deux en amont du site de clivage et six en aval): AG↓GT(A/G)AGT (pour revue, Green, 1986), mais des variations importantes par rapport à la séquence consensus sont tolérées, avec néanmoins comme déjà mentionné, une très forte conservation du dinucléotide GT. Une interaction non Watson et Crick, nécessaire pour la deuxième étape de la réaction d'épissage, a été mise en évidence entre les résidus G aux deux extrémités des introns (Parker et Siliciano, 1993; Chanfreau *et al.*, 1994; Deirdre *et al.*, 1995). Des expériences de mutagenèse, appuyées sur les prédictions d'interactions possibles entre les nucléotides, montrent qu'une interaction très similaire peut être formée entre un résidu A et un résidu C, les bases et les riboses étant placés de manière très similaire dans l'interaction G/G et dans l'interaction A/C (Parker et Siliciano, 1993; Chanfreau *et al.*, 1994). Ceci doit expliquer le fait que des introns portant les dinucléotides AT et AC à leurs extrémités (introns AT-AC de classe II), peuvent être épissés

Gène	Organisme	Numéro de l'intron	Nombre total d'intron	Dinucléotides terminaux	Fonction de la protéine	Gènes homologues dans lesquels un intron U12-dépendant est retrouvé	Dinucléotides terminaux des introns U12-dépendants homologues
CACNA1F	homme	2	47	GC-AG	Sous-unité α de canaux Ca^{2+}	Gènes humains CACNL1A1 CACNL1 A2 CACNL1 A3 CACNL1 A4 Gène CACNL de poisson (<i>fugu</i>)	AT-AG GT-AG GT-AG AT-AC GT-AG
CACNA1F	homme	16	47	GT-AG	Sous-unité α de canaux Ca^{2+}	Gènes humains CACNL1A1 CACNL1 A2 CACNL1 A3 CACNL1 A4 Gène CACNL de poisson (<i>fugu</i>)	GT-AG
CMP	poulet	7	7	AT-AC	Protéine de la matrice du cartilage	Homme, souris	AT-AC
DMC1	<i>Arabidopsis</i>	9	14	GT-AG	Recombinaison (Homologue RecA)	/	/
DMC1	<i>Arabidopsis</i>	14	14	AT-AC	Recombinaison (Homologue RecA)	/	/
MSH3	homme	6	23	AT-AA	Réparation de l'ADN (homologue MutS)	Gène REP3 de souris	AT-AC
RPB5	homme	5	6	GT-AG	RNA polymérase II (protéine de 23kDa)	/	/
RPL1a	<i>X. laevis</i>	3	9	GT-AG	Protéine ribosomique L1a	Gènes RPL1b de <i>X. laevis</i> , RPL1t de <i>X. tropicalis</i> , RPL1h humain	GT-AG

SCN4A	homme	2	23	AT-AC	Sous-unité α de canaux Na^+	Gènes humains SCN5A et SCN8A SCN8A de souris, SCN1 de <i>P. penicillatus</i>	AT-AC AT-AC GT-AG
SmE	homme	1	4	GT-AG	Protéine E des snRNP	/	/
SPS	souris	6	10	AT-AC	Spermine synthase	SPS humain	AT-AC
SPS	souris	10	10	GT-AG	Spermine synthase	SPS humain	GT-AG
XPG	souris	1	14	GT-AG	Pigment du Xéroderme	XPG humain	GT-AG
XPG	souris	13	14	AT-AC	Pigment du Xéroderme	/	/
E2F1	homme	4	6	AT-AC	Facteur de transcription	/	/
ERK2	souris	2	8	GT-AG	Protéine kinase	/	/
MYH9	homme	6	39	GT-AG	Myosine de type A non musculaire		

Tableau 1: Exemples de gènes contenant des introns épissés par les spliceosomes U12-dépendants.

Le nom du gène, contenant un ou des introns épissés par les spliceosomes U12-dépendants et l'organisme qui le porte sont indiqués, ainsi que la position de l'intron par rapport à l'ensemble des introns du gène et le nombre total d'introns dans le gène. Les gènes homologues, dans d'autres organismes, contenant au même endroit un intron épissé par les spliceosomes U12-dépendants, sont indiqués. La nature des dinucléotides terminaux des introns est précisée. La fonction de la protéine codée par le gène est indiquée.

D'après Burge *et al.*, 1998.

par les spliceosomes U2-dépendants. Il en est de même pour l'épissage, par les spliceosomes U12-dépendants, des introns AT-AC de classe I.

Chez les levures, et en particulier chez *S. cerevisiae*, la conservation de la séquence en 5' de l'intron est beaucoup plus stricte (GTATGT) que chez les vertébrés. Des divergences aux positions G₁, T₂, mais aussi G₅ ne sont pas tolérées pour la reconnaissance correcte de l'intron et/ou l'épissage (Jacquier *et al.*, 1985; Newman *et al.*, 1985; Fouser et Friesen, 1986; Vijayraghavan *et al.*, 1986; Rain et Legrain, 1997). On peut noter la différence entre la séquence 5' chez *S. cerevisiae* et les vertébrés en position 4. En effet, un résidu T chez *S. cerevisiae* est remplacé par un résidu A chez les vertébrés.

La séquence en 5' des introns U12-dépendants est très différente: ↓(A/G)TATCCTY. Dans tous les introns identifiés à ce jour, sa conservation est très stricte. Le nucléotide 5' terminal, observé naturellement le plus fréquemment, est le résidu G (80,5 % des cas). Les autres introns U12-dépendants portent un résidu A en 5' (Fig. 3). La mutation AT vers GT du dinucléotide 5' terminal de l'intron AT-AC I du gène humain P120 empêche *in vivo* l'utilisation du site 3' naturel AC et fait apparaître 5 sites 3' cryptiques correspondant à des dinucléotides AG. Ces 5 produits d'épissage sont générés soit par les spliceosomes U12-dépendants, soit par les spliceosomes U2-dépendants (Dietrich *et al.*, 1997). De plus, la mutation de chacun des résidus en positions 4 à 7 du site 5' des introns U12-dépendants abolit l'épissage *in vivo*, la mutation de la position 8 diminuant d'un tiers cet épissage (Kolossova et Padgett, 1997). L'identité du nucléotide en position 8 semble donc moins importante que l'identité des autres nucléotides du site 5' pour l'efficacité d'épissage de ces introns.

1.2. Les boîtes de branchement

Quels que soient l'organisme et le type d'intron, c'est, à de rares exceptions près, une adénosine qui réalise la première attaque nucléophile et constitue le point de branchement. Un très petit nombre d'introns U2-dépendants utilise les trois autres nucléotides comme point de branchement (Hartmuth et Barta, 1988; Adema *et al.*, 1990; Dyhr-Mikkelsen et Kjems, 1995; Damier *et al.*, 1997). Chez *S. cerevisiae*, la substitution du résidu A au point de branchement par un autre nucléotide conduit à une inhibition quasi totale de la réaction d'épissage (Jacquier *et al.*, 1985; Vijayraghavan *et al.*, 1986; Fouser et Friesen, 1986; Rain et Legrain, 1997).

Le résidu constituant le point de branchement est important pour la première réaction d'épissage, mais nous verrons qu'il est également reconnu lors de la formation du spliceosome et lors de la deuxième étape de la réaction d'épissage.

Pour les introns U2-dépendants de vertébrés, l'adénosine réalisant l'attaque nucléophile est située entre le 18^{ième} et le 40^{ième} nucléotides, en amont du site 3' d'épissage, dans une région (YNYTRA*Y) dont le degré de conservation de séquence est faible (A* correspond au point de branchement dans la Fig. 3, Y: pyrimidine, R: purine) (Keller et Noon, 1984; Ruskin *et al.*, 1984; Shapiro et Senapathy, 1987).

La séquence de la boîte de branchement (TACTAA*C) des introns de la levure *S. cerevisiae* est, comme la séquence en 5', très fortement conservée (pour revue, Rymond et Rosbash, 1992; Wu et Krainer, 1999). Seuls quelques introns de *S. cerevisiae* présentent une déviation par rapport à la séquence consensus, comme les introns des gènes des snoRNA U3A et U3B, dont la boîte de branchement a la séquence GACTAAC (Myslinski *et al.*, 1990). Les sept nucléotides de la boîte de branchement ont été mutés individuellement dans l'intron du gène de l'actine (Fouser et Friesen, 1986; Vijayraghavan *et al.*, 1986), dans celui du gène de la protéine RP51 de *S. cerevisiae* (Jacquier *et al.*, 1985; Pascolo et Séraphin, 1997) et dans un intron artificiel (Rain et Legrain, 1997). Suivant la position et l'identité de la substitution réalisée, l'efficacité d'épissage est affectée de manière différente. En général, les effets peuvent être corrélés avec les possibilités d'appariement avec le snRNA U2, bien qu'en plus de l'interaction avec ce UsnRNA, ces nucléotides semblent également impliqués dans des interactions ARN-protéines lors de l'assemblage du spliceosome (Rain et Legrain, 1997; Pascolo et Séraphin, 1997).

Des contraintes semblent également exister chez *S. cerevisiae* au niveau des nucléotides bordant la boîte de branchement. Ainsi, les nucléotides G et C sont rarement observés à la position immédiatement en aval de cette boîte, alors qu'une sélection positive d'un résidu T est observée (Rain et Legrain, 1997). La mutation de ce résidu T dans un intron artificiel entraîne un défaut de reconnaissance de l'intron (Rain et Legrain, 1997).

La séquence de la boîte de branchement chez les levures répond parfaitement au consensus des boîtes de branchement des introns U2-dépendants de vertébrés. Chez les vertébrés, la boîte de branchement des introns épissés par les spliceosomes U12-dépendants (TCCTTRA*Y) ne diffère que légèrement par rapport à celle des introns épissés par les spliceosomes U2-dépendants. Elle inclut un résidu T supplémentaire en 5', qui est essentiel puisque la mutation de ce résidu en un résidu A inhibe l'épissage (Hall et Padgett, 1996). Le degré de conservation de la boîte de branchement est aussi très fort pour ces introns U12-dépendants. Etant donné que la boîte de branchement des introns U12-dépendants répond au consensus des boîtes de branchement des introns U2-dépendants, la séquence en 5' doit jouer un plus grand rôle dans la discrimination des deux types d'introns par les deux machineries d'épissage.

1.3. Le site d'épissage en 3'

Chez les métazoaires, une séquence riche en pyrimidines (70 à 93%), de longueur variable, est fréquemment présente entre les positions -15 et -5 en amont du dinucléotide AG. Cette séquence polypyrimidine est importante, comme nous le verrons par la suite, pour la fixation de facteurs protéiques participant à la constitution des spliceosomes.

La présence d'une séquence polypyrimidine n'est pas indispensable chez la levure (Rymond et Rosbash, 1985). Deux classes d'introns coexistent chez la levure. La première classe (introns 3'S), à laquelle appartiennent les introns des gènes de snoRNA U3, possède une distance courte entre la boîte de branchement et le site 3' d'épissage (5 à 15 nucléotides) (Miller, 1984; Myslinski *et al.*, 1990). Ces introns 3'S ne présentent pas de séquence polypyrimidine. La deuxième classe d'introns de levure (3'L), la plus fréquente, est caractérisée par une distance plus élevée entre le site de branchement et le site 3' (20 à 150 nts, avec un optimum de fréquence à 40 nts) (Parker et Patterson, 1987). Le dinucléotide AG est précédé d'une séquence N(U/A)AY, elle-même précédée d'une séquence riche en uridines (58 à 100%) entre les positions -7 et -13.

Il existe aussi deux classes d'introns chez les métazoaires, selon que le dinucléotide 3' AG est impliqué ou non dans la première réaction d'épissage (Reed, 1989). En effet, ce dinucléotide est nécessaire à la première réaction d'épissage lorsqu'il est précédé d'une séquence polypyrimidine courte (14 nts) (introns AG-dépendants) et il ne l'est pas lorsqu'il est précédé d'une séquence polypyrimidine longue (26 nts) (introns AG-indépendants) (Reed, 1989). Chez *S. cerevisiae*, tous les introns sont AG-indépendants pour la première étape (Séraphin et Kandels-Lewis, 1993) alors que, chez *S. pombe*, ils sont tous AG-dépendants (Reich *et al.*, 1992).

Dans une large majorité des cas, pour les introns U2-dépendants de mammifères, le site 3' correspond au premier dinucléotide AG en aval de la boîte de branchement (Mount, 1982; Reed et Maniatis, 1985; pour revue, Green, 1986). Un modèle a été proposé, basé sur la recherche du premier dinucléotide AG, en aval de la boîte de branchement, comme site 3' ("scanning model") (Smith *et al.*, 1989). En effet, la délétion d'un site 3' naturel ou l'insertion de nouveaux dinucléotides AG dans l'intron, en aval du point de branchement, résulte dans la sélection d'un nouveau dinucléotide AG, celui le plus proche de la boîte de branchement (Smith *et al.*, 1993; Deirdre *et al.*, 1995). De plus, l'insertion d'une structure secondaire stable entre le site 3' et la boîte de branchement diminue très fortement l'épissage (Smith *et al.*, 1993; Deirdre *et al.*, 1995). Ce type de structure semble donc interférer avec le processus de recherche du premier AG par le spliceosome.

La situation semble différente chez la levure. En effet, la sélection du site 3' dépend de la richesse en uridines en aval du site 3' (Patterson et Guthrie, 1991). De plus, de multiples AG peuvent être présents entre la boîte de branchement et le site 3' de certains introns de levure (Schatz *et al.*, 1986) et la présence de structures tige-boucles entre le site 3' et la boîte de branchement ne bloque pas la deuxième réaction d'épissage (Halfter et Gallwitz, 1988; Deshler et Rossi, 1991).

Les sites 5' et 3' d'épissage et la boîte de branchement sont les segments qui, comme nous le verrons plus loin, jouent un rôle essentiel dans la reconnaissance et la fixation des éléments constituant le spliceosome. Les autres régions de l'intron ne présentent pas de conservation de séquence. Néanmoins, ces séquences peuvent jouer un rôle dans la formation du spliceosome et dans la sélection des sites d'épissage. Une distance minimale entre les 3 séquences conservées est requise pour que l'épissage ait lieu (Cellini *et al.*, 1986; Thompson-Jager et Domdey, 1987; Luukkonen et Séraphin, 1997). De plus, la structuration secondaire de ces régions, peut intervenir dans l'efficacité d'épissage et ceci tout particulièrement chez *S. cerevisiae*, comme cela a été montré au laboratoire pour les introns de gènes de snoRNA U3 (Mougin *et al.*, 1996). De même, l'équipe de M. Rosbash a montré que l'efficacité d'épissage de l'ARN pré-messager RP51B dépend de la complémentarité de 2 séquences situées, l'une en aval du site 5' d'épissage et l'autre en amont de la boîte de branchement (Goguel et Rosbash, 1993; Libri *et al.*, 1995; Charpentier et Rosbash, 1996). L'appariement formé rapproche les séquences fonctionnelles.

Les introns U12-dépendants, comme les introns de *S. cerevisiae*, n'ont pas de séquences polypyrimidines (Sharp et Burge, 1997). Le nucléotide en 3' des introns répertoriés peut être soit un A, soit un G ou un C (Fig. 3). La mutation du dinucléotide terminal AC en AG, de l'intron AT-AC I du gène P120 humain, permet l'épissage correct de cet intron par les spliceosomes U12-dépendants, mais induit également l'utilisation, par les spliceosomes U12-dépendants, d'un autre dinucléotide AG 3' terminal cryptique (Dietrich *et al.*, 1997).

2. LES UsnRNA

Chaque UsnRNA est codé par un seul gène chez *S. cerevisiae* (pour revue, Guthrie et Patterson, 1988), alors que les UsnRNA des spliceosomes U2-dépendants chez les vertébrés sont codés par des familles de gènes répétés (pour revue, Dahlberg et Lund, 1988). Les snRNA U11 et U12 des spliceosomes U12-dépendants de cellules HeLa semblent être codés par un seul gène (Tarn *et al.*, 1995, Jacobs *et al.*, 1999), mais, dans le génome humain, il existe un pseudogène pour le snRNA U12 (Tarn *et al.*, 1995). Les UsnRNA des spliceosomes

UsnRNA	<i>S. cerevisiae</i>	spliceosomes U2- dépendants de vertébrés	UsnRNA	spliceosomes U12- dépendants de vertébrés
U1	568	164	U11	134
U2	1175	187	U12	150
U4	160	141 à 145	U4atac	131
U5	U5L 214 U5S 178	116	U5	116
U6	112	106 à 107	U6atac	125

Tableau 2: Comparaison de la taille des UsnRNA de *S. cerevisiae* et de vertébrés.

Les tailles indiquées ont été extraites de la revue de Solymosy et Pollack (1993) pour les UsnRNA des spliceosomes U2-dépendants de vertébrés et les UsnRNA de *S. cerevisiae* et de Montzka et Steitz (1988) et Tarn et Steitz (1996b) pour les UsnRNA des spliceosomes U12-dépendants de vertébrés.

U12-dépendants sont 100 à 500 fois moins abondants que leurs homologues des spliceosomes U2-dépendants (Montzka et Steitz, 1988, Tarn et Steitz, 1996b).

Les premiers séquençages d'UsnRNA de spliceosomes U2-dépendants, isolés de vertébrés et de plantes (pour revue, Solymosy et Pollak, 1993), ont révélé une conservation peu élevée des séquences primaires, à l'exception de certains motifs fonctionnels, mais une conservation élevée des structures secondaires adoptées par ces UsnRNA. Les premières séquences et structures secondaires de ces ARN ont été établies au début des années 80, pour la plupart par l'équipe de C. Branlant: Branlant *et al.* (1980) et Mount et Steitz (1981) pour le snRNA U1, Branlant *et al.* (1982) pour le snRNA U2, Krol *et al.* (1981a) pour le snRNA U4, Branlant *et al.* (1983) pour le snRNA U5 et Epstein *et al.* (1980) pour le snRNA U6.

Les UsnRNA de *S. cerevisiae* U1, U2 et U5 se distinguent de ceux des autres espèces par leur taille élevée (Tableau 2). Cependant, les régions excédentaires dans les snRNA U1 et U2 peuvent être délétées sans affecter leur fonction dans l'épissage (Igel et Ares, 1988; McPheeters *et al.*, 1989; Siliciano *et al.*, 1991) et le snRNA U2 humain peut compléter, *in vivo* et *in vitro*, le snRNA U2 de *S. cerevisiae* (McPheeters *et al.*, 1989; Shuster et Guthrie, 1990)

Contrairement aux autres UsnRNA spliceosomaux, les snRNA U6 et U6atac se distinguent par la présence d'un phosphate γ monométhylé en 5' (Singh et Reddy, 1989; Tarn et Steitz, 1996b; pour revue, Reddy *et al.*, 1992). Tous les autres UsnRNA des spliceosomes U2 ou U12-dépendants portent une coiffe 5' terminale, qui comporte une guanosine tri-méthylée ($m_3^{2,2,7}G$) (Montzka et Steitz, 1988; Tarn et Steitz, 1996b; pour revue, Reddy *et al.*, 1992). La différence, entre le snRNA U6 et les autres UsnRNA, s'explique par le fait que le snRNA U6 est transcrit par l'ARN polymérase III (Kunkel *et al.*, 1986; Reddy *et al.*, 1987; Krol *et al.*, 1987), alors que les autres UsnRNA spliceosomaux sont des produits de l'ARN polymérase II (Hellung-Larsen *et al.*, 1980; Chandrasekharappa *et al.*, 1983; pour revue, Dahlberg et Lund, 1988). Bien qu'aucune étude n'ait été menée sur la transcription des gènes des UsnRNA des spliceosomes U12-dépendants, sur la base des caractéristiques de leurs extrémités 5', on peut supposer que les snRNA U11, U12 et U4atac sont transcrits par l'ARN polymérase II et le snRNA U6atac par l'ARN polymérase III. Ceci est en accord avec le fait que le promoteur du gène du snRNA U12 présente les caractéristiques de promoteurs utilisés par l'ARN polymérase II (Tarn *et al.*, 1995). Il est à noter que les UsnRNA transcrits par l'ARN polymérase II possèdent une coiffe en 5' qui se distingue de celle des ARNm par la présence de deux groupements méthyle supplémentaires en position 2. La genèse de cette hyperméthylation sera discutée plus loin dans la revue insérée à la fin de ce chapitre.

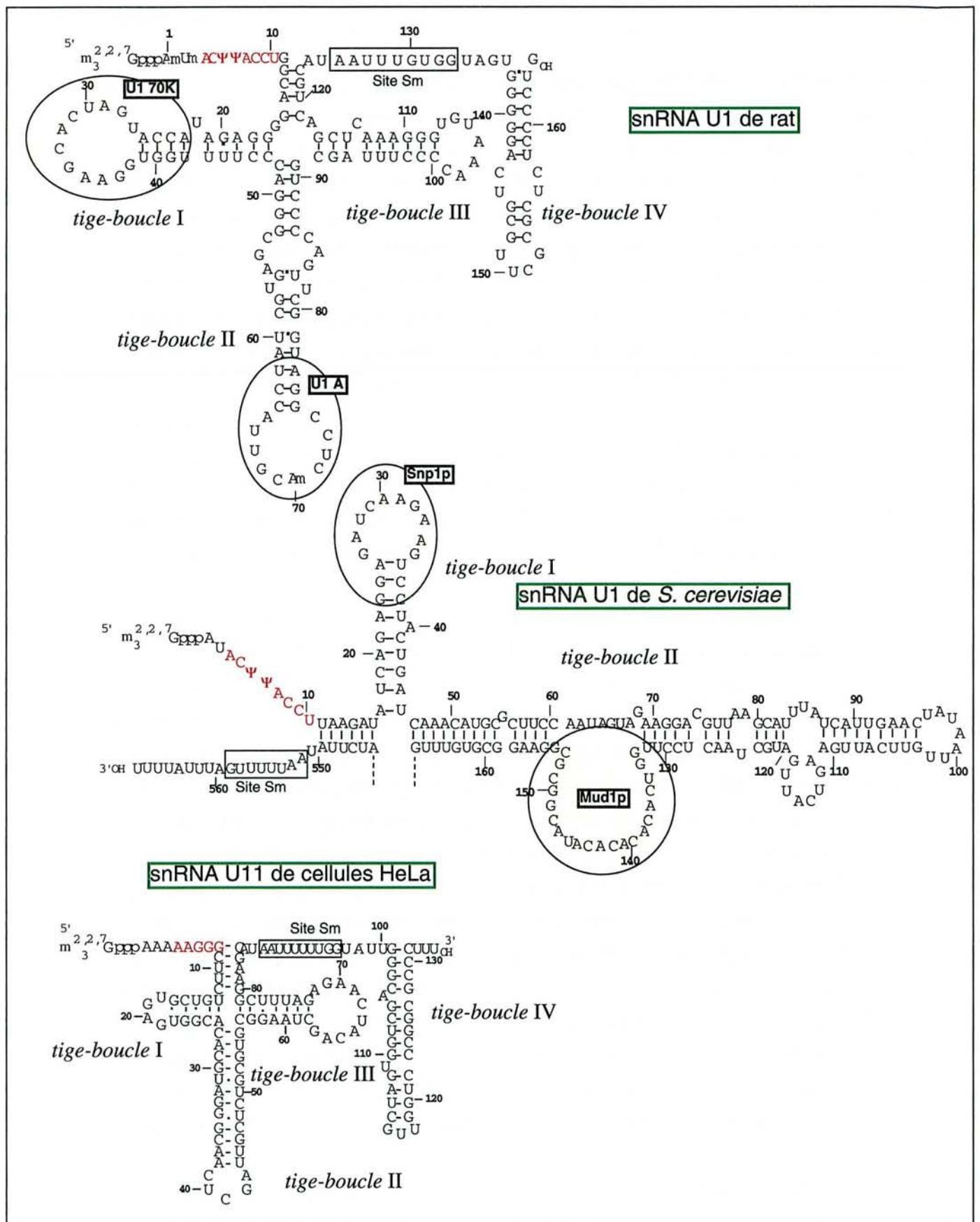


Figure 4: Structures secondaires des snRNA U1 de rat et de la levure *S. cerevisiae*, ainsi que du snRNA U11 de cellules HeLa.

Ces différentes structures sont extraites de Kretzner *et al.* (1990) pour le snRNA U1 de levure, de Branlant *et al.* (1981a) et Mount et Steitz (1981) pour le snRNA U1 de rat et de Wassarman et Steitz (1992a) pour le snRNA U11 de cellules HeLa. Pour le snRNA U1 de *S. cerevisiae*, seule la région correspondant aux séquences présentes dans le snRNA U1 de rat est montrée (Kretzner *et al.*, 1990).

Les sites Sm des différents snRNA sont encadrés. Les séquences complémentaires aux jonctions exon 1/intron, dans les 3 ARN, sont en rouge. Les sites de fixation des protéines U1A et U1 70K, ainsi que des protéines homologues de *S. cerevisiae*, Mud1p et Snp1p, respectivement, sont indiqués en jaune (pour revue, Will et Lührmann, 1997).

2.1. Un motif commun aux UsnRNA: le site Sm

Les UsnRNA U1, U2, U4, U5, U11, U12 et U4atac de toutes les espèces étudiées partagent un motif commun: la séquence A(U)_nG appelée site Sm, localisée dans une région en simple brin de l'ARN et, dans la plupart des cas, entourée de deux structures tige-boucles (Branlant *et al.*, 1982; Montzka et Steitz, 1988; Tarn et Steitz, 1996b; pour revue, Guthrie et Patterson, 1988). Le résidu adénosine en 5' du motif est le nucléotide le plus conservé. Il est, en effet, trouvé à plus de 95% dans les UsnRNA (pour revue, Guthrie et Patterson, 1988). Dès l'observation de ce domaine commun, Branlant *et al.* (1982) avaient proposé qu'il permettrait la fixation des protéines communes aux différents UsnRNA spliceosomaux, les protéines Sm (Liautard *et al.*, 1982). Cette hypothèse a ensuite été vérifiée expérimentalement (Mattaj et De Robertis, 1985; Mattaj *et al.*, 1986). Les snRNA U6 et U6atac ne possèdent pas ce motif mais, comme nous le verrons plus loin, ils fixent un autre ensemble de protéines Sm.

2.2. Les motifs spécifiques à chaque UsnRNA

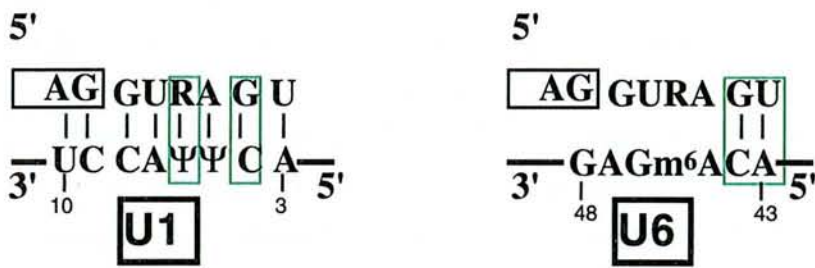
2.2.1. Les snRNA U1 et U11

Dans tous les organismes, à l'exception de la levure *S. cerevisiae*, la taille et la structure du snRNA U1 sont très conservées (pour revue, Guthrie et Patterson, 1988). Il est à noter qu'en dépit de sa séquence complètement différente de celle du snRNA U1 humain, le snRNA U11 peut se replier selon une structure 2D très similaire à celle du snRNA U1 (Fig. 4) (Montzka et Steitz, 1988; Wassarman et Steitz, 1992a).

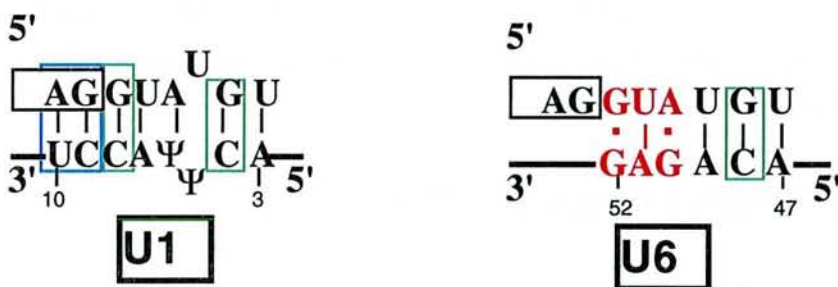
a) L'interaction entre le snRNA U1 ou U11 et le site 5' d'épissage

La séquence en 5' du snRNA U1, A₁UACUUACCU₁₀, est conservée chez tous les eucaryotes. Un seul cas de déviation a été observé chez *P. polycephalum*, où le résidu U₅ est remplacé par un A (Szkukalek *et al.*, 1996). Cette séquence 5' terminale est en simple brin, aussi bien dans le snRNA U1 en solution (Branlant *et al.*, 1980), que dans la snRNP U1 (Rinke *et al.*, 1984). C'est en 1980 qu'une interaction a été proposée pour la première fois entre le snRNA U1 et l'ARN pré-messager. Les auteurs se sont basés sur l'existence d'une complémentarité de séquences entre l'extrémité 5' du snRNA U1 et le site 5' d'épissage (Lerner *et al.*, 1980; Rogers et Wall, 1980). L'interaction alors proposée était beaucoup plus longue que celle existant réellement. Les analyses de la structure secondaire du snRNA U1 (Branlant *et al.*, 1981b; Mount et Steitz, 1981) ont montré que l'appariement devait en fait se limiter à la région 5' en simple brin du snRNA U1 (Fig. 5). Ce n'est que quelques années plus

A/ spliceosomes U2-dépendants de cellules HeLa



B/ spliceosomes de *S. cerevisiae*



C/ spliceosomes U12-dépendants de cellules HeLa

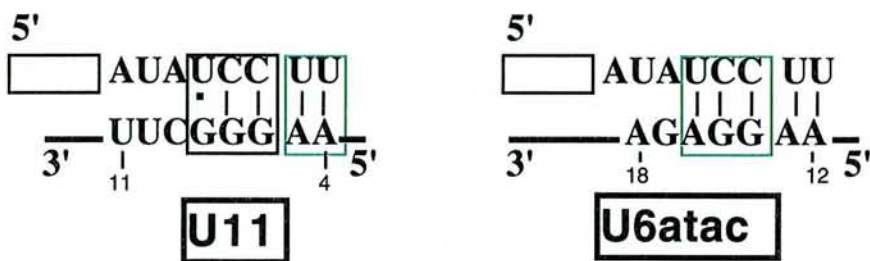


Figure 5: Les interactions des sites 5' d'épissage avec les snRNA U1 et U6, dans le cas d'ARN pré-messagers épissés par les spliceosomes U2-dépendants, et avec les snRNA U11 et U6atac, dans le cas d'ARN pré-messagers épissés par les spliceosomes U12-dépendants. D'après les revues, Madhani et Guthrie (1994b) et Wu et Krainer (1999).

L'existence des appariements de bases encadrés en bleu a été démontrée par mutations compensatoires. Les appariements de bases sont respectivement encadrés en noir et en vert, quand le défaut d'épissage, créé par la mutation d'un des nucléotides de l'appariement, n'est pas ou est seulement partiellement compensé par des mutations compensatoires dans l'autre nucléotide de l'appariement. D'après Zhuang et Weiner, 1986; Séraphin *et al.*, 1988; Siliciano et Guthrie, 1988; Kandels-Lewis et Séraphin, 1993; Séraphin et Kandels-Lewis, 1993; Lesser et Guthrie, 1993; Cohen *et al.*, 1994; Hwang et Cohen, 1996; Kolossova et Padgett, 1997; Luukkonen et Séraphin, 1998a et Incorvaia et Padgett, 1998.

L'extension, récemment proposée par Luukkonen et Séraphin (1998a), dans l'interaction entre le snRNA U6 de *S. cerevisiae* et le site 5' est indiquée en rouge.

tard que la fixation de la snRNP U1 au site 5' a pu être démontrée (Mount *et al.*, 1983), ainsi que la nécessité de l'intégrité de la séquence en 5' du snRNA U1 pour qu'il y ait épissage *in vitro* (Krämer *et al.*, 1984).

Alors que, chez les vertébrés, le degré de complémentarité entre le snRNA U1 et la région 5' est parfois faible (Jacob et Gallinaro, 1989), chez la levure *S. cerevisiae*, à l'exception du défaut d'appariement en position 4 de l'intron (Fig. 5), une conservation absolue des complémentarités aux autres positions est observée (pour revue, Rymond et Rosbash, 1992). Cependant, chaque appariement de bases entre le snRNA U1 et la région 5' ne semble pas avoir la même importance. En effet, des mutations dans le snRNA U1 en positions 4 et 8, qui empêchent son appariement au site 5', sont létales chez *S. cerevisiae*, alors que des mutations en positions 3, 9 et 10 ont peu d'effet sur la croissance (Siliciano et Guthrie, 1988; Séraphin *et al.*, 1988; Séraphin et Rosbash, 1990; Séraphin et Kandels-Lewis, 1993).

Dans la plupart des cas, chez les mammifères et chez *S. cerevisiae*, les défauts d'épissage observés lors de la création de mutations dans le site 5' ne sont que partiellement corrigés par la création de mutations dans le snRNA U1 restaurant l'appariement U1/site 5', indiquant que le site 5' interagit avec d'autres facteurs que le snRNA U1 lors de l'épissage (Fig. 5) (Zhuang et Weiner, 1986; Séraphin *et al.*, 1988; Siliciano et Guthrie, 1988; Séraphin et Rosbash, 1990; Séraphin et Kandels-Lewis, 1993; Alvarez *et al.*, 1996). Nous verrons, en effet, qu'un de ces facteurs est le snRNA U6, qui s'apparie également avec le site 5' des introns.

Chez *S. cerevisiae*, le nucléotide consensus à la quatrième position de l'intron est un résidu U et des mutations dans le snRNA U1 qui permettent un appariement de base avec ce nucléotide entraînent un défaut de croissance cellulaire (Siliciano et Guthrie, 1988; Séraphin et Rosbash, 1989b). Chez *S. pombe* et chez les mammifères, le nucléotide en position 4 de l'intron est généralement un A, qui peut s'apparier au snRNA U1, et des substitutions de ce nucléotide empêchant l'appariement avec le snRNA U1 à cette position sont létales chez *S. pombe* (Alvarez *et al.*, 1996). Chez *P. polycephalum*, la nécessité d'un appariement entre la position 4 de l'intron et le snRNA U1 semble également requise car cet organisme présente naturellement des mutations compensatoires dans ces deux ARN afin de préserver la complémentarité (U→A en position 5 du snRNA U1 et A→U en position 4 de l'intron) (Szkukalek *et al.*, 1996). Cette différence paradoxale entre *S. cerevisiae* et les autres espèces peut s'expliquer de deux manières:

⇒ une première explication peut être que l'épissage dans toutes les espèces est facilité par l'appariement U1/site 5', aussi longtemps que les régions complémentaires ne forment pas une hélice trop stable, diminuant ou prévenant le déplacement du snRNA U1 du site 5' dans une étape suivante de l'assemblage du spliceosome (voir ci-après). Chez *S. cerevisiae*, la

déstabilisation de l'hélice se ferait de la même manière pour tous les introns à partir du mésappariement existant en position 4 de l'intron, alors que chez les autres organismes, la ou les position(s) de mésappariement générant la dissociation seraient plus variables en fonction de la dégénérescence des séquences introniques au site 5' d'épissage.

⇒ la deuxième explication possible serait que le mésappariement interviendrait au contraire dans la stabilisation de l'interaction U1/site 5', en constituant par exemple un site de fixation de protéine(s).

Le snRNA U11 joue le même rôle que le snRNA U1, pour l'épissage des introns U12-dépendants, et la première évidence de cette homologie a été l'observation d'une complémentarité entre la séquence A₄AGGG₈, localisée dans une région en simple brin du snRNA U11 (Montzka et Steitz, 1988; Wassarman et Steitz, 1992a), et le site 5' d'épissage des ARN pré-messagers (Fig. 4 et 5) (Hall et Padgett, 1994). Depuis, l'existence de cette interaction a été démontrée par des expériences de mutations compensatoires et par la création de pontages covalents entre ces ARN (Kolossova et Padgett, 1997; Yu et Steitz, 1997). Cependant l'analogie d'action entre les snRNA U1 et U11 n'est pas complète. En effet, le snRNA U1 peut s'apparier entièrement au site 5' d'épissage, alors que l'appariement entre le snRNA U11 et le site 5' d'épissage n'implique que les nts 4 à 8 des introns U12-dépendants (Kolossova et Padgett, 1997; Yu et Steitz, 1997). L'appariement U1/site 5' est donc plus long que l'appariement U11/site 5' et, ceci, de 2 à 3 paires de bases (Fig. 5).

Il est à noter qu'un défaut d'épissage dû à des mutations dans les résidus 4 à 6 de l'intron n'est pas diminué par l'introduction de mutations compensatoires dans le snRNA U11, restaurant l'appariement (Fig. 5) (Kolossova et Padgett, 1997). Cette observation impliquerait donc, comme c'est le cas pour les introns épissés par les spliceosomes U2-dépendants, que le site 5' des introns épissés par les spliceosomes U12-dépendants soit reconnu durant l'épissage par un autre élément que le snRNA U11. Comme nous le verrons, il s'agit du snRNA U6atac.

b) *Les régions d'interaction des protéines spécifiques au snRNA U1*

En plus du site Sm et de la séquence en 5', les séquences des boucles terminales des structures tige-boucles I et II sont également très conservées dans les snRNA U1 de métazoaires. Ces structures tige-boucles I et II constituent, respectivement, les sites de fixation des protéines U1 70K et U1A (Fig. 4) (Scherly *et al.*, 1989; Lutz-Freyermuth *et al.*, 1990; pour revue, Will et Lührmann, 1997). Ces protéines et, par voie de conséquence les structures tige-boucles I et II, sont nécessaires à l'épissage chez les métazoaires (Hamm *et al.*, 1990b). Les homologues respectifs des protéines U1 70K et U1A, chez *S. cerevisiae*, Snp1p et Mud1p, se fixent respectivement à la structure tige-boucle I et à une large boucle de la

structure tige-boucle II du snRNA U1 de levure (Fig. 4) (Smith et Barrell, 1991; Kao et Siliciano, 1992; Tang et Rosbash, 1996).

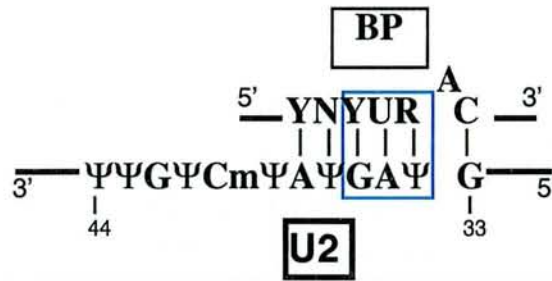
2.2.2. Les snRNA U2 et U12

a) L'interaction entre le snRNA U2 ou U12 et la boîte de branchement

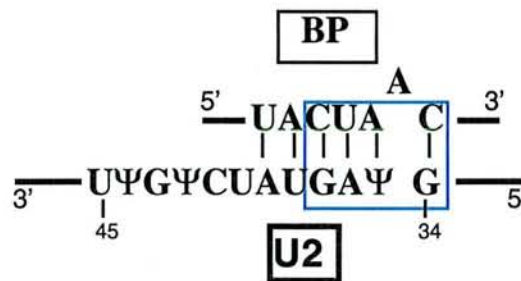
Le snRNA U2, comme le snRNA U1, forme un appariement de bases avec l'intron. En effet, l'interaction postulée sur la base d'une complémentarité de séquences entre le snRNA U2 et la boîte de branchement (Ares, 1986) a ensuite été démontrée par création de mutations compensatoires entre les deux ARN, aussi bien chez *S. cerevisiae* que chez les mammifères (Fig. 6) (Parker *et al.*, 1987; Wu et Manley, 1989; Zhuang et Weiner, 1989; Pascolo et Séraphin, 1997). Chez les eucaryotes supérieurs, le degré de complémentarité entre le snRNA U2 et la boîte de branchement est souvent très faible et le type d'interaction parfaite montré dans la Figure 6, n'est que très rarement observé, alors que chez la levure, à de très rares exceptions près, un appariement de 6 paires de bases peut être formé. Des mutations de l'un des 4 résidus en 5' de la séquence GΨAGUA, qui entraînent une déstabilisation de l'interaction entre le snRNA U2 et la boîte de branchement, sont létales chez *S. cerevisiae* (Miraglia *et al.*, 1991; Pascolo et Séraphin, 1997). La déstabilisation de l'hélice d'ARN formée entre le snRNA U2 de *S. cerevisiae* et l'intron peut diminuer l'efficacité d'épissage *in vitro* de 50 à 99%, suivant la position de la mutation dans le snRNA U2 (McPheeters et Abelson, 1992). Le caractère essentiel de l'interaction entre le snRNA U2 et la boîte de branchement est liée au fait qu'elle permet de placer l'adénosine, qui réalise l'attaque nucléophile lors de la première réaction de l'épissage, en boucle latérale, c'est-à-dire dans sa conformation active (Fig. 6) (Query *et al.*, 1994; 1996).

Le résidu réalisant la première attaque nucléophile est également positionné en boucle latérale dans l'appariement formé par le snRNA U12 et les introns U12-dépendants. En effet, le snRNA U12 est l'homologue du snRNA U2 pour l'épissage de ces introns. La première évidence de cette homologie, comme pour l'homologie entre les snRNA U1 et U11, a été l'observation d'une complémentarité de la séquence GΨAAGGA du snRNA U12 et de la boîte de branchement des introns U12-dépendants (Hall et Padgett, 1994). L'existence de cette interaction a été démontrée par des expériences de pontage et de mutagenèse compensatoire (Fig. 6) (Hall et Padgett, 1996; Tarn et Steitz, 1996a). Elle est nécessaire à la formation de spliceosomes U12-dépendants actifs. Comme pour les spliceosomes de *S. cerevisiae*, la complémentarité entre le snRNA U12 et la boîte de branchement des introns U12-dépendants est très fortement conservée. De plus, elle implique la formation de 7 paires de base, contre 6 pour l'interaction entre le snRNA U2 et la boîte de branchement des introns U2-dépendants

A/ spliceosomes U2-dépendants de cellules HeLa



B/ spliceosomes de *S. cerevisiae*



C/ spliceosomes U12-dépendants de cellules HeLa

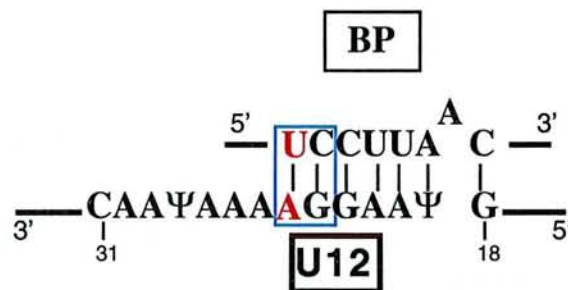


Figure 6: Les interactions mises en place entre les snRNA U2 ou U12 et la boîte de branchement des ARN pré-messagers lors de l'épissage.

D'après les revues, Madhani et Guthrie (1994b) et Wu et Krainer (1999).

Les appariements de bases encadrés en bleu ont été démontrés par mutations compensatoires (Parker *et al.*, 1987; Wu et Manley, 1989; Zhuang et Weiner, 1989; Hall et Padgett, 1996; Pascolo et Séraphin, 1997). L'appariement de bases supplémentaire, dans l'interaction entre le snRNA U12 et la boîte de branchement, est indiqué en rouge.

Y: pyrimidine

R: purine

(Fig. 6). La paire de bases supplémentaire, nécessaire dans le cas des spliceosomes U12-dépendants, implique l'uridine en 5' de la boîte de branchement. La création de mutations dans le snRNA U12, empêchant la formation de cette paire de bases, inhibe complètement l'épissage *in vitro* (Hall et Padgett, 1996). Une interaction plus stable est donc formée entre le snRNA U12 et la boîte de branchement des introns U12-dépendants, qu'entre le snRNA U2 et la boîte de branchement des introns U2-dépendants. Le 7^{ième} appariement de bases qui peut être formé entre la boîte de branchement des introns U12-dépendants et le snRNA U12 peut donc augmenter l'affinité de ce UsnRNA par rapport au snRNA U2 et ainsi jouer un rôle de discrimination entre ces deux UsnRNA. Un autre paramètre intervient sans doute dans la sélectivité de reconnaissance. En effet, les snRNA U11 et U12 sont associés au sein de la même UsnRNP. Leur appariement à l'intron se fait de manière concomitante, la fixation du snRNA U11 au site 5' des introns U12-dépendants pourrait donc imposer la fixation du snRNA U12 à leur boîte de branchement ou, inversement, la fixation du snRNA U12 à la boîte de branchement pourrait imposer la fixation du snRNA U11 au site 5' des introns U12-dépendants.

b) *La structure primaire et secondaire des snRNA U2 et U12*

D'une manière générale, les 100 premiers résidus des snRNA U2 et U12, bien que de nature très différente, ont une séquence phylogénétiquement conservée (Tarn *et al.*, 1995; Shukla et Padgett, 1999; pour revue, Guthrie et Patterson, 1988). Ils jouent un rôle important dans la fonction de ces ARN lors de l'épissage en permettant l'appariement avec la boîte de branchement, mais également, comme nous le verrons en détail par la suite, en s'appariant, respectivement, aux snRNA U6 et U6atac dans les spliceosomes actifs. Cette région du snRNA U2 porte de nombreuses modifications post-transcriptionnelles, comme nous le verrons dans la deuxième partie de l'introduction (Fig. 7).

Comme nous l'avons déjà mentionné, le snRNA U2 de *S. cerevisiae* est très long par rapport aux snRNA U2 d'autres espèces (Tableau 2). Cependant, sa région 5' fonctionnelle adopte une structure identique à celle des snRNA U2 de vertébrés (Fig. 7) (Branlant *et al.*, 1982; Igel et Ares, 1988). De plus, les snRNA U2, quelle que soit leur origine, semblent pouvoir se replier sous 2 formes alternatives, A et B (Fig. 8). Le passage d'une forme à l'autre est facilité par la formation d'une structure en pseudo-noeud. Des variations compensatoires de bases préservent la complémentarité entre les séquences A et A' et les séquences B et B' (Fig. 8) (pour revues, Guthrie et Patterson, 1988; Ares et Igel, 1989). Des mutations, empêchant la formation de l'hélice IIa du snRNA U2 de levure, entraînent une accumulation de la forme A dans la cellule. Cette accumulation a pour conséquence l'inhibition de la fixation de la snRNP U2 au pré-messager et donc l'inhibition de l'épissage (Zavanelli et Ares, 1991; Zavanelli *et*

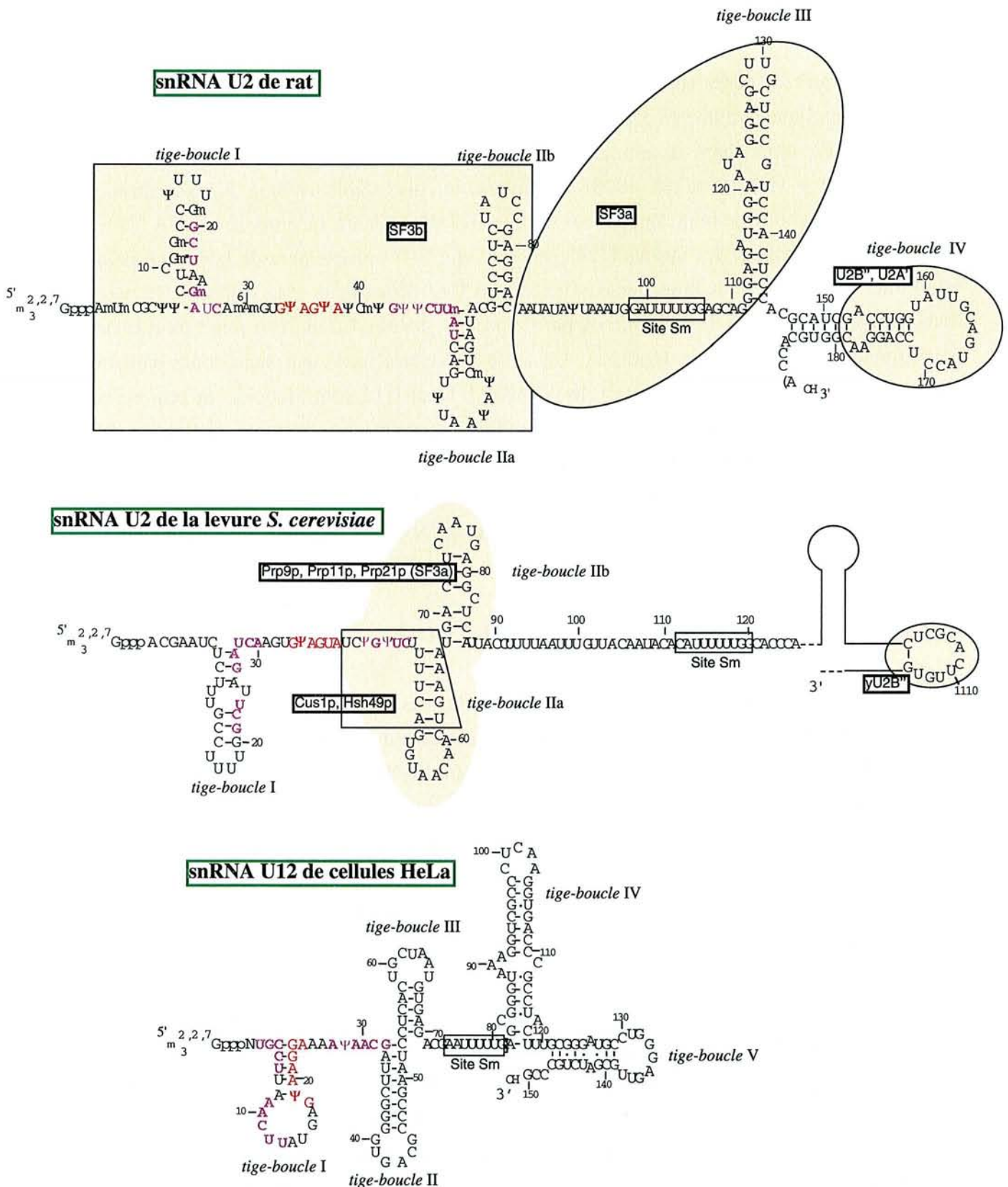


Figure 7: Structures secondaires des snRNA U2 de la levure *S. cerevisiae* et de rat, ainsi que du snRNA U12 de cellules HeLa.

Ces différentes structures sont extraites de Igel et Ares (1988) pour le snRNA U2 de *S. cerevisiae*, de Reddy *et al.*, (1981) et Branlant *et al.* (1982) pour le snRNA U2 de rat et de Wassarman et Steitz (1992a) pour le snRNA U12 de cellules HeLa. Pour le snRNA U2 de *S. cerevisiae*, seule la région correspondant aux séquences présentes dans le snRNA U2 de rat est montrée.

Les sites Sm des différents snRNA sont encadrés. Les séquences complémentaires aux boîtes de branchement des introns sont indiquées en rouge et les séquences des snRNA U2 et U12 complémentaires, respectivement, aux snRNA U6 et U6atac sont en violet. Les sites de protection des facteurs SF3a et SF3b (Behrens *et al.*, 1993b; Krämer *et al.*, 1999) et des protéines U2B'' et U2A' sur le snRNA U2 de rat (pour revue, Will et Lüthmann, 1997), ainsi que de la protéine yU2B'' sur le snRNA U2 de *S. cerevisiae* sont indiqués (Tang *et al.*, 1996). Les sites de fixation des facteurs SF3a et SF3b de *S. cerevisiae* ont été proposés sur la base de liens génétiques mis en évidence entre les protéines et les régions indiquées en jaune du snRNA U2 de *S. cerevisiae* (Yan et Ares, 1996).

al., 1994). La forme B semble donc être la forme dans laquelle le snRNA U2 s'associe à l'ARN pré-messager pour constituer le spliceosome. La création de mutations dans le snRNA U2 de levure et l'utilisation *in vitro* d'oligonucléotides complémentaires de la structure tige-boucle IIa du snRNA U2 de cellules HeLa ont aussi permis de montrer la nécessité de l'intégrité de cette structure tige-boucle pour l'insertion de la snRNP U2 dans le spliceosome (Ares et Igel, 1990; Barabino *et al.*, 1992). La structure tige-boucle IIa est le domaine principal de fixation des facteurs protéiques SF3a et SF3b (Fig. 7). Le fait que ces facteurs soient indispensables à l'épissage explique sans doute la nécessité de cette structure tige-boucle. Par contre, bien que la possibilité de former la structure tige-boucle IIb soit conservée phylogénétiquement (pour revue, Guthrie et Patterson, 1988), cette structure n'est indispensable ni à la croissance, ni à l'épissage (Ares et Igel, 1990).

La délétion de la séquence immédiatement en aval du site Sm (séquence B') du snRNA U2 de levure n'a aucun effet sur la croissance cellulaire (Ares et Igel, 1990). Cependant, l'existence *in vivo* d'un équilibre entre les formes A et B du snRNA U2 pourrait apporter un avantage sélectif à la cellule (Zavanelli et Ares, 1991; Zavanelli *et al.*, 1994). Il faut noter que la formation d'un appariement entre B et B', du moins partiel, n'est pas incompatible avec l'appariement A/A'. Ce double appariement conduit à la formation d'une structure pseudo-nœud. On peut imaginer que la formation de cette structure pourrait favoriser certaines transitions structurales auxquelles le snRNA U2 est soumis au cours de l'assemblage ou la dissociation du spliceosome.

La structure du snRNA U12 de cellules HeLa, dans un premier temps établie par ordinateur, a ensuite été confirmée par l'utilisation de sondes chimiques et enzymatiques. Bien que la séquence primaire du snRNA U12 diffère de celle du snRNA U2, la structure du snRNA U12 est très similaire à la structure active B du snRNA U2 de cellules HeLa (Fig. 7) (Montzka et Steitz, 1988; Wassarman et Steitz, 1992a). Elle est conservée chez la souris, le poulet et chez *A. thaliana* (Tarn *et al.*, 1995; Shukla et Padgett, 1999). Nous pouvons remarquer que le snRNA U12 ne contient pas de séquence complémentaire à la boucle de la structure tige-boucle II et ne peut donc pas se structurer sous la forme A, ni contenir une structure en pseudo-nœud.

c) *Les régions d'interaction des protéines spécifiques du snRNA U2*

La région 5' terminale conservée des snRNA U2 de mammifères constitue, comme nous l'avons mentionné ci-dessus, le site de fixation de protéines essentielles à l'épissage formant les facteurs SF3a et SF3b. Les zones de fixation de ces deux facteurs, indiquées dans la Figure 7, ont été définies par des études de protection du snRNA U2 par ces protéines, vis-à-vis de sondes chimiques et enzymatiques, et par des études de microscopie électronique (Behrens *et*

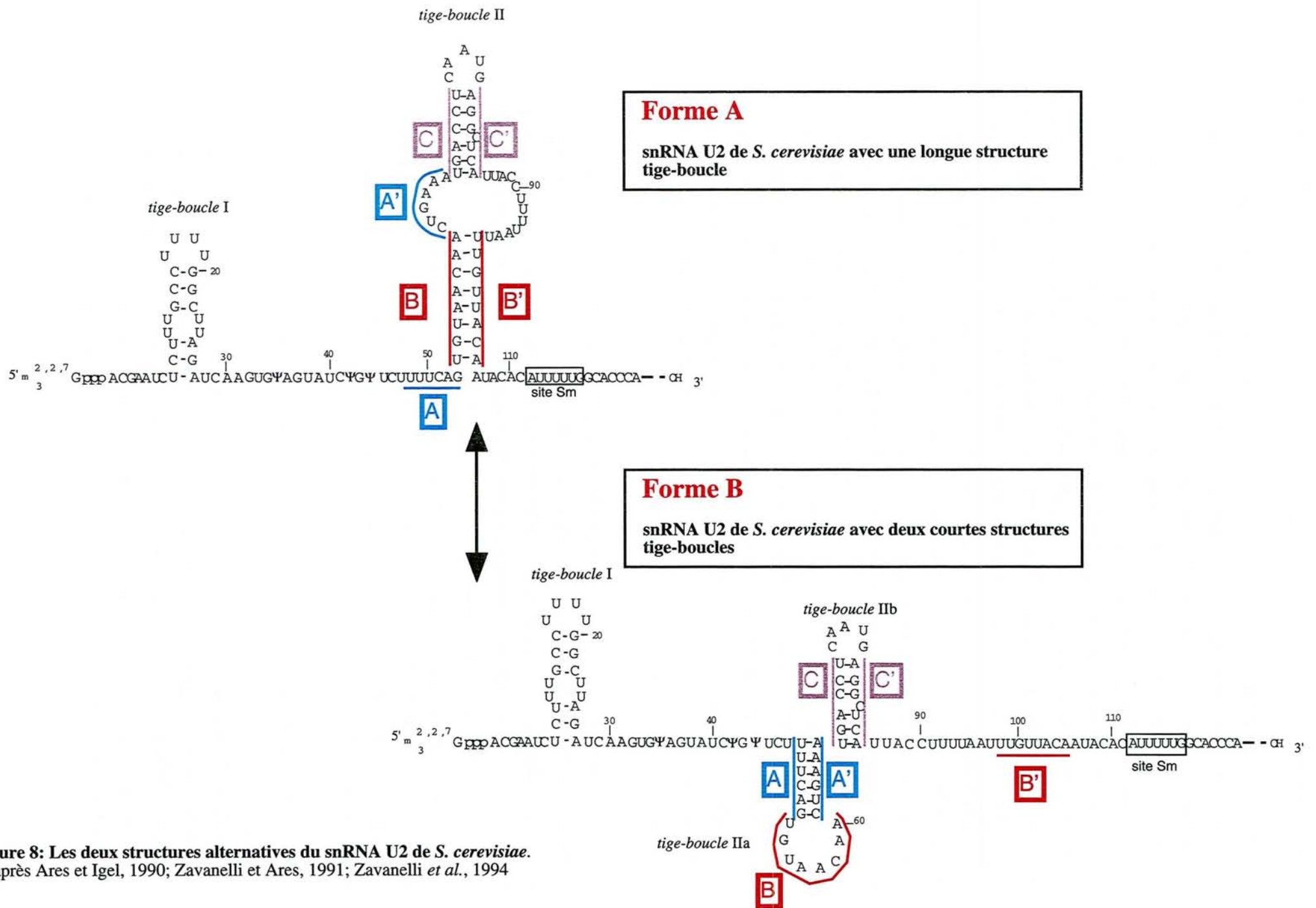


Figure 8: Les deux structures alternatives du snRNA U2 de *S. cerevisiae*.
D'après Ares et Igel, 1990; Zavanelli et Ares, 1991; Zavanelli *et al.*, 1994

al., 1993a; Krämer *et al.*, 1999). Les facteurs SF3 de *S. cerevisiae* semblent également interagir avec la région 5' du snRNA U2 de levure. En effet, des liens fonctionnels ont été mis en évidence par voie génétique entre certaines protéines, composant ces deux facteurs, et les structures tige-boucles IIa et IIb du snRNA U2 (Fig. 7) (Yan et Ares, 1996).

En plus de ces facteurs, d'autres protéines sont en contact avec le snRNA U2. Ainsi, la protéine U2B'' se fixe spécifiquement à la tige-boucle IV du snRNA U2, ceci en présence de U2A' (Scherly *et al.*, 1990; Boelens *et al.*, 1991). L'homologue chez la levure de U2B'', yU2B'' ou MSL1, se fixe également à la tige-boucle IV du snRNA U2, sans nécessiter la présence d'une autre protéine (Tang *et al.*, 1996). Le gène codant cette protéine n'est pas indispensable à la croissance cellulaire (Tang *et al.*, 1996), ce qui est en accord avec le fait que la tige-boucle IV, aussi bien chez la levure que chez les métazoaires, n'est nécessaire ni à la croissance ni à l'épissage (McPheeters *et al.*, 1989; Hamm *et al.*, 1989; Pan et Prives, 1989; Shuster et Guthrie, 1990).

2.2.3. Le snRNA U5

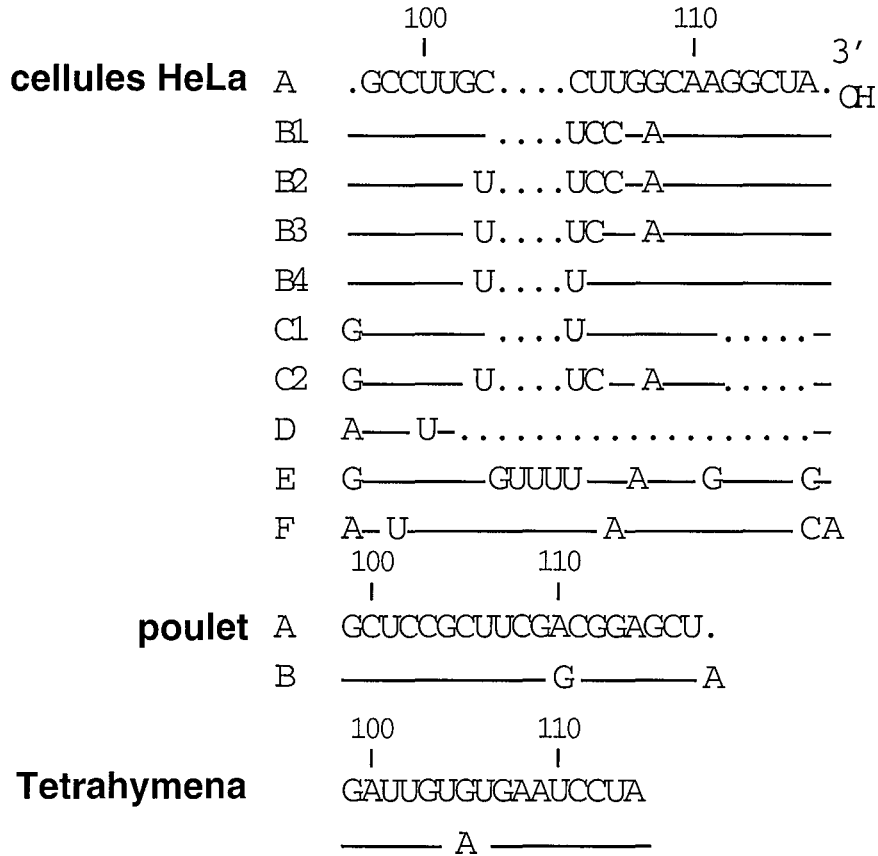
a) Les différents variants du snRNA U5

Le snRNA U5 est le seul snRNA commun aux deux types de spliceosomes, U12-dépendants et U2-dépendants. Néanmoins, chez les vertébrés et les plantes, différentes formes du snRNA U5 sont présentes dans une même cellule (Fig. 9) (Krol *et al.*, 1981b; Branlant *et al.*, 1983; Sontheimer et Steitz, 1992). Au moins 10 variants ont été répertoriés dans les cellules HeLa. Ils diffèrent en séquence et quelquefois en taille, notamment dans leur domaine 3' terminal. Ceci s'explique par l'existence de plusieurs types de gènes codant ces différents snRNA U5. A l'heure actuelle, on ne sait pas si certains de ces variants sont plus particulièrement incorporés dans les spliceosomes U2-dépendants et si d'autres sont utilisés pour les spliceosomes U12-dépendants. Bien qu'il n'y ait pas de spliceosomes U12-dépendants chez *S. cerevisiae*, deux snRNA U5 différents en taille, mais pas en séquence, sont produits à partir d'un gène unique: une forme courte (U5S) et une forme longue (U5L) (Tableau 2 et Fig. 9) (Patterson et Guthrie, 1987, Frank *et al.*, 1994). Ces deux formes sont générées par des mécanismes alternatifs de maturation à l'extrémité 3' (Chanfreau *et al.*, 1997).

b) L'interaction du snRNA U5 avec les deux exons (Fig. 10)

En plus du site Sm, la conservation en séquence primaire, au niveau du snRNA U5, est restreinte à la boucle terminale de la structure tige-boucle I. Les 9 nucléotides de cette boucle, y compris les nucléotides modifiés, sont invariants dans l'ensemble des organismes étudiés, à

A



B

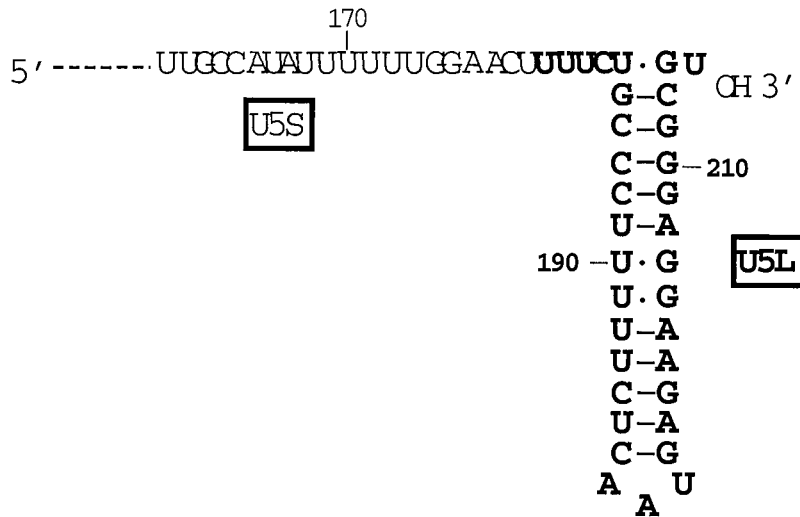


Figure 9: Séquence 3' terminale de différents variants du snRNA U5 chez les métazoaires (A) et les deux formes du snRNA U5 de *S. cerevisiae* (B).

A/ La séquence complète du fragment 3' terminal indiquée est celle du snRNA U5 le plus abondant de chaque organisme. Pour les autres variants, seules les modifications sont montrées. Un point indique une délétion.

D'après Branlant *et al.*, 1983 et Sontheimer et Steitz, 1992.

B/ La séquence supplémentaire présente dans le snRNA U5L de *S. cerevisiae*, par rapport au snRNA U5S, est indiquée en gras.

D'après Patterson et Guthrie, 1987.

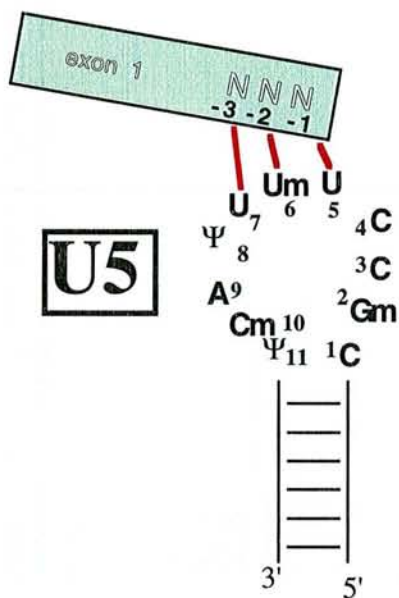
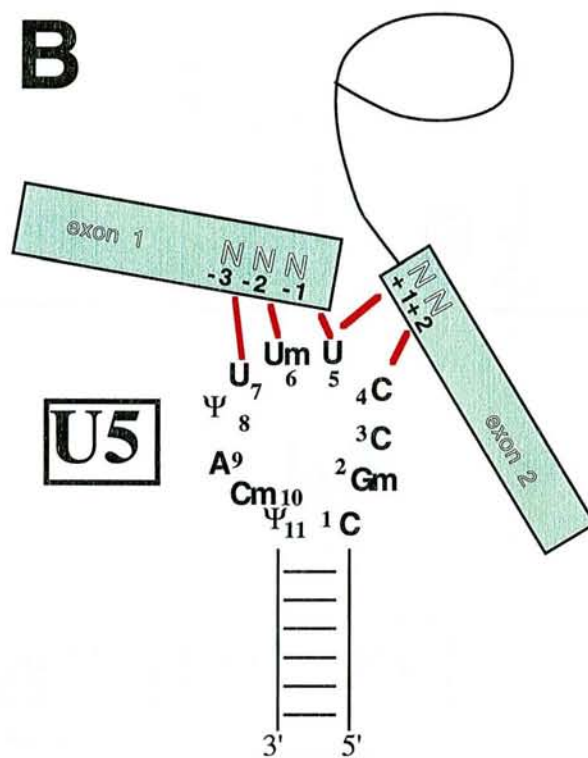
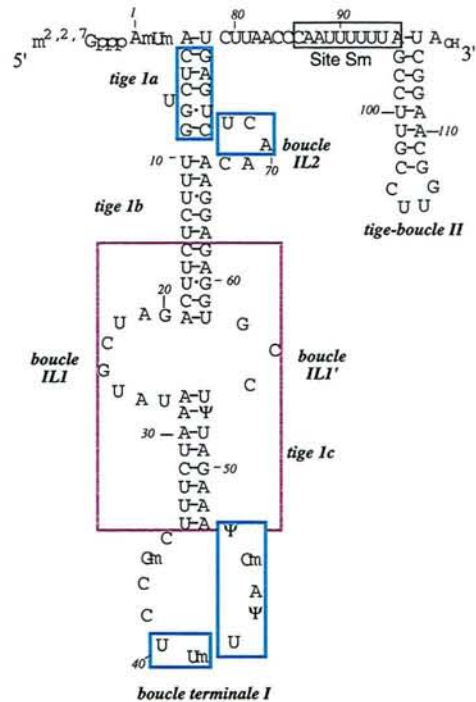
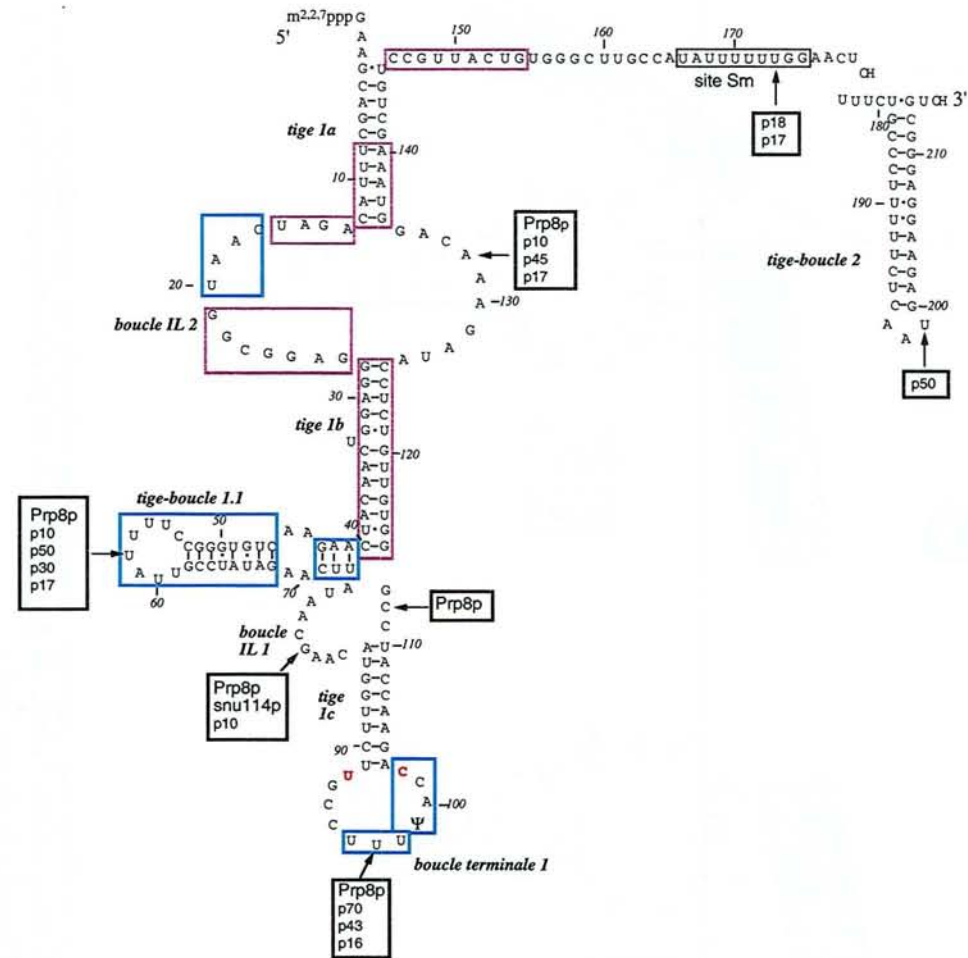
A**B**

Figure 10: Mise en évidence par pontage covalent des interactions hétérologues, entre le snRNA U5 et les deux exons, mises en jeu au cours de la réaction d'épissage par les spliceosomes U2-dépendants de vertébrés .

Les panneaux A et B correspondent à des interactions détectées entre le snRNA U5 et les exons avant la première réaction d'épissage (A) ou avant la deuxième réaction d'épissage (B).

Les traits rouges entre la boucle I conservée du snRNA U5 et les exons indiquent les résidus pontés (pour revues, Madhani et Guthrie, 1994b; Nilsen, 1998).

A**snRNA U5 humain****B****snRNA U5 de *S. cerevisiae*****Figure 11: Association du snRNA U5 avec les protéines spécifiques de la tri-snRNP [U4/U6.U5].**

Les régions des snRNA U5 humain (A) ou de *S. cerevisiae* (B), totalement ou partiellement protégées dans les tri-snRNP [U4/U6.U5] purifiées, sont, respectivement, en violet et en bleu (d'après Mougin *et al.*, résultats non publiés). Les régions 3' terminales de chaque UsnRNA n'ont pas été étudiées.

Dans le panneau B, les sites de pontages covalents entre les protéines de la snRNP U5 et le snRNA U5 de *S. cerevisiae* sont indiqués par des flèches, (Dix *et al.*, 1998).

Les nucléotides de la boucle du snRNA U5 de levure, non conservés dans les snRNA U5 des autres espèces, sont indiqués en rouge.

l'exception des levures *S. cerevisiae* et *S. pombe*. Dans ces deux cas, la boucle présente des variations de séquence aux première et dernière positions (Fig. 11) (Patterson et Guthrie, 1987; Szkukalek *et al.*, 1995; Gu *et al.*, 1996).

Il avait été proposé que cette boucle permette l'alignement des deux exons pendant la réaction d'épissage et intervienne dans la sélection des sites 3' et 5' d'épissage. En effet, en utilisant un intron portant la mutation G→A en position 1, mutation qui inhibe la réaction d'épissage chez *S. cerevisiae* (Newman *et al.*, 1985; Fouser et Friesen, 1986; Vijayraghavan *et al.*, 1986), l'équipe de Newman a pu sélectionner un gène suppresseur. Il s'agissait du gène du snRNA U5 portant une simple substitution dans la boucle conservée. En présence de ce snRNA U5 muté, un site 5' cryptique était utilisé. Ce site de coupure était sélectionné du fait de l'existence d'une complémentarité entre la boucle terminale du snRNA U5 et la bordure exonique, située en 5' de ce site (Newman et Norman, 1991; 1992). Des expériences similaires, réalisées *in vitro* pour le système mammifère, ont conduit aux mêmes conclusions (Cortes *et al.*, 1993). De la même manière, des souches mutées dans la partie du gène codant la boucle du snRNA U5 sont capables d'épisser des introns portant des mutations du dinucléotide AG 3' terminal. Les sites 3' cryptiques utilisés étaient sélectionnés par la complémentarité de la bordure exonique, en 3' de l'intron, avec la boucle terminale du snRNA U5 (Newman et Norman, 1992).

Des confirmations de l'existence de telles interactions, dans des spliceosomes assemblés autour d'introns non mutés, ont été obtenues, chez *S. cerevisiae* et chez les mammifères, par des expériences de pontage au psoralène et de pontage obtenu après incorporation de nucléotides photo-activables dans les séquences exoniques en bordure de l'intron. La réalisation de cinétique de pontage a permis de montrer que la boucle terminale du snRNA U5 interagit avec l'extrémité 5' de l'exon 1, avant la première étape d'épissage. Ce n'est qu'après la première étape de la réaction qu'une interaction est mise en place avec les extrémités des deux exons (Fig. 10) (Wyatt *et al.*, 1992; Sontheimer et Steitz, 1993; Newman *et al.*, 1995).

c) Interaction du snRNA U5 avec les protéines de la tri-snRNP [U4/U6.U5]

La boucle du snRNA U5 n'est que partiellement protégée dans la tri-snRNP humaine et de *S. cerevisiae* (Fig. 11), ce qui permet son interaction avec les deux exons.

En plus de la séquence Sm qui fixe les protéines Sm, la boucle IL2 dans le snRNA U5 humain semble jouer un rôle important pour la fixation des protéines spécifiques de la snRNP U5: les protéines de 220 kDa et 116 kDa (Hinz *et al.*, 1996; Ségault *et al.*, 1999), la protéine de 220 kDa étant au cœur d'un complexe regroupant plusieurs protéines de la snRNP U5 (les protéines de 40, 116 et 200 kDa) (Achsel *et al.*, 1998). La délétion de la boucle IL2 du snRNA U5 de cellules HeLa inhibe l'épissage *in vitro* (Hinz *et al.*, 1996; Ségault *et al.*, 1999).

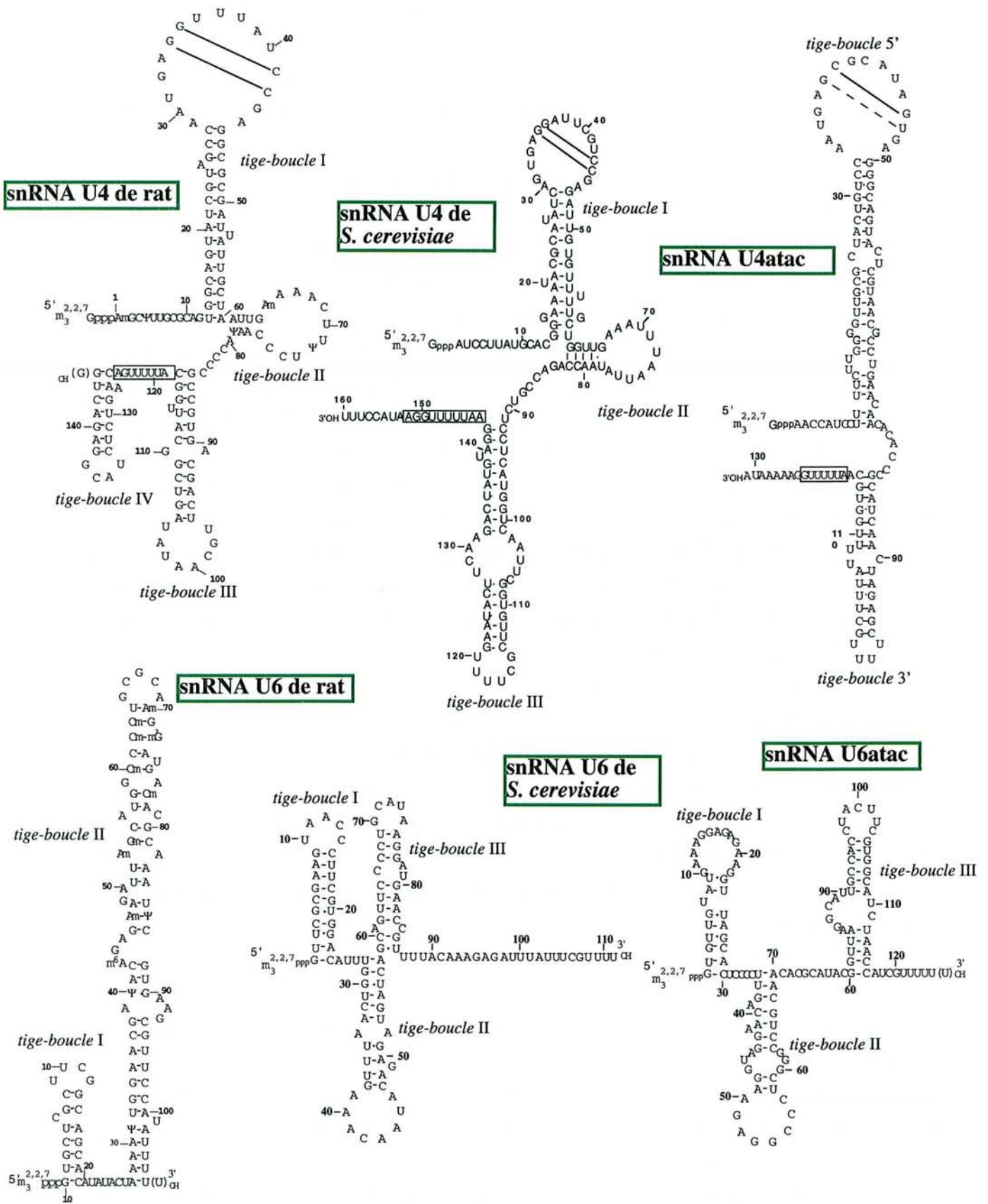


Figure 12: Structures des snRNA U4 et U6 de rat et de la levure *S. cerevisiae* et des snRNA U4atac et U6atac de cellules HeLa.

Ces différentes structures sont extraites de Myslinski et Branlant (1991) pour les snRNA U4 de rat et de *S. cerevisiae*, de Mougin *et al.* (résultats non publiés) et de Fortner *et al.* (1994) pour les snRNA U6 de rat et de *S. cerevisiae*, respectivement, de Tarn et Steitz (1996b) pour le snRNA U6atac humain. La structure du snRNA U4atac est proposée sur la base des structures des snRNA U4 de rat et de *S. cerevisiae*.

Chez *S. cerevisiae*, l'utilisation de nucléotides photoactivables a permis de détecter des sites de pontage covalent entre le snRNA U5 et différentes protéines dont la protéine Prp8p, homologue à la protéine de 220 kDa humaine (Fig. 11) (Dix *et al.*, 1998). Les sites de pontage de Prp8p ne sont plus détectables après délétion des éléments structuraux IL1 et IL2. Ces deux éléments sont les seuls nécessaires à la fixation stable de la protéine Prp8p au snRNA U5 (Dix *et al.*, 1998). Cependant, d'autres sites de pontage au snRNA U5 ont été détectés (Fig. 11), impliquant que cette protéine recouvre une large partie de cet ARN. Il est intéressant de noter que la protéine Prp8p peut être pontée à la boucle terminale I du snRNA U5. Chez l'homme, il a été montré que cette boucle est importante pour la fixation de la protéine de 220 kDa au snRNA U5 (Hinz *et al.*, 1996). Donc, en plus de son rôle dans l'alignement des deux exons lors de l'épissage, cette boucle intervient également dans la fixation de protéines de la snRNP U5.

2.2.4. Les snRNA U4 et U6 et les snRNA U4atac et U6atac

a) Les séquences primaires et secondaires des snRNA U4, U4atac, U6 et U6atac

Les snRNA U4 et U6 sont, parmi les cinq UsnRNA spliceosomaux, les plus conservés en taille, et ceci même chez la levure *S. cerevisiae* (Tableau 2). Alors que le snRNA U4 est peu conservé en séquence d'une espèce à l'autre, le snRNA U6 est le plus conservé de tous les UsnRNA spliceosomaux (Brow et Guthrie, 1988). En particulier, son domaine central présente au moins 80% d'identité de séquence, chez toutes les espèces étudiées (pour revue, Guthrie et Patterson, 1988). En dépit du fort degré de conservation en séquence primaire, des structures secondaires très différentes ont été proposées pour les snRNA U6 humain et de *S. cerevisiae* (Fig. 12) (Fortner *et al.*, 1994; Mougins *et al.*, résultats non publiés). Les snRNA U4atac et U6atac humains, identifiés très récemment, présentent 40% d'homologie de séquence avec les snRNA U4 et U6 humains (Tarn et Steitz, 1996b). La structure proposée pour le snRNA U6atac libre a plus de point commun avec celle du snRNA U6 de levure, qu'avec celle du snRNA U6 humain (Fig. 12) (Tarn et Steitz, 1996b) et une structure comparable peut être formée pour le snRNA U6atac de la plante *A. thaliana*, ceci en dépit des variations de séquence (Shukla et Padgett, 1999).

b) Les duplex d'ARN U4/U6 et U4atac/U6atac

Les variations de structure secondaire des snRNA U6 reflètent le fait que les différentes formes actives de cet ARN ne correspondent pas à la forme libre. En effet, très tôt, il a été montré que les snRNA U4 et U6 étaient associés au sein d'une même UsnRNP (Bringmann *et al.*, 1984; Hashimoto et Steitz, 1984). Différents modèles d'interaction ont été

proposés, aucun n'était compatible avec les variations de séquences entre espèces. Ce n'est qu'avec la détermination de la séquence des snRNA U4 et U6 de *S. cerevisiae* qu'un modèle d'interaction entre ces deux ARN a pu être établi (Brow et Guthrie, 1988). Il a été validé par des comparaisons phylogénétiques (Brow et Guthrie, 1988), des expériences de mutagenèse (Hamm et Mattaj, 1989; Pikielny *et al.*, 1989; Bindereif *et al.*, 1990; Vankan *et al.*, 1990; Shannon et Guthrie, 1991; Wolff et Bindereif, 1992) et des études par des sondes chimiques et enzymatiques (Mougin *et al.*, résultats non publiés). L'interaction proposée comporte deux hélices intermoléculaires, I et II, séparées par une structure tige-boucle du snRNA U4 (Fig. 13). L'hélice I avait déjà été mise en évidence antérieurement par des expériences de pontage au psoralène (Rinke *et al.*, 1985). Chacune des hélices est requise pour la formation du duplex U4/U6 et donc pour l'épissage, à la fois chez les vertébrés et chez *S. cerevisiae* (Hamm et Mattaj, 1989; Pikielny *et al.*, 1989; Bindereif *et al.*, 1990; Vankan *et al.*, 1990; Shannon et Guthrie, 1991; Wolff et Bindereif, 1992).

Le duplex d'ARN U4/U6 humain est instable après extraction phénolique. L'équipe de Brow (Brow et Vidaver, 1995) et notre équipe ont montré que ceci était dû à la possibilité d'un appariement interne au snRNA U6 au sein du duplex (entre les séquences A et A', Fig. 13). Cet appariement, retrouvé dans la structure du snRNA U6 libre, initie la dissociation du duplex U4/U6, lorsqu'il est formé dans ce dernier. Ce type d'appariement n'existe pas dans la forme libre du snRNA U6 de *S. cerevisiae*, d'où la plus grande stabilité du duplex d'ARN U4/U6 chez cette levure.

Sur la base d'expériences de pontage et de mutagenèse dirigée, une structuration similaire a pu être proposée pour le duplex d'ARN U4atac/U6atac, malgré la faible conservation des séquences primaires entre, respectivement, les snRNA U4 et U4atac et les snRNA U6 et U6atac (Fig. 13) (Tarn et Steitz, 1996b; Shukla et Padgett, 1999). Il est à noter que le snRNA U6atac présente 2 structures tige-boucles à son extrémité 3' dans le duplex, contrairement au snRNA U6, et que le snRNA U4atac, comme le snRNA U4 de *S. cerevisiae*, ne porte qu'une structure tige-boucle en 3'.

Sur la base de comparaisons de séquences, notre équipe a proposé l'existence d'une troisième hélice (hélice III). La possibilité de la former est conservée phylogénétiquement, aussi bien dans le duplex d'ARN U4/U6 que dans le duplex d'ARN U4atac/U6atac (Fig. 13) (Jakab *et al.*, 1997). L'utilisation de sondes enzymatiques et chimiques pour l'analyse des duplex d'ARN, au sein d'extraits nucléaires et dans des tri-snRNP [U4/U6.U5] purifiées de cellules HeLa ou des tri-snRNP [U4atac/U6atac.U5] semi-purifiées, a montré que l'existence de cette hélice III n'est effective que dans le cas du duplex U4atac/U6atac, l'hélice étant présente dans la tri-snRNP des spliceosomes U12-dépendants (A. Mougin *et al.*, résultats non publiés). Il reste maintenant à définir le rôle de cette troisième hélice intermoléculaire.

c) *L'interaction des snRNA U6 et U6atac avec le site 5' d'épissage*

Lors de la formation du spliceosome actif, l'interaction U4/U6 est dissociée au profit d'une interaction U2/U6, comme nous le verrons plus loin. Actuellement, le snRNA U4 est surtout considéré comme un transporteur du snRNA U6 qui, lui, jouerait un rôle durant la catalyse.

Par des techniques de pontage, la formation d'un appariement de bases, lors de l'assemblage du spliceosome actif, a été mise en évidence entre la séquence ACA, située en 5' de la séquence conservée ACAGAG du snRNA U6, et les résidus aux positions 5 et 6 de l'intron chez les mammifères, et aux positions 4, 5 et 6 de l'intron chez *S. cerevisiae*. Cet appariement serait maintenu au cours des deux étapes de la réaction d'épissage (Fig. 5) (Sawa et Shimura, 1992; Wassarman et Steitz, 1992b; Sawa et Abelson, 1992; Sontheimer et Steitz, 1993; Kim et Abelson, 1996). L'interaction entre le snRNA U6 et le site 5' met en jeu des nucléotides de l'intron qui, au début de l'assemblage des spliceosomes, sont appariés avec le snRNA U1 (Fig. 5). Aussi bien chez la levure que chez les mammifères, le caractère essentiel pour la réaction d'épissage de cette interaction U6/extrémité 5' de l'intron a pu être mis en évidence par création de mutations dans les séquences impliquées (Fig. 5) (Kandels-Lewis et Séraphin, 1993; Lesser et Guthrie, 1993; Hwang et Cohen, 1996).

Plus récemment, il a été proposé, toujours sur la base de résultats obtenus par mutagenèse, que l'appariement entre le snRNA U6 et le site 5' soit plus étendu que l'appariement précédemment proposé chez *S. cerevisiae*. Ainsi, des interactions de type non Watson et Crick prolongeraient l'appariement de bases classique et mettraient en jeu les résidus G₅₀ et G₅₂ du snRNA U6 et les résidus en position +3 et +1 de l'intron (Fig. 5) (Luukkonen et Séraphin, 1998a).

Par homologie avec l'interaction U6/site 5', une interaction du même type a été proposée entre le snRNA U6atac et la région 5' terminale des introns U12-dépendants (Tarn et Steitz, 1996b). Elle met en jeu les 5 résidus en 5' de la séquence AGGAAGA, conservée entre l'homme et *A. thaliana* (Tarn et Steitz, 1996b; Shukla et Padgett, 1999), et les positions +4 à +8 des introns U12-dépendants (Fig. 5). Cette interaction a pu être démontrée par formation de pontages covalents et son caractère indispensable pour l'épissage par les spliceosomes U12-dépendants a ensuite été prouvé par des expériences de mutagenèse compensatoire entre le snRNA U6atac et le site 5' (Fig. 5) (Yu et Steitz, 1997; Incorvaia et Padgett, 1998). On peut noter, une fois encore, l'analogie entre l'hélice formée entre le snRNA U6atac et le site 5' des introns U12-dépendants, d'une part, et entre la région 5' des introns de *S. cerevisiae* et le snRNA U6 de cet organisme, d'autre part. Ces hélices comportent 5 et 6 appariements de bases, contrairement aux deux appariements de bases mis en jeu entre la région 5' des introns U2-dépendants et le snRNA U6 chez l'homme. Donc, les appariements formés sont beaucoup

plus stables dans le cas des spliceosomes U12-dépendants et des spliceosomes de *S. cerevisiae*, que dans le cas des spliceosomes U2-dépendants de vertébrés, situation que nous avons déjà observée pour l'interaction U2 ou U12/boîte de branchement.

d) *L'interaction du duplex d'ARN U4/U6 avec les protéines qui lui sont associées*

En plus des séquences des snRNA U4 et U6 impliquées dans la formation du duplex d'ARN U4/U6, de la séquence de reconnaissance des régions 5' des introns du snRNA U6 et du site Sm du snRNA U4, d'autres éléments de ces deux ARN ont un rôle fonctionnel important. Ils interviennent dans la fixation de facteurs protéiques. Ainsi, la structure tige-boucle 5' du snRNA U4 est très conservée (76% de conservation entre *S. cerevisiae* et mammifères) (Brow et Guthrie, 1988). Chez *S. cerevisiae*, il a été montré que cette structure n'est pas nécessaire à la stabilité de la snRNP U4/U6, mais est indispensable à sa fonction (Bordonné *et al.*, 1990; Wersig et Bindereif, 1990). Un snRNA U4 de *S. cerevisiae*, dans lequel la structure tige-boucle 5' est délétée, ne fixe plus la protéine Prp4p, qui reconnaît directement ou indirectement les 65 nucléotides à l'extrémité 5' du snRNA U4 dans le duplex d'ARN U4/U6 (Xu *et al.*, 1990b; Bordonné *et al.*, 1990), ce qui inhibe la formation de la tri-snRNP [U4/U6.U5] (Bordonné *et al.*, 1990). De plus, les 15 résidus à l'extrémité 5' du snRNA U4 sont aussi indispensables à l'association de Prp3p à la snRNP U4/U6 (Anthony *et al.*, 1997). La structure tige-boucle 5' du snRNA U4 est également le site de fixation de la protéine Snu13p chez *S. cerevisiae*, ou de son homologue, la protéine de 15,5 kDa, chez l'homme (Fig. 13) (Notrott *et al.*, soumis).

Comme nous l'avons vu, de nombreuses régions des UsnRNA sont impliquées dans des interactions avec l'ARN pré-messager et/ou avec d'autres UsnRNA. Ces régions sont, pour la plupart, conservées phylogénétiquement et, comme nous le verrons dans la deuxième partie de l'introduction, contiennent en général de nombreuses modifications post-transcriptionnelles. De plus, les UsnRNA sont associés avec des protéines et, là encore, certains nucléotides impliqués dans les interactions UsnRNA/protéines sont modifiés.

C/ LES UsnRNP

Les UsnRNA spliceosomaux ne sont pas biologiquement actifs en tant que tels, mais associés à un certain nombre de protéines avec lesquelles ils forment des complexes ribonucléoprotéiques, appelés UsnRNP. Comme nous l'avons déjà mentionné, les UsnRNA spliceosomaux constituent quatre UsnRNP distinctes: les snRNP U1, U2, U5 et U4/U6, ces

deux dernières étant associées sous la forme d'une tri-snRNP. La production des UsnRNP matures actives implique plusieurs étapes, dont certaines sont concomitantes: maturation des extrémités 5' et 3' des UsnRNA, modifications post-transcriptionnelles de certains résidus et fixation des protéines. Ce processus est complexe et seules certaines étapes ont déjà été bien caractérisées.

Dans cette partie de l'introduction, nous donnerons une vision globale du processus. Nous détaillerons plus particulièrement les aspects concernant l'identité des protéines des UsnRNP, l'état des connaissances sur leurs fonctions et sur le processus d'assemblage. La partie concernant les nucléotides modifiés des UsnRNA, leur identité et leur genèse sera décrite dans la revue insérée dans la deuxième partie de l'introduction.

1. BIOGENESE DES UsnRNP

La biogenèse des UsnRNP est un processus complexe, ayant lieu à la fois dans le cytoplasme et dans le noyau chez les vertébrés (pour une revue récente, voir Mattaj et Englmeier, 1998). Les snRNA U1, U2, U4 et U5, transcrits par l'ARN polymérase II, sont tout d'abord exportés dans le cytoplasme grâce à la présence de la coiffe m⁷G, qui est reconnue par un complexe protéique spécifique. Dans le cytoplasme, ils s'associent aux protéines Sm. Les particules ribonucléoprotéiques formées sont appelées pré-particules. La formation de ces pré-particules est requise pour l'hyperméthylation de la coiffe dans le cytoplasme, à la position 2 du résidu G terminal (Mattaj, 1986; Plessel *et al.*, 1994). La coiffe triméthylée m₃^{2,2,7}G ainsi générée et la structure formée par les protéines Sm, associées au site Sm, sont les déterminants de la reconnaissance des pré-particules par le système récepteur permettant leur réimport dans le noyau (Fischer et Lührmann, 1990; Hamm *et al.*, 1990a; Fischer *et al.*, 1991). Après retour dans le noyau, les protéines spécifiques à chaque UsnRNP sont associées. Leur fixation repose, à la fois, sur des interactions avec les UsnRNA, avec les protéines Sm et sur des interactions entre protéines spécifiques. La snRNP U4/U6, puis la particule [U4/U6.U5], sont assemblées. Le snRNA U6, transcrit par l'ARN polymérase III, reste dans le noyau et pourrait transiter dans le nucléole (Ganot *et al.*, 1999). A l'heure actuelle, on ne sait pas si la biogenèse des UsnRNPs chez *S. cerevisiae* procède par les mêmes étapes que chez les vertébrés.

De même, le processus de biogenèse des UsnRNP des spliceosomes U12-dépendants n'est pas connue. Cependant, les caractéristiques communes des coiffes, portées par les UsnRNA des deux classes de spliceosomes, laissent supposer que l'assemblage des deux types d'UsnRNP procède par le même mécanisme. Les snRNP U11 et U12 sont trouvées dans le noyau sous forme d'une particule unique, la snRNP U11/U12.

Les compositions protéiques des UsnRNP des spliceosomes U2-dépendants de vertébrés ont été intensivement étudiées, notamment après purification de chacune des UsnRNP, à partir d'un extrait nucléaire de cellules HeLa. Ces purifications sont réalisées par chromatographie d'affinité, en utilisant des anticorps dirigés contre la structure coiffe 5' des UsnRNA, suivie d'autres types de chromatographies (pour revues, Lührmann *et al.*, 1990; Will et Lührmann, 1997). L'analyse des UsnRNP des spliceosomes U12-dépendants de vertébrés a débuté plus récemment (Will *et al.*, 1999). Les études sur la composition protéique des UsnRNP de *S. cerevisiae* sont également récentes. Cependant, grâce à la combinaison d'approches génétiques et biochimiques, la composition en protéines de l'ensemble des UsnRNP de *S. cerevisiae* est maintenant connue.

2. LES PROTEINES COMMUNES

2.1. Les protéines Sm

2.1.1. Nature et conservation phylogénétique

Très tôt, des protéines communes aux UsnRNP U1, U2, U5 et U4/U6 ont été mises en évidence. Ces protéines sont les antigènes majeurs d'anticorps (anti Sm) produits lors de maladies auto-immunes, comme le lupus érythémateux. Sur le plan biochimique, ces protéines ont d'abord été caractérisées par purification à partir de cellules HeLa puis clonage chez *E. coli*. Il s'agit de huit protéines, de poids moléculaires compris entre 9 et 29 kDa: les protéines B, B', D1, D2, D3, E, F et G (Tableau 3) (pour revue, Lührmann *et al.*, 1990). Les protéines B et B' sont produites par épissage alternatif d'un même gène (Chu et Elkon, 1991). Une autre protéine commune de 69 kDa a été identifiée. Elle se fixe seulement de manière transitoire à la pré-particule et pourrait donc jouer un rôle dans l'assemblage et/ou le transport de cette particule (Hackl *et al.*, 1994; Hackl et Lührmann, 1996).

La présence des protéines Sm dans les snRNP U11/U12 et [U4atac/U6atac.U5] a été révélée par le fait que ces UsnRNP sont immunoprécipitées par les anticorps anti Sm (Montzka et Steitz, 1988; Tarn et Steitz, 1996b). La récente purification de la snRNP U11/U12 de cellules HeLa (Will *et al.*, 1999) montre que les mêmes protéines Sm sont présentes dans les spliceosomes U2- et U12-dépendants.

Plus récemment, chez *S. cerevisiae*, 7 protéines homologues ont été caractérisées (Tableau 3). Leurs gènes ont été identifiés par homologie de séquences avec celles des protéines Sm de vertébrés. En effet, la présence de deux enchaînements d'acides aminés, les motifs Sm1 et Sm2, est partagée par toutes les protéines Sm (Fig. 14) (Cooper *et al.*, 1995, Hermann *et al.*, 1995, Séraphin, 1995).

mammifères	Masse moléculaire en kDa
protéines Sm	
G	9
F	11
E	12
D1	16
D2	16,5
D3	18
B	28
B'	29

<i>S. cerevisiae</i>	Masse moléculaire en kDa	Références
protéines Sm		
SmG	9	<i>Séraphin, 1995</i>
SmF	10	<i>Séraphin, 1995</i>
SmE	10	<i>Bordonné et Terassov, 1996</i>
SmD1	16	<i>Rymond, 1993</i>
SmD2	13	<i>Neubauer et al., 1997</i>
SmD3	11	<i>Roy et al., 1995</i>
SmB	22	<i>Gottschalk et al., 1998</i>
-	-	-

<i>S. cerevisiae</i>	Masse moléculaire en kDa	Références
protéines Lsm		
Lsm2p/Smx5p	11	<i>Gottschalk et al., 1999</i>
Lsm3p/Smx4p	10	<i>Séraphin, 1995</i>
Lsm4p/Uss1p	21	<i>Cooper et al., 1995</i>
Lsm5p (YER146w)	10	<i>Gottschalk et al., 1999</i>
Lsm6p (YDR378c)	14	<i>Gottschalk et al., 1999</i>
Lsm7p (YNL147w)	12	<i>Gottschalk et al., 1999</i>
Lsm8p (YJR022w)	15	<i>Pannone et al., 1998</i>

Tableau 3: Les protéines Sm et Lsm chez les mammifères et chez *S. cerevisiae*.
Toutes ces protéines contiennent les motifs Sm1 et Sm2.

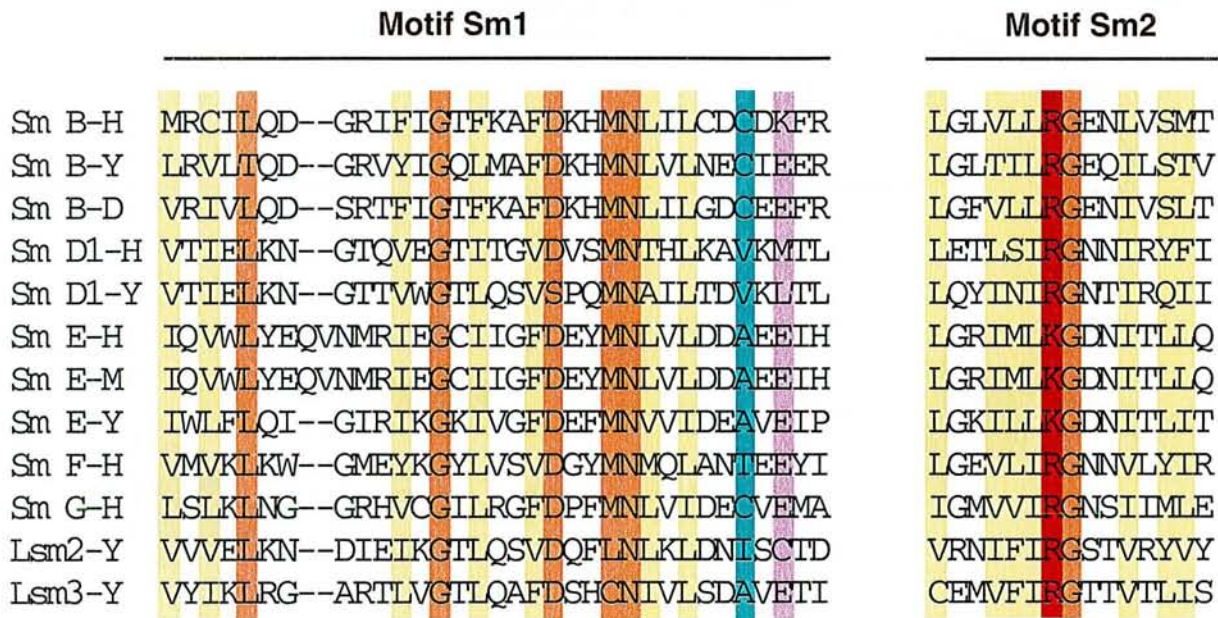
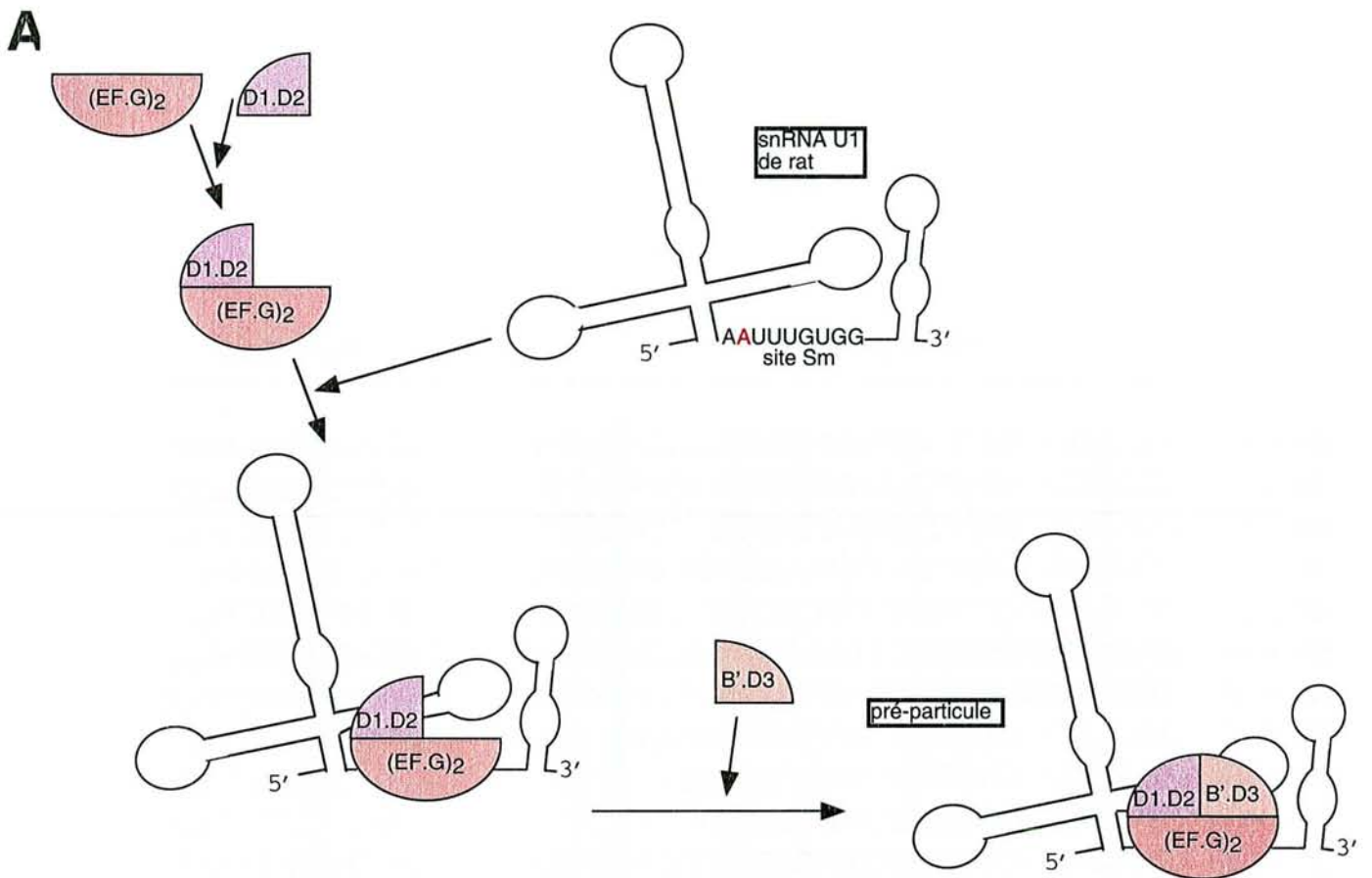


Figure 14: Alignement des motifs Sm1 et Sm2 des protéines Sm et Lsm de différentes espèces.

D'après les alignements de Cooper *et al.*, 1995 et Séraphin, 1995.

Chaque séquence est précédée du nom de la protéine et de l'organisme dont elle est issue, les annotations suivantes sont utilisées pour désigner les espèces: H, *H. sapiens*; M, *M. musculus*; D, *D. melanogaster*; Y, *S. cerevisiae*. Les acides aminés strictement conservés sont indiqués en orange. Lorsqu'une propriété physico-chimique est conservée à une position donnée, le code couleur suivant est utilisé: jaune, acides aminés hydrophobes; rouge, acides aminés basiques; violet, acides aminés acides et vert, acides aminés hydrophiles.



B

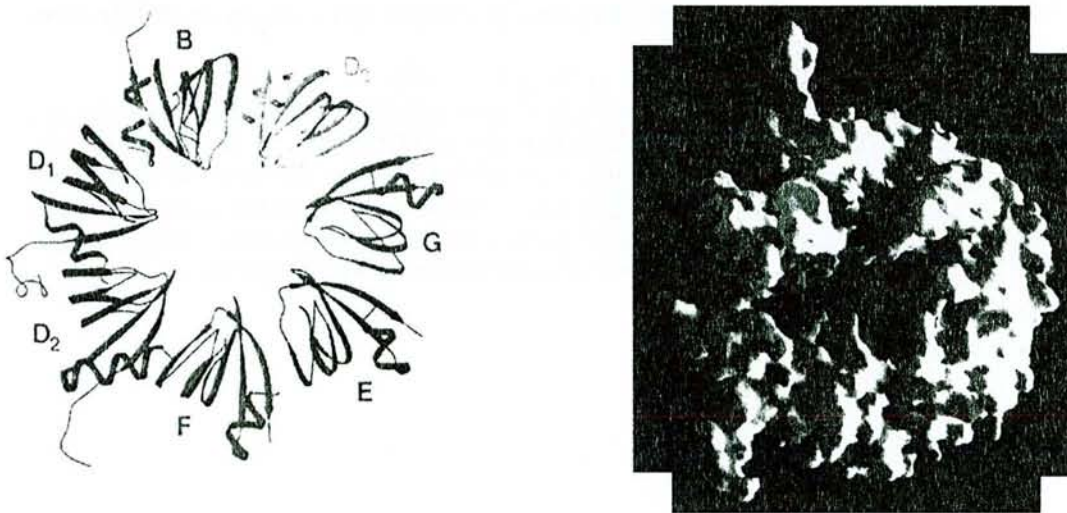


Figure 15: Assemblage des pré-particules

Le panneau A représente les différentes étapes de formation des pré-particules (d'après Raker *et al.*, 1996). Le site Sm du snRNA U1 de rat est indiqué. L'adénosine du site Sm, réactive au DMS après formation de la première particule, est marquée en rouge (Hartmuth *et al.*, 1999).

Le panneau B représente un modèle d'interaction des protéines Sm dans la pré-particule. Le site Sm serait en contact avec la région centrale de l'anneau formé par les protéines Sm. Ce schéma est extrait de la revue de Kambach *et al.* (1999b), d'après Kambach *et al.* (1999a).

2.1.2. Les protéines Sm sont essentielles pour la constitution de spliceosomes actifs

Des expériences de reconstitution de snRNP U2 et U5 de mammifères, *in vitro*, à partir des ARN et des protéines, démontrent que les protéines Sm sont interchangeable entre particules mais qu'en leur absence il n'est pas possible de former des UsnRNP fonctionnelles (Ségault *et al.*, 1995; Will *et al.*, 1996). Le caractère essentiel des gènes codant les protéines Sm D1, D3, E et G a été montré chez *S. cerevisiae* (Roy *et al.*, 1995; Rymond, 1993; Séraphin, 1995; Bordonné et Tarassov, 1996). Une diminution du taux d'expression des protéines D1 et D3 de *S. cerevisiae* conduit à une diminution du taux des snRNA U1, U2, U4 et U5 dans la cellule et à une sous-méthylation de la coiffe 5' de ces UsnRNA (Roy *et al.*, 1995; Rymond, 1993). Le grand degré de conservation des protéines Sm a été démontré par le fait que les protéines humaines D1 et E peuvent compenser l'absence de leurs homologues chez *S. cerevisiae* (Rymond *et al.*, 1993; Bordonné et Tarassov, 1996).

2.1.3. La formation des pré-particules

L'ordre d'addition des protéines Sm de cellules HeLa aux UsnRNA a pu être établi *in vivo* (Fisher *et al.*, 1985) et, par la suite, de façon plus exhaustive, *in vitro* (Raker *et al.*, 1996). Deux complexes protéiques stables sont formés entre les protéines D1 et D2 (Lehmeier *et al.*, 1994), d'une part, et entre les protéines E, F et G (Raker *et al.*, 1996), d'autre part. Ensuite, ces deux complexes s'associent aux UsnRNA pour former une première particule (Fig. 15A). Enfin, les protéines B/B' et D3 viennent s'associer sous forme de complexe (Hermann *et al.*, 1995). Ce schéma de formation de la pré-particule semble également valable chez la levure, car des interactions protéine/protéine ont été mises en évidence, entre les protéines E et F, E et G et les protéines D3 et B (Camasses *et al.*, 1998). Très récemment, la structure 3D des complexes D1/D2 et D3/B a été résolue, par diffraction aux rayons X (Kambach *et al.*, 1999a). Les quatre protéines ont une structuration commune: une courte hélice α côté N-terminal, suivie de 7 brins β anti-parallèles. Les brins β_1 , β_2 et β_3 forment le motif Sm1 et les brins β_4 et β_5 le motif Sm2 (pour revue, Kambach *et al.*, 1999b). Ces protéines sont liées 2 à 2 par l'intermédiaire d'interactions hydrophobes (Camasses *et al.*, 1998) entre le brin β_4 d'une protéine et le brin β_5 de l'autre protéine (Kambach *et al.*, 1999a). Ainsi, comme cela avait été proposé antérieurement, le motif Sm2 est impliqué dans la liaison entre les protéines Sm (Hermann *et al.*, 1995). Sur la base des structures 3D établies, un modèle d'interaction entre les 8 protéines Sm a été proposé. Selon ce modèle, les protéines Sm forment un anneau, dont le centre basique constituerait le site de fixation à la séquence Sm des UsnRNA (Fig. 15B).

2.1.4. Reconnaissance entre protéines Sm et site Sm

En dépit des avancées sur la structure des protéines Sm, le mécanisme de reconnaissance entre protéines Sm et site Sm reste encore largement inconnu. Une étude génétique, *in vivo*, du site Sm du snRNA U5 de levure, a montré qu'en dépit de la grande conservation de séquence du site Sm au cours de l'évolution, la création de mutations au sein de la série de résidus U a peu d'effet sur la fixation des protéines Sm (Jones et Guthrie, 1990). Ces résultats suggèrent qu'il existe peu de contacts spécifiques entre les protéines Sm et leur site de fixation. Cette hypothèse est confortée par le fait que, dans une snRNP U1 de cellules HeLa reconstituée *in vitro*, seule la protéine G peut être pontée au site Sm (Heinrichs *et al.*, 1992). La structure en anneau proposée pour les protéines Sm (Fig. 15B) pourrait expliquer la mise en jeu d'interactions essentiellement de type électrostatique entre le site Sm et les protéines Sm (Kambach *et al.*, 1999a).

L'interaction des protéines Sm au site Sm a un effet marqué sur les propriétés chimiques des bases des nucléotides du site Sm. En effet, des travaux menés au laboratoire, en collaboration avec l'équipe de R. Lührmann, ont permis de mettre en évidence un changement de sensibilité vis-à-vis du DMS de la séquence Sm des UsnRNA de vertébrés, après fixation des protéines Sm. Ainsi, de manière inattendue, la position N₇ du résidu A, localisé en 5' du site Sm, est méthylée après fixation des protéines Sm (Hartmuth *et al.*, 1999). La modification d'un résidu A par le DMS n'est pas attendue sur le plan chimique. L'interaction ARN/protéine modifie donc l'électronégativité du résidu A. D'après des essais réalisés sur un snRNA U1 muté, cette sensibilité est une signature d'un assemblage correct des protéines Sm au site Sm. En effet, des snRNA U1, portant des mutations empêchant l'assemblage correct avec les protéines Sm et donc le réimport des pré-particules U1 dans le noyau d'œufs de *X. laevis*, ne présentaient pas l'hyperréactivité au DMS (Hartmuth *et al.*, 1999).

Les sites Sm des différents UsnRNA ne sont pas forcément interchangeables *in vivo*: le site Sm du snRNA U5 de *X. laevis* peut remplacer le site Sm du snRNA U1 du même organisme, mais le site Sm du snRNA U1 n'est pas fonctionnel dans le snRNA U5 (Jarmolowski et Mattaj, 1993). Ceci peut s'expliquer par des interactions additionnelles entre les protéines Sm et des séquences autres que le site Sm des UsnRNA. En effet, l'utilisation de sondes enzymatiques et chimiques, sur des pré-particules de vertébrés naturelles ou reconstituées *in vitro*, montre la protection, par les protéines Sm, de fragments des UsnRNA comprenant le site Sm, mais également de régions plus éloignées comme la région 5' de l'hélice IV du snRNA U1 ou la boucle IL2 du snRNA U5 (Liautard *et al.*, 1982; Hartmuth *et al.*, 1999; A. Mougouin *et al.*, résultats non publiés). Donc, des régions distinctes du site Sm dans les snRNA U1, U2, U4 et U5 pourraient être impliquées dans des interactions avec les protéines Sm.

2.1.5. Les protéines « Sm like »

Récemment, en utilisant les motifs signatures Sm1 et Sm2 des protéines Sm, une nouvelle classe de protéines Sm a été découverte. Neuf de ces protéines, appelées Lsm (pour "Sm like"), ont été détectées chez *S. cerevisiae* (Séraphin, 1995; Fromont-Racine *et al.*, 1997; Salgado-Garrido *et al.*, 1999).

Sept de ces protéines sont présentes dans la tri-snRNP [U4/U6.U5] de *S. cerevisiae* (Tableau 3) (Gottschalk *et al.*, 1999; Stevens et Abelson, 1999) et constitueraient un complexe protéique spécifique du snRNA U6 (Salgado-Garrido *et al.*, 1999; Mayes *et al.*, 1999). Les protéines Lsm sont nécessaires à la stabilité du snRNA U6 et à la croissance cellulaire (Cooper *et al.*, 1995; Pannone *et al.*, 1998; Mayes *et al.*, 1999). La découverte, par approche génétique, d'une interaction directe ou indirecte entre Lsm8p et Lh1p, la protéine chaperonne associée à tous les transcrits produits par l'ARN polymérase III, semble indiquer que Lsm8p se fixerait au snRNA U6, à une étape précoce de l'assemblage de la snRNP U6 (Pannone *et al.*, 1998). Elle est vraisemblablement la première protéine de la UsnRNP à se fixer, protégeant de la dégradation le snRNA U6 et aidant au recrutement des autres protéines Lsm. Sept protéines Lsm, homologues aux protéines Lsm de levure, ont également été mises en évidence dans la tri-snRNP [U4/U6.U5] humaine. D'après des études par microscopie électronique, les protéines Lsm formeraient une structure en anneau, similaire à la structure formée par les protéines Sm classiques. Elles se fixeraient à la séquence polyU à l'extrémité 3' du snRNA U6 humain (Achsel *et al.*, soumis).

Deux des protéines Lsm détectées chez la levure ne sont pas associées au snRNA U6. Par ailleurs, des protéines de type Lsm ont aussi été mises en évidence chez les *Archaea*, ce qui suggère un rôle additionnel de ces protéines dans la cellule, en plus de leur rôle dans l'épissage (Salgado-Garrido *et al.*, 1999). Il a été proposé que cet ensemble de protéines intervienne dans les mécanismes de dégradation des ARNm (Boeck *et al.*, 1998; Mayes *et al.*, 1999).

2.2. Le complexe SMN/SIP1

L'assemblage des pré-particules est un processus complexe qui doit impliquer des facteurs externes participant à la mise en place de ces particules et au contrôle du processus. Un de ces facteurs semble être la protéine de 69 kDa, citée auparavant. Des études récentes suggèrent que les protéines SMN et SIP1 (pour Survival of Motor Neurons et SMN Interacting Protein 1, respectivement) joueraient un rôle dans ce processus d'assemblage, ainsi que dans la régénération des UsnRNP spliceosomales. Les protéines SMN et SIP1 forment un complexe. Ces deux protéines sont présentes à la fois dans le cytoplasme et dans le noyau (Liu et Dreyfuss, 1996; Liu *et al.*, 1997a).

Les protéines SMN et SIP1 sont associées aux UsnRNP dans le cytoplasme. Elles semblent intervenir dans l'assemblage des pré-particules dans ce compartiment cellulaire (Fischer *et al.*, 1997; Pellizzoni *et al.*, 1998). Elles pourraient participer à un mécanisme de contrôle de l'assemblage, ce qui éviterait l'import dans le noyau de particules non fonctionnelles.

Dans le noyau, elles sont concentrées dans des régions particulières, au niveau de structures appelées "gems". Ces "gems" sont, elles-mêmes, associées à d'autres structures nucléaires, les "coiled bodies", contenant les UsnRNP spliceosomales et un certain nombre d'autres protéines (pour revue, Huang et Spector, 1997). Dans ces structures "gems", les protéines SMN et SIP1 appartiennent à un complexe contenant également les protéines Sm (Liu *et al.*, 1997a). Des études *in vivo* chez *X. laevis* et des études par épissage *in vitro* suggèrent que la protéine SMN est requise pour l'épissage des ARN pré-messagers nucléaires, probablement en intervenant dans la régénération ou le recyclage des UsnRNP dans le noyau (Pellizzoni *et al.*, 1998). En effet, il a été proposé que les UsnRNP existent sous deux formes dans le noyau, les formes active et inactive, et que la conversion des formes inactives en formes actives requiert la présence de la protéine SMN (Pellizzoni *et al.*, 1998).

Chez l'homme, la délétion ou la mutation du gène télomérique de SMN est la cause de la majorité des cas d'atrophie musculaire spinale (SMA), une maladie génétique dégénérative des neurones moteurs (Lefebvre *et al.*, 1995). La SMA est donc liée à un défaut d'activité de la protéine SMN. Il pourrait s'agir d'un défaut de production d'ARN messagers matures dans les neurones moteurs (Pellizzoni *et al.*, 1998).

Aucun homologue de la protéine SMN n'a été détecté dans le génome de *S. cerevisiae* (Li *et al.*, 1997). Par contre, une homologie existe entre la protéine humaine SIP1 et la protéine Brr1 de *S. cerevisiae*, qui est également associée aux UsnRNA. Il a été montré que cette protéine est requise pour la formation des UsnRNP, mais pas pour leur stabilité. Elle est liée génétiquement à la protéine SmD1 (Noble et Guthrie, 1996). La délétion du gène codant Brr1 n'est pas létale, mais engendre une thermosensibilité de la souche (Noble et Guthrie, 1996).

3. MATURATION DES EXTREMITES 3' DES UsnRNA ET FORMATION DE LA PRE-PARTICULE

Les ARN transcrits par l'ARN polymérase II portent en général des extensions en 3' qui sont éliminées par un processus de maturation. Les homologues de la RNase III bactérienne, RNT1 chez *S. cerevisiae* et PAC 1 chez *S. pombe*, sont impliquées dans le processus. En effet, il a été montré que l'enzyme PAC1 joue un rôle dans la formation des extrémités 3' des snRNA U2, U4 et U5 (Potashkin et Frenthewey, 1990; Rotondo *et al.*, 1995).

D'après des études *in vitro*, chez *S. cerevisiae*, l'enzyme produirait des clivages à des sites spécifiques au sein des précurseurs des UsnRNA, générant des intermédiaires, ensuite convertis en UsnRNA matures, par l'action d'exonucléases (Chanfreau *et al.*, 1997; Abou Elela et Ares, 1998; Seipelt *et al.*, 1999). Une ARN hélicase ATP dépendante, SEN 1, serait liée à ce processus de maturation (Ursic *et al.*, 1997).

La fixation des protéines Sm et l'hyperméthylation de la coiffe semblent être des événements concomitants avec la formation des extrémités 3' matures du snRNA U5 de *S. cerevisiae* (Chanfreau *et al.*, 1997). L'assemblage de la pré-particule serait nécessaire à la maturation des extrémités 3' des UsnRNA de levure. En effet: i) le site Sm, et au moins, la protéine SmD1 seraient indispensables à la maturation de l'extrémité 3' du snRNA U1 (Seipelt *et al.*, 1999), ii) l'inactivation de la protéine Brr1 conduit à un ralentissement de la conversion du précurseur du snRNA U2 vers sa forme mature (Noble et Guthrie, 1996). Par contre, des mutations générées dans le site Sm du snRNA U5 entraînent la production de transcrits plus courts que le snRNA U5S *in vivo* (Jones et Guthrie, 1990). Une explication possible est que la fixation des protéines Sm bloque l'action d'exonucléases et définit l'extrémité de l'ARN (Chanfreau *et al.*, 1997).

4. LES PROTEINES SPECIFIQUES A CHAQUE UsnRNP

4.1. Les domaines fonctionnels des protéines spécifiques de chaque UsnRNP

On peut répertorier différents types de domaines fonctionnels dans les protéines des UsnRNP. Ils peuvent être regroupés en trois catégories:

⇒ *les domaines de fixation à l'ARN*

En général, ces domaines sont du type RRM pour "RNA Recognition Motif" (ou nommé également RBD pour "RNA Binding Domain") (pour revue, Nagai, 1996). Le domaine RRM localisé à l'extrémité N-terminale de la protéine U1A de la snRNP U1 humaine, a été le premier domaine de ce type à être mis en évidence et à être caractérisé. La structure 3D de ce domaine seul ou en interaction avec le snRNA U1 a été établie, à la fois par diffraction des rayons X et par RMN (Nagai *et al.*, 1990; Oubridge *et al.*, 1994; pour revue, Nagai, 1996). Depuis, ce domaine a été mis en évidence dans beaucoup de protéines se liant à l'ARN, dont de nombreuses protéines des UsnRNP. Il consiste en une séquence d'environ 80 acides aminés qui contient deux petites régions très conservées, RNP1 (RGFGFVXF) et RNP2 (LFVGNL), situées dans un ensemble de quatre feuillets β anti-parallèles en contact avec l'ARN. Ce domaine renferme deux hélices α sur la face opposée à l'ARN.

<i>S. cerevisiae</i>		Mammifères		Type de motif ARN hélicase	Étapes d'intervention dans la réaction d'épissage	Fonctions putatives
Nom de la protéine	UsnRNP à laquelle elle appartient	Nom de la protéine	UsnRNP à laquelle elle appartient			
?		UAP 56	/	DECD	1ère	addition de la snRNP U2
Prp5p	/	?		DEAD	1ère	addition de la snRNP U2
Prp28p	/	100 kD	U5	DEAD	1ère	addition de la tri-snRNP/transition U1/site 5' → U6/site 5'
Brr2/Snu246p/ Rss2/Slt22	[U4/U6.U5]	200 kD	U5	DEIH pour Brr2, DExH pour 200kDa*	1ère	formation des interactions U2/U6, rupture des interactions U4/U6
Prp2p	/	?	/	DEAH	1ère	réarrangements structuraux avant la première réaction d'épissage
Prp16p	/	hPrp16p	/	DEAH	2ème	réarrangements structuraux avant la seconde étape/reconnaissance du site 3'
Prp22p	/	HRH1	/	DEAH	après la 2ème	relargage de l'ARN messager mature
Prp43p	/	mDEAH9	/	DEAH	après la 2ème	dissociation du spliceosome

Tableau 4: Protéines spliceosomales, présentant les motifs signatures des ARN hélicases ATP dépendantes, identifiées chez *S. cerevisiae* et chez les mammifères.

D'après la revue, Staley et Guthrie, 1998.

*La protéine de 200 kDa contient deux motifs, un motif DEIH et un motif DEVH (Lauber *et al.*, 1996).

⇒ *les domaines d'interactions protéine/protéine*

Certaines protéines des UsnRNP contiennent des domaines riches en enchaînement sérine/arginine (les domaines RS). Elles font partie de la famille des protéines SRrp (pour SR protein Related Polypeptide). Ces protéines peuvent établir des contacts protéine/protéine, par l'intermédiaire de leur domaine RS (pour revues, Fu, 1995; Manley et Tacke, 1996). Nous verrons que les domaines RS des protéines des UsnRNP permettent leur interaction avec de nombreux facteurs d'épissage et sont nécessaires à la mise en place d'un spliceosome actif.

D'autres motifs connus pour participer aux interactions protéine/protéine contiennent 7 répétitions de séquences WD (motif WD) (Neer *et al.*, 1994) ou des répétitions TPR (motif TRP, Tetratricopeptide Repeat) (Ordway *et al.*, 1994; Lamb *et al.*, 1995). Ces motifs permettent d'établir des interactions entre protéines d'une même UsnRNP. C'est le cas pour la protéine Prp4p de la snRNP U4/U6 de *S. cerevisiae*, qui contient un domaine WD, nécessaire à son interaction avec une autre protéine de la snRNP U4/U6, la protéine Prp3p (Hu *et al.*, 1994; Ayadi *et al.*, 1998). C'est également le cas des protéines Prp39p et Prp42p, qui interagiraient entre elles au sein de la snRNP U1 de *S. cerevisiae* par leur motif TPR respectif (Gottschalk *et al.*, 1998).

⇒ *les domaines intervenant dans des variations de conformation des ARN ou des protéines*

Comme nous le verrons en détail plus loin, le spliceosome est une machinerie dynamique. Chaque étape de son assemblage et chaque étape de la réaction d'épissage requiert la formation ou la dissociation d'interactions ARN/ARN, ARN/protéine et protéine/protéine. De nombreux facteurs d'épissage vont jouer un rôle dans ces transitions structurales. Certaines des protéines des UsnRNP interviennent aussi dans ces processus.

- les domaines de type ARN hélicase

La plupart des étapes conduisant à la formation d'un spliceosome actif nécessitent l'hydrolyse de l'ATP. Chacune de ces étapes ATP dépendantes requiert une ou plusieurs protéines de type ARN hélicase ATP dépendante. Toutes portent les séquences en acides aminés caractéristiques de ce type de protéines, en particulier le motif conservé DExD/H et le motif de fixation à l'ATP, G/A(x)₄GKT/S (pour revue, Chiara *et al.*, 1996). Les différentes ARN hélicases identifiées, intervenant lors de l'épissage, sont indiquées dans le Tableau 4. La plupart sont des facteurs d'épissage non liés aux UsnRNP, seules trois d'entre elles font partie intégrante d'une UsnRNP.

Les premières ATPases ARN dépendantes mises en évidence étaient Prp2p, Prp5p, Prp16p et la protéine Snu24p de la snRNP [U4/U6.U5] de *S. cerevisiae* (Schwer et Guthrie, 1991; Kim et Lin, 1993; O'Day *et al.*, 1996; Xu *et al.*, 1996). Très récemment, il a été montré que les protéines Prp22p, Prp16p et la protéine de 200 kDa de la snRNP U5 humaine ont la capacité d'induire la dissociation de duplex d'ARN *in vitro* (Wagner *et al.*, 1998; Schwer et Gross, 1998; Wang *et al.*, 1998; Laggerbauer *et al.*, 1998). Il est donc vraisemblable que ces

Mammifères	Poids moléculaire en kDa	Fixation directe à l'ARN	Type de domaine	Fonction	Références	<i>S. cerevisiae</i>	Poids moléculaire en kDa	Fixation directe à l'ARN	Type de domaine	Phénotype après délétion	Fonction	Références
U1-70K	70	+	RRM, RS	assemblage du spliceosome	<i>Kohtz et al., 1994</i>	Snp1p	34	+	RRM	létal	?	<i>Smith et Barell, 1991</i>
U1A	34	+	RRM	?	<i>Scherly et al., 1991</i>	Mud1p	34	+	RRM	viable	configuration active de la snRNP U1	<i>Liao et al., 1993</i>
U1C	22	-	doigt à zinc	assemblage du spliceosome	<i>Heinrichs et al., 1990</i> <i>Nelissen et al., 1991</i> <i>Jamison et al., 1995</i> <i>Will et al., 1996</i>	yU1C	27			létal	assemblage du spliceosome	<i>Tang et al., 1997;</i> <i>Neubauer et al., 1997</i>
						Snu56p/Mud10p	56	+		létal	assemblage du spliceosome	<i>Gottschalk et al., 1998</i>
						Nam8p/Mud15p	57	+	3 RRM	viable	stabilité de la snRNP, assemblage du spliceosome	<i>Gottschalk et al., 1998</i>
						Snu65p/Mud16p/Prp42p	65	-	5 TPR	létal	stabilité de la snRNP, assemblage du spliceosome	<i>Gottschalk et al., 1998;</i> <i>McLean et Rymond, 1998</i>
						Prp39p	75	-	6 TPR	létal	stabilité de la snRNP, assemblage du spliceosome	<i>Lockhart et Rymond, 1994;</i> <i>McLean et Rymond, 1998</i>
						Prp40p	69	-	2 WWP	létal	1ère étape	<i>Kao et Siliciano, 1996</i>
						Snu71p	71	-	RS	létal	?	<i>Gottschalk et al., 1998</i>
						Npl3p	45	+	2RRM+RS (protéine SR)	létal	?	<i>Gottschalk et al., 1998</i>
						Exm2p/Luc7p/Snu30p	30	-	GAR doigt à zinc	létal	?	<i>Gottschalk et al., 1999;</i> <i>Rigaut et al., 1999</i> <i>Fortes et al., soumis</i>

Tableau 5: Protéines spécifiques des snRNP U1 de mammifères et de *S. cerevisiae*.

Les protéines homologues des snRNP U1 humaine et de *S. cerevisiae* sont indiquées sur une même ligne.

TPR: Tetratric Peptide Repeat

RRM: RNA Recognition Motif

protéines catalysent la mise en place et/ou la rupture d'interactions ARN/ARN au cours du processus d'assemblage et de catalyse du spliceosome (pour revue, Staley et Guthrie, 1998).

- les domaines d'hydrolyse du GTP

Un seul membre de la famille des GTPases a été recensé parmi les protéines connues des UsnRNP. Il s'agit de la protéine de 116 kDa de la snRNP U5 humaine et de son homologue chez *S. cerevisiae*, la protéine Snu114p (Fabrizio *et al.*, 1997). Il est à noter que ces protéines présentent des homologies de séquence avec le facteur d'élongation de la traduction EF2 (Fabrizio *et al.*, 1997), requis pour la translocation de l'ARNm dans les ribosomes (Abel et Jurnak, 1996). L'intégrité du motif de fixation du GTP de la protéine Snu114p est essentielle à la croissance, indiquant que la fixation du GTP, et probablement son hydrolyse, sont nécessaires à la fonction des protéines Snu114p/116 kDa (Fabrizio *et al.*, 1997).

- les domaines à activité proline isomérase

Une protéine de la tri-snRNP [U4/U6.U5] humaine, la protéine de 20 kDa, appartient à la famille des cyclophilines, catalysant l'isomérisation des résidus proline (Horowitz *et al.*, 1997; Teigelkamp *et al.*, 1998). Cette observation suggère donc que la protéine de 20 kDa joue un rôle dans des variations de conformation de protéines lors de l'association du spliceosome et/ou la catalyse de la réaction d'épissage.

Toutes les protéines identifiées dans les différentes UsnRNP de *S. cerevisiae* et de cellules HeLa sont présentées dans les Tableaux 5 à 8. Les motifs protéiques spécifiques de chacune de ces protéines sont indiqués.

4.2. La snRNP U1

La snRNP U1 de mammifères a un coefficient de sédimentation de 12S et renferme 3 protéines spécifiques, 70K, U1A et U1C (Hinterberger *et al.*, 1983) (Tableau 5). La protéine 70K joue, comme nous le verrons, un rôle important dans les étapes précoces de l'assemblage du spliceosome. La protéine U1C stabilise le complexe formé entre la snRNP U1 et l'ARN pré-messager (Heinrichs *et al.*, 1990; Jamison *et al.*, 1995).

La purification récente de la snRNP U1 de *S. cerevisiae* (Fabrizio *et al.*, 1994; Tang *et al.*, 1997; Neubauer *et al.*, 1997) a montré qu'elle comporte un nombre beaucoup plus élevé de protéines que la snRNP U1 humaine (Tableau 5). Des expériences de délétion de gènes ont montré que seules deux de ces protéines ne sont pas nécessaires à la croissance cellulaire, dont Mud1p, l'homologue de la protéine U1A humaine (Gottschalk *et al.*, 1998; Liao *et al.*, 1993). En général, les protéines de la snRNP U1 de *S. cerevisiae* semblent avoir un rôle stabilisateur de l'association snRNP U1-site 5' (Tableau 5) (Tang *et al.*, 1997; Gottschalk *et al.*, 1998; McLean et Rymond, 1998).

mammifères	Poids moléculaire en kDa	Fixation directe à l'ARN	Type de domaine	Fonction	Références	<i>S. cerevisiae</i>	Poids moléculaire en kDa	Fixation directe à l'ARN	Type de domaine	Phénotype après délétion	Fonction	Références	
B''	28,5	nécessite A'	2RRM	Fixation SF3a au snRNA U2	<i>Scherly et al., 1990; Boelens et al., 1991; Fresco et al., 1991; Brosi et al., 1993a; Krämer et al., 1999</i>	yU2B''/MSL1	12,8	+	RRM	viable		<i>Tang et al., 1996</i>	
A'	31		riche en Leu										
	33												
	35												
	92												
SAP61/SF3a ⁶⁰	60	+	doigt à zinc	Association de la snRNP U2 à l'ARN pré-messager	<i>Brosi et al., 1993a et b, Behrens et al., 1993b; Staknis et Reed, 1994a; Champion-Arnaud et Reed, 1994; Krämer et al., 1999</i>	Prp9	63	-	motif leucines répétées 2 doigts à zinc	létal		<i>Lin et al., 1987; Abovich et al., 1990; Chapon et Legrain, 1992; Legrain et al., 1993; Arena et Abelson, 1993; Ruby et al., 1993; Wells et Ares, 1994; Rain et al., 1996; Wells et al., 1996; Fromont-Racine et al., 1997; Igel et al., 1998; Chen et al., 1998; Gottschalk et al., 1999</i>	
SAP62/SF3a ⁶⁶	SF3a 66	-	doigt à zinc			Prp11	SF3a 29,9	-	doigt à zinc	létal			
SAP114/SF3a ¹²⁰	110	-	motifs SURP (famille SWAP)			Prp21/Spp91P	33	-	motifs SURP (famille SWAP)	létal			
SAP49/SF3b ⁵⁰	53	+	2 RRM Pro/Gly			Hsh49p	24,5	+	2 RRM	létal	Association de la snRNP U2 à l'ARN pré-messager		
SAP130/SF3b ¹³⁰	120					Snup154p/Rse1p	154			?			
SAP145/SF3b ¹⁴⁵	SF3b 150					Cus1p	SF3b 50	+	domaine superenroulé (coiled-coil domain)	létal			
SAP155/SF3b ¹⁵⁵	160					?							

Tableau 6: Les protéines spécifiques des snRNP U2 de mammifères et de *S. cerevisiae*.

Les protéines homologues des snRNP U2 humaine et de *S. cerevisiae* sont indiquées sur une même ligne.

RRM: RNA Recognition Motif

SWAP: Suppressor of White Apricot Protein

Les snRNP U1 humaine et de *S. cerevisiae* se distinguent donc, à la fois, par la taille de leur UsnRNA et par le nombre de protéines spécifiques qu'elles contiennent. Une explication possible de cette différence est que le rôle joué par les facteurs de type protéines SR dans l'assemblage des spliceosomes humains, soit joué par des protéines directement liées à la snRNP U1 chez *S. cerevisiae*. Cela pourrait être le cas des protéines Snu71p et Npl3p (associée de manière transitoire à la snRNP U1) qui possèdent un domaine de type RS (Gottschalk *et al.*, 1998).

4.3. La snRNP U2

La composition en protéines de la snRNP U2 humaine purifiée dépend de la concentration en sels utilisée pour sa purification. A haute concentration saline, seules deux protéines spécifiques U2A' et U2B'' sont présentes (Behrens *et al.*, 1993a) (Tableau 6). A faible concentration saline, une UsnRNP 17S, contenant au moins 12 protéines, est obtenue. Sept de ces protéines correspondent aux protéines des facteurs d'épissage SF3a et SF3b (Brosi *et al.*, 1993b, pour revue, Krämer, 1996) (Tableau 6). Comme nous l'avons vu, la majorité des protéines, constituant ces deux facteurs, sont associées à l'extrémité 5' conservée du snRNA U2 qui porte de nombreuses modifications post-transcriptionnelles (Fig. 7) (Behrens *et al.*, 1993a; pour revue, Will et Lührmann, 1997). La fixation du facteur SF3a à l'ARN précède la fixation du facteur SF3b (Brosi *et al.*, 1993b). Ces deux facteurs sont requis pour la fixation de la snRNP U2 à la boîte de branchement de l'ARN pré-messager (Krämer et Utans, 1991; Brosi *et al.*, 1993a; Behrens *et al.*, 1993b). Toutes les protéines des facteurs SF3, à l'exception de SAP 130, sont en contact direct ou indirect avec la région située en aval de la boîte de branchement de l'ARN pré-messager (Staknis et Reed, 1994a; Champion-Arnaud et Reed, 1994; Gozani *et al.*, 1994; 1996) et ces interactions entre les facteurs SF3a et SF3b et l'ARN pré-messager sont indispensables à la fixation de la snRNP U2 sur l'ARN pré-messager (Gozani *et al.*, 1996).

Chez *S. cerevisiae*, l'homologue du facteur SF3a est constitué par un complexe formé entre les protéines Prp9p, Prp11p et Prp21p (Legrain *et al.*, 1993; Legrain et Chapon, 1993; Ruby *et al.*, 1993; Wells et Ares, 1994; Yan et Ares, 1996; Rain *et al.*, 1996) et l'homologue du facteur SF3b par un complexe formé par les protéines Cus1p, Hsh49p et Snu154p (Wells *et al.*, 1996; Igel *et al.*, 1998; Gottschalk *et al.*, 1999) (Tableau 6). Des liens génétiques ont été mis en évidence entre ces complexes protéiques et les régions du snRNA U2 correspondant aux structures tige-boucles IIa et IIb et à la région immédiatement en amont de la séquence complémentaire à la boîte de branchement (Fig. 7) (Ruby *et al.*, 1993; Yan et Ares, 1996; Wells et Ares, 1994). Chez *S. cerevisiae*, les facteurs SF3a et SF3b sont aussi indispensables à l'association de la snRNP U2 à l'ARN pré-messager (Abovich *et al.*, 1990; Ruby *et al.*, 1993; Wells *et al.*, 1996), ceci en association avec des facteurs protéiques non associés à la snRNP

U2: une ARN hélicase ATP dépendante, Prp5p (Ruby *et al.*, 1993; Wells et Ares, 1994, Yan et Ares, 1996) et un facteur d'épissage non essentiel à la croissance, Cus2p (Yan *et al.*, 1998).

4.4. La snRNP U6

Aucune protéine, à part les protéines Lsm, ne semble liée de manière durable au snRNA U6. Chez *S. cerevisiae*, la protéine Prp24 est associée au snRNA U6 dans sa forme libre et transitoirement au duplex d'ARN U4/U6 (Shannon et Guthrie, 1991; Ghetti *et al.*, 1995), mais n'est pas présente dans la tri-snRNP [U4/U6.U5] (Gottschalk *et al.*, 1999; Stevens et Abelson, 1999). Cette protéine semble être un facteur de recyclage du duplex d'ARN U4/U6. En effet, associée au snRNA U6, elle participerait après chaque réaction d'épissage à la reformation du duplex d'ARN U4/U6 (Ghetti *et al.*, 1995; Raghunathan et Guthrie, 1998). Prp24p est dissociée au moment de l'association entre la snRNP U4/U6 et la snRNP U5 (Raghunathan et Guthrie, 1998). Une autre protéine, non associée aux UsnRNP, Sad1, interviendrait, elle, dans la formation des duplex d'ARN U4/U6, à partir de snRNA U4 nouvellement synthétisés (Lygerou *et al.*, 1999).

4.5. La snRNP U4/U6

Les protéines Prp3p et Prp4p sont les seules protéines de la snRNP U4/U6 identifiées, à ce jour, chez *S. cerevisiae* (Tableau 7) (Banroques et Abelson, 1989; Bjorn *et al.*, 1989; Anthony *et al.*, 1997). Prp3p et Prp4p interagissent entre elles et cette interaction est nécessaire à la viabilité des cellules (Ayadi *et al.*, 1998). Ces protéines sont également présentes dans la tri-snRNP [U4/U6.U5] et sont requises pour la première réaction de transestérification (Banroques et Abelson, 1989; Bjorn *et al.*, 1989; Anthony *et al.*, 1997). Prp4p est requise pour l'association de la snRNP U4/U6 avec la snRNP U5 (Banroques et Abelson, 1989; Bjorn *et al.*, 1989) et les deux protéines sont nécessaires à l'association de la tri-snRNP au pré-spliceosome (Banroques et Abelson, 1989; Bordonné *et al.*, 1990; Anthony *et al.*, 1997). Prp4p pourrait être seulement associée de façon transitoire au pré-spliceosome, et pourrait intervenir dans le relargage de la snRNP U4 avant la première réaction d'épissage (Ayadi *et al.*, 1997).

Récemment, les homologues humains des protéines Prp3p et Prp4p, les protéines de 90 et 60 kDa, ont été identifiées (Tableau 7) (Lauber *et al.*, 1997; Horowitz *et al.*, 1997)

4.6. La snRNP U5

Neuf protéines spécifiques de la snRNP U5 humaine ont été identifiées après purification de cette UsnRNP (Bach *et al.*, 1989; pour revue, Will et Lührmann, 1997)

(Tableau 7). La région du snRNA U5, protégée par les protéines, est de taille limitée par rapport au nombre de ces protéines (Fig. 11), ce qui impliquerait de fortes interactions protéine/protéine au sein de cette UsnRNP. En effet, des complexes stables, en absence d'ARN, sont formés entre plusieurs de ces protéines (Achsel *et al.*, 1998). La snRNP U5 de *S. cerevisiae* n'a pas été purifiée individuellement, seule la tri-snRNP [U4/U6.U5] a été purifiée. Seules 3 protéines spécifiques de la snRNP U5 de *S. cerevisiae* ont été caractérisées à ce jour (Tableau 7).

Les protéines identifiées dans la snRNP U5 humaine ont des rôles fonctionnels divers et importants pour l'activité du spliceosome. En effet, trois d'entre elles (les protéines de 100, 116 et 200 kDa) possèdent des domaines structuraux spécifiques (Tableau 7), dont les activités pourraient être impliquées dans des changements de conformation des complexes spliceosomaux durant l'épissage des ARN pré-messagers (Lauber *et al.*, 1996; Lagerbauer *et al.*, 1996; Teigelkamp *et al.*, 1997; Fabrizio *et al.*, 1997). Les protéines de 110 kDa et 116 kDa ont été pontées au site 3' d'épissage au niveau de la séquence polypyrimidine (Chiara *et al.*, 1997). Le pontage entre la protéine de 116 kDa et une structure tige-boucle interne à l'intron bloque la deuxième réaction de transestérification, de même que l'utilisation dans des expériences d'épissage *in vitro* d'anticorps dirigés contre la protéine de 200 kDa (Liu *et al.*, 1997b; Lauber *et al.*, 1996). Cette dernière est aussi capable *in vitro* de dissocier le duplex U4/U6 et pourrait donc être également impliquée dans le relargage de la snRNP U4 avant la première réaction d'épissage (Lagerbauer *et al.*, 1998).

La protéine Prp18p, qui fait partie intégrante de la snRNP U5 de *S. cerevisiae*, et son homologue chez les mammifères hPrp18p, interviennent également lors de la deuxième réaction d'épissage (voir §D.5) (Vijayraghavan *et al.*, 1989; Vijayraghavan et Abelson, 1990; Horowitz et Abelson, 1993a; Horowitz et Krainer, 1997).

La protéine de 220 kDa et son homologue chez la levure, Prp8p, sont au cœur de la fonction de la snRNP U5. Le rôle central joué par cette protéine de la snRNP U5 est souligné par son haut degré de conservation phylogénétique (Hodges *et al.*, 1995; Lucke *et al.*, 1997). Chez l'homme, la protéine de 220 kDa forme un complexe stable avec les protéines de 49 et 116 kDa de la snRNP U5, ainsi probablement qu'avec la protéine de 200 kDa. Ce complexe protéique interagit avec l'ARN pré-messager au niveau du site 5', avant la première réaction d'épissage (Wyatt *et al.*, 1992; Teigelkamp *et al.*, 1995a; Reyes *et al.*, 1996) et au niveau du site 3' (site polypyrimidique et exon 2) après cette première étape (Teigelkamp *et al.*, 1995b; Umen et Guthrie, 1995a; Chiara *et al.*, 1997). De plus, ces protéines peuvent également être pontées au point de branchement (MacMillan *et al.*, 1994; Teigelkamp *et al.*, 1995b). Ainsi, la protéine de 220 kDa humaine, comme son homologue Prp8p, est le seul facteur protéique intervenant dans l'épissage connu pour être en contact avec les trois éléments de l'ARN pré-messager impliqués dans l'épissage. Ce facteur joue clairement un

Mammifères					<i>S. cerevisiae</i>							
Poids moléculaire en kDa	Fixation directe à l'ARN UsnRNP	Type de domaine	Fonction	Références	Poids moléculaire UsnRNP	Fixation directe à l'ARN	Type de domaine	Phénotype après délétion	Fonction	Références		
15	U5	WD		<i>Behrens et Lührmann, 1991</i>	Snu16p /Dib1p	?	16	-	létal	<i>Gottschalk et al., 1999; Stevens et Abelson, 1999; Reuter et al., soumis</i>		
40				<i>Achsel et al., 1998</i>	-							
52				<i>Behrens et Lührmann, 1991</i>	-							
100	+	motif ARN hélicase DEAD+ motif RS		<i>Teigelkamp et al., 1997</i>	Prp28p	facteur d'épissage	66,6	+	motif ARN hélicase DEAD	létal	assemblage du spliceosome	<i>Strauss et Guthrie, 1991, 1994; Staley et Guthrie, 1999</i>
102				<i>Behrens et Lührmann, 1991</i>	Prp6p	[U4/U6.U5]	104		motifs TPR répétés doigt à zinc	létal	assemblage et stabilité de la snRNP [U4/U6.U5]	<i>Abovich et al., 1990; Legrain et al., 1991; Gallison et Legrain, 1993</i>
110				<i>Behrens et Lührmann, 1991</i>								
116	+	domaine GTPase 2 ^{ème} étape (EF2"like")		<i>Fabrizio et al., 1997; Liu et al., 1997</i>	Snu114p	U5	114	+	domaine GTPase (EF2-like)	létal		<i>Fabrizio et al., 1997</i>
200	+	2 motifs ARN hélicase DExH	dissociation du duplex U4/U6, 2 ^{ème} étape	<i>Lauber et al., 1996; Lagerbauer et al., 1996</i>	Snu246p/Brr2p	U5	246	+	2 motifs ARN hélicase DExH	létal	2 ^{ème} étape	<i>Lauber et al., 1996; Noble et Guthrie, 1996; Dix et al., 1998; Lin et Rossi, 1996; Xu et al., 1996</i>
220	?		1 ^{ère} et 2 ^{ème} étapes	<i>Anderson et al., 1989</i>	Prp8p	U5	280	?	-	létal	assemblage du spliceosome, 1 ^{ère} et 2 ^{ème} étapes, dissociation du duplex U4/U6	<i>Lossky et al., 1987; Jackson et al., 1988; Whittaker et al., 1990; Brown et Beggs, 1992; Kuhn et al., 1999</i>
60	?	domaines WD		<i>Gozani et al., 1994; Lauber et al., 1997; Horowitz et al., 1997</i>	Prp4p	U4/U6	52	?	domaines WD	létal	assemblage de la tri-snRNP et du spliceosome, relargage de la snRNP U4	<i>Banroques et Abelson, 1989; Bjorn et al., 1989; Dalrymple et al., 1989; Bordonné et al., 1990; Ayadi et al., 1997</i>
90	?	-			Prp3p	U4/U6	56	?	-	létal	assemblage du spliceosome	<i>Last et Woolford, 1986; Lustig et al., 1986; Anthony et al., 1997</i>

Mammifères				<i>S. cerevisiae</i>								
Poids moléculaire en kDa	Type de domaine	Fonction	Références	UsnRNP	Poids moléculaire en kDa	Fixation directe à l'ARN	Type de domaine	Phénotype après délétion	Fonction	Références		
15,5	motif des cyclophilines [U4/U6.U5] domaine RS		<i>Behrens et Lührmann, 1991</i>		Snu13p (YDL0988c)	?	13	+	motif RRM possible	létal	<i>Gottschalk et al., 1999; Stevens et Abelson, 1999</i>	
20			<i>Horowitz et al., 1997; Teigelkamp et al., 1998</i>		-							
27			<i>Fetzer et al., 1997</i>		-							
61			<i>Behrens et Lührmann, 1991</i>		Prp31p	?	56	-		létal	assemblage du spliceosome	<i>Weidenhammer et al., 1997</i>
63			<i>Behrens et Lührmann, 1991</i>									
					Snu34p/Spp381p	[U4/U6.U5]	34		motif protéolyse PEST	viable	relargage de la snRNP U4	<i>Gottschalk et al., 1999, Lybarger et al., 1999</i>
hPrp18	facteur d'épissage	2 ^{ème} étape	<i>Horowitz et Krainer, 1997</i>		Prp18p	U5	28	-		létal	2 ^{ème} étape	<i>Vijayraghavan et al., 1989; Vijayraghavan et Abelson, 1990; Horowitz et Abelson, 1993a et 1993b; Gottschalk et al., 1999</i>
					Snu23p	[U4/U6.U5]	23		doigt à zinc C2H2	viable		<i>Gottschalk et al., 1999; Stevens et Abelson, 1999</i>
					Snu66p (YOR308c)	[U4/U6.U5]	66			viable		<i>Gottschalk et al., 1999; Stevens et Abelson, 1999</i>
					Prp38p	[U4/U6.U5]	28		riche en sérines, acide	létal	1 ^{ère} étape, dissociation du duplex U4/U6, relargage de la snRNP U4	<i>Blanton et al., 1992; Xie et al., 1998</i>

Tableau 7: Les protéines spécifiques des tri-snRNP [U4/U6.U5] de mammifères et de *S. cerevisiae*.

Les protéines homologues des tri-snRNP [U4/U6.U5] de *S. cerevisiae* et de mammifères sont indiquées sur une même ligne. Xie *et al.* (1998) ont proposé que la protéine Prp38p soit l'homologue de la protéine de 27 kDa de mammifères, mais Gottschalk *et al.* (1999) indiquent l'absence d'homologue de la protéine de 27 kDa chez *S. cerevisiae*.

Les UsnRNP auxquelles appartiennent les différentes protéines de *S. cerevisiae* (snRNP U5, U4/U6 ou [U4/U6.U5]) sont indiquées, quand elles sont connues. Les facteurs d'épissage hPrp18p et Prp28p ne sont pas associés, respectivement, aux tri-snRNP de mammifères et de *S. cerevisiae*. Elles sont les homologues respectives des protéines Prp8p et 200 kDa et jouent les mêmes rôles dans l'épissage.

TPR: Tetratricopeptide Repeat

RRM: RNA Recognition Motif

rôle central dans la coordination des événements moléculaires nécessaires à la formation d'un spliceosome actif, aussi bien pour la première étape que pour la deuxième étape de transestérification. Son action est complémentaire à celle du snRNA U5. Il peut même jouer le rôle de celui-ci lorsque la boucle terminale I est mutée (Ségault *et al.*, 1999), cette boucle et les protéines Prp8/220 kDa permettant l'alignement des deux exons et la définition du site 3'. Il a été proposé que la protéine Prp8p intervienne aussi dans la dissociation du duplex d'ARN U4/U6 (Kuhn *et al.*, 1999).

4.7. La tri-snRNP [U4/U6.U5]

La tri-snRNP [U4/U6.U5] humaine a été purifiée, à partir d'un extrait nucléaire de cellules HeLa, par chromatographie d'affinité en utilisant des anticorps spécifiques de la protéine de 100 kDa de la snRNP U5, suivie d'un fractionnement en gradient de glycérol (Behrens et Lührmann, 1991). Cette purification a permis de mettre en évidence la présence de 5 protéines spécifiques de la tri-snRNP (Tableau 7). Une de ces protéines, la protéine de 20 kDa forme un complexe stable avec les protéines de 60 et 90 kDa de la snRNP U4/U6 (Horowitz *et al.*, 1997; Teigelkamp *et al.*, 1998) et pourrait être associée à la snRNP U4/U6 avant formation de la tri-snRNP (Horowitz *et al.*, 1997). C'est une protéine de la famille des cyclophilines (Horowitz *et al.*, 1997; Teigelkamp *et al.*, 1998). Cette observation suggère un rôle possible de cette protéine dans l'assemblage de la tri-snRNP et/ou du spliceosome. La protéine de 27 kDa porte un domaine de type domaine RS. Ce domaine peut être phosphorylé *in vitro* (Fetzer *et al.*, 1997). De manière intéressante, la tri-snRNP purifiée possède une activité kinase qui phosphoryle la protéine de 52 kDa de la snRNP U5 (Behrens et Lührmann, 1991). La protéine portant cette activité n'a pas été identifiée.

La purification de la tri-snRNP [U4/U6.U5] de *S. cerevisiae* est très récente et les premières analyses de sa composition protéique indiquent la présence des protéines Sm et Lsm, ainsi que de 14 autres protéines, dont certaines avaient déjà été caractérisées en tant que protéines spécifiques des snRNP U4/U6 ou U5 (Tableau 7) (Fabrizio *et al.*, 1994; Gottschalk *et al.*, 1999; Stevens et Abelson, 1999). Certaines des protéines spécifiques de la tri-snRNP, pourraient être nécessaires à l'assemblage et/ou la stabilité de la tri-snRNP [U4/U6.U5], à l'association de cette tri-snRNP au pré-spliceosome ou, encore, à la dissociation du duplex d'ARN U4/U6 avant la première réaction d'épissage (Abovich *et al.*, 1990; Galisson et Legrain, 1993; Weidenhammer *et al.*, 1997; Xie *et al.*, 1998; Gottschalk *et al.*, 1999).

5. LES PROTEINES DES UsnRNP DES SPLICEOSOMES U12-DEPENDANTS

Peu d'informations sont disponibles sur les protéines associées aux UsnRNA des spliceosomes U12-dépendants. Des études antérieures de l'équipe de J. Steitz avaient montré que les snRNP U11 et U12 interagissent sous forme d'un complexe, cette interaction étant probablement médiée par des protéines (Wassarman et Steitz, 1992a). Très récemment, cette snRNP U11/U12 a été purifiée, en 3 étapes: chromatographie d'immunoaffinité, fractionnement en gradient de glycérol, suivie de l'utilisation d'oligonucléotides biotinylés complémentaires à des séquences des snRNA U11 ou U12 avec sélection du complexe sur billes d'agarose/streptavidine (Will *et al.*, 1999). Cette technique de purification a permis de pallier la faible concentration des snRNP U11/U12. Vingt et une protéines ont été détectées dans le complexe, dont les huit protéines Sm des UsnRNP des spliceosomes U2-dépendants (Tableau 8). Les protéines du facteur SF3a de la snRNP U2 17S étaient également présentes. Les autres protéines détectées semblent spécifiques à la snRNP U11/U12, en particulier une protéine de 35 kDa associée spécifiquement au snRNA U11, qui présente des homologies de séquence avec la protéine U1 70K de la snRNP U1.

Il a également été montré que la snRNP U5 était associée à la snRNP U4atac/U6atac, sous la forme d'une tri-snRNP contenant une large majorité des protéines de la tri-snRNP des spliceosomes U2-dépendants (R. Lührmann et coll., résultats non publiés).

Il apparaît donc que certaines des protéines des UsnRNP des spliceosomes U12-dépendants sont communes aux deux types de spliceosomes, en particulier les protéines Sm. D'autres protéines sont, par contre, spécifiques des spliceosomes U12-dépendants.

D/ LES COMPLEXES SPLICEOSOMAux ET LES INTERACTIONS ARN-ARN MISES EN JEU LORS DE L'ASSEMBLAGE ET DE L'ACTION DU SPLICEOSOME

L'assemblage du spliceosome est un processus séquentiel (Fig. 16 et 17). Différents complexes, correspondants à des intermédiaires d'assemblage et des intermédiaires de la réaction d'épissage des spliceosomes U2-dépendants de vertébrés, ont été mis en évidence après fractionnement par électrophorèse en gel de polyacrylamide non dénaturant ou par centrifugation en gradient de glycérol. Ces différents complexes ont été purifiés par chromatographie de gel filtration, suivie d'une colonne d'affinité (pour revue, Reed et Palandjian, 1997). Cette méthode de purification repose sur l'utilisation d'un ARN pré-messenger biotinylé, ce qui permet de purifier les différents complexes sur colonne d'avidine-

protéines Sm	Poids moléculaire en kDa		snRNP U11/U12 de cellules HeLa	Poids moléculaire en kDa
G	9		G	9
F	11		F	11
E	12		E	12
D1	16	=	D1	16
D2	16,5		D2	16,5
D3	18		D3	18
B	28		B	28
B'	29		B'	29

protéines spécifiques snRNP U2 de mammifères	Poids moléculaire en kDa	facteur	snRNP U11/U12 de cellules HeLa	Poids moléculaire en kDa	facteur	
B''	28,5					
A'	31					
-	33					
-	35					
-	92					
SAP61/SF3a ⁶⁰	60	SF3a				
SAP62/SF3a ⁶⁶	66					
SAP114/SF3a ¹²⁰	110					
SAP49/SF3b ⁵⁰	53	SF3b	=	-	49	
SAP130/SF3b ¹³⁰	120			-	130	SF3b
SAP145/SF3b ¹⁴⁵	150			-	150	
SAP155/SF3b ¹⁵⁵	160			-	160	

protéines spécifiques snRNP U1 de mammifères	Poids moléculaire en kDa	snRNP U11/U12 de cellules HeLa	Poids moléculaire en kDa
U1-70K	70	-	35 (homologue de U1 70K)
U1A	34		
U1C	22		
		-	20
		-	25
		-	28,5
		-	35
		-	48
		-	59
		-	64
		-	65

Tableau 8: Comparaison de la composition en protéines des snRNP U1 et U2 de mammifères avec celle de la snRNP U11/U12 de cellules HeLa.

D'après Will *et al.*, 1999.

= signifie que les protéines indiquées ont la même séquence dans la snRNP U11/U12 et dans les snRNP U1 et U2.

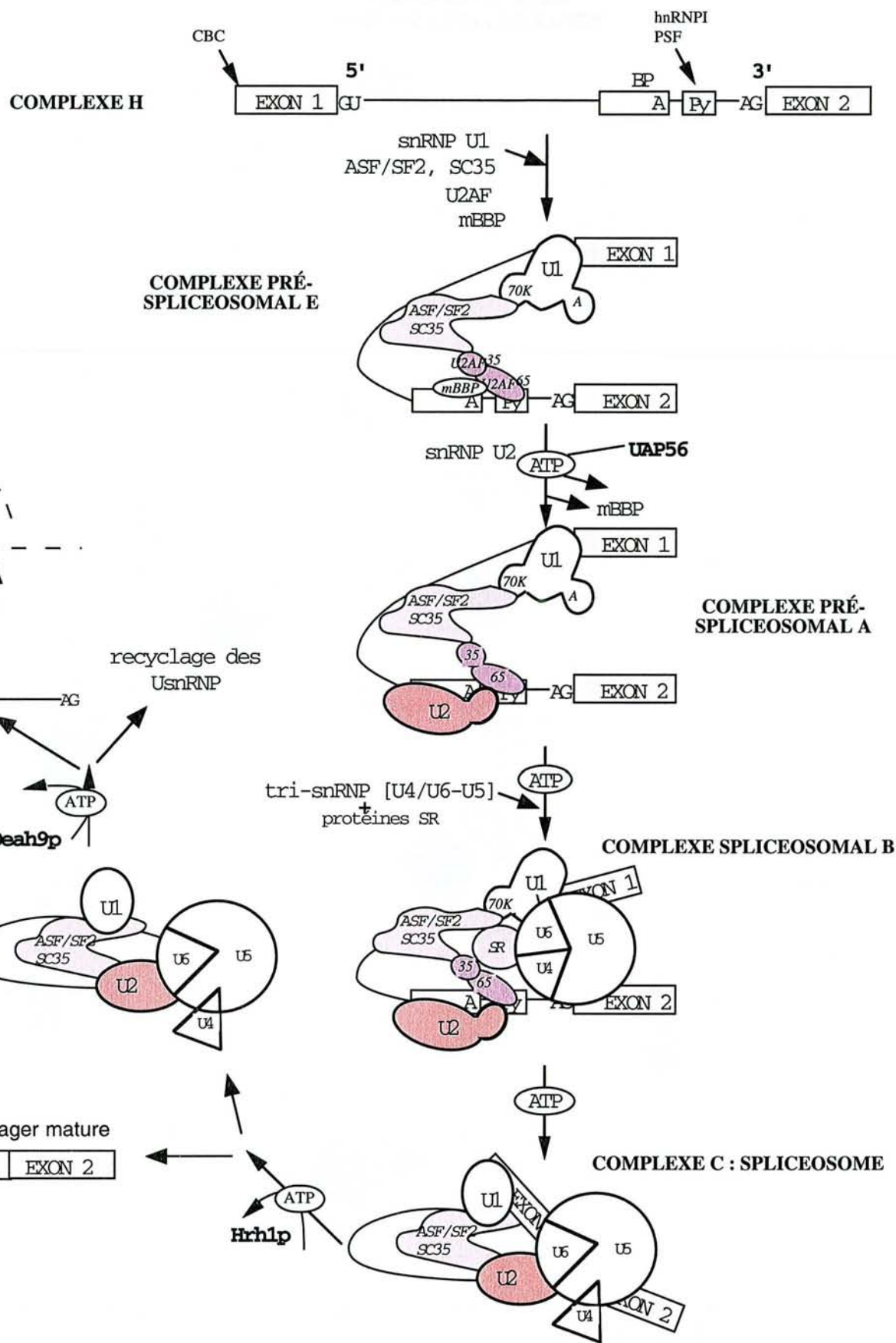


Figure 16: Les étapes d'assemblage et d'action des spliceosomes U2-dépendants de vertébrés.

D'après les revues, Reed et Paladjian (1997) et Staley et Guthrie (1998).
BP désigne la boîte de branchement.

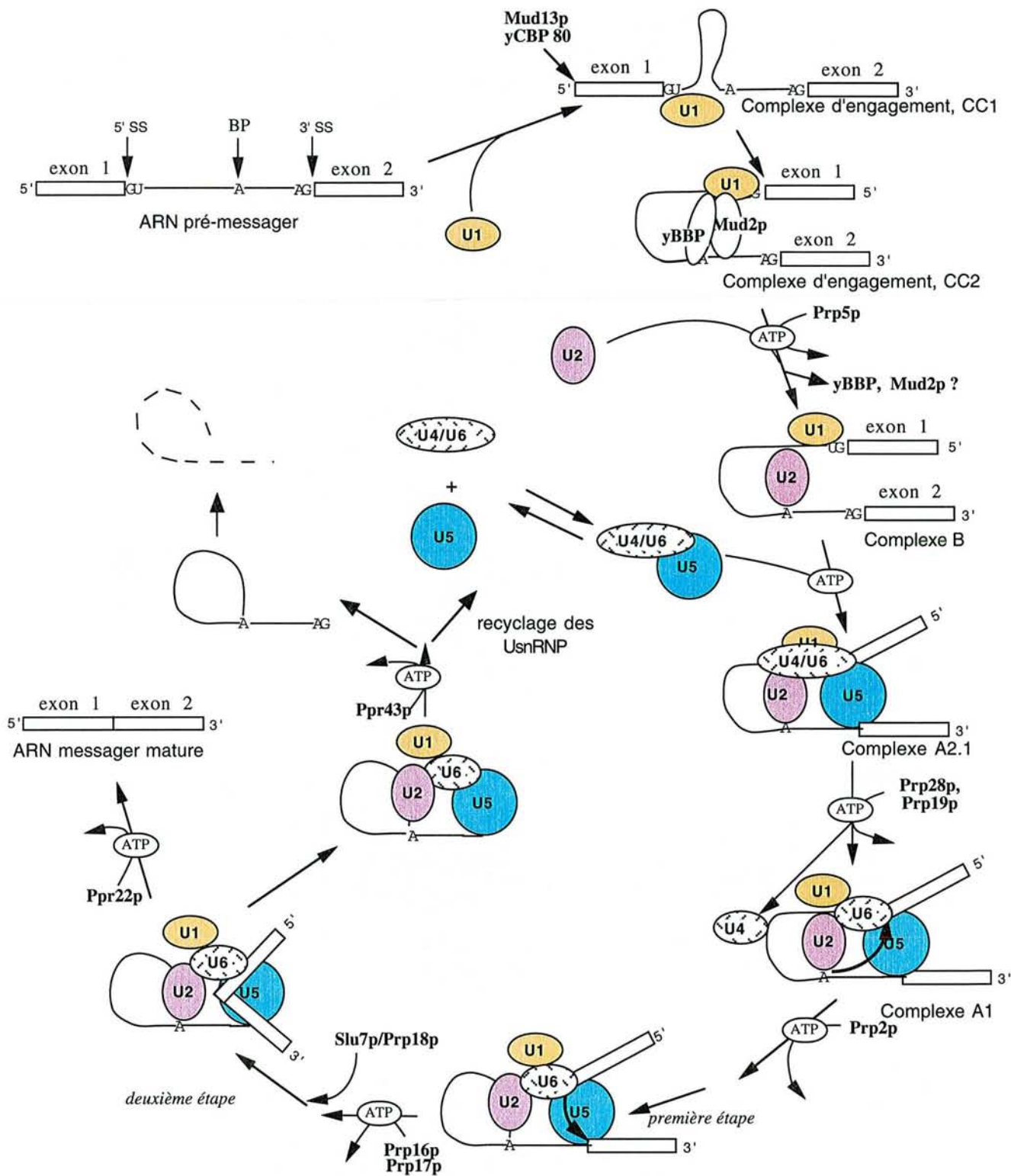


Figure 17: Représentation schématique du cycle d'assemblage, d'action et de dissociation du spliceosome chez *S. cerevisiae*.

5' ss et 3' ss représentent, respectivement, les sites 5' et 3' d'épissage. BP est la boîte de branchement. Les étapes ATP dépendantes sont indiquées par ATP encerclé. Les étapes où interviennent des protéines Prp et l'identité de celles-ci sont indiquées. Voir le texte pour les références.

agarose (Reed, 1990). Le contenu en ARN et en protéines de ces complexes spliceosomaux a ensuite été étudié. Chez la levure *S. cerevisiae*, le processus d'assemblage a été étudié par voie génétique, en utilisant des mutants des différentes protéines Prp (pour "Precursor RNA processing"). Ainsi, il est possible de bloquer, successivement et de manière dirigée, chaque étape d'assemblage du spliceosome ou chaque étape de la réaction d'épissage. Ces protéines Prp ont été caractérisées comme étant essentielles à la production d'ARNm et un nombre élevé de ces protéines est indispensable à l'assemblage du spliceosome et/ou à la réaction d'épissage (pour revue, Moore *et al.*, 1993).

L'épissage médié par les spliceosomes U12-dépendants semble également faire intervenir un processus séquentiel, puisque des complexes spliceosomaux intermédiaires, similaires à ceux obtenus pour les spliceosomes U2-dépendants, ont également été détectés par électrophorèse sur gel de polyacrylamide natif. Néanmoins, des différences existent dans les étapes précoces d'assemblage des spliceosomes U2-dépendants et U12-dépendants (Tarn et Steitz, 1996a; Frilander et Steitz, 1999).

Nous allons maintenant décrire comment les différentes interactions entre les UsnRNA et l'ARN pré-messager se mettent en place lors de l'assemblage des spliceosomes et comment les protéines des UsnRNP ou les différents facteurs d'épissage interviennent dans ce processus. Nous allons voir que, même si les interactions UsnRNA/ARN pré-messager sont essentielles pour l'épissage, souvent les protéines du spliceosome peuvent suppléer un défaut de ces interactions.

1. LE COMPLEXE H OU hnRNP

1.1. La composition en protéines

Chez les eucaryotes supérieurs, les ARN pré-messagers lors de leur transcription reflètent une grande hétérogénéité en taille et sont appelés hnRNA (pour "heterogenous Nuclear RNA"). Ils sont associés à une multitude de protéines sous forme de particules hnRNP (pour "heterogenous nuclear Ribonucleic Particle"), formant ainsi le complexe H (pour revue, Dreyfuss *et al.*, 1993; Krecic et Swanson, 1999). Ces particules hnRNP contiennent une vingtaine de protéines appelées protéines hnRNP, dont les protéines A1, A2, B1, B2, C1 et C2, qui constituent 90% de ces protéines (pour revue, McAfee *et al.*, 1997), et le complexe CBC (pour "Cap Binding Protein") composé des deux facteurs CBP20 et CBP80.

Il n'y a pas d'évidence montrant que le complexe H est un intermédiaire dans l'assemblage du spliceosome et seules quelques protéines hnRNPs sont connues pour jouer un rôle dans l'épissage, ce qui corrobore le fait que le taux de protéines hnRNP détecté dans les complexes spliceosomaux est beaucoup plus faible que dans le complexe H (Bennett *et al.*, 1992; Staknis

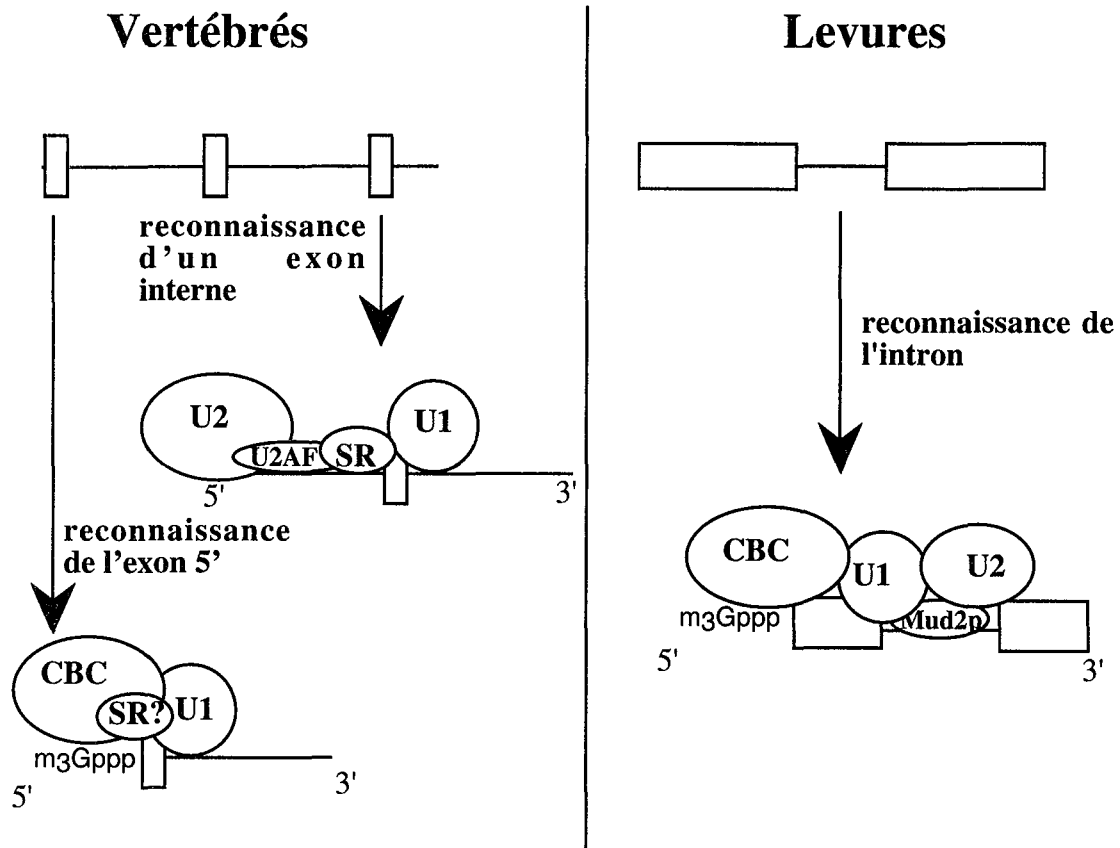


Figure 18: Formation des complexes pré-spliceosomaux par la définition d'exon pour les vertébrés (à gauche) et par la définition d'intron pour les levures (à droite). Les mécanismes de définition de l'exon en 5' et des exons internes sont représentés, dans le cas des vertébrés.

D'après Talerico et Berget, 1994; Berget, 1995; Izaurralde *et al.*, 1994 et 1995; Lewis *et al.*, 1996a et 1996b; Colot *et al.*, 1996.

et Reed, 1994a). Parmi les protéines hnRNP jouant un rôle dans l'épissage, la protéine hnRNPA1 intervient dans le choix des sites 5' d'épissage (Mayeda et Krainer, 1992). La protéine PTB (pour "Polypyrimidine Track Protein"), tout d'abord identifiée comme une protéine se fixant à la région riche en pyrimidine (Patton *et al.*, 1991), s'est avérée être la protéine hnRNPI (Ghetti *et al.*, 1992). Cette protéine joue un rôle important dans l'assemblage des spliceosomes mais est aussi, comme la protéine hnRNPA1, un régulateur de l'épissage alternatif (pour revue, Krecic et Swanson, 1999). La protéine hnRNPC a une affinité pour l'extrémité 3' du snRNA U6 et pourrait, elle, jouer un rôle dans la rupture de l'appariement entre les snRNA U4 et U6 dans des snRNP isolées (Forne *et al.*, 1995).

1.2. Reconnaissance des exons et des introns

Chez les vertébrés, la plupart des gènes contient de multiples introns qui sont en général longs, jusqu'à 10^4 à 10^5 bases, alors que les exons sont relativement courts (Hawkins, 1988; Sterner *et al.*, 1996). La première étape du processus d'épissage est alors une étape de définition des exons (Fig. 18) (Berget, 1995). En effet, comme l'a montré l'équipe de Berget, la mutation d'un site 5' aval inhibe l'épissage de l'intron en amont (Talerico et Berget, 1994) et, au contraire, le renforcement d'un site 5' aval augmente l'épissage de l'intron en amont (Kuo *et al.*, 1991). Dans un deuxième temps, par rapprochement des composants fixés aux extrémités des introns, le spliceosome est assemblé. Dans le cas de l'exon en 5' de l'ARN pré-messager, la reconnaissance se fait entre la coiffe terminale et la jonction en 5' et pour l'exon en 3' de l'ARN pré-messager, la reconnaissance se fait entre la jonction 3' et le site de polyadénylation (Ohno *et al.*, 1987; Niwa et Berget, 1991; Izaurralde *et al.*, 1994; Lewis *et al.*, 1996a; 1996b). Le complexe CBC reconnaît spécifiquement la structure coiffe 5' des ARN pré-messagers (Izaurralde *et al.*, 1994; 1995). Ce complexe reste associé à cette structure tout au long de la réaction d'épissage. Il semble faciliter la fixation de la snRNP U1 à l'ARN pré-messager. Il interviendrait donc dans la définition du premier exon en 5' de l'ARN pré-messager (Fig. 18) (Lewis *et al.*, 1996a).

Chez *S. cerevisiae* et chez beaucoup d'eucaryotes unicellulaires, les introns sont courts et, en général, un ARN pré-messager ne contient qu'un seul intron. Dans ce cas, l'intron unique est reconnu directement par le spliceosome, par un mécanisme de définition d'intron (Talerico et Berget, 1994). Chez *S. cerevisiae*, seules quelques protéines homologues aux protéines des complexes hnRNP des vertébrés sont connues, dont les protéines Mud13 et yCBP80, homologues au complexe CBC et participant également à la reconnaissance de l'intron unique des ARN pré-messagers (Fig. 18) (Lewis *et al.*, 1996b; Colot *et al.*, 1996).

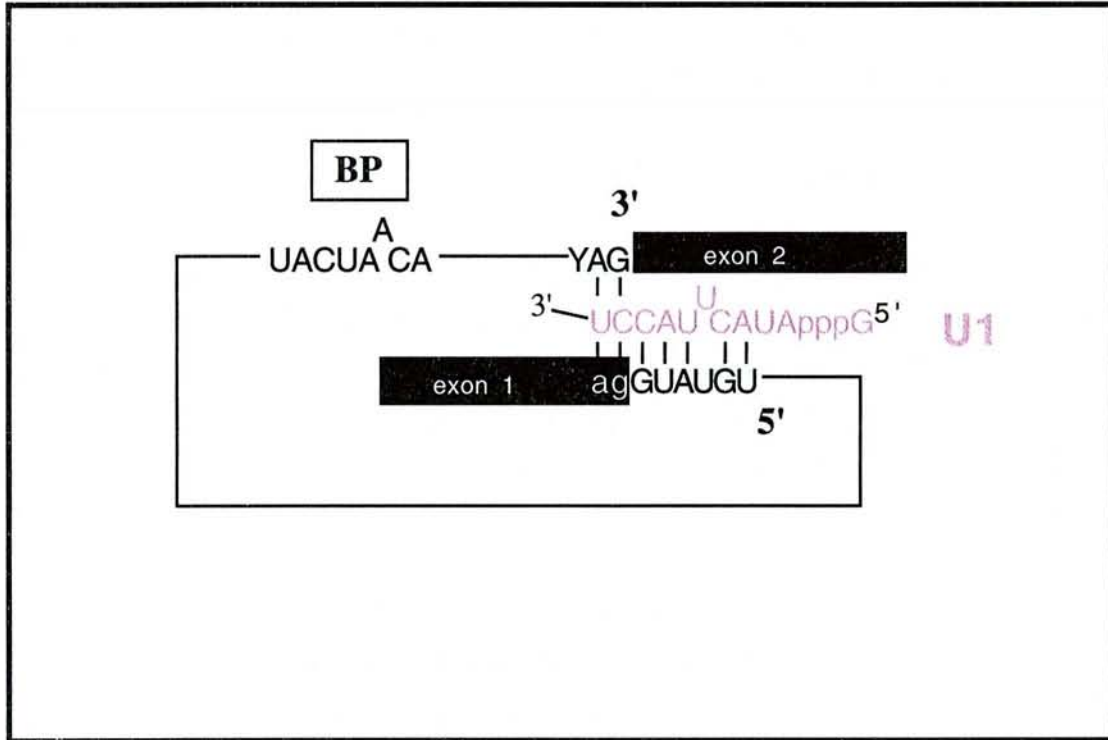


Figure 19: L'interaction du snRNA U1 humain avec les sites 3' et 5' d'épissage.
 D'après la revue, Madhani et Guthrie, 1994.

2. LE COMPLEXE D'ENGAGEMENT

Des études *in vitro*, par électrophorèse en gel de polyacrylamide non dénaturant, ont permis de définir, chez *S. cerevisiae*, les complexes d'engagement, ou "commitment complex", comme étant les premiers complexes stables formés entre l'ARN pré-messager et les constituants du spliceosome (Legrain *et al.*, 1988). Leur formation permet l'entrée de l'ARN pré-messager dans le processus d'épissage. Ils ne nécessitent ni la présence d'ATP, ni la présence de la snRNP U2. Par contre, l'intégrité des séquences constituant le site 5' d'épissage et la boîte de branchement, ainsi que la snRNP U1, sont indispensables (Legrain *et al.*, 1988; Séraphin et Rosbash, 1989a). En effet, la première interaction ARN/ARN mise en place lors de l'assemblage du spliceosome est l'interaction entre le snRNA U1 et le site 5' d'épissage. Bien que la complémentarité entre le snRNA U1 et l'intron soit essentielle à l'efficacité de formation du complexe d'engagement, cet appariement n'est pas le seul paramètre définissant la force d'un site 5' d'épissage. En effet, les structures, emprisonnant ce site 5', interviennent dans l'efficacité des sites 5'. C'est également le cas de certaines séquences, avoisinant le site 5', car elles peuvent fixer, comme nous le verrons, des facteurs protéiques facilitant ou, au contraire, limitant l'assemblage du spliceosome.

Dans le modèle original d'interaction entre le snRNA U1 et le pré-messager, il avait été proposé que le dinucléotide AG en 3' de l'intron s'apparie à la séquence invariante CU en amont de la séquence 5' du snRNA U1 s'appariant à la séquence 5' de l'intron (Fig. 19) (Lerner *et al.*, 1980; Rogers et Wall, 1980). Ce modèle est attractif car il offre la possibilité d'un mécanisme concerté d'identification des sites 3' et 5' d'épissage par le snRNA U1. Chez *S. pombe*, l'existence de cet appariement a été démontrée par production de mutations compensatoires (Reich *et al.*, 1992). Comme nous l'avons déjà indiqué, les introns AG-dépendants chez les mammifères requièrent l'intégrité du dinucléotide AG pour la première réaction d'épissage (Reed, 1989) et l'interaction U1/ site 3' serait nécessaire pour ces introns. Par contre, chez *S. cerevisiae*, cette interaction ne semble pas jouer un rôle important car des mutations dans le snRNA U1 qui préviennent la formation de cette interaction n'affectent pas la croissance cellulaire (Séraphin et Kandels-Lewis, 1993). Ceci va dans le sens de l'observation que le dinucléotide AG chez *S. cerevisiae* n'est nécessaire qu'à la deuxième étape d'épissage (Vijayraghavan *et al.*, 1986).

L'appariement U1/site 3', quand il existe, ne peut avoir lieu que dans une étape tardive du complexe d'engagement. En effet, dans un premier temps, la snRNP U1 interagit uniquement avec le site 5' (Mount *et al.*, 1983; Black *et al.*, 1985). Ensuite, la mise en place d'interactions entre la snRNP U1 et des facteurs d'épissage liés à la boîte de branchement ou à la séquence polypyrimidine provoquerait un repliement de l'intron, ce qui amènerait le site 3' à proximité

du site 5' et de la boîte de branchement, permettant alors une interaction entre la snRNP U1 et le site 3'.

Le complexe E, purifié par chromatographie d'affinité à partir d'un extrait cellulaire de cellules HeLa, contient le facteur U2AF (composé des protéines U2AF⁶⁵ et U2AF³⁵, Zamore et Green, 1989), la snRNP U1, ainsi qu'un certain nombre de protéines appelées SAP (pour "Spliceosome Associated Proteins") (Michaud et Reed, 1991; Bennett *et al.*, 1992) et des protéines SR dont ASF/SF2 et SC35 (Staknis et Reed, 1994b). La snRNP U2 pourrait également être présente dans ce complexe, mais elle serait faiblement associée (Hong *et al.*, 1997). Le facteur U2AF⁶⁵ est lié à la séquence polypyrimidine par son domaine RRM (Zamore et Green, 1989; Zamore *et al.*, 1992; Valcarcel *et al.*, 1993; Singh *et al.*, 1995) et interagit avec un autre facteur d'épissage associé à la boîte de branchement, mBBP (pour "Binding Branch-point Protein"), également appelé SF1 (Arning *et al.*, 1996), (Abovich *et al.*, 1994; Chiara *et al.*, 1996; Abovich et Rosbash, 1997; Berglund *et al.*, 1997).

Un modèle a été établi dans lequel les protéines SR, et certainement plus spécifiquement SC35 et/ou ASF/SF2, joueraient un rôle dans le recrutement du facteur U2AF à la séquence polypyrimidine. Elles permettraient le rapprochement des deux sites d'épissage en interagissant avec la snRNP U1 fixée au site 5' et U2AF³⁵ fixé au site 3' (pour revues, Fu, 1995 et Manley et Tacke, 1996). Des analyses par la technique du double hybride indiquent que chacune des protéines SR, ASF/SF2 et SC35, peut interagir avec la protéine U1 70K de la snRNP U1 et le facteur U2AF³⁵ (Wu et Maniatis, 1993). L'interaction entre ASF/SF2 et U1 70K se fait, *in vitro*, via les domaines RS de ces deux protéines (Kohtz *et al.*, 1994). Les protéines SR faciliteraient donc l'interaction de la snRNP U1 au site 5' et participeraient au repliement de l'ARN pré-messager, repliement nécessaire aux étapes ultérieures d'assemblage du spliceosome. Ces protéines, lorsqu'elles sont ajoutées en grande quantité dans des expériences d'épissage *in vitro*, peuvent même suppléer l'absence de la snRNP U1 (Crispino *et al.*, 1994; Tarn et Steitz, 1994).

La protéine Urp (pour UAF³⁵ Related Protein), essentielle à l'épissage, interagit à la fois avec U2AF⁶⁵, SC35 et ASF/SF2, mais son rôle est inconnu, et on ne sait pas si son association aux complexes d'épissage est stable ou transitoire (Tronchere *et al.*, 1997).

La formation du complexe E a pu être décomposée en deux étapes chez *S. cerevisiae* (Fig. 17). Un premier complexe d'engagement formé (CC1) implique la fixation de la snRNP U1 au site 5' d'épissage. Le complexe CC2 est ensuite formé (Séraphin et Rosbash, 1989a) par la fixation, à l'ARN pré-messager, de deux facteurs d'épissage, qui sont les homologues respectifs des protéines de mammifères mBBP et U2AF⁶⁵: les protéines yBBP et Mud2p (Rutz et Séraphin, 1999). La fonction de cette dernière n'est pas essentielle dans un contexte génétique WT, cependant la délétion du gène codant cette protéine est létale en présence de

mutations déstabilisant le snRNA U1 (Abovich *et al.*, 1994). Cette protéine semble être le facteur de reconnaissance de la boîte de branchement (Abovich *et al.*, 1994) et cette reconnaissance se ferait, *in vivo*, via le nucléotide précédant la boîte de branchement (Rain et Legrain, 1997). Le facteur yBBP, quant à lui, est essentiel à l'épissage et est également associé à la boîte de branchement (Berglund *et al.*, 1997). La protéine yBBP interagit avec Mud2 et la protéine Prp40p de la snRNP U1, formant ainsi un pont entre les deux extrémités de l'intron, comme c'est le cas chez les mammifères (Abovich *et al.*, 1994; Abovich et Rosbash, 1997; Fromont-Racine *et al.*, 1997).

En absence d'ATP, aucun complexe d'engagement n'a été détecté dans le cas des spliceosomes U12-dépendants (Tarn et Steitz, 1996a; Frilander et Steitz, 1999).

3. LES PRE-SPLICEOSOMES (complexe A chez les mammifères ou B chez *S. cerevisiae*)

L'étape suivante de l'assemblage consiste en l'association de la snRNP U2 avec obtention du complexe A. Cette étape est ATP dépendante (Ruskin et Green, 1985) et met en jeu, en plus de l'interaction entre le snRNA U2 et la boîte de branchement, de nombreuses interactions ARN/protéine et protéine/protéine. Chez les mammifères, dix protéines supplémentaires, par rapport au complexe E, sont présentes dans ce complexe (Bennett *et al.*, 1992; Staknis et Reed, 1994a). La plupart sont des protéines de la snRNP U2 et notamment les facteurs SF3a et SF3b. Une des protéines du facteur SF3b, SAP155, peut être pontée en amont et en aval de la boîte de branchement. Elle interagit directement avec U2AF (Gozani *et al.*, 1998). Dans le complexe A, U2AF⁶⁵ est fixée à la séquence polypyrimidine et interagit également avec la boîte de branchement (Gaur *et al.*, 1995) par son domaine RS (Valcarcel *et al.*, 1996). Ce domaine RS est nécessaire à la fixation de la snRNP U2 à l'ARN pré-messager (Zamore *et al.*, 1992; Valcarcel *et al.*, 1996). L'interaction U2AF/mBBP, mise en place dans le complexe E, est donc remplacée, dans le complexe A, par une interaction U2AF/SAP155 (Gozani *et al.*, 1998). Ces observations suggèrent que les facteurs U2AF et SAP155 interviennent ensemble dans le recrutement de la snRNP U2.

Chez *S. cerevisiae*, la protéine Mud2p permettrait, comme son homologue chez les mammifères U2AF⁶⁵, de recruter la snRNP U2 en interagissant directement avec la protéine Prp11p (Abovich *et al.*, 1994). Cependant les protéines Mud2p et yBBP ne sont pas détectables dans le complexe B. Leur départ pourrait, donc, être concomitant avec l'arrivée de la snRNP U2 (Rutz et Séraphin, 1999).

Comme nous l'avons vu, l'adénosine au point de branchement est située en boucle latérale dans l'appariement formé entre l'ARN pré-messager et le snRNA U2, ce qui doit favoriser sa réactivité (Fig. 6) (Query *et al.*, 1994; 1996). Néanmoins, il faut noter que, chez *S. cerevisiae*, le point de branchement semble être sélectionné avant l'appariement du snRNA U2 à la boîte de branchement. Cette reconnaissance aurait lieu durant l'assemblage du complexe CC2, car des mutations du point de branchement inhibent le passage CC1→CC2 (Pascolo et Séraphin, 1997). Cette hypothèse est en accord avec le fait que, chez les mammifères, le point de branchement est reconnu par différentes protéines, avant l'arrivée de la snRNP U2, dans le complexe E, puis tout au long de l'assemblage du spliceosome et de la réaction d'épissage (MacMillan *et al.*, 1994; Chiara *et al.*, 1996). Il existe donc une action conjuguée du snRNA U2 et des facteurs d'épissage, pour la sélection de l'adénosine constituant le point de branchement. Les interactions protéine/protéine jouent sans doute un rôle prédominant chez les vertébrés et doivent compenser le très faible degré d'appariement existant entre le snRNA U2 et la boîte de branchement.

Des complexes A, contenant les snRNP U11 et U12, ont été détectés lors de l'assemblage des spliceosomes U12-dépendants (Tarn et Steitz, 1996a; Frilander et Steitz, 1999). Comme nous l'avons déjà indiqué, ces UsnRNP sont associées au sein d'une même particule. Il a aussi été montré que l'appariement des snRNA U11 et U12 au site 5' d'épissage et à la boîte de branchement, respectivement, sont des phénomènes coopératifs (Frilander et Steitz, 1999). Ainsi, dans les spliceosomes U12-dépendants, le pont qui connecte les extrémités 5' et 3' de l'intron est déjà établi par la snRNP U11/U12, alors que, dans les spliceosomes U2-dépendants, l'assemblage de composés individuels est requis.

4. LE COMPLEXE B (nomenclature vertébrés)

La dernière étape de l'assemblage des spliceosomes U2- et U12-dépendants va consister, respectivement, dans l'association, ATP dépendante, au pré-spliceosome des tri-snRNP [U4/U6.U5] (pour revue, Reed et Palandjian, 1997) et [U4atac/U6atac.U5] (Tarn et Steitz, 1996a; Frilander et Steitz, 1999),.

Cette association est suivie de nombreux changements conformationnels des UsnRNA, conduisant à la formation du complexe C, portant l'activité catalytique pour la première étape de la réaction d'épissage:

⇒ dissociation des interactions U4/U6 et U4atac/U6atac, au profit de la mise en place d'interactions U2/U6 et U12/U6atac, respectivement,

⇒ dissociation des interactions U1/extrémité 5' de l'intron et U11/extrémité 5' de l'intron, qui sont, respectivement, remplacées par une interaction U6/extrémité 5' de l'intron et U6atac/extrémité 5' de l'intron,
 ⇒ mise en place d'une interaction entre la bordure 3' de l'exon 1 et la boucle terminale de la snRNP U5.

De nombreuses preuves expérimentales de l'existence de ces changements ont été obtenues. Les premières étaient la mise en évidence d'une labilité des snRNA U1 et U4 dans les complexes actifs (Pikielny *et al.*, 1986; Cheng et Abelson, 1987; Konarska et Sharp, 1987; Lamond *et al.*, 1988). La rupture de l'appariement entre la séquence en 5' de l'intron et l'extrémité 5' du snRNA U1 est requise pour la catalyse (Konforti *et al.*, 1993). La snRNP U1 interviendrait donc dans la reconnaissance des sites d'épissage et dans l'initiation de l'assemblage du spliceosome, mais pas dans la catalyse.

Nous avons déjà cité un ensemble de résultats d'études phylogénétiques, génétiques et par pontage montrant la mise en place d'une interaction entre les snRNA U6 ou U6atac et l'extrémité 5' des introns, ainsi que la mise en place de l'interaction entre la boucle terminale I conservée du snRNA U5 et les bordures exoniques. Nous détaillerons donc, plus particulièrement, dans cette partie, les arguments expérimentaux en faveur de la mise en place d'interactions entre les snRNA U2 ou U12 et les snRNA U6 ou U6atac, ainsi que l'importance de chacune de ces interactions pour l'épissage.

4.1. Les interactions U2/U6 et U12/U6atac

En dépit des différences de séquence, des interactions peuvent être formées entre les snRNA U2 et U6, d'une part, et les snRNA U12 et U6atac, d'autre part (Fig. 20). Nous verrons que l'importance de chacun des appariements formés varie suivant le type de spliceosomes.

Le premier appariement de bases, qui a été proposé entre les snRNA U2 et U6, implique l'extrémité 3' du snRNA U6 et l'extrémité 5' du snRNA U2 (Fig. 20) (Hausner *et al.*, 1990). La possibilité de former cet appariement, maintenant nommé hélice II, est indispensable pour qu'il y ait épissage *in vitro* dans un extrait nucléaire de cellules HeLa (Datta et Weiner, 1991; Wu et Manley, 1991). De plus, des mutations dans le snRNA U6 humain, qui inhibent la formation de cette hélice, empêchent l'assemblage du complexe B (Wolff et Bindereif, 1992). Chez *S. cerevisiae*, le même type d'interaction peut être formé, d'après la conservation de la complémentarité des séquences, mais cet appariement n'est pas indispensable à la croissance cellulaire (Fabrizio *et al.*, 1989; Madhani *et al.*, 1990) et donc à l'épissage. Cette hélice est seulement requise chez *S. cerevisiae* quand la formation d'une

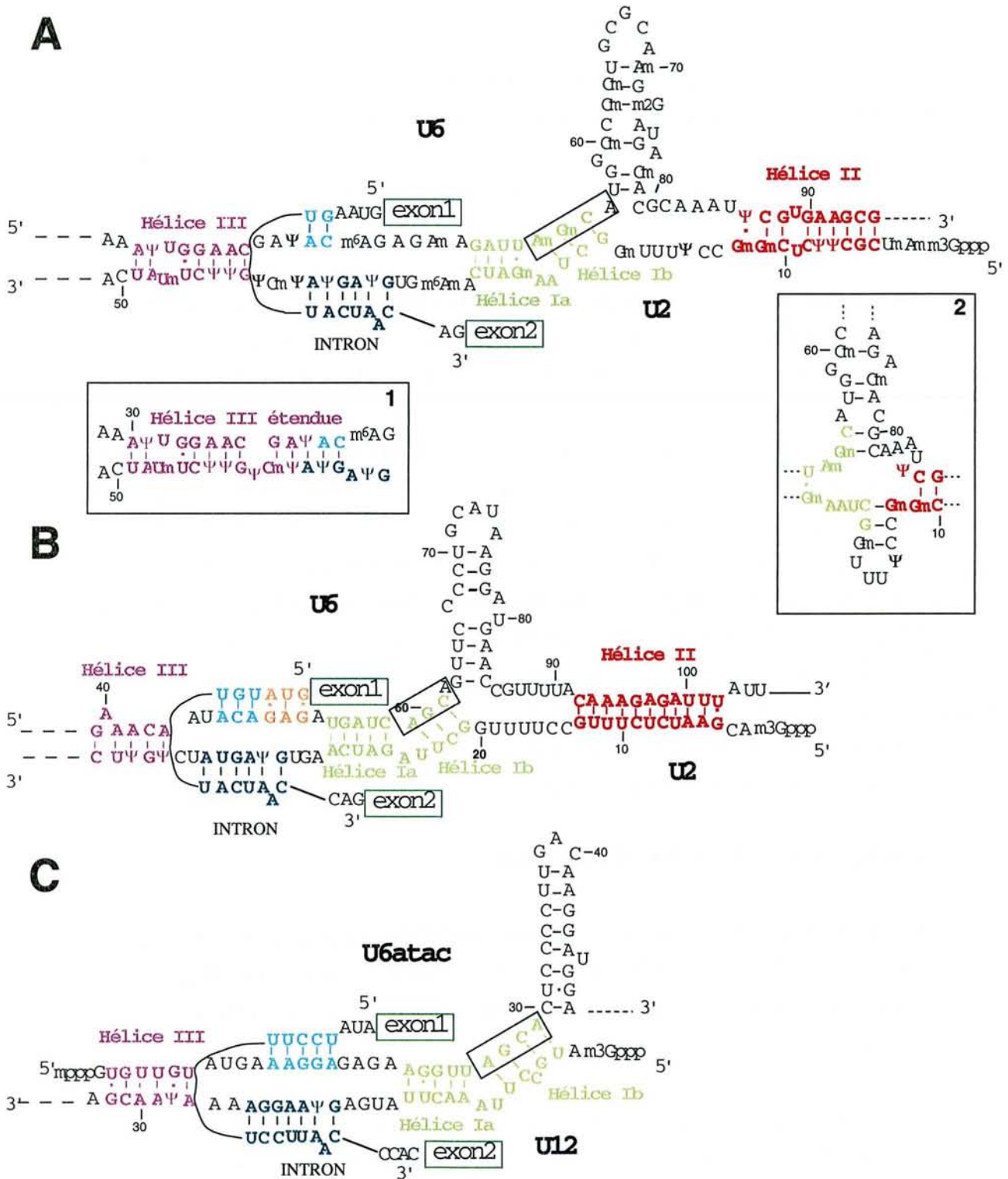


Figure 20: Les interactions ARN/ARN impliquant les snRNA U6 et U2 de mammifères (A) ou de *S. cerevisiae* (B) et les snRNA U6atac et U12 (C), lors de la réaction d'épissage.
 Voir le texte pour les références.

L'hélice III étendue, ainsi que la structure alternative de l'hélice Ib, proposées chez les mammifères sont indiquées respectivement dans les cadres 1 et 2, panneau A. L'extension proposée de l'appariement entre le snRNA U6 de *S. cerevisiae* et le site 5' d'épissage est indiquée en jaune dans le Panneau B. La séquence AGCA conservée dans les snRNA U6 et U6atac est encadrée.

Les hélices I, II et III sont respectivement en vert, rouge et violet. L'appariement de bases entre le snRNA U2 ou U12 et la boîte de branchement est en bleu foncé et l'appariement de bases entre le snRNA U6 ou U6atac et l'ARN pré-messager est en bleu clair.

autre hélice impliquant les snRNA U2 et U6 (hélice Ib) est inhibée par mutagenèse dirigée. La réciproque est aussi vraie (Field et Friesen, 1996). Ceci suggère une redondance des fonctions de ces deux hélices. Dans les spliceosomes U12-dépendants, l'hélice II ne peut être formée, car l'extrémité 5' du snRNA U12 est tronquée par rapport au snRNA U2 (Fig. 20).

Une autre interaction entre les snRNA U2 et U6 a tout d'abord été mise en évidence chez *S. cerevisiae* par Madhani et Guthrie (1992). Elle correspond à la formation de deux courtes hélices contiguës Ia et Ib, interrompues par deux nucléotides du snRNA U2 en boucle latérale. Des mutations dans les nucléotides conservés du snRNA U6 formant l'hélice Ib (séquence AGCA, encadrée dans la Fig. 20) inhibent soit la première, soit la deuxième étape de la réaction d'épissage (Fabrizio et Abelson, 1990a; 1990b; Wolff *et al.*, 1994; Yu *et al.*, 1995). Cependant, la nécessité de l'hélice Ib n'a pas été démontrée chez *S. cerevisiae* de manière individuelle, mais seulement en présence de mutations inhibant la formation de l'hélice II (Field et Friesen, 1996). Chez les mammifères, une structure alternative mettant en jeu les nucléotides des snRNA U2 et U6, impliqués dans la formation de l'hélice Ib, a été proposée (Encart 2, Fig. 20) (Sun et Manley, 1995). La formation de l'hélice Ia est, quant à elle, indispensable à la formation d'un spliceosome actif dans un extrait cellulaire de cellules HeLa (Sun et Manley, 1995) ou chez *S. cerevisiae* (Madhani et Guthrie, 1992; 1994a; Field et Friesen, 1996). L'hélice I semble donc avoir un rôle fonctionnel important dans les spliceosomes U2-dépendants. La possibilité de former une structure équivalente est conservée dans les spliceosomes U12-dépendants (Tarn et Steitz, 1996b) et son existence dans ces spliceosomes a été montrée par pontage au psoralène (Shukla et Padgett, 1999).

Enfin, la possibilité de former une troisième hélice entre la séquence du snRNA U2 située en aval de la séquence appariée à la boîte de branchement et un segment du snRNA U6 comportant une partie de la séquence très conservée ACAGAG, a été mise en évidence plus récemment (Sun et Manley, 1995) (Fig. 20). Son importance pour la formation de spliceosomes actifs a été démontrée *in vitro* dans un extrait de cellules HeLa. Cette interaction n'est pas requise chez *S. cerevisiae* (Yan et Ares, 1996). Elle n'a pas été détectée à ce jour dans les spliceosome U12-dépendants de cellules HeLa (Tarn et Steitz, 1996b). Par ailleurs, sur la base des séquences des UsnRNA, elle ne peut pas être formée dans les spliceosomes U12-dépendants d'*A. thaliana* (Shukla et Padgett, 1999). Il est vraisemblable que dans les spliceosomes U2-dépendants, où l'hélice III peut être formée, cette hélice ne soit que partielle puisqu'elle implique une partie de la séquence du snRNA U2 s'appariant à la boîte de branchement et une partie du snRNA U6 s'appariant au site 5' (Fig. 20). Cependant, l'appariement complet peut être requis pour des étapes transitoires de mise en place ou de dissociation de la structure active.

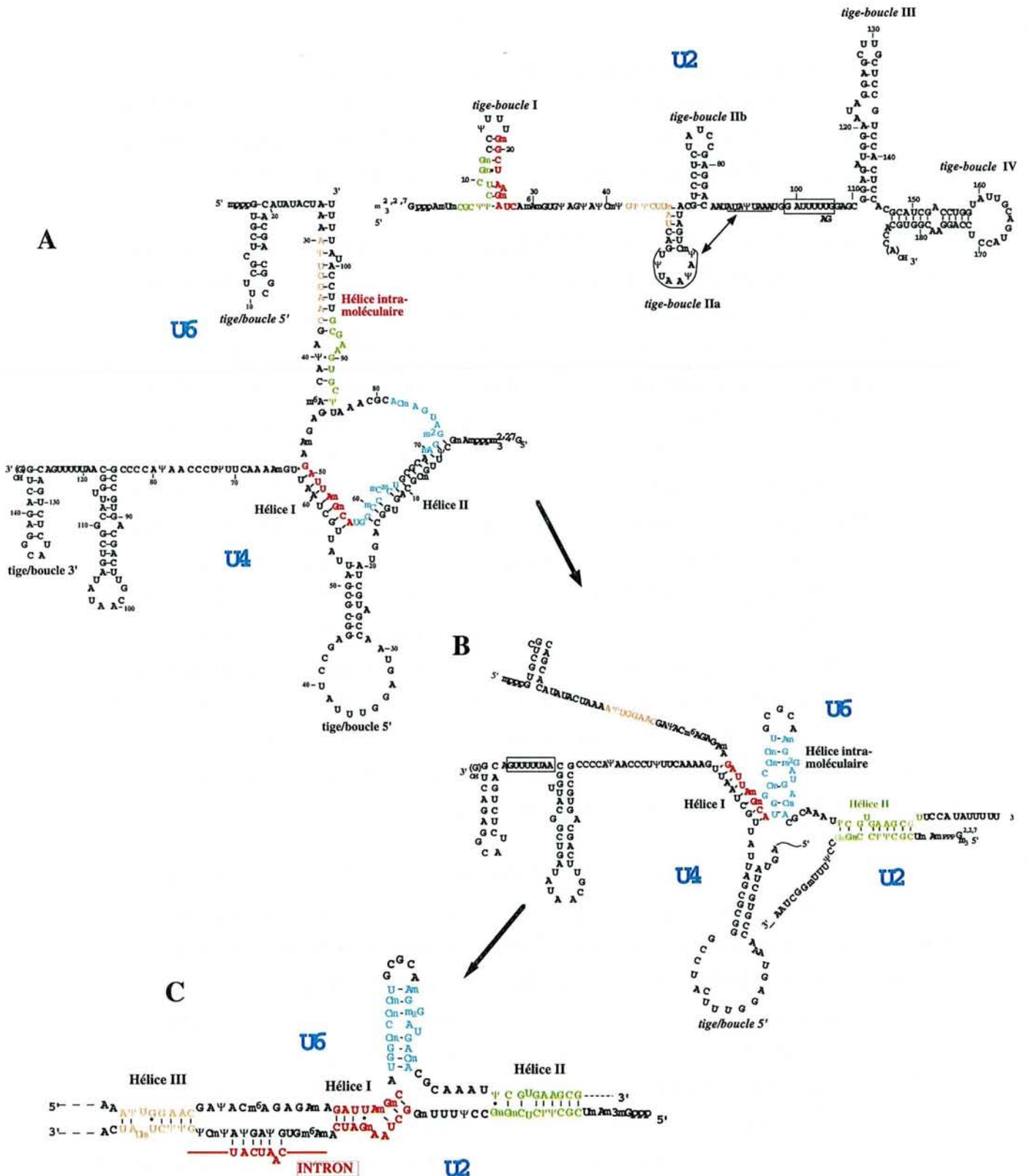


Figure 21: Mécanisme proposé pour le passage de l'interaction U4/U6 à l'interaction U2/U6, chez les mammifères.

Mougin *et al.*, résultats non publiés.

A. Structure du snRNA U2 et interactions des snRNA U4 et U6 au sein de la tri-snRNP [U4/U6.U5].

B. Dissociation de l'interaction U4/U6 au profit d'une interaction U6/U2 au sein du spliceosome B.

C. Interactions ARN/ARN au coeur du spliceosome.

La protéine NS1 du virus *Influenza* se fixe sur la région centrale de la structure tige-boucle II du snRNA U6 et inhibe la formation du duplex d'ARN U4/U6 et l'interaction U2/U6 (Qiu *et al.*, 1995; Lu *et al.*, 1994). Cette protéine inhibe également la formation du duplex d'ARN U12/U6atac, mais pas du duplex d'ARN U4atac/U6atac (Wang et Krug, 1998). Or, cette protéine se fixe sur la structure tige-boucle 3' terminale du snRNA U6atac, qui n'existe pas dans le snRNA U6 (Fig. 12). Sur la base de cette observation, les auteurs ont proposé que cette structure tige-boucle du snRNA U6atac participe à la formation d'une interaction supplémentaire entre les snRNA U6atac et U12.

4.2. Les transitions $U4/U6 \Rightarrow U2/U6$ et $U4atac/U6atac \Rightarrow U12/U6atac$

La formation des interactions entre les snRNA U2 et U6 implique la dissociation du duplex d'ARN formé entre les snRNA U6 et U4, la dissociation de la structure tige-boucle I du snRNA U2, ainsi que la formation d'une nouvelle hélice intramoléculaire dans le snRNA U6, le snRNA U2 étant toujours fixé à la boîte de branchement. Comme nous l'avons vu, un certain nombre de facteurs protéiques pourraient être impliqués dans la dissociation du duplex U4/U6.

Les études réalisées par A. Mougin, au laboratoire, sur les snRNP U4/U6, les tri-snRNP [U4/U6.U5], les duplex d'ARN U4/U6 et le snRNA U6 isolé d'extraits nucléaires de cellules HeLa ou de *S. cerevisiae*, conduisent à proposer un modèle de transition $U4/U6 \leftrightarrow U2/U6$ qui apporte une explication aux différences d'importance observées pour les hélices hétérologues du complexe U2/U6, suivant les espèces et le type de spliceosome.

En effet, A. Mougin a montré que le snRNA U6 humain à l'état libre se replie sur lui-même en une structure très stable qui peut être initiée, au sein du duplex d'ARN U4/U6, par le repliement interne du snRNA U6 (Fig. 13 et 21). Lors de l'assemblage du spliceosome U2-dépendant humain, la transition $U4/U6 \rightarrow U2/U6$ serait initiée par la formation de l'hélice II entre les snRNA U2 et U6, ce qui évite le repliement interne du snRNA U6 (Fig. 21). L'hélice III dans sa configuration partielle, compatible avec l'appariement au site de branchement, pourrait aussi être mise en place en même temps, puisqu'elle est compatible avec les hélices I et II des snRNA U4 et U6. De la même manière, lors de la transition inverse, $U2/U6 \rightarrow U4/U6$, les hélices I et III entre snRNA U2 et U6 pourraient être maintenues lors de la reformation de l'hélice II entre U4 et U6. Ainsi le snRNA U6 ne serait jamais sous forme libre (Fig. 21).

Chez la levure, la structuration du snRNA U6 libre est différente. Elle ne peut pas être initiée par un repliement interne du snRNA U6 au sein du duplex U4/U6 (Brow et Vidaver, 1995). La nécessité d'éviter la formation d'un repliement interne du snRNA U6 au sein du duplex U4/U6 n'est donc pas un impératif aussi important chez *S. cerevisiae*, expliquant la moins grande dépendance, vis-à-vis de la formation des hélices II et III entre les snRNA U2 et U6,

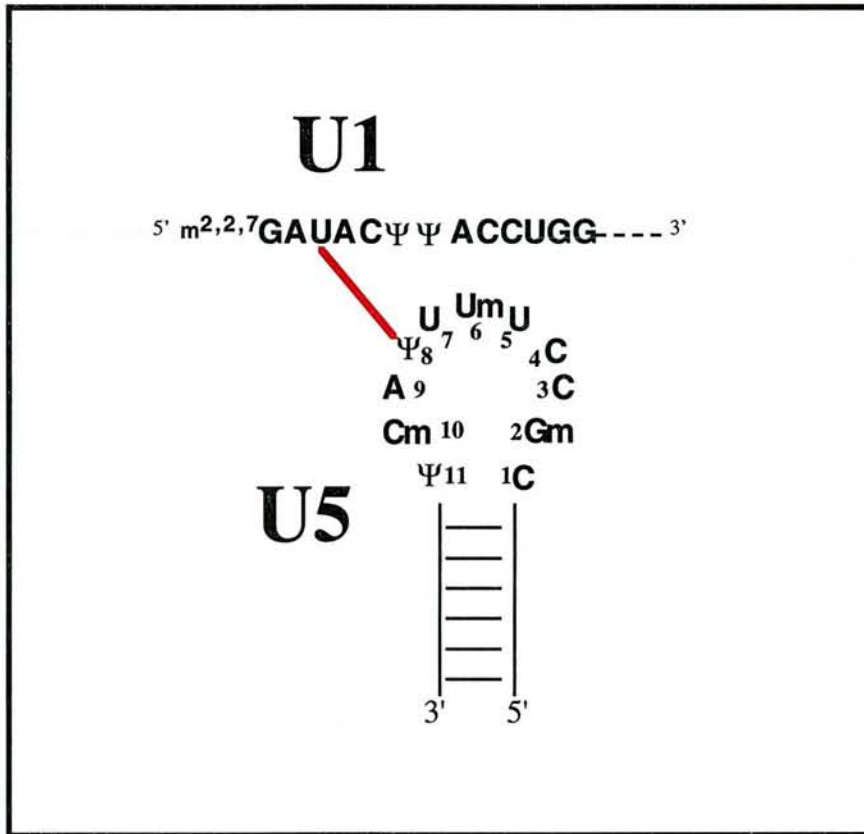


Figure 22: Mise en évidence par pontage covalent d'une interaction entre les snRNA U1 et U5, au cours de l'épissage.
 Les résidus pontés dans les deux ARN sont reliés par un trait rouge (Ast et Weiner, 1997b).

chez la levure que dans le système humain. La même situation est retrouvée pour les spliceosomes U12-dépendants, puisque le snRNA U6atac sous forme libre a une structure voisine de celle du snRNA U6 de *S. cerevisiae*. Ceci expliquerait l'absence de l'hélice II entre les snRNA U6atac et U12 et l'absence de dépendance vis-à-vis de l'hélice III.

4.3. Les transitions U1/site 5' ⇒ U6/site 5' et U11/site 5' ⇒ U6atac/site 5'

Les transitions U4/U6 → U2/U6 et U4atac/U6atac → U12/U6atac s'accompagnent de la génération d'interactions U6/site 5' et U6atac/site 5'. Comme les interactions des snRNA U1 ou U11 et U6 ou U6atac avec le site 5' sont mutuellement exclusives, les snRNA U6 ou U6atac déplacent respectivement les snRNA U1 ou U11 du site 5'. Ceci est en accord avec le fait que l'hyperstabilisation de l'interaction U6/site 5' supprime le défaut d'épissage conféré par une hyperstabilisation de l'interaction U1/site 5' (Staley et Guthrie, 1999). Un autre argument est l'observation que la séquestration de la région du snRNA U6, qui se fixe au site 5', bloque la rupture de l'interaction U1/site 5' (Kuhn *et al.*, 1999).

4.4. La formation du premier complexe actif est un mécanisme concerté, faisant intervenir à la fois les UsnRNA et les protéines des UsnRNP.

Des mutations des nucléotides 72 à 75 du snRNA U4 de *S. cerevisiae*, entraînant un phénotype cryosensible de croissance, empêchent la fixation du snRNA U6 au site 5' (Li et Brow, 1996), le relargage de la snRNP U1 et la rupture du duplex d'ARN U4/U6 (Kuhn *et al.*, 1999), suggérant que ces trois mécanismes soient couplés.

La snRNP U5 peut être pontée de façon covalente à l'extrémité 3' de l'exon 1, par l'intermédiaire de la protéine Prp8p chez la levure et de la protéine de 220 kDa chez les mammifères (Wyatt *et al.*, 1992; Teigelkamp *et al.*, 1995a; Reyes *et al.*, 1996), ainsi que par l'intermédiaire de la boucle terminale conservée du snRNA U5 (pour revue, Newman, 1997). Dans un extrait de cellules HeLa, un oligonucléotide complémentaire aux nucléotides 68 à 88 du snRNA U5 n'abolit pas l'épissage, mais conduit à la dissociation de la tri-snRNP [U4/U6.U5] avec la formation d'un complexe [U1/U4/U5] (Ast et Weiner, 1996; Ast et Weiner, 1997a). Il avait été proposé que ce complexe artificiel corresponde à un complexe transitoire, avant le départ du snRNA U1 du site 5'. Deux observations vont dans ce sens: i) l'extrémité 5' du snRNA U1 humain peut être pontée à un résidu modifié (Ψ_8) de la boucle invariante du snRNA U5 (Fig. 22) (Ast et Weiner, 1997b) et ii) chez la levure, des interactions ont été mises en évidence par la technique du double hybride entre les protéines Prp8p de la snRNP U5 et Prp40p de la snRNP U1 (Abovich et Rosbash, 1997), ainsi qu'entre les protéines Snu246p de la tri-snRNP [U4/U6.U5] et la protéine U1 70K (Fromont-Racine *et al.*, 1997). La snRNP U1 pourrait donc intervenir pour diriger la snRNP U5 au site 5'. De

même, une tétra-snRNP [U2.U4/U6.U5] peut être purifiée chez *S. cerevisiae* et des interactions entre la snRNP U2 et la tri-snRNP [U4/U6.U5] pourraient donc aider à l'intégration de la tri-snRNP au pré-spliceosome (Gottschalk *et al.*, 1999).

A part les composants des UsnRNP, peu de facteurs, nécessaires à la mise en place du complexe C actif pour la première réaction catalytique, sont connus. Ils ont essentiellement été mis en évidence chez *S. cerevisiae*. La protéine Prp19p s'associe aux complexes spliceosomaux de façon concomitante ou juste après la dissociation du duplex d'ARN U4/U6 (Tarn *et al.*, 1993a, 1993b). Elle est présente dans un complexe protéique, dont une seule autre protéine a été caractérisée, Snt 309 (Chen *et al.*, 1998). Deux ARN hélicases ATP dépendante potentielles, les protéines Prp2p et Prp28p, pourraient aussi intervenir à cette étape (Tableau 4) (Kim et Lin, 1993; 1996; Staley et Guthrie, 1999).

5. LE COMPLEXE C: LE SPLICEOSOME (nomenclature vertébré)

Dans le complexe C, après la première réaction d'épissage, des transitions structurales vont permettre de rapprocher les deux exons pour conduire à leur ligation. La boucle terminale I du snRNA U5 va jouer un rôle important dans cette étape. En effet, chez *S. cerevisiae*, la délétion ou des mutations importantes dans cette boucle bloque l'épissage après la première réaction (O'Keefe *et al.*, 1996). De larges délétions ou insertions de nucléotides dans cette boucle entraînent un mauvais alignement des exons, impliquant que la taille de la boucle est importante pour l'alignement correct des exons pour la deuxième réaction catalytique (O'Keefe et Newman, 1998). Par contre, la délétion de cette boucle, dans un snRNA U5 humain produit par transcription *in vitro*, n'a pas d'effet sur la capacité de ce UsnRNA à être incorporé dans des spliceosomes et à compléter un extrait nucléaire de cellules HeLa, dont le snRNA U5 endogène a été éliminé par emploi d'oligonucléotides biotinylés (Ségault *et al.*, 1999). Cette observation montre une moins grande dépendance du mécanisme vis-à-vis de cette boucle terminale chez l'homme. Il faut cependant noter que de petites délétions et insertions de nucléotides dans cette boucle terminale du snRNA U5 de *S. cerevisiae*, n'empêchent pas la réalisation de la deuxième réaction d'épissage (O'Keefe et Newman, 1998). Ces résultats suggèrent que d'autres éléments du spliceosome, en particulier la protéine Prp8p/220 kDa, peuvent compenser les défauts de la boucle terminale I du snRNA U5. En effet, si la boucle du snRNA U5 est très conservée phylogénétiquement, il n'en est pas de même pour les bordures exoniques des ARN pré-messagers. De ce fait, l'interaction possible avec le snRNA U5, même dans un contexte génétique sauvage, est souvent de faible énergie et les protéines de la snRNP U5 doivent donc jouer un rôle essentiel dans le rapprochement des deux exons.

Les protéines vont également jouer un rôle important pour permettre à la deuxième étape d'épissage d'avoir lieu. En effet, chez *S. cerevisiae*, beaucoup de progrès ont été faits pour identifier et caractériser ces protéines (pour revue, Umen et Guthrie, 1995b). Elles sont au nombre de trois: Prp17p, Prp16p et Slu7p. Les deux dernières sont essentielles à la viabilité cellulaire. Les protéines Prp8p et Prp18p de la snRNP U5 sont également requises pour cette deuxième étape. Le site 3' semble être reconnu au moins en deux phases durant la seconde partie de la réaction d'épissage:

⇒ Dans un premier temps, la protéine Prp16p, membre de la famille des ARN hélicases ATP dépendantes, interagit directement avec le site 3' (Umen et Guthrie, 1995a), de façon concomitante avec Prp17p (Jones *et al.*, 1995). L'hydrolyse de l'ATP par Prp16p induit une variation conformationnelle qui entraîne une protection du site 3' vis-à-vis de sondes enzymatiques (Schwer et Guthrie, 1992). Cette variation structurale et la dissociation de Prp16p du site 3' sont simultanées (Umen et Guthrie, 1995a).

⇒ Cette hydrolyse d'ATP est également accompagnée de la fixation au site 3' des protéines Slu7p et Prp8p (Umen et Guthrie, 1995a). Prp18p aide sûrement Slu7p à rejoindre le site 3' d'épissage (Horowitz et Abelson, 1993b). Slu7p, de même que Prp18p, sont seulement requises pour la deuxième étape d'épissage dans les cas où la distance entre la boîte de branchement et le dinucléotide AG 3' terminal est supérieure à 7 nts (Brys et Schwer, 1996; Zhang et Schwer, 1997). Un autre facteur, SSF1, serait également nécessaire à la deuxième réaction d'épissage et agirait, comme Slu7p, après Prp16p (Ansari et Schwer, 1995).

La protéine Prp16p, en plus de son rôle dans la transition structurale qui permet la deuxième réaction d'épissage, jouerait également un rôle de contrôle de la fidélité du processus: la ligation des deux exons ne se ferait que si le complexe réactionnel a été correctement formé. L'hydrolyse de l'ATP par Prp16p entraînerait vers une voie de dégradation tous les complexes mal structurés, au lieu de permettre la transition permettant la deuxième réaction d'épissage (Burgess et Guthrie, 1993; pour revue, Umen et Guthrie, 1995b). En effet, un défaut dans l'activité ATPase ARN dépendante de Prp16p, supprime un défaut d'épissage dû à une mutation dans le point de branchement, ceci en limitant la dégradation des complexes aberrants (Burgess et Guthrie, 1993).

Chez les mammifères, des spliceosomes, bloqués après la première réaction d'épissage, peuvent être accumulés puis purifiés en utilisant un ARN pré-messager muté dans la séquence AG 3' terminale de l'intron. Ces complexes comportent 14 nouvelles protéines, en plus de celles déjà présentes avant la première étape d'épissage (Gozani *et al.*, 1994). Certaines de ces protéines ont été identifiées comme étant les homologues humaines des facteurs Prp16p, Prp17p, Prp18p et Slu7p de *S. cerevisiae*. Il s'agit respectivement des protéines hPrp16p (Zhou et Reed, 1998), hPrp17p (Zhou et Reed, 1998; Lindsey et Garcia-

Blanco, 1998) hPrp18p (Horowitz et Krainer, 1997) et hSlu7p (Chua et Reed, 1999). Contrairement à Prp18p, hPrp18p n'est pas associée à une UsnRNP (Horowitz et Krainer, 1997). L'existence de ces protéines homologues, dans les systèmes humain et de *S. cerevisiae*, montre que le mécanisme conduisant à la seconde étape de l'épissage a été très conservé au cours de l'évolution, ceci d'autant plus que certaines protéines peuvent être échangées entre les deux systèmes. Ainsi, les protéines humaines hPrp17p et hPrp18p peuvent respectivement pallier l'absence des protéines Prp17p et Prp18p chez *S. cerevisiae* (Horowitz et Krainer, 1997; Zhou et Reed, 1998).

Chez les mammifères, avant la deuxième réaction d'épissage, l'interaction entre le facteur U2AF et la séquence polypyrimidine est remplacée par une interaction avec trois composants de la snRNP U5, les protéines de 110, 116 et 220 kDa (Chiara *et al.*, 1997). La protéine de 220 kDa, comme son homologue Prp8p chez la levure, interagit à ce stade avec la boîte de branchement et l'extrémité 5' de l'exon 2, tout en restant liée à l'extrémité 3' de l'exon 1 (MacMillan *et al.*, 1994; Chiara *et al.*, 1997; Teigelkamp *et al.*, 1995a; 1995b; Umen et Guthrie, 1995a). Le snRNA U5 contacte les extrémités des deux exons, ce qui, avec l'action des protéines de la snRNP U5 (Prp8p/220 kDa en particulier), permet d'aligner les extrémités des exons, afin de conduire à leur ligation.

Un autre facteur, la protéine PSF, est également indispensable à la deuxième réaction d'épissage chez les mammifères (Gozani *et al.*, 1994). Elle est associée à la protéine hnRNPI/PTB et se fixe également à la séquence polypyrimidine (Patton *et al.*, 1993). Il est à noter également, qu'au cours de la deuxième étape d'épissage, le dinucléotide AG 3' terminal est reconnu par la protéine p70 (Wu et Green, 1997). Cette protéine serait la protéine AG⁷⁵ déjà détectée par Chiara *et al.* (Chiara *et al.*, 1996) et correspondrait à la protéine de 69 kDa, interagissant transitoirement avec les pré-particules, dont nous avons parlé précédemment (§C. 2. 1. 1.).

Un certain nombre d'observations suggèrent que la mise en place du site catalytique pour la deuxième réaction d'épissage se ferait, au moins en partie, par un mécanisme impliquant des contacts directs ou indirects entre les snRNP U2 et U5. En effet, en plus de leur proximité dans l'espace, des expériences génétiques chez la levure suggèrent des interactions fonctionnelles entre les composants de ces deux UsnRNP. Notamment, Xu *et al.* (1996, 1998) ont montré que des mutations dans les protéines Snu246p et Prp8p et dans le snRNA U5 étaient létales en présence de mutations dans le snRNA U2. De plus, des mutations dans les snRNA U2 et U5 entraînent une létalité en présence de mutations dans la protéine Prp8p ou dans les facteurs d'épissage nécessaires à la deuxième réaction de transestérification, Prp17p et Slu7p (Frank *et al.*, 1992; Xu *et al.*, 1996; 1998). Chiara *et al.*

(1997) ont proposé que la snRNP U2 joue un rôle dans le positionnement de la snRNP U5 au site 3'.

Un complexe C a également été détecté par électrophorèse en gel de polyacrylamide natif pour les spliceosomes U12-dépendants, mais aucune étude approfondie n'a encore été menée sur ce complexe (Tarn et Steitz, 1996a; Frilander et Steitz, 1999).

6. LES INTERACTIONS ARN-ARN AU SEIN DES COMPLEXES SPLICEOSOMAUX ACTIFS

Aucune expérience ne démontre encore actuellement de manière définitive quelles sont, parmi toutes les interactions citées, celles ayant lieu à l'une ou l'autre des étapes depuis l'assemblage du complexe C (dans la nomenclature vertébrés), jusqu'à la dissociation du spliceosome après la réaction. Néanmoins, il est raisonnable de penser, d'après i) la conservation phylogénétique, ii) les études basées sur la cinétique de formation des pontages covalents et iii) les études de l'effet de mutations dans les ARN, que les éléments suivants constituent, avec les protéines qui stabilisent ces interactions et en modulent la réactivité, les éléments essentiels du site catalytique des deux étapes successives de la réaction d'épissage (Fig. 23 et 24):

- l'hélice formée entre les snRNA U2 et U12 et la boîte de branchement,
 - les hélices hétérologues Ia et Ib formées entre les snRNA U2 et U6 ou U12 et U6atac,
 - la structure tige-boucle du snRNA U6 qui jouxte cette hélice hétérologue,
 - l'interaction entre le snRNA U6 ou U6atac et la séquence en 5' de l'intron
- et
- les interactions entre le snRNA U5 et la jonction exon 1-intron, dans un premier temps, puis les deux exons, dans un deuxième temps.

D'après la conservation phylogénétique de l'hélice I formée entre les snRNA U2 et U6 et l'importance des nucléotides qui la composent, cette hélice semble être directement impliquée dans la catalyse. Il en est de même pour la structure tige-boucle du snRNA U6 située entre les hélices I et II. Cette hypothèse est étayée par les résultats de nombreuses expériences réalisées chez *S. cerevisiae*.

Ainsi, des mutations du résidu U₂₃ du snRNA U2, déstabilisant l'hélice Ib, affectent le choix du site 5' d'épissage *in vivo* (Luukkonen et Séraphin, 1998b). La restauration d'un appariement entre les positions 23 du snRNA U2 et 59 du snRNA U6 (dans l'hélice Ib) (Fig. 23) supprime l'utilisation de sites 5' aberrants, suggérant que l'hélice Ib est impliquée dans le choix du site 5' d'épissage (Luukkonen et Séraphin, 1998b). Ce mécanisme pourrait également impliquer l'interaction tertiaire mise en évidence entre les résidus G₃₂ du snRNA

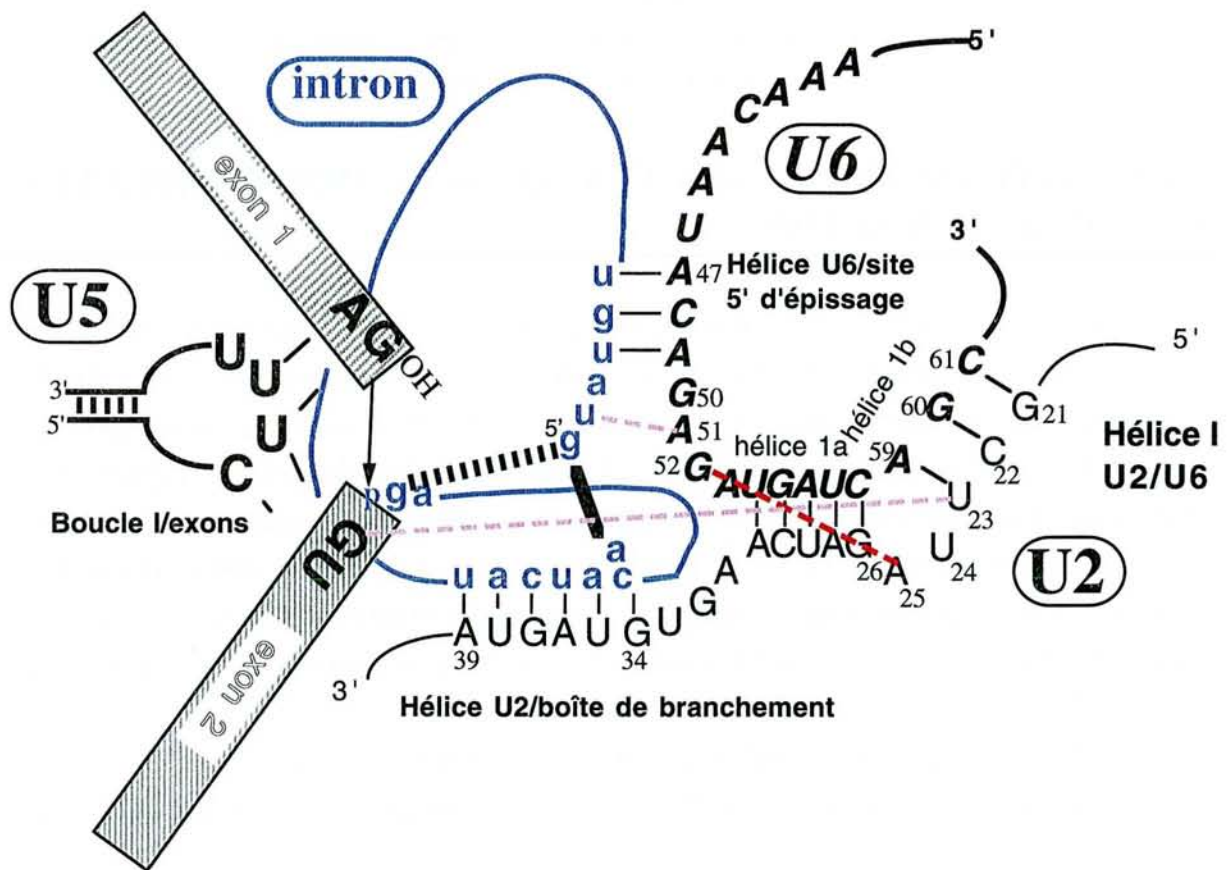


Figure 23: Réseau d'interactions entre ARN lors de la deuxième étape de la réaction d'épissage.

La numérotation des séquences est celle utilisée chez *S. cerevisiae*. Les exons sont représentés sous forme de rectangles gris hachurés, l'intron est coloré en bleu.

Le trait épais noir représente la liaison 2'-5' A-G entre le point de branchement et l'extrémité 5' libre de l'intron (Query *et al.*, 1994).

Le trait épais et pointillé représente l'interaction non canonique de type Hoogsteen mise en évidence entre les deux résidus G aux extrémités de l'intron (Parker et Siliciano, 1993; Chanfreau *et al.*, 1994; Deirdre *et al.*, 1995).

L'interaction tertiaire entre les résidus A25 et G52 des snRNA U2 et U6 est représentée en pointillé rouge.

Les pontages entre le résidu A51 du snRNA U6 et le résidu +2 de l'intron (Sontheimer et Steitz, 1993) et entre le résidu U23 du snRNA U2 et le résidu +1 de l'exon 2 (Newman *et al.*, 1995) sont représentés en pointillé vert.

Ce schéma est une adaptation de ceux présentés par Madhani et Guthrie (1994a) et Umen et Guthrie (1995b).

U6 et A₂₅ du snRNA U2 (Madhani et Guthrie, 1994a). Cette interaction place l'hélice I à proximité de la position +1 de l'intron (Fig. 23).

La mutation de plusieurs résidus des snRNA U2 et U6, impliqués dans la formation de l'hélice I, bloque la seconde réaction d'épissage (Fabrizio et Abelson, 1990a; Madhani et Guthrie, 1992). Ici encore, une interaction tertiaire mise en évidence entre le résidu U₂₃ du snRNA U2 et la position +1 de l'exon 2 doit permettre le rapprochement entre l'hélice I et le site de coupure en 3' (Fig. 23) (Newman *et al.*, 1995).

De plus, comme nous l'avons déjà mentionné, chez *S. cerevisiae* une relation fonctionnelle a pu être mise en évidence, par voie génétique, entre certains résidus du snRNA U2 impliqués dans la formation de l'hélice I et les résidus 97 à 99 de la boucle conservée du snRNA U5. A certaines étapes, les fonctions de l'hélice hétérologue I et de la boucle terminale I du snRNA U5 pourraient donc être liées (Xu *et al.*, 1998).

L'hélice intramoléculaire du snRNA U6 (Fig. 24), située entre les hélices I et II dans l'interaction U2/U6, est nécessaire à l'épissage (Wolff et Bindereif, 1993; Sun et Manley, 1997). Il faut aussi noter que la mutation de certains résidus du snRNA U6 formant cette hélice, ainsi que la mutation du résidu en position 53, qui est impliqué dans la formation de l'hélice Ia, peuvent pallier au défaut d'épissage généré par des mutations du nucléotide au point de branchement (McPheeters, 1996). Le phénotype observé est similaire à celui observé pour les mutants de la protéine Prp16p, dont l'activité ATPase est réduite (Burgess et Guthrie, 1993).

L'ensemble de ces études montre donc que l'hélice hétérologue I formée entre les snRNA U2 et U6 et l'hélice intramoléculaire du snRNA U6 jouent un rôle central pour l'activité du spliceosome.

7. LES REGULATIONS DE L'ÉPISSAGE ET RÔLE DES PROTÉINES SR

De nombreux ARN pré-messagers peuvent être soumis à un épissage différent suivant le type de tissu dans lequel ils sont épissés, le stade du développement, l'état physiologique de la cellule, etc...(pour revue, Lopez, 1998). Plusieurs paramètres vont jouer dans ces choix alternatifs des sites 3' ou 5' d'épissage: i) la force relative des sites, définie par la complémentarité avec les UsnRNA, ii) la structuration secondaire de l'ARN pré-messager qui peut masquer ou favoriser certains sites d'épissage, et iii) la fixation de facteurs inhibiteurs ou activateurs de l'épissage, dont les protéines SR. Ces protéines contiennent un ou deux domaines RRM de fixation à l'ARN côté N-terminal et un domaine RS, riche en enchaînements sérine/arginine, côté C-terminal. Les protéines SR sont essentiellement présentes chez les vertébrés et les plantes. Huit de ces protéines sont connues chez l'homme,

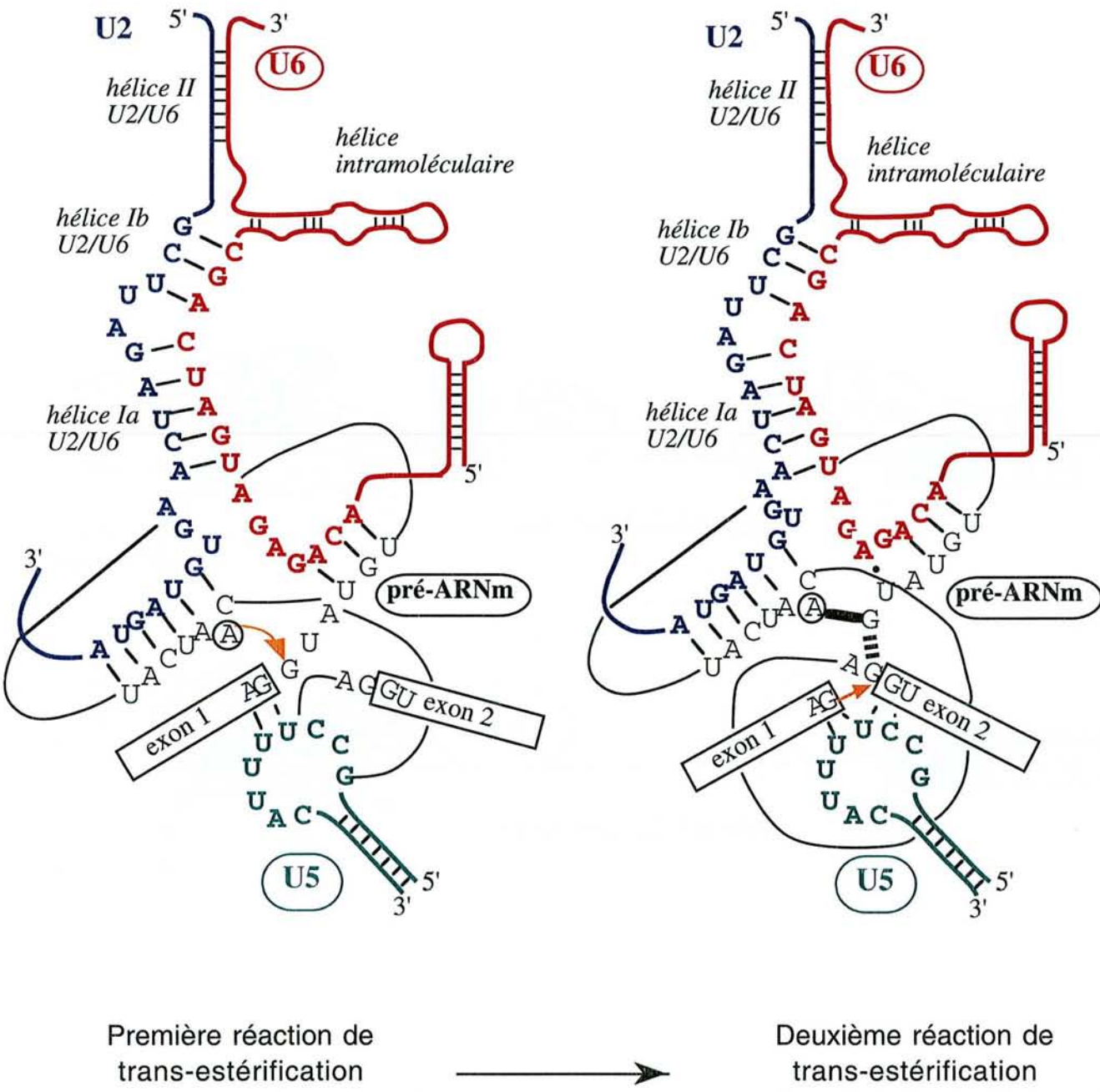


Figure 24: Les interactions hétérologues entre ARN supposées exister au sein des complexes spliceosomaux assurant la première et la deuxième étapes de la réaction d'épissage.

Le pré-ARN messenger est indiqué en noir, les snRNA U2, U5 et U6 en bleu, vert et rouge, respectivement.

Les séquences nucléotidiques sont celles de *S. cerevisiae*.

Le trait épais noir indique la liaison covalente 2'-5' s'établissant entre le premier nucléotide de l'intron et le site de branchement au cours de la première réaction de trans-estérification. Le trait épais en pointillé entre les deux résidus G aux extrémités de l'intron indique l'interaction de type Hoogsteen proposée entre ces deux résidus (Parker et Siliciano, 1993; Chanfreau *et al.*, 1994; Deirdre *et al.*, 1995).

Ce schéma est une adaptation de celui de Legrain et Chanfreau (1994).

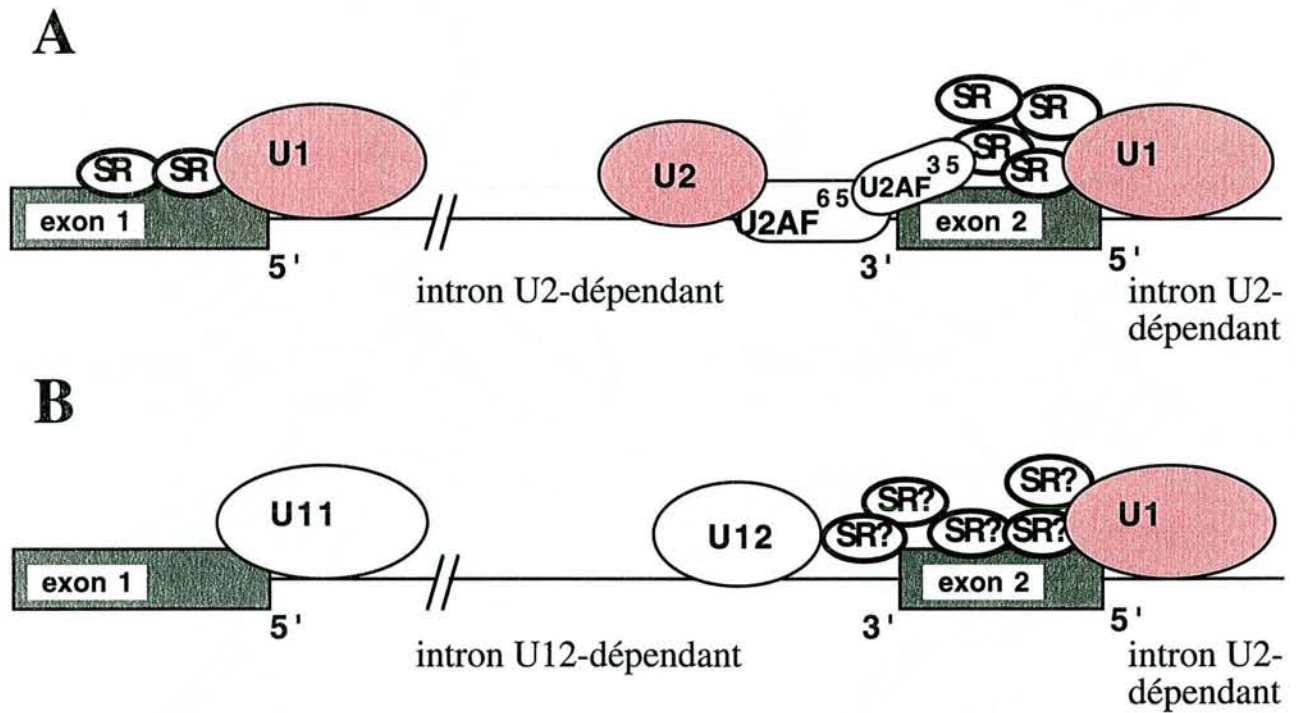


Figure 25: Comparaison des reconnaissances des jonctions intron-exon dans le cas des spliceosomes U2-dépendants (A) et dans le cas des spliceosomes U12-dépendants, l'intron U12-dépendant étant encadré d'introns U2-dépendants (B).

D'après la revue, Wu et Krainer, 1999.

dont ASF/SF2 et SC35, deux ont été identifiées chez *S. pombe* et une seule chez *S. cerevisiae* (Gross *et al.*, 1998; Lützelberger *et al.*, 1999; pour revues, Fu, 1995; Manley et Tacke, 1996; Tacke et Manley, 1999). Cette dernière est la protéine Npl3p faiblement associée à la snRNP U1. Les protéines SR sont également en contact avec l'ARN polymérase II et pourraient permettre le couplage entre transcription et épissage (pour revue, Bentley, 1999).

Les protéines SR participent à la régulation de l'épissage en modulant la sélection de sites d'épissage alternatif, par fixation sur des éléments ESE (pour "Exonic Splicing Enhancers"). Ces éléments sont, en général, riches en purines (pour revue, Bentley, 1999). Le degré de phosphorylation des protéines SR joue un rôle important dans la fonction et/ou la régulation de l'activité de ces protéines.

Nous avons vu que les protéines SR participent de manière primordiale à la formation du complexe d'engagement. Elles pourraient également intervenir dans les étapes ultérieures d'assemblage du spliceosome en jouant un rôle dans l'incorporation de la tri-snRNP [U4/U6.U5] au spliceosome (Rosigno et Garcia-Blanco, 1995) et dans l'établissement de l'interaction entre les snRNA U2 et U6 (Tarn et Steitz, 1995). Les protéines SR stabiliseraient donc les interactions entre les différentes UsnRNP et l'ARN pré-messager. Le nombre élevé de protéines SR chez les vertébrés pourrait contribuer à pallier la faible stabilité des interactions entre les UsnRNA des spliceosomes U2-dépendants et l'ARN pré-messager. Ces protéines semblent également intervenir dans l'épissage des introns U12-dépendants car l'insertion d'éléments ESE dans l'exon aval stimule l'épissage de l'intron U12-dépendant situé en amont (Wu et Krainer, 1998) et l'ajout de protéines SR, dans un extrait S100 de cellules HeLa, permet l'épissage *in vitro* des introns U12-dépendants (A. Krainer et coll., résultats non publiés), comme c'est le cas pour les introns U2-dépendants.

La présence des protéines SR essentiellement chez les vertébrés et chez les plantes (Lopato *et al.*, 1996; pour revue, Manley et Tacke, 1996) pourrait être liée au fait qu'elles interviennent dans le mode de reconnaissance à partir des exons qui est plus spécifique à ces organismes (Fig. 18 et 25). Les protéines SR pourraient également jouer un rôle dans la définition d'exons placés entre un intron U12-dépendant et un intron U2-dépendant (Fig. 25). En effet, il a été montré que l'épissage U12-dépendant de l'intron 2 du gène SCN4A humain est stimulé *in vitro* lorsque l'exon 3 est suivi d'un site 5' U2-dépendant. Cette stimulation est dépendante de la présence de la snRNP U1 (Wu et Krainer, 1996; 1998). La définition d'exon entre les deux systèmes se ferait donc par contact direct entre la snRNP U1, fixée au site 5' aval de l'intron U12-dépendant et la machinerie d'épissage U12-dépendante située en amont, et ceci probablement par l'intermédiaire de protéines SR (Fig. 25) (pour revue, Wu et Krainer, 1999).

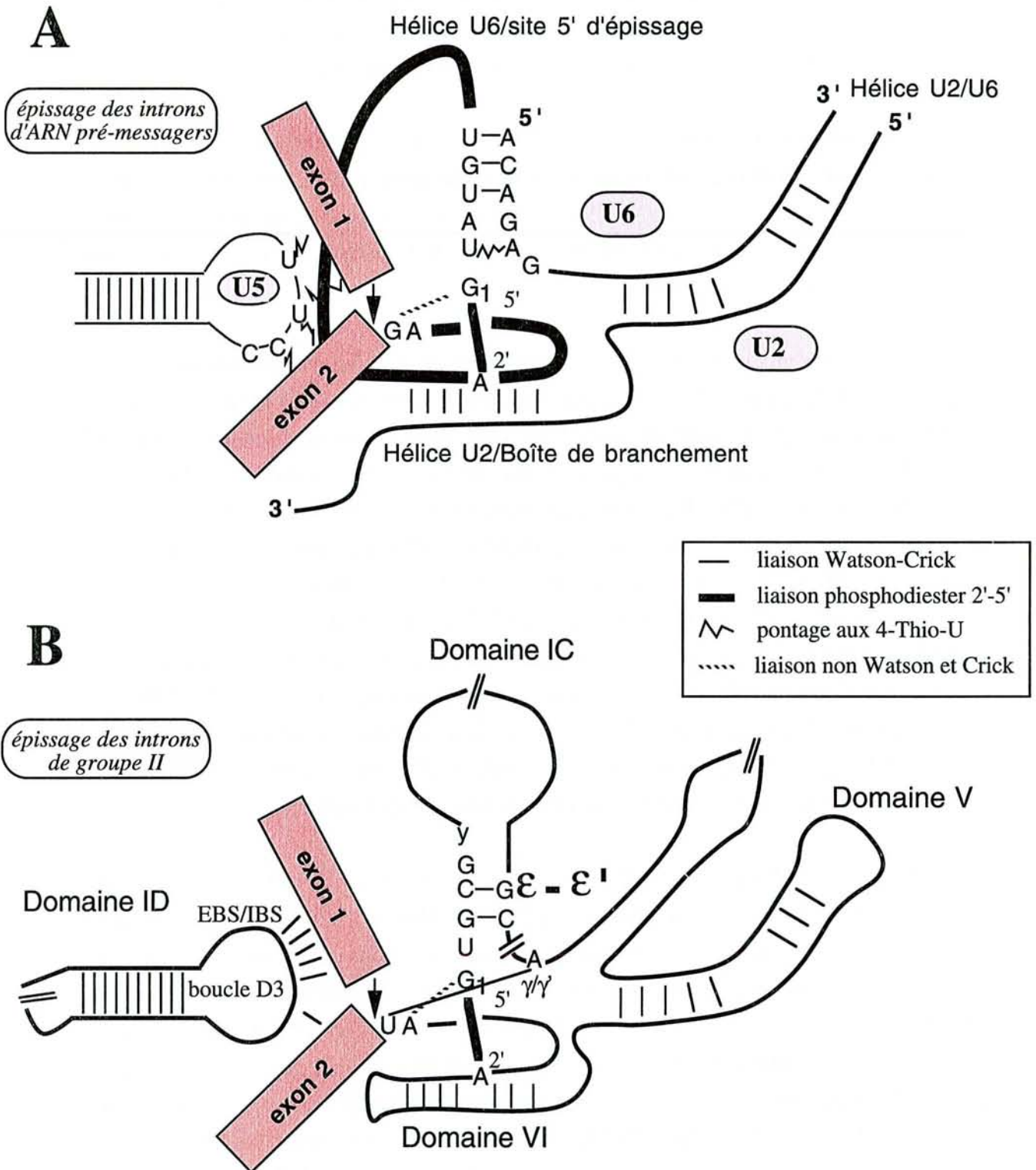


Figure 26: Comparaison des interactions ARN/ARN mises en jeu dans l'épissage des introns des ARN pré-messagers (A) et des introns de groupe II (B) (pour revues, Wise, 1993; Jacquier, 1997).

A. Interactions au niveau du site catalytique du spliceosome au cours de la deuxième réaction de transestérification, les séquences utilisées sont les séquences consensus chez *S. cerevisiae*, **B.** Interactions au site catalytique des introns de groupe II s'établissant au cours de la deuxième réaction de transestérification. Les séquences conservées des introns de la forme IIA sont indiquées.

E/ LE POINT DE VUE PHYLOGENETIQUE

1. UN ANCETRE COMMUN AUX INTRONS DES ARN PRE-MESSAGERS NUCLEAIRES ET AUX INTRONS DE GROUPE II?

Comme nous l'avons vu, la structure des ARN au sein du site catalytique semble très conservée entre les spliceosomes U2- et U12-dépendants. De plus, elle présente de fortes homologies avec la structure du site catalytique des introns de groupe II (Fig. 26) (pour revue, Jacquier, 1997). Comme nous l'avons déjà mentionné, l'épissage des introns de groupe II et l'épissage des ARN pré-messagers implique les mêmes attaques nucléophiles, d'où l'idée que les introns des ARN pré-messagers pourraient avoir un ancêtre commun avec les introns de groupe II.

Cette hypothèse est étayée par le fait qu'une structure équivalente à celle formée par le snRNA U2 et la boîte de branchement est retrouvée dans les introns de groupe II (domaine VI). Dans ce cas aussi, le nucléotide de branchement, qui est également une adénosine, est situé en boucle latérale.

Le domaine V des introns de groupe II, situé immédiatement en amont du domaine VI, joue un rôle essentiel dans l'épissage de ces introns (Koch *et al.*, 1992; Chanfreau et Jacquier, 1994). Ce domaine présente certaines homologies avec la structure formée entre les snRNA U2 et U6 (pour revue, Wise, 1993).

L'appariement ε - ε' à longue distance, entre la séquence en 5' du site d'épissage et un dinucléotide du domaine IC dans les introns de groupe II (Jacquier et Michel, 1990), peut être considéré comme l'homologue de l'interaction entre le snRNA U6 et la séquence en 5' de l'intron, dans les introns épissés dans un spliceosome.

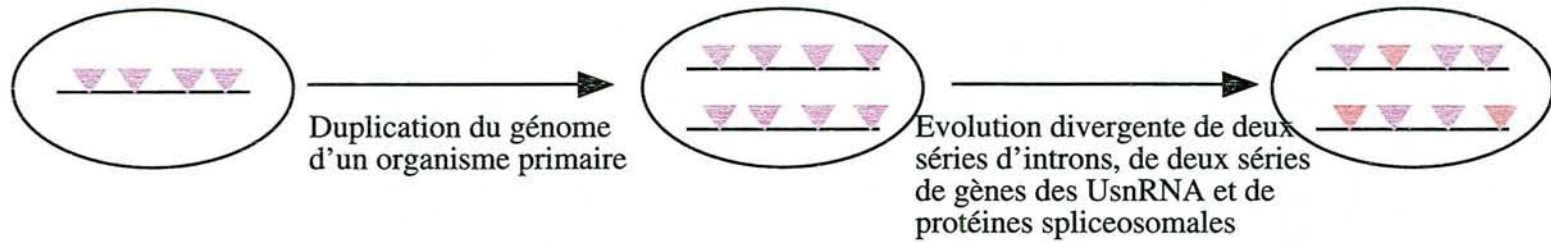
De plus, la boucle D3 du domaine ID des introns de groupe II joue, pour ces introns, un rôle similaire à celui qui est proposé pour le snRNA U5 dans le cas des introns épissés dans un spliceosome. En effet, la séquence EBS de cette boucle s'apparie à l'exon 1 lors de la première étape de la réaction et servirait à aligner les deux exons pour la deuxième étape de la réaction (Jacquier et Michel, 1987; Augustin *et al.*, 1990; Morl et Schmelzer, 1990; Morl *et al.*, 1992). De plus, le snRNA U5 humain, ajouté en trans dans des expériences d'épissage *in vitro* d'un intron de groupe II, peut remplacer, avec cependant une efficacité moindre, la structure tige-boucle D3 en interagissant avec les extrémités des exons de l'intron de groupe II utilisé (Hetzer *et al.*, 1997). Ceci montre une fonction parallèle de la boucle du snRNA U5 et de la boucle D3 du domaine I pour l'épissage des ARN pré-messagers nucléaires et des introns de groupe II, respectivement.

Enfin, le premier nucléotide de l'intron (G1) interagit avec le site 3' des introns de groupe II (Chanfreau et Jacquier, 1993), mimant l'interaction entre les deux résidus G aux extrémités des introns des ARN pré-messagers nucléaires.

La découverte de ces similitudes a donc permis de suggérer que les introns de groupe II et les introns des ARN pré-messagers nucléaires pourraient avoir une origine commune (pour revue, Jacquier, 1997). Les introns de groupe II ancestraux auraient pénétré dans le génome nucléaire d'eucaryotes ancestraux, au cours d'événements qui ont abouti à l'endosymbiose des ancêtres des mitochondries et des chloroplastes dans les cellules eucaryotes primitives (Cavalier-Smith, 1991). Une fois intégrés dans le génome nucléaire, ces introns se seraient progressivement fragmentés pour fournir les introns nucléaires et les gènes des UsnRNA. Dans le sens de cette hypothèse, des introns de groupe II ont en effet été découverts dans les ancêtres des mitochondries et des chloroplastes actuels, les cyanobactéries et les bactéries pourpres, respectivement (Ferat et Michel, 1993). De même, le transfert de gènes d'organites vers le noyau a été prouvé expérimentalement (Thorsness et Fox, 1990) et des introns de groupe II fragmentés ont été découverts (Goldschmidt-Clermont *et al.*, 1991; Wissinger *et al.*, 1991). Le nombre de protéines impliquées, ainsi que leurs rôles, se sont sans doute diversifiés au cours du temps, jusqu'à l'obtention des spliceosomes actuels.

Il n'y a pas de résultats démontrant l'implication directe des UsnRNA dans la catalyse, au sein des spliceosomes. Cependant, en plus de l'analogie structurale entre spliceosomes et introns de groupe II, les arguments en faveur d'une catalyse de la réaction d'épissage par les UsnRNA sont nombreux (pour revue, Nilsen, 1998): i) les UsnRNA sont absolument requis pour l'épissage, ii) la réaction catalytique procède par deux étapes de transestérification analogues à celles observées pour les introns de groupe II, iii) certaines mutations des UsnRNA ont des effets sur la catalyse, incluant des effets spécifiques à la deuxième réaction d'épissage et, iv) la première réaction catalytique, au moins, est dépendante d'un ion magnésium suggérant une catalyse par l'ARN (Steitz et Steitz, 1993; Sontheimer *et al.*, 1997). Néanmoins, certaines protéines doivent jouer un rôle essentiel pour la constitution de l'entité catalytique.

A/ Codivergence



B/ Fission/Fusion

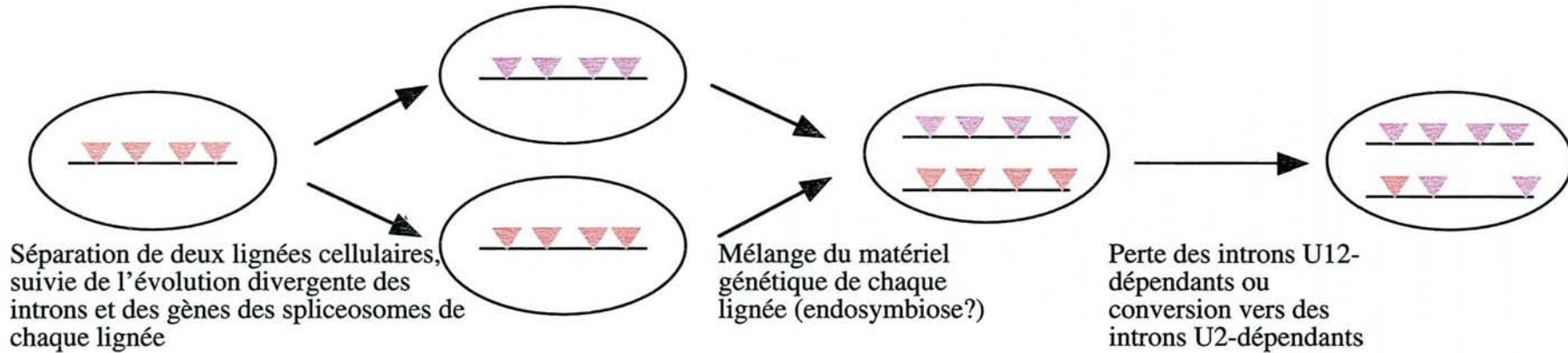


Figure 27: Les différents modèles proposés pour l'évolution des spliceosomes
Cette figure est tirée de Burge *et al.* (1998).

▼ introns U2-dépendants ▼ introns U12-dépendants

2. MODELE PROPOSE POUR LA GENERATION DES DEUX SYSTEMES DISTINCTS D'ÉPISSAGE DES ARN PRE-MESSAGERS NUCLEAIRES

Depuis la récente découverte des introns U12-dépendants chez les métazoaires et les plantes, de nombreux auteurs ont posé la question de la relation phylogénétique entre les deux systèmes d'épissage et de la, ou des, raisons pour lesquelles les organismes maintiennent une deuxième machinerie d'épissage pour l'élimination d'un nombre d'introns très réduit. L'observation d'introns U12-dépendants, chez la plante *A. thaliana* (Fig. 2), montre que les deux mécanismes d'épissage ont existé côte à côte depuis au moins un billion d'années (Wu *et al.*, 1996; Burge *et al.*, 1998).

Sur la base de l'examen des introns U2- et U12-dépendants dans les espèces où ils coexistent et d'une analyse statistique des séquences des sites 5' d'épissage et des boîtes de branchement, Burge *et al.* (1998) ont proposé plusieurs modèles possibles d'évolution des spliceosomes U2- et U12-dépendants (Fig. 27).

Selon un premier modèle (Fig. 27A), qui semble peu vraisemblable, il y aurait eu duplication du génome d'un organisme avec une évolution divergente de deux séries d'introns, de gènes de UsnRNAs et de gènes de protéines spliceosomales, toujours présents dans cet organisme.

Le second modèle de Fission/Fusion (Fig. 27B) est plus vraisemblable. Il implique que la divergence entre les deux systèmes se soit produite après séparation de deux lignées cellulaires, chacune avec un seul type de spliceosome et d'intron, suivi par un mélange du matériel génétique, peut-être par des mécanismes d'endosymbiose dans un ancêtre des eucaryotes supérieurs. Dans ce modèle, les spliceosomes U12-dépendants ont été retenus pour épisser les introns des gènes acquis pendant l'endosymbiose. Par la suite, la majorité des introns U12-dépendants aurait été soit perdue, soit convertie en introns U2-dépendants. Les UsnRNA et les facteurs protéiques qui n'auraient pas beaucoup divergés, par exemple le snRNA U5, seraient devenus redondants après la fusion des deux lignées et une seule version aurait donc été retenue. Une explication alternative de la présence de certains de ces facteurs communs est qu'ils seraient devenus des acteurs de la réaction d'épissage seulement après l'étape de fusion.

Ce modèle est en accord avec plusieurs observations. Il implique la présence de deux classes de gènes ayant des origines différentes: i) une classe dérivant des organismes à introns U2-dépendants et qui ne doivent pas contenir d'introns U12-dépendants et ii) une classe dérivant des organismes à introns U12-dépendants et qui doivent contenir, à la fois, des introns U2-dépendants et U12-dépendants du fait de la conversion possible des introns U12-dépendants en introns U2-dépendants, l'inverse n'étant pas possible. En accord avec ce

modèle, l'analyse de la localisation des différents introns dans les génomes montre bien l'existence des deux classes de gènes mentionnées ci-dessus. De plus, certains introns U12-dépendants sont conservés dans plusieurs espèces et/ou dans plusieurs gènes identiques d'une même espèce.

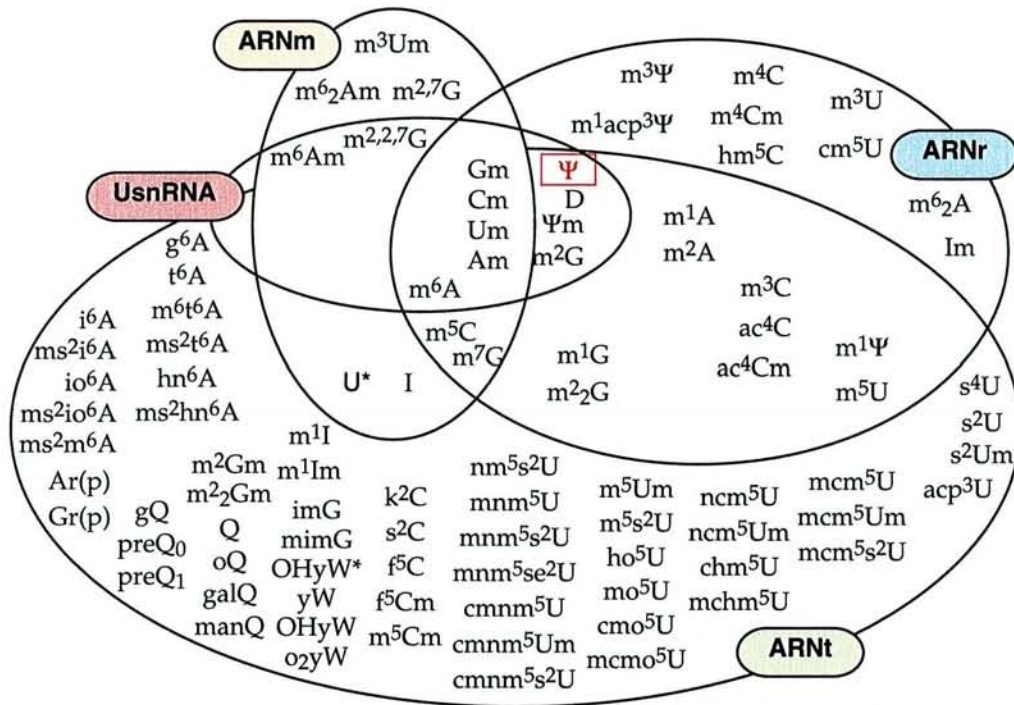
La comparaison des mêmes introns dans plusieurs espèces a montré que certains introns U12-dépendants, contenant des extrémités introniques GT-AG, sont remplacés dans d'autres espèces par des introns U2-dépendants, ce qui est en accord avec la conversion possible introns U12-dépendants → introns U2-dépendants. Toujours en accord avec cette hypothèse, Dietrich *et al.* (1997) ont montré qu'un intron épissé par les spliceosomes de type U12-dépendant pouvait être épissé *in vivo* par les spliceosomes de type U2-dépendant, par introduction de simple ou double mutations dans les sites 3' et 5' d'épissage.

En accord avec la disparition possible d'introns U12-dépendants, certains gènes d'espèces données ne contiennent pas d'intron, alors que les mêmes gènes dans d'autres espèces contiennent un intron U12-dépendant.

Les gènes impliqués dans la transmission de l'information génétique (réparation/réplication de l'ADN, transcription et traduction) sont souvent plus similaires à leurs homologues des *Archaea* qu'à leurs homologues des bactéries (pour revue, Brown et Doolittle, 1997). Il est intéressant de noter qu'une large proportion des introns U12-dépendants est trouvée dans ces gènes.

*PARTIE 2: LES RESIDUS PSEUDOURIDINES PORTES PAR LES
ARN; LOCALISATION, BIOSYNTHESE ET ROLES*

A



B

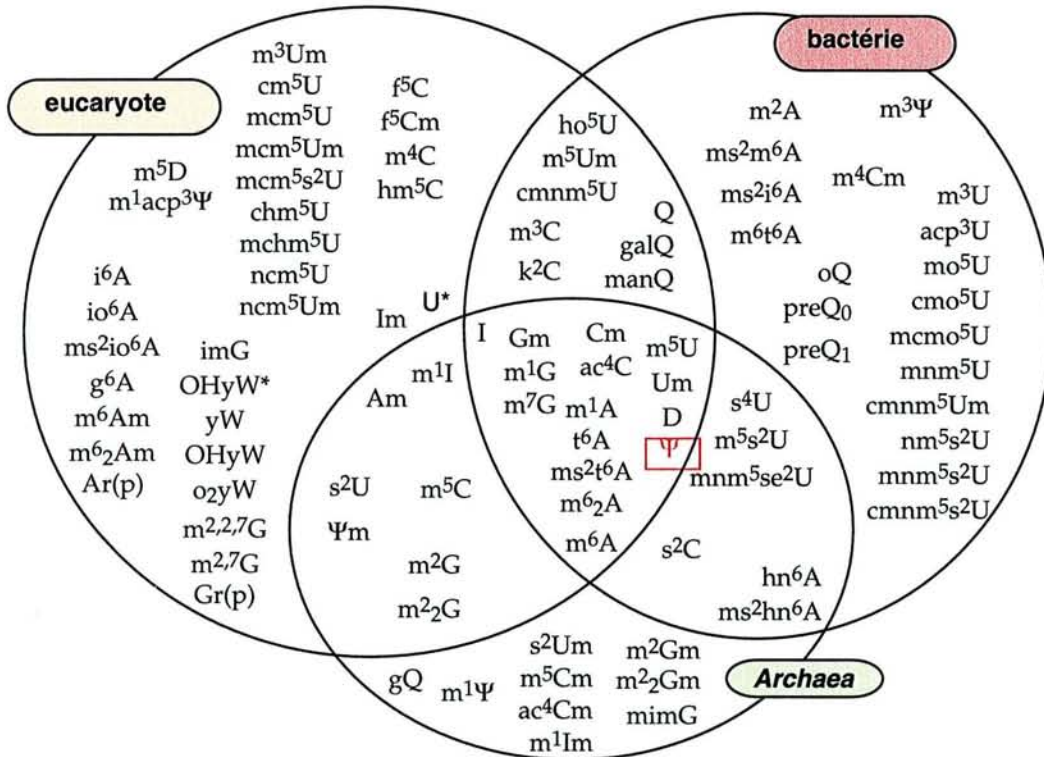


Figure 28: Distribution des nucléotides modifiés par type d'ARN (A) et dans les trois grands domaines: bactérie, eucaryote et Archaea (B)

U* correspond à l'uridine formée par déamination d'une cytosine par un mécanisme d' "editing" post-transcriptionnel.

Les résidus pseudouridines sont indiqués en rouge.

D'après la revue, Motorin et Grosjean, 1998.

A/ GENERALITES

Tous les ARN métaboliquement stables: ARNm, ARNt, ARNr, UsnRNA et snoRNA (pour "small nucleolar RNA") portent des modifications post-transcriptionnelles, soit au niveau des bases, soit au niveau des riboses. L'existence des nucléotides modifiés dans les ARN est une propriété commune aux organismes appartenant aux trois domaines: eucaryote, bactérie et *Archaea*. C'est en particulier le cas pour la présence de la pseudouridine (Ψ), dihydrouridine (D) ou encore de certaines méthylations sur le ribose (Fig. 28). Les nucléotides modifiés sont générés à partir des quatre nucléotides classiques, portant respectivement les quatre bases: adénine, cytosine, uridine et guanine. A l'heure actuelle, 95 nucléotides modifiés différents ont été recensés (pour revues, Limbach *et al.*, 1994; Motorin et Grosjean, 1998), dont au moins 81 appartiennent aux seuls ARNt (Fig. 28). L'annexe présente l'ensemble des nucléotides modifiés identifiés à ce jour.

La synthèse de beaucoup de modifications post-transcriptionnelles implique seulement l'action d'une ou deux enzymes. C'est le cas pour les réactions de méthylation sur les riboses ou sur les bases, ou les réactions de thiolation impliquées dans la formation de résidus tels que s^2U ou s^4U . C'est également le cas pour les réactions de déamination, d'isomérisation, d'alkylation et de réduction, impliquées respectivement dans la synthèse des résidus inosine (I), Ψ , N^6 -(Δ^2 -isopentenyl)-adénosine (i^6A) et D (Fig. 29) (pour revue, Garcia et Goodenough-Lashua, 1998). Par contre, la synthèse d'autres modifications post-transcriptionnelles, chimiquement plus complexes, nécessite une cascade d'activités enzymatiques, comme c'est le cas pour la formation des résidus mnm^5s^2U ou de la wybutosine (W). La formation des résidus queunine (Q) et archaeosine (gG) implique, quant à elle, la biosynthèse d'hétérocycles modifiés incorporés ensuite dans l'ARN par un mécanisme d'échange de bases (pour revue, Garcia et Goodenough-Lashua, 1998).

Malgré le fait que plus de 90 nucléotides modifiés soient recensés à ce jour, peu d'enzymes responsables de leur synthèse ont été identifiées et caractérisées. Il a été estimé que, dans un organisme comme *E. coli*, environ 45 enzymes de modification distinctes interviendraient dans la formation des nucléotides modifiés localisés dans les ARNt (pour revue, Björk, 1995). En posant l'hypothèse que la séquence codant pour une enzyme de modification est en moyenne de 1kb, cela signifierait que 1% (45kb) du génome bactérien coderait pour les enzymes de modification des ARNt, et ceci sans tenir compte des séquences codant pour les enzymes de modification des autres ARN.

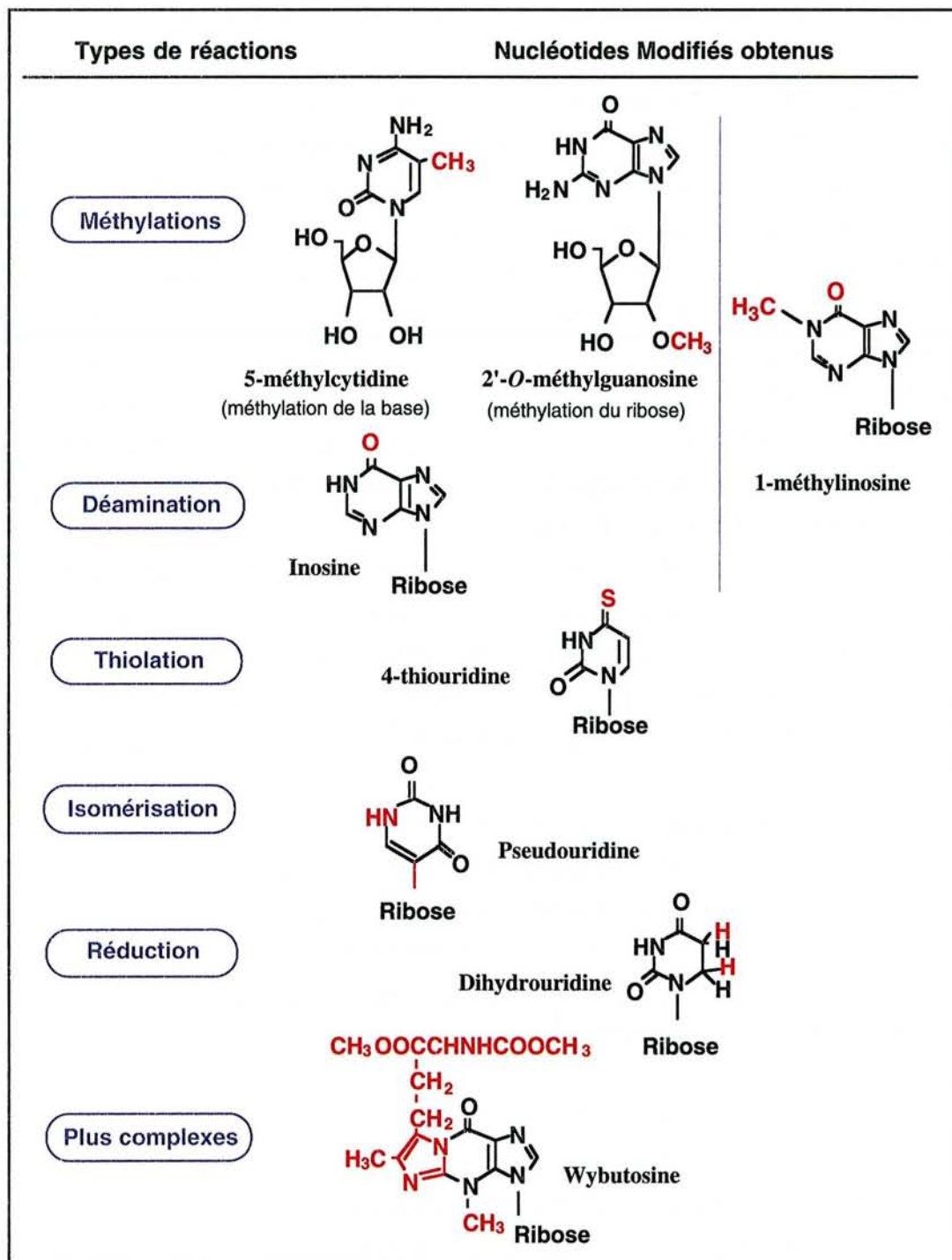


Figure 29: Structure chimique de quelques nucléotides modifiés.
D'après la revue, Limbach *et al.*, 1994.

Le résidu pseudouridine (5- β -D-ribofuranosyluracile, Ψ) et les résidus 2'-O-méthylés sont les plus abondants des nucléotides modifiés trouvés dans les ARN. Le résidu Ψ a été le premier nucléotide modifié identifié. Il a même été nommé le « cinquième nucléotide » (Davis et Allen, 1957). C'est au début des années 60 que Cohn décrivit la structure chimique de ce nouveau nucléotide (Cohn, 1959; 1960). Le résidu Ψ dérive du résidu U par une réaction d'isomérisation. Il y a rupture de la liaison N-glycosidique de l'uridine, rotation de 180° de la base uracile et enfin, création d'une liaison carbone-carbone entre le C₁' du ribose et le C₅ du cycle pyrimidique (Fig. 30) (Kammen *et al.*, 1988). Cette réaction est irréversible et ne requiert pas d'énergie. Les résidus Ψ sont synthétisés par une enzyme unique: la Ψ -synthase. Comme toutes les modifications post-transcriptionnelles, les résidus Ψ ne se forment pas aléatoirement le long de la séquence ARN, mais à des sites spécifiques.

Mon travail de thèse ayant porté principalement sur l'étude des résidus Ψ dans les ARN et surtout dans les UsnRNA, cette deuxième partie de l'introduction bibliographique sera consacrée presque essentiellement à cette modification post-transcriptionnelle particulière. Des exemples concernant d'autres modifications seront donnés, quand ils s'avèreront informatifs pour la problématique qui nous concerne.

Nous présenterons donc, dans un premier temps, la localisation des résidus Ψ dans les ARNt, les ARNr et dans les UsnRNA. Dans un deuxième temps, la nature, le mécanisme enzymatique et les mécanismes de reconnaissance de l'ARN des Ψ -synthases seront exposés. Enfin, nous aborderons la question des rôles prouvés ou putatifs de ce nucléotide modifié dans les différents ARN. Le cas des UsnRNA est développé de manière plus exhaustive dans la revue bibliographique que nous avons publiée dans ce domaine, et qui est insérée à la fin de cette partie de l'introduction.

B/ LOCALISATION DES RÉSIDUS Ψ ET DES AUTRES NUCLÉOTIDES MODIFIÉS DANS LES ARN: UN PREMIER PAS VERS LA CONNAISSANCE DE LEUR FONCTION

1. LOCALISATION DANS LES ARNt

Le taux de modification dans les ARNt est très variable d'un ARNt à un autre, et se situe entre 3 et 25 %. Ce sont les ARN contenant le plus de nucléotides modifiés, en tenant compte de leur taille réduite. Ce sont également les ARN présentant la plus grande variété de modifications, les nucléotides hypermodifiés étant principalement situés aux positions 34 (1^{er}

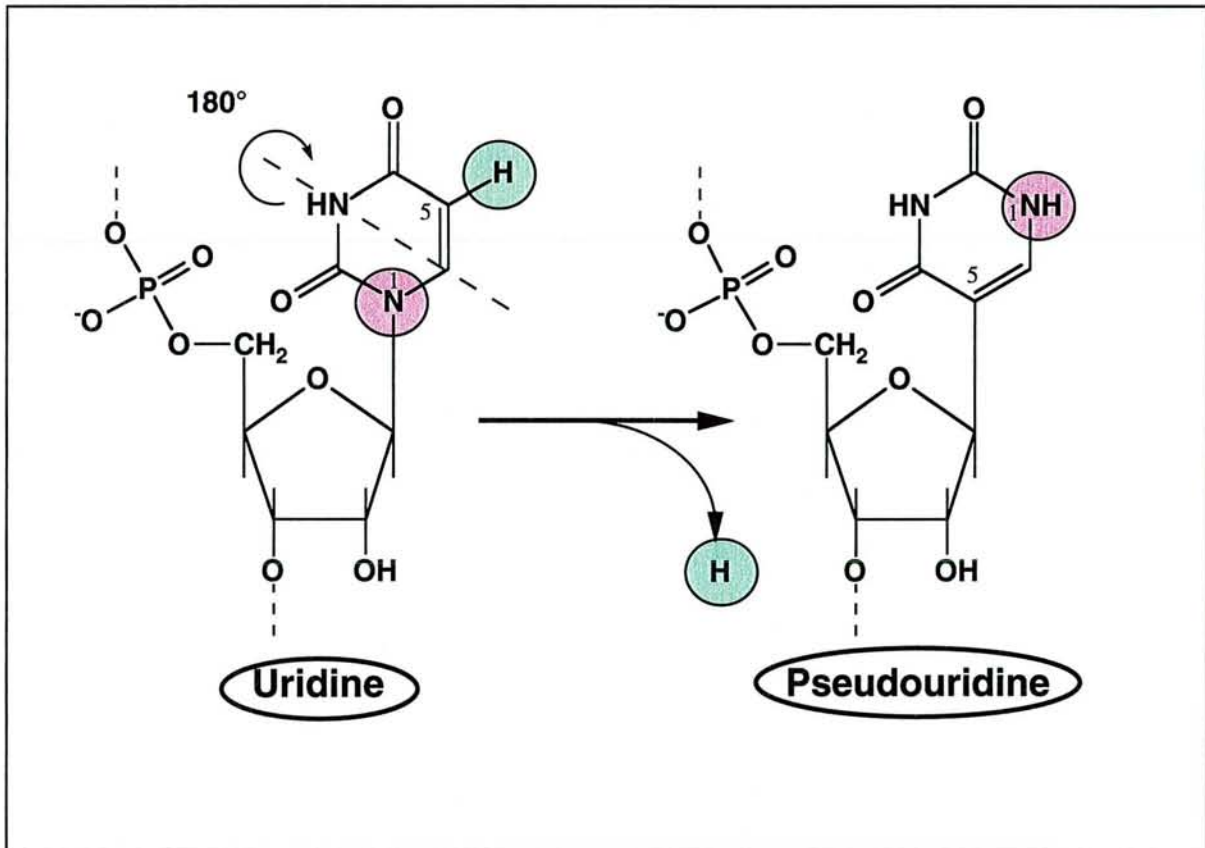


Figure 30: Isomérisation de l'uridine en pseudouridine

nt de l'anticodon) et 37 (nt 3' adjacent à l'anticodon) (Fig. 31). Les régions les plus modifiées sont les trois boucles des ARNt. Certaines modifications sont localisées à la même position dans la quasi-totalité des ARNt des trois domaines, bactérie, *Archaea* et eucaryote. C'est le cas des résidus $\Psi_{13, 38, 39 \text{ et } 55}$, Cm_{34} , t^6A_{37} , m^1G_{37} et m^1A_{58} (pour revue, Cermakian et Cedergren, 1998). Dans deux des boucles de l'ARNt, les résidus dihydrouridine (D_{20}), ribothymidine (T_{54} ou m^5U_{54}) et Ψ_{55} sont pratiquement toujours conservés et c'est pourquoi ces boucles ont été respectivement dénommées boucle D et boucle T Ψ (Fig. 31 et 32). Cependant, il existe des nucléotides modifiés spécifiques d'un domaine, par exemple les résidus $\text{mnm}^5\text{s}^2\text{U}$, $\text{m}^1\Psi$ et manQ ne se trouvent respectivement que chez les bactéries, les *Archaea* et les eucaryotes (Fig. 28) (pour revue, Grosjean *et al.*, 1995). On note une augmentation du nombre et de la variété de nucléotides modifiés au passage procaryote/eucaryote. Beaucoup des modifications portées par les ARNt cytoplasmiques d'eucaryotes sont également trouvées, aux mêmes positions, dans les ARNt mitochondriaux. Comme contre-exemple, on peut citer le cas du résidu inosine, qui est présent dans les ARNt cytoplasmiques, mais n'est pas détecté dans les ARNt mitochondriaux.

2. LOCALISATION DANS LES ARNr

L'état des connaissances est décrit dans certaines revues générales (Maden, 1990 et Maden *et al.*, 1995) et dans des revues plus spécifiques sur la localisation des résidus Ψ (Ofengand et Fournier, 1998) ou des résidus 2'-O-méthylés (Bachelierie et Cavaillé, 1998).

Les ARNr contiennent essentiellement, comme nucléotides modifiés, des résidus Ψ et des résidus méthylés soit sur le ribose (cas le plus fréquemment observé), soit sur les bases. Il existe en moyenne dix fois plus de sites de méthylation sur le ribose et de pseudouridylation chez les vertébrés que chez *E. coli* (Tableau 9). Par contre, les sites de méthylation sur les bases sont plus fréquents chez les bactéries.

Dans l'ARN de la petite sous-unité ribosomique, le seul résidu Ψ trouvé chez *E. coli* et *B. subtilis* n'a pas d'homologue chez les eucaryotes, ceci en dépit de la présence de beaucoup plus de sites pseudouridylés chez ces organismes: 14 chez *S. cerevisiae* et 36 chez l'homme. Il y a, par contre, une grande correspondance dans les localisations des résidus Ψ chez différents eucaryotes, bien que certains sites soient espèce-spécifiques. De plus, il est à noter qu'il n'existe pas de regroupement marqué des nucléotides modifiés dans un domaine particulier de l'ARNr de la petite sous-unité, contrairement à la situation observée dans la grande sous-unité ribosomique (Fig. 33).

En effet, dans cette sous-unité qui contient l'ARN 23S-28S (selon les espèces) et les ARNr 5S et 5,8S (spécifiques chez les eucaryotes), toutes les modifications portées par l'ARNr 23S-28S

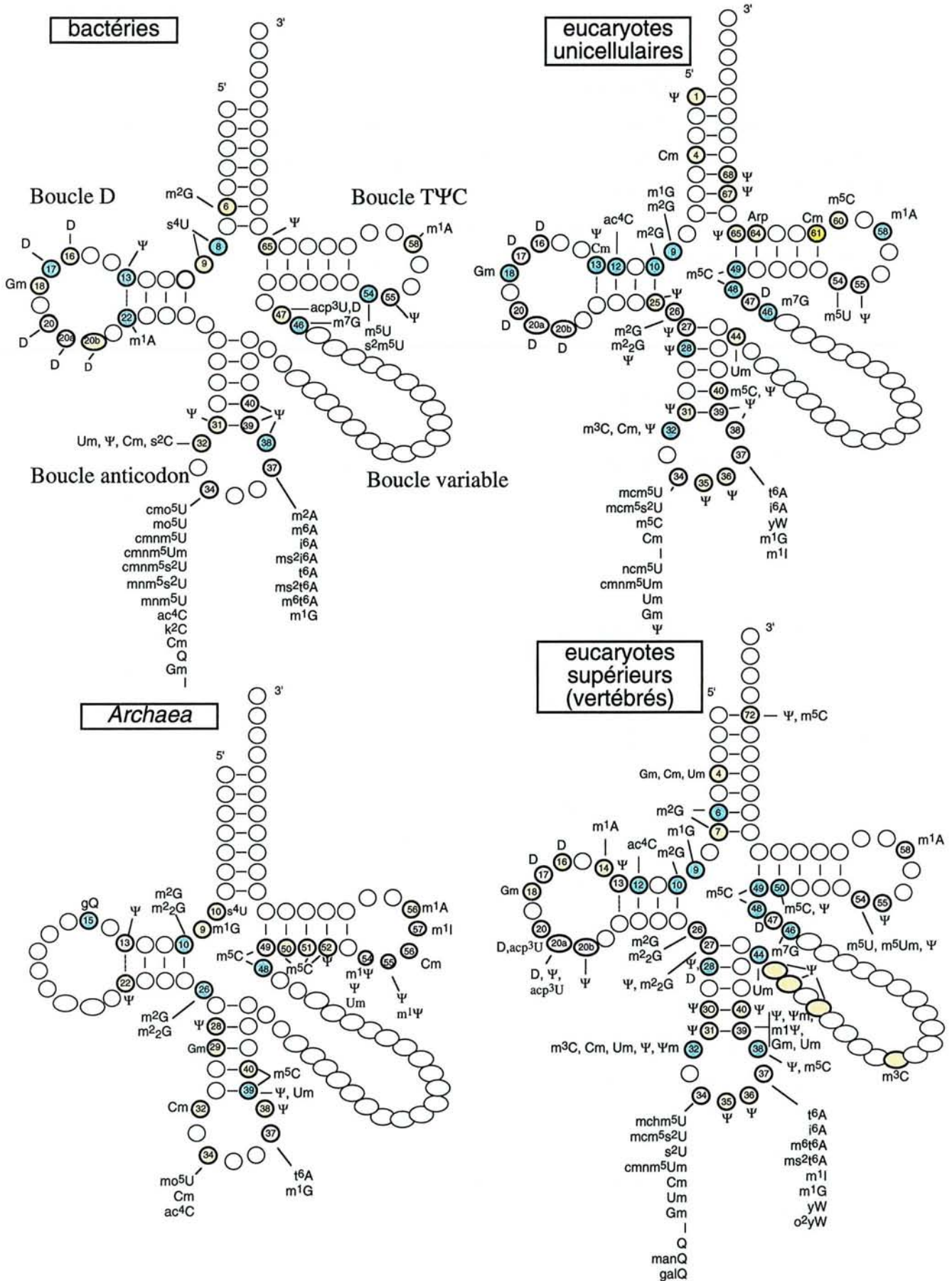


Figure 31: Les nucléotides modifiés dans les ARNt, leur localisation et leur fréquence.
D'après la revue, Grosjean *et al.*, 1995.

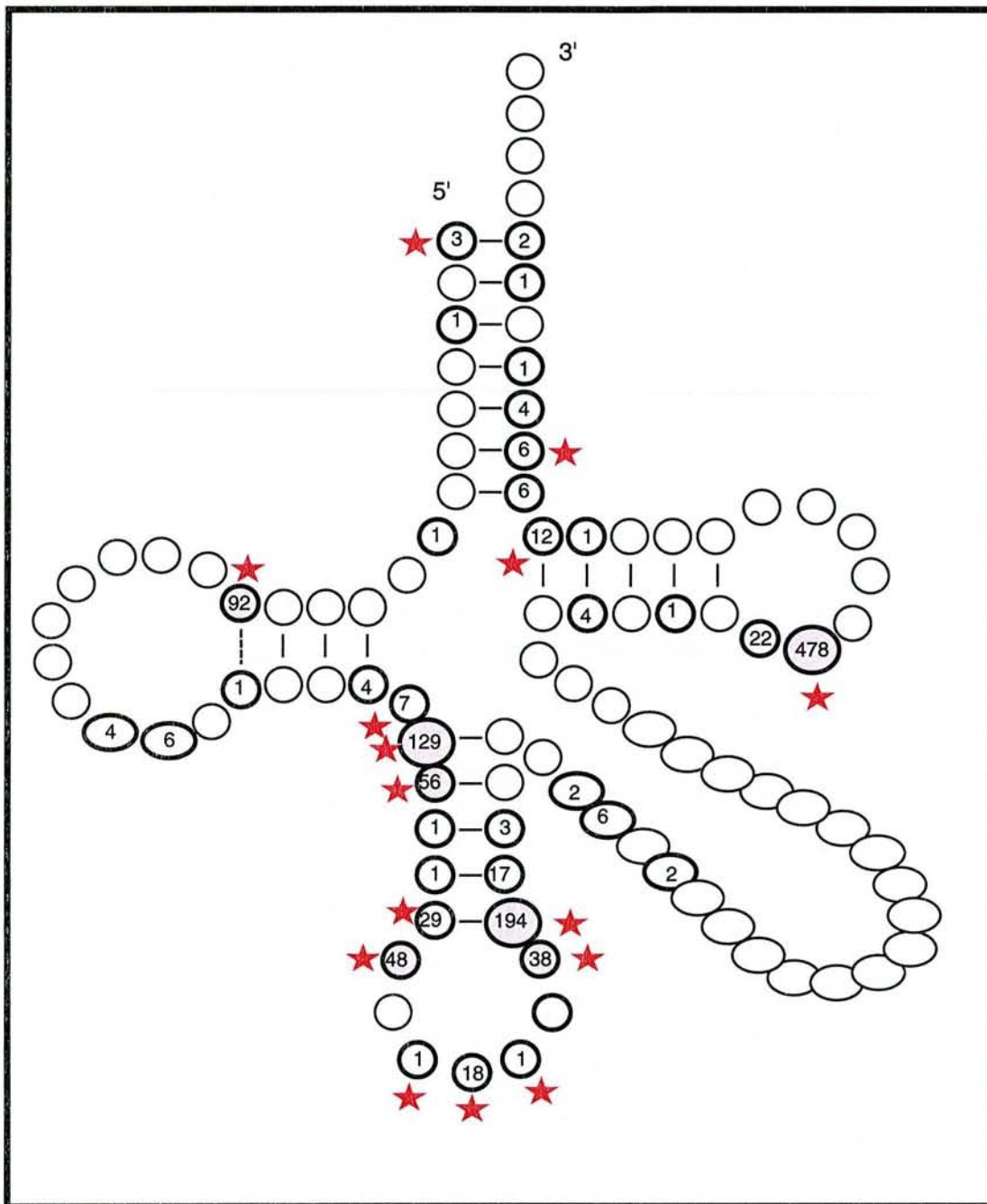


Figure 32: Sites et fréquence de formation des résidus Ψ dans les ARNt cytoplasmiques des eucaryotes et dans les ARNt des *Archaea* et bactéries.

La figure est établie à partir de 546 séquences d'ARNt. Le nombre de résidus Ψ trouvé en chacun des sites est indiqué dans les ronds. Les positions contenant des résidus Ψ chez *S. cerevisiae* sont indiquées par des étoiles.

D'après les revues, Grosjean *et al.*, 1995; Auffinger et Westhof, 1998.

A

Organisme	Nombre de moles de Ψ /mole d'ARN de la petite sous-unité ribosomique
<i>E. coli</i>	1
<i>B. subtilis</i>	1
<i>H. halobium</i>	0
<i>S. carlsbergensis</i>	11
<i>S. cerevisiae</i>	14
<i>X. laevis</i>	44
<i>M. musculus</i>	~36
<i>R. norvegicus</i>	~37
<i>H. sapiens</i>	~36

B

Organisme	Nombre de moles de Ψ /mole d'ARN de la grande sous-unité ribosomique
A. ARNr cytoplasmique	
<i>E. coli</i>	9
<i>B. subtilis</i>	5
<i>H. halobium</i>	4
<i>S. carlsbergensis</i>	~32
<i>S. cerevisiae</i>	30
<i>D. melanogaster</i>	57
<i>X. laevis</i>	~52
<i>M. musculus</i>	~57
<i>H. sapiens</i>	~57
B. ARNr mitochondrial	
<i>S. cerevisiae</i>	1
<i>M. musculus</i>	1
<i>H. sapiens</i>	1
<i>T. brucei</i>	6
C. ARNr chloroplastique	
<i>Z. mays</i>	4

Tableau 9: Nombre de résidus pseudouridylés présents dans les grands ARNr des ribosomes cytoplasmiques, mitochondriaux et chloroplastiques de différents organismes (Tableau A, ARN de la petite sous-unité, et Tableau B, grand ARN de la grande sous-unité).

D'après les revues, Maden, 1990 ; Ofengand et Fournier, 1998.

sont regroupées dans 3 domaines (II, IV et V) (Branlant *et al.*, 1981b, pour revue, Ofengand et Fournier, 1998), qui sont situés dans la partie supérieure active de la sous-unité 50-60S. Certaines régions des domaines IV et V font partie intégrante du centre peptidyl transférase (PTC) du ribosome (pour une revue récente, Noller, 1998). Le domaine II joue un rôle important dans la fixation des facteurs d'élongation (Moazed *et al.*, 1988). Bien que tous ces sites de modification soient localisés dans le centre actif du ribosome, dans tous les organismes étudiés, les positions exactes de chaque nucléotide modifié sont modérément conservées au cours de l'évolution (Fig. 34). On peut seulement noter la présence constante, dans toutes les espèces étudiées, de 2 résidus Ψ , sur les 55 trouvés chez l'homme et les 30 trouvés chez *S. cerevisiae*, aux positions 2257 et 2259 dans le domaine central (numérotation de l'ARN 26S de *S. cerevisiae*) (Fig. 34). La position des résidus Ψ chez les ARNr des *Archaea* n'est connue que pour *H. halobium*, et, comme nous allons le voir, nous avons mené une étude sur une *Archaea* hyperthermophile, *S. acidocaldarius*. Chez *H. halobium*, le nombre de résidus Ψ est plus similaire de celui trouvé chez les bactéries que de celui trouvé chez les eucaryotes, par contre les positions de ces modifications sont plus similaires à celles observées chez les eucaryotes que chez les bactéries.

Contrairement aux ARNt mitochondriaux, les ARNr des mitochondries sont très peu modifiés par rapport aux ARNr cytoplasmiques (Tableau 9). En effet, un seul résidu Ψ a été détecté dans le grand ARN de la grande sous-unité ribosomique mitochondriale chez l'homme et chez *S. cerevisiae*.

3. LOCALISATION DANS LES UsnRNA

Comme pour les ARNt et les ARNr, les nucléotides modifiés portés par les UsnRNA sont localisés dans les régions fonctionnelles de ces ARN. Nous avons montré que leur nombre augmente en passant des eucaryotes unicellulaires aux eucaryotes multicellulaires. Les données dans ce domaine sont présentées dans la revue qui suit.

Il semblerait donc que l'augmentation du nombre de résidus modifiés au cours de l'évolution soit une règle générale pour l'ensemble des ARN: ARNr, ARNt et UsnRNA. La présence de ces modifications dans des régions fonctionnelles très conservées de ces ARN suggère que ces modifications jouent un rôle dans la fonction de ces molécules. La conservation, au cours de l'évolution, de certains nucléotides modifiés à un site donné, laisse supposer que ces nucléotides apportent un fort avantage à la cellule.

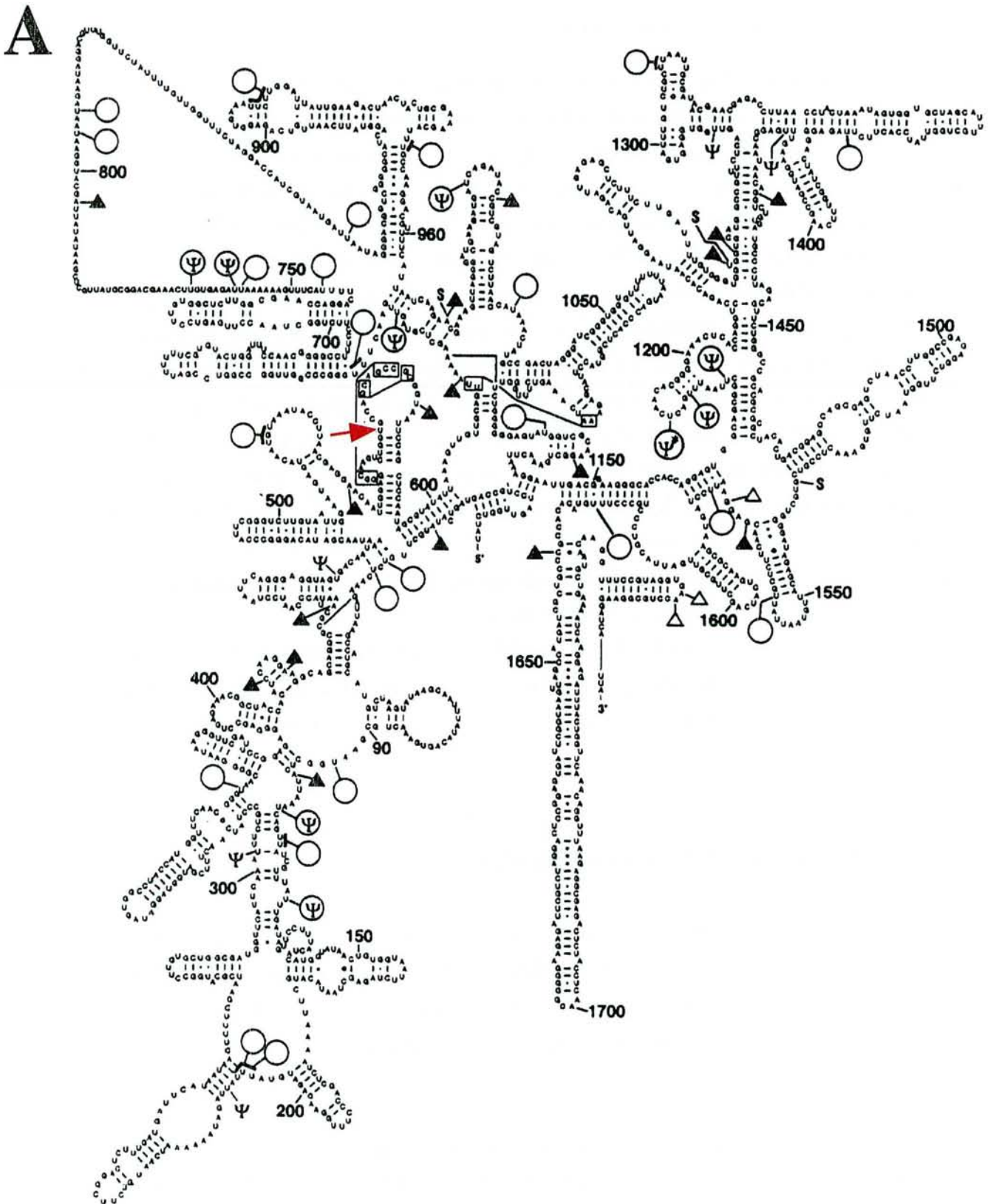


Figure 33A: Localisation des résidus Ψ et méthylés dans l'ARN de la petite sous-unité ribosomique .

La structure est celle de l'ARNr 18S de *S. cerevisiae*. Ψ indique les résidus Ψ dans l'ARNr de *S. cerevisiae*, \bigcirc les résidus Ψ dans les ARN de mammifères et \bigcirc les résidus communs à *S. cerevisiae* et aux mammifères. La flèche rouge indique la position de l'unique résidu Ψ dans l'ARN 16S d'*E. coli* et de *B. subtilis*. Les triangles blancs et noirs représentent, respectivement, les résidus méthylés sur les bases et les résidus méthylés sur le ribose.

D'après la revue, Ofengand et Fournier, 1998.

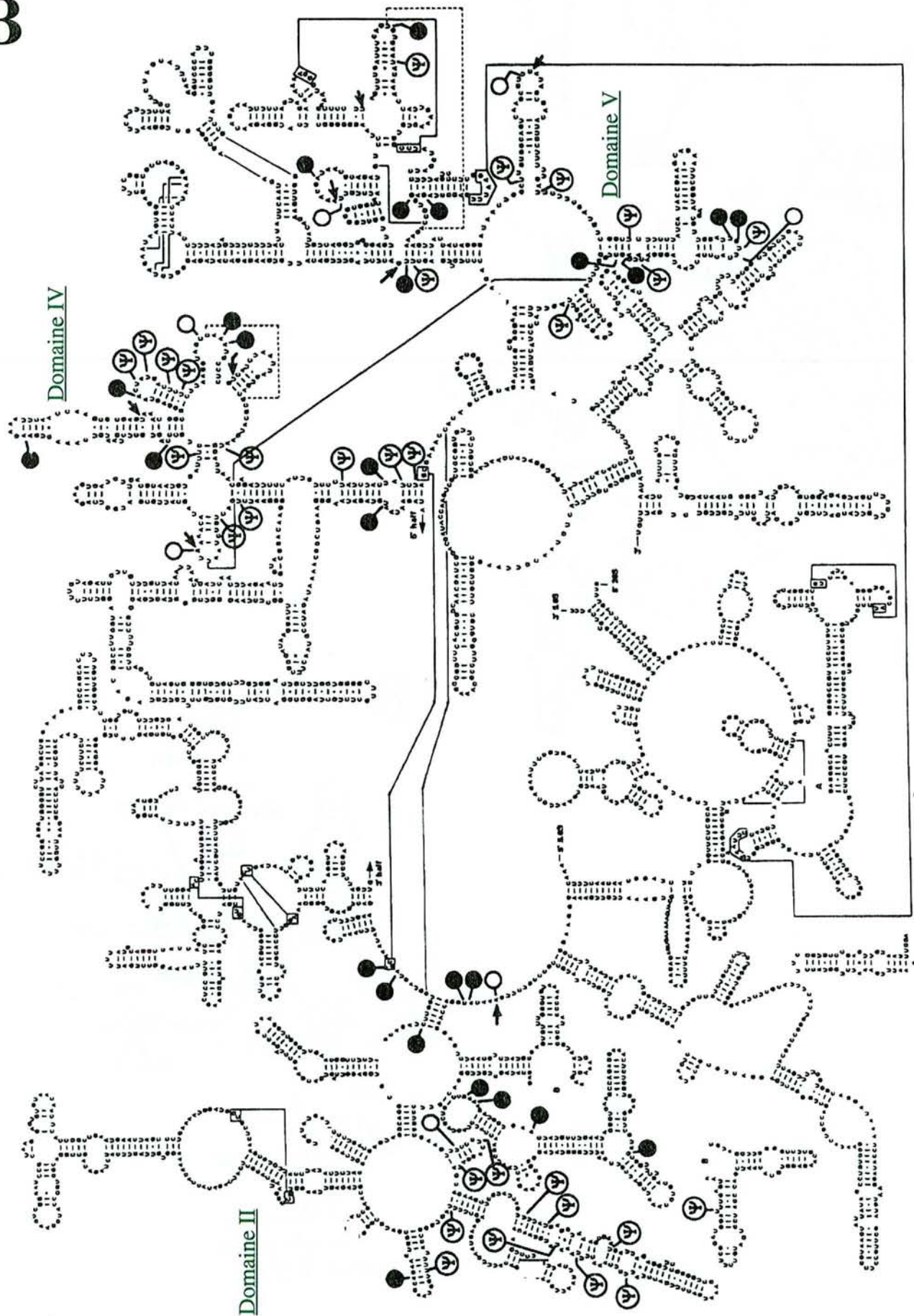
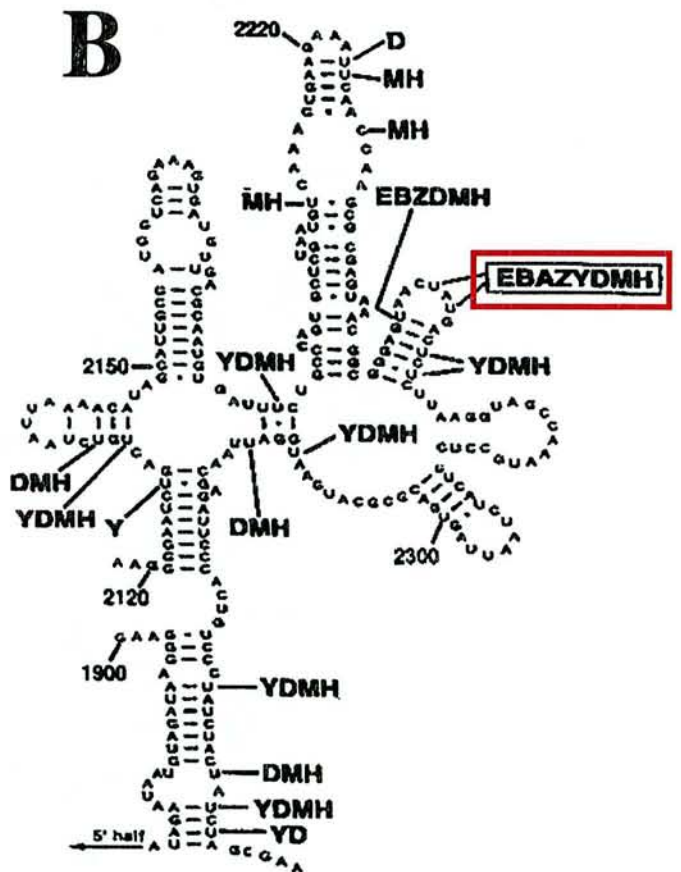
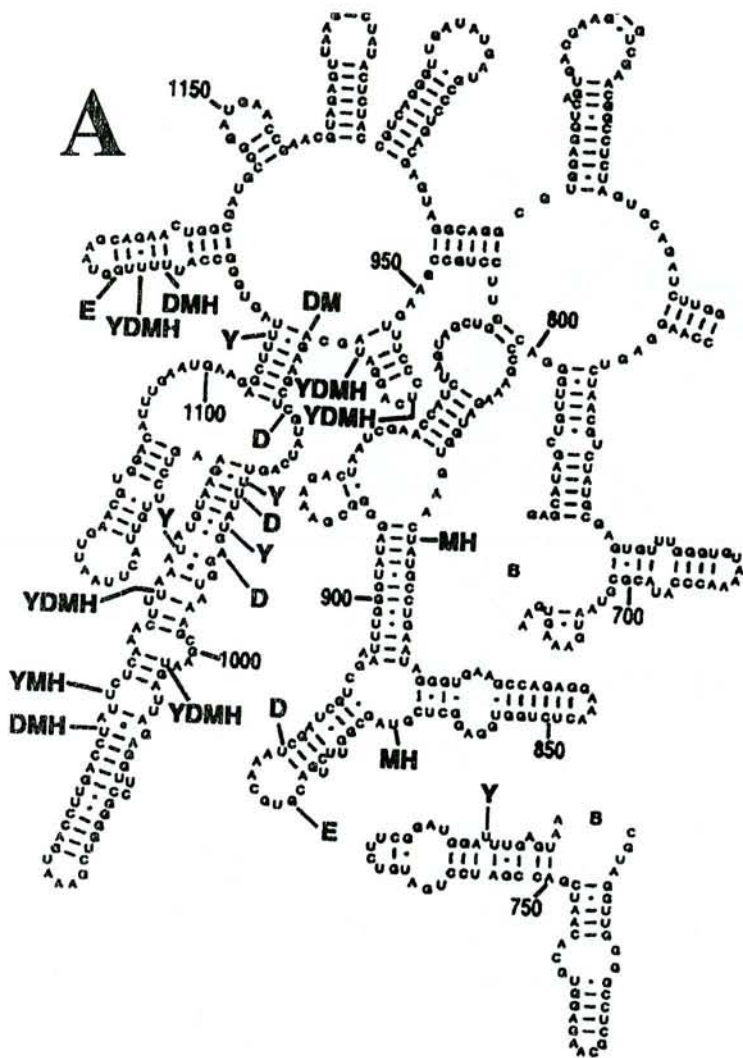
B

Figure 33B: Localisation des résidus Ψ et 2'-O-méthylés dans l'ARNr 26S de *S. cerevisiae*.

Ⓞ Ψ indique les résidus Ψ , les cercles blancs et noirs représentent les résidus méthylés sur les bases et les résidus méthylés sur le ribose, respectivement.

D'après Branlant *et al.*, 1981b et la revue, Ofengand et Fournier, 1998.



C

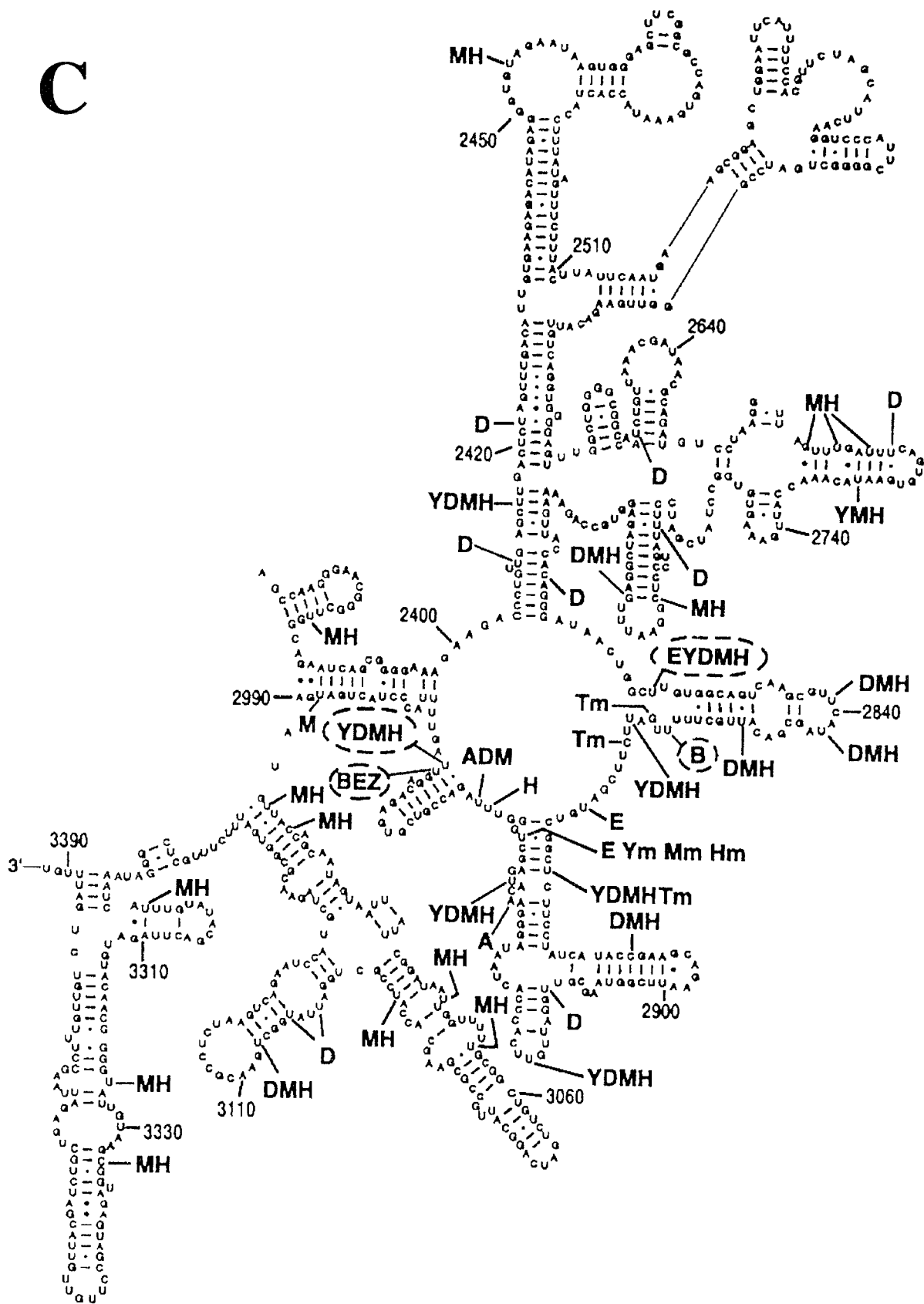


Figure 34: Comparaison chez plusieurs organismes des positions des résidus Ψ dans les domaines II (A), IV (B) et V (C) du grand ARNr de la grande sous-unité ribosomique.

La séquence est celle de l'ARNr 26S de *S. cerevisiae*. Les positions où sont localisées les résidus Ψ sont indiquées par des traits. Les symboles suivants représentent chacun une espèce étudiée: E, *E. coli*; B, *B. subtilis*; A, *H. halobium*; Z, chloroplastes de maïs; Y, *S. cerevisiae*; Ym, mitochondries de *S. cerevisiae*; D, *D. melanogaster*; M, *M. musculus*; Mm, mitochondries de *M. musculus*; H, *H. sapiens*; Hm, mitochondries de *H. sapiens*; T, mitochondries de *T. brucei*. Les deux résidus Ψ strictement conservés au cours de l'évolution sont indiqués en rouge, dans le panneau B.

D'après la revue, Ofengand et Fournier, 1998

Enzyme	Organisme	Localisation cellulaire chez les eucaryotes *	Substrat	Position(s) modifiée(s)	Références
TruB	<i>E.coli</i>		ARNt	55	Nurse <i>et al.</i> (1995)
Pus4p	<i>S.cerevisiae</i>	(noyau, mitochondrie)	ARNt	55	Becker <i>et al.</i> (1997a)
Cbf5p	<i>S.cerevisiae</i>	nucléole	ARNr	Formation de tous les résidus Ψ	Lafontaine <i>et al.</i> (1998a)
TruA	<i>E.coli</i>		ARNt	38, 39, 40	Kammen <i>et al.</i> (1988)
Pus3p	<i>S.cerevisiae</i>	noyau, mitochondrie et cytoplasme	ARNt	38, 39	Lecoite <i>et al.</i> (1998)
Pus1p	<i>S.cerevisiae</i>	noyau	ARNt snRNA U2	26, 27, 28, 65, 67, 34, 35, 36 44	Simos <i>et al.</i> (1996) Motorin <i>et al.</i> (1998) Massenet <i>et al.</i> (1999)
mPus1p	<i>M.musculus</i>	nd	ARNt snRNA U2	27, 28, 32?, 34, 36, ?	Chen et Patton (1999)
RluA	<i>E.coli</i>		ARNt ARN 23S	32 746	Wrzesinski <i>et al.</i> (1995a) Raychaudhuri <i>et al.</i> (1999)
RluC	<i>E.coli</i>		ARN 23S	955, 2504, 2580	Huang <i>et al.</i> (1998a) /Conrad <i>et al.</i> (1998)
RluD	<i>E.coli</i>		ARN 23S	1911, 1915, 1917	Huang <i>et al.</i> (1998a) Raychaudhuri <i>et al.</i> (1998)
RsuA	<i>E.coli</i>		ARN 16S	516	Wrzesinski <i>et al.</i> (1995b) /Conrad <i>et al.</i> (1999)
RluB	<i>B.subtilis</i>		ARN 23S	2633	Niu et Ofengand (1999)
RluE	<i>E.coli</i>		ARN 23S	2457	Conrad <i>et al.</i> (résultats non publiés)

Tableau 10: Liste des Ψ-synthases caractérisées et purifiées à ce jour.

Les quatre familles sont représentées par des couleurs différentes (vert: famille TruB; rose: famille TruA; bleu: famille RluA; noir: famille RsuA). * La localisation cellulaire des différentes enzymes a été étudiée par immunofluorescence, sauf celle de Pus4p qui a été déduite par l'étude directe de l'action de l'enzyme sur les ARNt provenant des différents compartiments cellulaires. nd : non déterminé.

C/ LES Ψ -SYNTHASES

1. IDENTIFICATION DES Ψ -SYNTHASES

Les premières identifications et purifications de Ψ -synthases ont été faites sur la base de techniques de chromatographie. Les activités étaient détectées grâce à une méthode mise au point par Cortese *et al.* (1974). Cette technique consiste à utiliser des ARN ayant incorporé des uraciles marqués au tritium (^3H) en position 5. Pendant la réaction d'isomérisation (Fig. 30), il y a relargage du tritium, qui peut être dosé. Cette approche biochimique a permis de purifier et de caractériser les premières Ψ -synthases: les enzymes TruA de *S. typhimurium* (Arena *et al.*, 1978) et d'*E. coli* (Kammen *et al.*, 1988), responsables de la formation des résidus Ψ en position 38, 39 et 40 dans les ARNt.

La plupart des Ψ -synthases d'*E. coli* peuvent être regroupées en quatre familles en fonction de leur séquence en acides aminés (Koonin, 1996): les familles RluA, RsuA, TruA et TruB. Le Tableau 10 présente ces différentes enzymes et indique la nature de l'ARN ou des ARN substrats et la ou les positions modifiées. Les Ψ -synthases responsables de la formation de 9 des 10 résidus Ψ portés par les ARNr d'*E. coli* ont été identifiées (voir tableau 10), ainsi que les enzymes modifiant les positions 32 (Wrzesinski *et al.*, 1995a), 55 (Nurse *et al.*, 1995), 38, 39 et 40 (Kammen *et al.*, 1988) des ARNt. A l'heure actuelle, seule la ou les enzymes responsables de la catalyse des résidus Ψ aux positions 13 et 65 des ARNt et 2605 dans l'ARNr 23S ne sont pas encore connues.

Par homologie de séquences avec les Ψ -synthases d'*E. coli*, les gènes correspondant de *S. cerevisiae* ont pu être identifiés et clonés, et les Ψ -synthases codées par ces gènes ont été purifiées et caractérisées par les équipes de E. C. Hurt, H. Grosjean et D. Tollervey (Tableau 10 et Fig. 35). Les sites modifiés par ces enzymes ont été pour la plupart identifiés, à la fois *in vitro* en utilisant les enzymes purifiées et des ARN cibles produits par transcription *in vitro*, et *in vivo* par analyse d'ARN issus de souches de levure dont les gènes correspondant à chacune de ces Ψ -synthases avaient été délétés.

L'enzyme Pus3p synthétise les résidus Ψ en position 38 et 39 (Lecoite *et al.*, 1998), alors que Pus4p est responsable de la formation du résidu Ψ_{55} (Becker *et al.*, 1997a) dans les ARNt. Pus1p convertit les résidus U en résidus Ψ à 8 positions distinctes dans les ARNt (Simos *et al.*, 1996; Motorin *et al.*, 1998) (Fig. 35). Toutefois, *in vivo*, une deuxième Ψ -synthase semble pouvoir compléter l'action de cette enzyme à la position 35 (Motorin *et al.*, 1998). Comme nous le verrons, nous avons montré que Pus1p est également responsable de la formation du résidu Ψ_{44} dans le snRNA U2. Un homologue de Pus1p a été identifié chez la souris et modifie *in vitro* les mêmes sites dans les ARNt que l'enzyme de *S. cerevisiae* (Chen

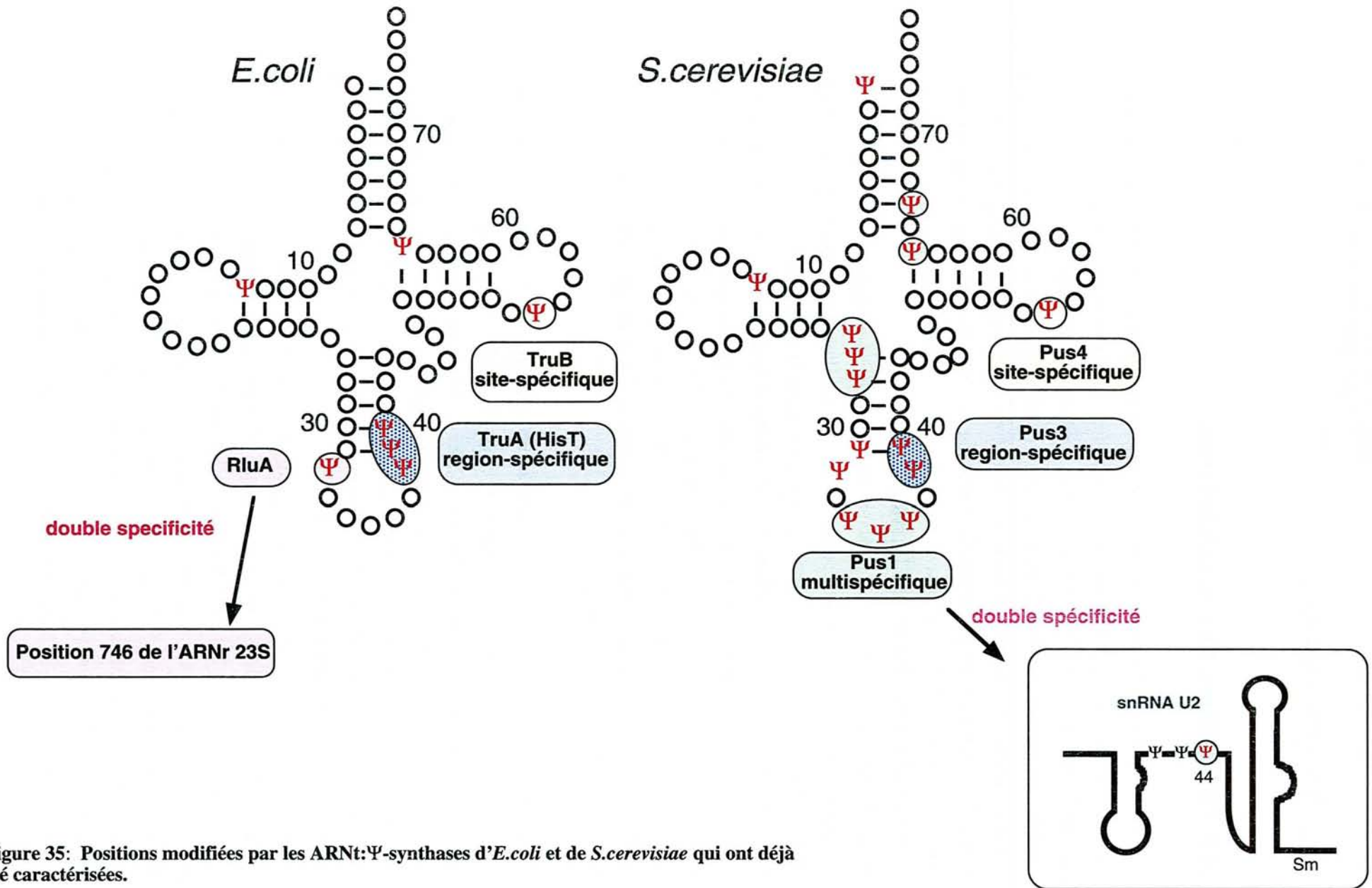


Figure 35: Positions modifiées par les ARNt:Ψ-synthases d'*E. coli* et de *S. cerevisiae* qui ont déjà été caractérisées.

Les références sont celles citées dans le Tableau 10

et Patton, 1999). Cbf5p semble être la Ψ -synthase responsable de la formation des résidus Ψ dans les ARNr cytoplasmiques de levure (Lafontaine *et al.*, 1998a). Nous reviendrons plus loin dans ce chapitre sur les caractéristiques particulières de double spécificité de substrats présentées par Pus1p et de multispécificité de substrats présentées par Cbf5p.

Chez la levure, les enzymes responsables de la formation, dans les ARNt cytoplasmiques, des résidus Ψ_{13} et Ψ_{32} , ainsi que la deuxième enzyme modifiant la position 35, ne sont pas connues. Cependant, trois activités Ψ -synthases, pour la modification des sites 13, 32 et 55, avaient pu être initialement séparées par chromatographie (Samuelsson et Olsson, 1990), indiquant donc la présence d'au moins deux autres Ψ -synthases spécifiques des ARNt chez *S. cerevisiae*. De même, les Ψ -synthases responsables de la formation des résidus Ψ peu fréquents en position 1 et 31 ne sont pas connues (Fig. 31, 32 et 35), ainsi que la plupart des ARNt: Ψ -synthases modifiant les ARNt mitochondriaux.

En ce qui concerne la formation des résidus Ψ dans les UsnRNA, peu de données sont disponibles. Comme nous l'avons déjà mentionné, nous avons montré que l'enzyme Pus1p modifie le snRNA U2. C'est la seule Ψ -synthase agissant sur les UsnRNA, caractérisée à ce jour. De nombreux travaux ont été effectués par l'équipe de Patton, sur les activités Ψ -synthases de cellules HeLa (pour revue, Patton, 1994a). En utilisant un extrait nucléaire ou un extrait S100 de cellules HeLa, cette équipe a montré l'existence d'activités Ψ -synthases distinctes agissant sur chaque UsnRNA. Ce point est présenté dans la revue qui suit.

2. MÉCANISME D'ACTION DES Ψ -SYNTHASES

(pour revue, Garcia et Goodenough-Lashua, 1998)

Bien que la comparaison des séquences protéiques des Ψ -synthases connues (Gustafsson *et al.*, 1996; Koonin, 1996) révèle peu d'homologies en acides aminés, les Ψ -synthases des familles RluA et RsuA contiennent 3 motifs conservés, dont 2 sont également présents dans la séquence protéique des Ψ -synthases de la famille TruB (motifs I et II) (Fig. 36).

Le motif II est également présent dans les déoxycytidine triphosphate déaminases (Dcd) et uridine triphosphatases (Dut) (Koonin, 1996) (Fig. 36). Pour les 3 familles d'enzymes, il existe une conservation totale d'un résidu acide aspartique et d'un résidu glycine dans ce domaine II (Dut, Dcd et Ψ -synthase) (Fig. 36). La structure tridimensionnelle de l'enzyme Dut d' *E. coli* a été déterminée par cristallographie (Cedergren-Zeppezauer *et al.*, 1992). La glycine et l'acide aspartique conservés sont trouvés dans une boucle entre 2 brins β , ce qui pourrait correspondre au contexte structural de ces 2 acides aminés dans les Ψ -

A	<u>MOTIF I</u>	<u>MOTIF II</u>	<u>MOTIF III</u>
consensus	UUUU.KP.GUU...	U.RLD..T\$GUUU	U.U...GR..QUR..U..UG..U
FAMILLE RluA			
RluA, <i>E. coli</i>	IMVVKPSGLLSVP	VHRLDMATSGVIVV	VVLKPTITGRSHQLRVHMLALGHPI
RluC, <i>E. coli</i>	ILVLNKPSGTAVHG	VHRLDRDTSGVLLV	VRCSPVTIGRTHQIRVHTQYAGHPI
RluD, <i>E. coli</i>	IIVVDKPTGIVCQP	AHRIDRNTSGIVIG	ITIRLISGRTHQIRACLNLFSNQL
structure secondaire	eeeeel1111eee1	eeeee11111eeee	eeeeel1111hhhhhhhhhhhh1111
FAMILLE RsuA			
RsuA, <i>E. coli</i>	YFMLNKPOGYVCST	AGRLDIDDTGLVLM	VRLTISEGRYHQVKRMFAAVGNHV
RluB, <i>B. subtilis</i>	YFLLYKPRGVISAA	IGRLDYDTSGLLLL	IQLTITHEGRNRQVRRMFEATGHEV
structure secondaire	eeeeel1111eee1	l1eee11111eeee	eeeeel1111hhhhhhhhhhhhhhhh
FAMILLE TruB			
TruB, <i>E. coli</i>	VINLDKPSNPSSHE	TGALDPLATGMLPI	
Cbf5p, <i>S. cerevisiae</i>	VLLLDKPOGMSSND	SGTLDPKVTGCLIV	
structure secondaire	eeee1111111hhh	l111111111eeee	
Dcd, <i>E. coli</i>		AHRIDPGWSGCIVL	
Dut, <i>E. coli</i>		VGLIDSDYQGQLMI	
structure secondaire		eeee1111111eeee	

B

Famille RluA	--HRLD--
Famille RsuA	--GRLD--
Famille TruA 1	--GRTD--
Famille TruA 2	--XRTD--
Famille TruB	--GLD--

Figure 36: A/ Les motifs conservés dans les trois familles de Ψ -synthases RluA, RsuA et TruB

Les séquences consensus sont en bleu. U indique un résidu hydrophobe et \$, soit un T, soit un S. Le motif II est le domaine putatif de fixation à l'uridine, conservé dans les triphosphate déaminases (Dcd) et les uridine triphosphatases (Dut). La région très conservée, comprenant l'acide aspartique, est indiquée en jaune. Elle est également indiquée dans la partie B de la figure. Les résidus acide aspartique et glycine conservés dans le motif II sont indiqués par des flèches. Les structures secondaires possibles sont indiquées: e représente un résidu dans un brin β , h dans une hélice α et l dans une boucle.

B/ Le motif conservé dans toutes les Ψ -synthases.

L'acide aspartique supposé catalytique est indiqué en rouge, les acides aminés en lettres capitales sont invariants ou quasi-invariants. X indique un résidu quelconque. Les enzymes de la famille TruA peuvent être regroupées en 2 sous-familles (1 et 2) en fonction de la séquence des 4 acides aminés du site catalytique ou en 8 sous-familles sur la base de la séquence entière de leur domaine II (Conrad *et al.*, 1999).

D'après Koonin, 1996; Huang *et al.*, 1998b et Conrad *et al.*, 1999.

synthases des familles RluA et RsuA (Gustafsson *et al.*, 1996; Koonin, 1996) (Fig. 36). Le dénominateur commun entre ces 3 familles de protéines est leur fixation à l'uridine. Il a donc été suggéré que le motif II des Ψ -synthases soit une partie très conservée du site de fixation de l'uridine (Koonin, 1996).

Le produit du gène *hist*, l'enzyme TruA, a été la Ψ -synthase la plus étudiée en terme de mécanisme enzymatique. L'observation que le traitement avec des réactifs des groupements thiols (comme l'acide iodoacétique) inactivait complètement la Ψ -synthase TruA d'*E. coli* et de *S. typhimurium*, avait conduit à l'hypothèse d'une implication d'un ou plusieurs résidus cystéines dans le mécanisme catalytique de pseudouridylation (Mullenbach *et al.*, 1976; Kammen *et al.*, 1988). Des études récentes par mutagenèse dirigée, portant sur les 3 cystéines de la Ψ -synthase TruA d'*E. coli* (Zhao et Horne, 1997), ont conduit à réfuter cette hypothèse. En effet, la conversion des 3 cystéines en alanines provoque seulement une perte de l'activité enzymatique de 10% par rapport à l'enzyme sauvage. De plus, il a été démontré que l'interaction d'un réactif des thiols avec des résidus cystéines, proches ou à l'intérieur du site actif, peut inhiber la catalyse, en dépit de l'absence d'un rôle catalytique direct de ces cystéines (Giorgianni *et al.*, 1995; Garcia et Chong, 1997). Il est donc possible que la perte d'activité catalytique de TruA, observée par Kammen *et al.* (1988) et Mullenbach *et al.* (1976) soit dûe à ce phénomène.

Bien que les Ψ -synthases de la famille TruA ne présentent pas de similitudes significatives avec les Ψ -synthases des 3 autres familles (Koonin, 1996), elles contiennent un motif conservé G/h R L/t D (Huang *et al.*, 1998b; Conrad *et al.*, 1999), également présent dans les 3 autres familles et faisant partie du motif II (Fig. 36). La mutation de l'acide aspartique conservé dans la Ψ -synthase TruA d'*E. coli* entraîne une inhibition totale de l'activité catalytique de cette enzyme (Huang *et al.*, 1998b). La nécessité, pour la catalyse, de la présence de l'acide aspartique conservé a également été démontrée dans le cas des Ψ -synthases RluA (Raychaudhuri *et al.*, 1998), RsuA (Conrad *et al.*, 1999), TruB (Ramamurthy *et al.*, 1999) et RluC d'*E. coli* (Ofengand et coll., résultats non publiés) et semble donc être une règle commune à toutes les Ψ -synthases.

Huang *et al.* (1998) ont proposé deux mécanismes catalytiques, conduisant à la formation d'un résidu Ψ à partir d'un résidu U:

⇒ **le premier mécanisme catalytique** implique l'attaque nucléophile du C₆ de l'uridine de l'acide aspartique conservé de l'enzyme (Fig. 37A). Une attaque nucléophile de ce type a déjà été proposée pour une autre enzyme de modification: RUMT (ARNt m⁵U₅₄ méthyltransférase), mais cette enzyme utilise une cystéine comme acide aminé catalytique (Santi et Hardy, 1987; pour revue, Garcia et Goodenough-Lashua, 1998).

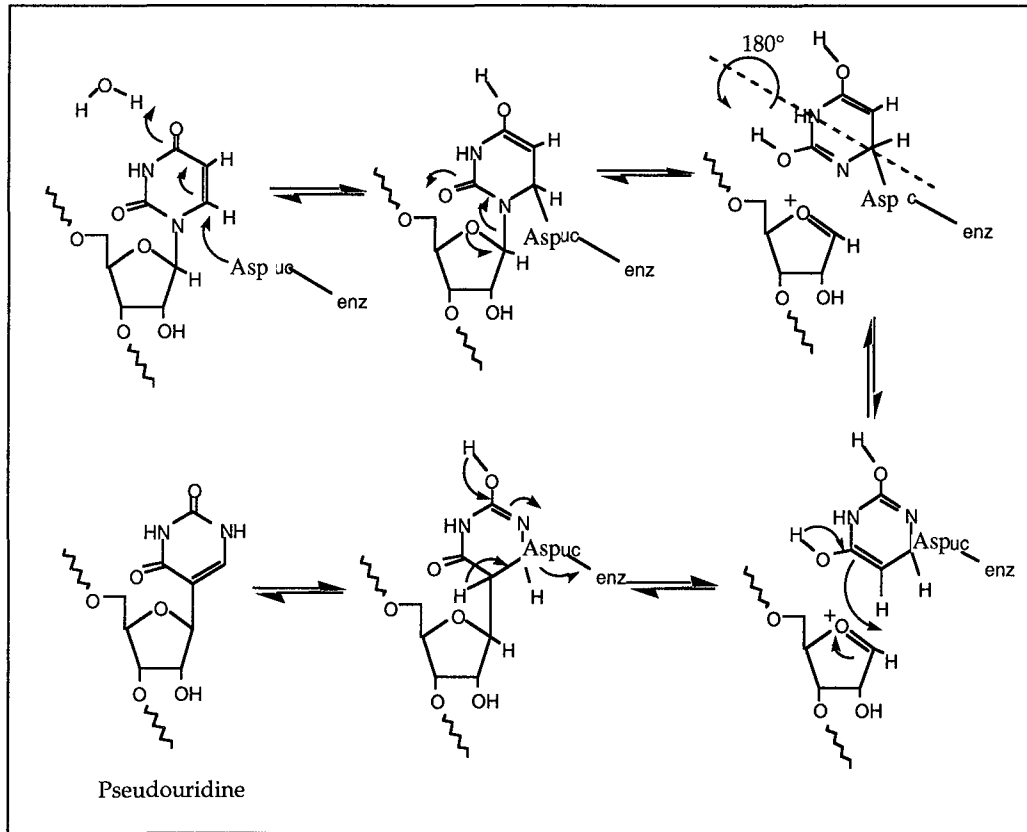
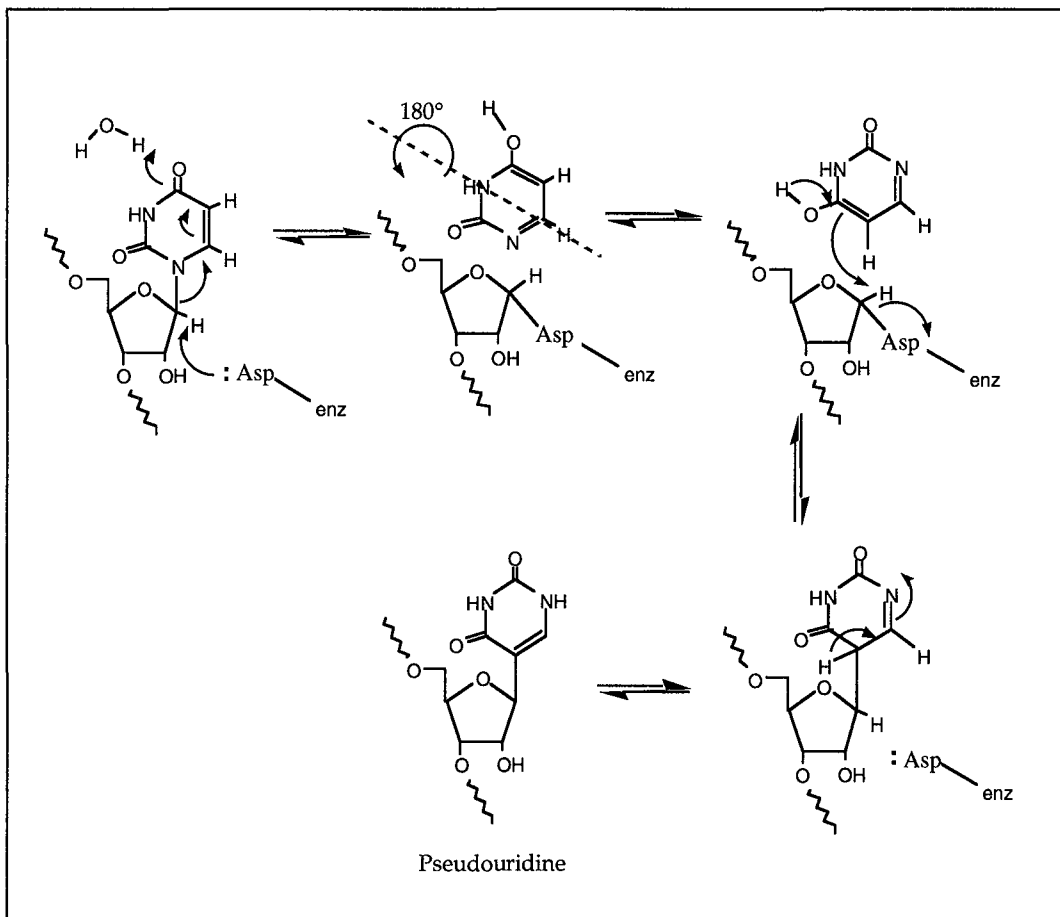
A**B**

Figure 37: Mécanismes enzymatiques hypothétiques des Ψ -synthases, basés sur l'attaque nucléophile en C₆ de l'uridine (A) et en C_{1'} du ribose de l'uridine (B) (d'après Kammen *et al.*, 1988; Huang *et al.*, 1998b; et pour revue, Garcia et Goodenough-Lashua, 1998)

⇒ le deuxième mécanisme catalytique possible (Fig. 37B) implique, cette fois, l'attaque nucléophile du carbone C_{1'} du ribose de l'uridine par le carboxylate de l'acide aspartique catalytique. Cette réaction a été suggérée sur la base du mécanisme enzymatique proposé pour la première étape de la formation de la queunine à partir de la guanine par l'ARNt guanine transglycosylase (pour revue, Romier *et al.*, 1998).

Dés résultats récents ont montré que la Ψ -synthase Pus1p de *S. cerevisiae* était une zinc-métalloenzyme (Arluisson *et al.*, 1998). Cet atome de zinc jouerait un rôle conformationnel en maintenant l'association entre les domaines structuraux. Il serait nécessaire à l'intégrité globale de la protéine.

En conclusion, il semblerait donc que toutes les Ψ -synthases partagent un même processus enzymatique, pouvant impliquer un acide aspartique conservé. Cependant, les similitudes de séquence entre ces enzymes se réduit à de petits motifs (ou seulement au motif G/h R L/t D pour la famille TruA). Cette observation signifierait que la réaction d'isomérisation de l'uridine ne requiert la conservation que d'éléments de séquence de taille limitée. Une partie des divergences de séquence doit être liée à la nécessité de reconnaître des éléments structuraux distincts dans l'ARN, ce qui assure la spécificité de ces enzymes.

3. SPECIFICITE DE RECONNAISSANCE DES Ψ -SYNTHASES

La définition des bases moléculaires de la spécificité de reconnaissance des ARN par les Ψ -synthases, ainsi que celles des autres enzymes de modification de l'ARN, est à ce jour une grande préoccupation des équipes qui s'intéressent aux nucléotides modifiés. Elle a d'ailleurs conduit récemment à une découverte importante dans le domaine des ARN, par la mise en évidence de petits ARN nucléolaires guides de modification: des snoRNA servant de guides pour la formation des résidus Ψ et 2'-O-méthylés dans les ARNr des eucaryotes supérieurs. Nous pouvons donc distinguer, à ce jour, deux types de Ψ -synthases, celles nécessitant et celles ne nécessitant pas d'ARN guides pour la reconnaissance de leur substrat.

3.1. La spécificité de reconnaissance des ARN par les Ψ -synthases ne nécessitant pas de guides de modification.

3.1.1. Les différentes classes d'enzymes

La spécificité des Ψ -synthases va d'un site unique de modification dans une classe d'ARN donnée, ou dans un ARN donné, à plusieurs sites de modification dans plusieurs classes d'ARN (Tableau 10).

A l'heure actuelle, il est possible de regrouper les Ψ -synthases en trois classes:

⇒ la première classe d'enzymes regroupe les Ψ -synthases les plus spécifiques qui ne modifient qu'un résidu U dans un type d'ARN donné. C'est le cas de l'ARNr: Ψ -synthase d'*E. coli* RsuA (Wrzesinski *et al.*, 1995b), de l'ARNt: Ψ -synthase TruB d'*E. coli* (Nurse *et al.*, 1995) et de son homologue Pus4p chez la levure *S. cerevisiae* (Becker *et al.*, 1997a), ainsi que de la Ψ -synthase RluB de *B. subtilis* (Niu et Ofengand, 1999) (Tableau 10 et Fig. 35).

⇒ la deuxième classe d'enzymes regroupe les Ψ -synthases qui modifient plusieurs sites dans un même type d'ARN, sites de modification proches les uns des autres dans la séquence nucléotidique (5 à 6 nts de distance au maximum). Les enzymes connues, qui présentent simultanément cette multispécificité et cette régio-spécificité, sont l'ARNr: Ψ -synthase RluD d'*E. coli* (Raychaudhuri *et al.*, 1998; Huang *et al.*, 1998a) et les ARNt: Ψ -synthases TruA d'*E. coli* et Pus3p de *S. cerevisiae* (Kammen *et al.*, 1988; Lecointe *et al.*, 1998)(Tableau 10 et Fig. 35).

⇒ la troisième classe d'enzymes inclut les Ψ -synthases qui modifient plusieurs résidus U dispersés dans un même ARN. C'est le cas des ARNt: Ψ -synthase Pus1p de *S. cerevisiae* et de souris (Simos *et al.*, 1996; Motorin *et al.*, 1998; Chen et Patton, 1999), ainsi que de l'ARNr: Ψ -synthase RluC d'*E. coli* (Conrad *et al.*, 1998; Huang *et al.*, 1998a) (Tableau 10 et Fig. 35).

Deux des Ψ -synthases connues présentent, en plus, une caractéristique de double spécificité de substrats, c'est-à-dire qu'elles modifient deux types d'ARN. Cette double spécificité a aussi été démontrée, *in vitro*, pour une autre enzyme de modification des ARN: la m⁵U méthyltransférase (RUMT) d'*E. coli*, responsable de la formation des nucléotides modifiés m⁵U₅₄ dans les ARNt et m⁵U₇₈₈ dans l'ARNr 16S (Gu *et al.*, 1994). La Ψ -synthase RluA d'*E. coli* présente également cette caractéristique, car elle modifie, *in vitro* et *in vivo*, à la fois la position 32 dans les ARNt et la position 746 dans l'ARNr 23S (Wrzesinski *et al.*, 1995a;

Raychaudhuri *et al.*, 1999). Le dernier exemple est, comme nous le verrons, le cas de la Ψ -synthase Pus1p de *S. cerevisiae*.

D'autres enzymes de modification des ARN déjà caractérisées pourraient également avoir une double spécificité de substrats, ce type de propriété ayant rarement été testé par le passé. Pourquoi ces cas de double spécificité de substrats? Deux cas de figures peuvent être envisagés: soit le second ARN cible de l'enzyme est aussi un véritable substrat, soit il est reconnu par l'enzyme car il porte un motif structural proche de celui du vrai substrat.

3.1.2. Reconnaissance d'une séquence ou d'une structure?

Avec l'observation initiale que certains nucléotides modifiés sont communs à tous les types d'ARNt d'un même organisme, on pouvait penser que les enzymes responsables de la synthèse de ces modifications étaient sensibles, soit à la structure 3D globale de l'ARNt, soit à celle d'un élément de cette structure. L'étude des éléments reconnus par les ARNt et les ARNr: Ψ -synthases a été abordée en testant la formation de nucléotides modifiés *in vitro*, après incubation de l'enzyme purifiée avec des transcrits produits *in vitro*, correspondant à des ARN mutés ou tronqués.

a) Cas des ARNr: Ψ -synthases d'*E. coli*.

A ce jour, le mécanisme de reconnaissance des sites de modification des enzymes RluA, RluC et RluD n'a pas été étudié expérimentalement. L'enzyme RluA présente une double spécificité de substrats. Une recherche d'homologie entre les deux sites reconnus (Wrzesinski *et al.*, 1995a) a permis de dégager une séquence commune: UUNNAAA (pour revue, Ofengand et Fournier, 1998), où le résidu U en 5' est converti en résidu Ψ . RluA semble donc être une Ψ -synthase séquence-spécifique. RluD modifie 3 sites dans l'ARNr 23S, qui sont situés dans une structure tige-boucle. Cet élément structural pourrait être l'élément de reconnaissance de cette enzyme (Raychaudhuri *et al.*, 1998; Huang *et al.*, 1998a). Il est à noter également que chaque résidu U modifié est suivi par un résidu A en 3'. RluC modifie également 3 sites dans l'ARNr 23S, éloignés les uns des autres dans la séquence primaire (Conrad *et al.*, 1998; Huang *et al.*, 1998a). Cependant les résidus U en positions 2504 et 2580 sont proches l'un de l'autre dans la structure secondaire (positions 2869 et 2945 dans la Fig. 34C) et la position 955 (position 1125 dans la Fig. 34A) semble également proche des 2 autres sites modifiés dans la structure tertiaire. En effet, par des expériences de pontage, Dokudovskaya *et al.* (1996) ont montré la proximité dans l'espace des résidus A₉₆₀ et C₂₄₇₅. RluC pourrait donc reconnaître un élément structural dans l'ARNr 23S et modifier dans cette région des résidus U qui sont toujours suivis par un résidu G en 3'. En effet, RluC, comme RluD, semble sensible à la nature du nucléotide en 3' du site à modifier.

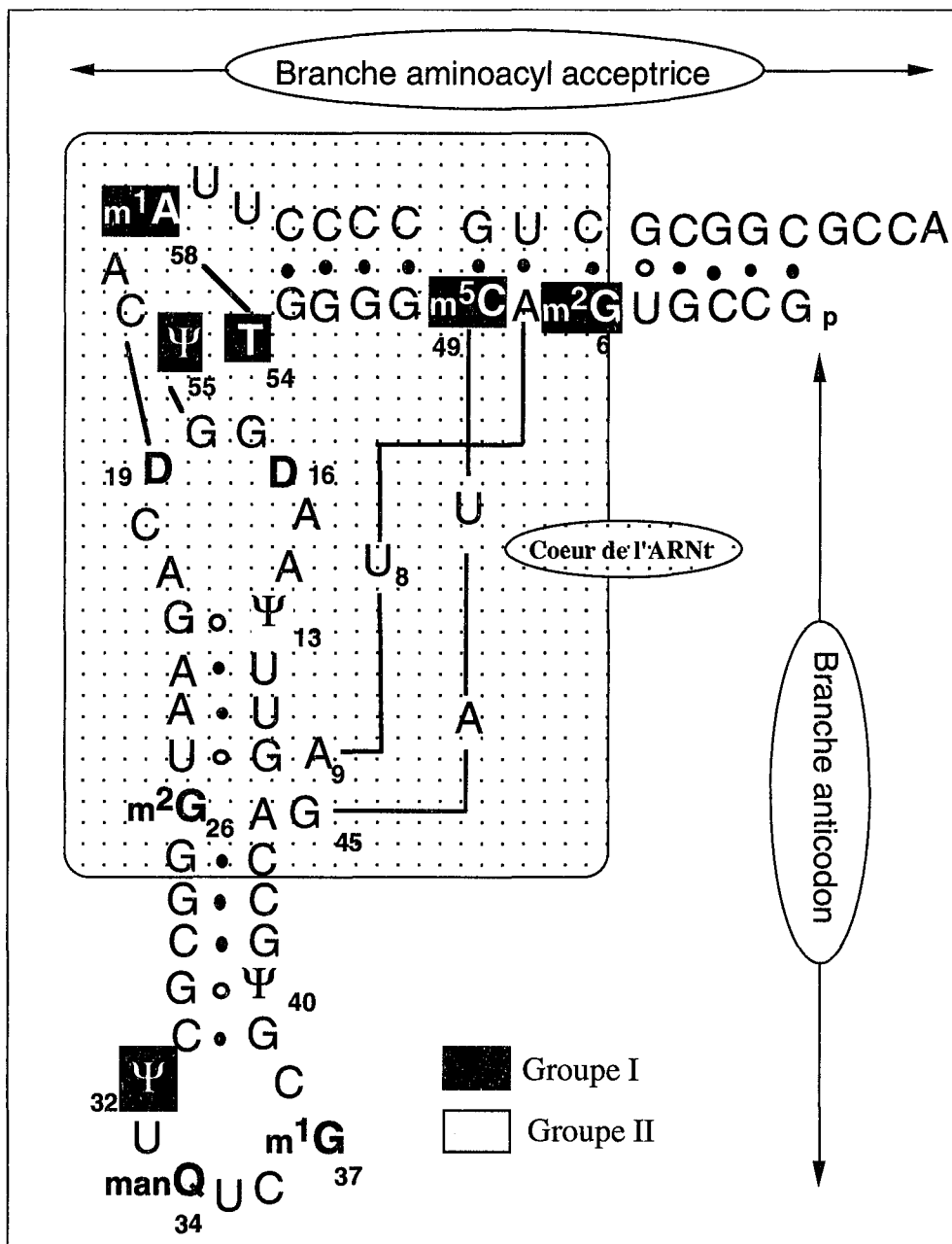


Figure 38: Classification des enzymes de modification des ARNt en 2 groupes, suivant leur sensibilité vis-à-vis de perturbations de la structure tertiaire de l'ARNt.

Les enzymes du groupe I sont insensibles aux perturbations portées à l'architecture globale de l'ARNt, tandis que les enzymes du groupe II dépendent de l'architecture tridimensionnelle de cet ARN. La figure représente l'ARNt^{Asp} de *S. cerevisiae* et les modifications sont celles qui sont formées dans cet ARNt (ou ses variants) lorsqu'ils sont microinjectés dans le cytoplasme d'oocytes de *X. leavis* (d'après Grosjean *et al.*, 1996).

L'enzyme RsuA catalyse la formation du seul résidu Ψ dans l'ARNr 16S d'*E. coli*. On peut se demander ce qui distingue le résidu U_{516} qui va être modifié des 313 autres résidus U de l'ARN 16S. Des transcrits, produits *in vitro*, correspondant à l'ARNr 16S entier (nts 1 à 1542) ou tronqué (nts 1 à 526 ou 1 à 678) ne sont pas des substrats pour RsuA. Une particule ribonucléoprotéique assemblée *in vitro*, en utilisant un transcrit comprenant les nucléotides 1 à 678 et les protéines ribosomiques, est modifiée. Une particule 30S entière assemblée *in vitro* est un substrat médiocre. Enfin, une particule assemblée à partir d'un transcrit *in vitro* contenant les nucléotides 1 à 526 n'est pas un substrat de l'enzyme (Wrzesinski *et al.*, 1995b). Ces résultats semblent montrer que l'enzyme requiert une structure définie de l'ARNr qui n'est observée que dans des conditions très particulières. Il a été proposé que la particule ribonucléoprotéique, qui est un substrat de l'enzyme *in vitro*, mime un stade intermédiaire de l'assemblage du ribosome, stade auquel le résidu Ψ_{516} serait formé (Wrzesinski *et al.*, 1995b).

b) Cas des ARNt et UsnRNA: Ψ -synthases procaryotes et eucaryotes. Comparaison avec les autres enzymes de modification des ARNt.

Edqvist *et al.* (1993) et Grosjean *et al.* (1996) ont utilisé la même approche expérimentale afin d'étudier les éléments structuraux de la molécule d'ARNt nécessaires à la formation des différents nucléotides modifiés. Pour cela, des variants d'ARNt de levure ont été microinjectés dans des oocytes de *X. laevis* et les modifications post-transcriptionnelles formées ont été analysées. Les résultats obtenus ont amené Grosjean *et al.* (1996) à proposer une séparation des enzymes de modification des ARNt en deux groupes, en fonction de leur sensibilité aux perturbations de structure de la molécule d'ARNt. L'existence de deux groupes distincts a été, par la suite, confirmée par l'étude spécifique de certaines de ces enzymes. On trouve des Ψ -synthases dans chacun de ces groupes d'enzymes de modification (Fig. 38).

\Rightarrow Les enzymes du groupe I reconnaissent un élément structural local de l'ARNt.

C'est le cas pour la formation du résidu Ψ_{55} . En effet, un mini-substrat mimant le bras T de l'ARNt est un bon substrat de l'enzyme Pus4p de levure (Becker *et al.*, 1997b; 1997a) et de l'enzyme TruB d'*E. coli* (Gu *et al.*, 1998). Pour que l'activité soit observée, aucun des nucléotides conservés de la boucle T ne peut être muté et il semblerait que la formation du résidu Ψ_{55} dépende de la conservation de la structure particulière de cette boucle qui a été établie par cristallographie (Kim *et al.*, 1974; Suddath *et al.*, 1974; Giegé *et al.*, 1977). Notamment la conservation de la paire de bases reverse Hoogsteen, entre la position 54 et la position 58, est nécessaire à la formation du résidu Ψ_{55} .

La reconnaissance d'éléments structuraux localisés dans le bras T est également une nécessité

pour qu'il y ait formation d'autres nucléotides modifiés, comme le résidu m⁵U₅₄ chez la levure (Becker *et al.*, 1997b) et chez les procaryotes (Gu et Santi, 1991; Gu *et al.*, 1994; Kealey et Santi, 1995), ainsi que les résidus m⁵C₄₉ et m¹A₅₈ chez *X. laevis* (Edqvist *et al.*, 1993; Grosjean *et al.*, 1996).

De la même façon, la méthylation sur le ribose du nucléotide Gm₃₄ dans les ARNt de *X. laevis* (Droogmans *et al.*, 1986), la formation du résidu queunine à partir de la guanine en position 34 chez *E. coli* (Curnow *et al.*, 1993; Mueller et Slany, 1995) et la formation du résidu Ψ₃₂ chez *X. laevis* (Grosjean *et al.*, 1996), ne dépendent que de la présence de la structure tige-boucle anticodon. La méthylation sur le ribose de la position 18 chez *Thermus thermophilus* ne nécessite qu'un mini-substrat mimant le bras D (Matsumoto *et al.*, 1990), et la formation du nucléotide modifié m²G₆ ne dépend que du bras accepteur des ARNt de levure (Grosjean *et al.*, 1996). Cependant, les enzymes procaryotes et eucaryotes nécessaires pour la formation du résidu queunine en position 34 ont des spécificités de reconnaissance différentes. En effet, chez les eucaryotes, la structure tridimensionnelle de l'ARNt est nécessaire à l'ARNt guanine transglycosylase (Edqvist *et al.*, 1993; Grosjean *et al.*, 1996; pour revue, Romier *et al.*, 1998).

Il est beaucoup moins évident de définir un motif consensus de reconnaissance de l'ARNt par les Ψ-synthases Pus1p de levure et de souris. L'analyse de l'enzyme Pus1p de *S. cerevisiae* faisant partie de mon sujet de thèse, ce cas sera discuté plus en détail dans la partie Résultats/Discussion de ce manuscrit. Cependant, on peut noter que la formation de certains résidus Ψ, catalysée par Pus1p, est intron-dépendante, c'est-à-dire qu'elle a lieu dans le précurseur non épissé des ARNt. En effet, il a été démontré *in vivo* et *in vitro* pour la levure *S. cerevisiae* (Johnson et Abelson, 1983, Szweykowska-Kulinska *et al.*, 1994; Motorin *et al.*, 1998), et *in vitro* chez les plantes (Szweykowska-Kulinska et Beier, 1992; Szweykowska-Kulinska *et al.*, 1995) et chez la souris (Chen et Patton, 1999), que la formation des résidus Ψ_{34, 35, 36} nécessite la présence de l'intron. C'est également le cas pour les résidus m⁵C₄₀ et m⁵C₃₄ (Jiang *et al.*, 1997; Strobel et Abelson, 1986; Motorin et Grosjean, 1999). Motorin *et al.* (1998) ont montré qu'*in vitro*, un mini-substrat correspondant à la boucle anticodon et à l'intron, suffisait à la formation de résidus Ψ en position 34 et 36 par la Ψ-synthase Pus1p de levure produite chez *E. coli*. La même conclusion avait été émise par Szweykowska *et al.* (1994) en utilisant un extrait S100 de levure.

⇒ Les enzymes du groupe II sont sensibles à la structure tridimensionnelle globale de l'ARNt.

Les Ψ-synthases TruA et Pus3p, catalysant les Ψ en position 38, 39 et 40 chez *E. coli* et chez *S. cerevisiae*, respectivement, peuvent être classées dans ce groupe II (Edqvist *et al.*, 1993; Grosjean *et al.*, 1996; Lecointe *et al.*, 1998). Cependant, Pus3p semble présenter des

différences subtiles de reconnaissance de l'ARNt par rapport à TruA. En effet, l'enzyme de levure modifie les positions 38 et 39, mais pas 40 (contrairement à TruA), *in vitro* comme *in vivo* (Lecoïnte *et al.*, 1998).

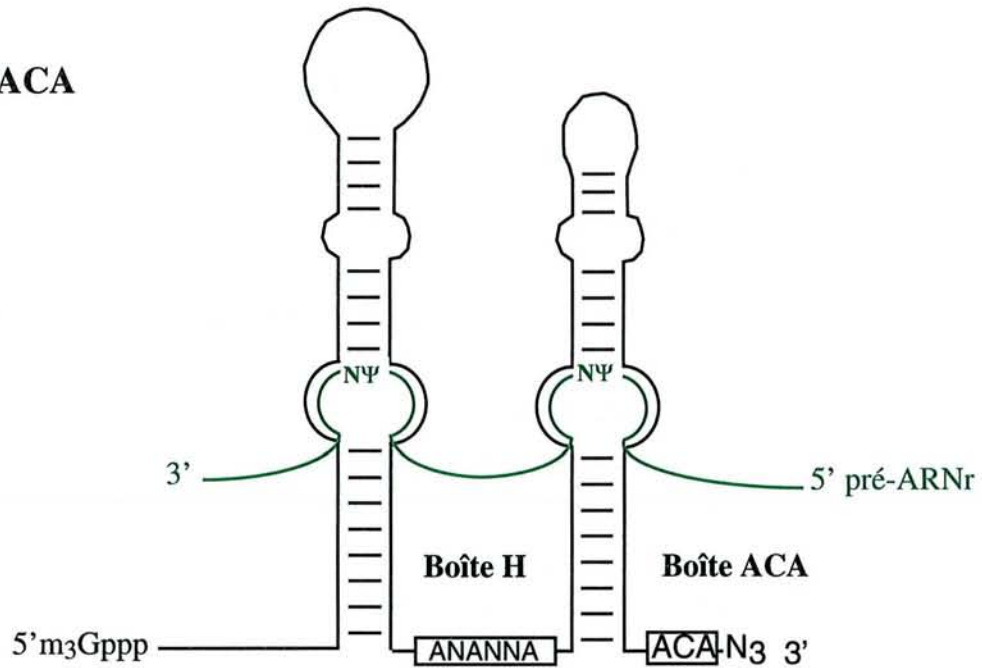
Un certain nombre d'autres enzymes de modification des ARN peuvent être classées dans ce groupe. On peut citer, par exemple, les cas des enzymes les plus étudiées formant les nucléotides modifiés m^2G_{26} (Edqvist *et al.*, 1992; 1994), m^1G_{37} (Holmes *et al.*, 1992; Redlak *et al.*, 1997; Brulé *et al.*, 1998), I_{34} (Auxilien *et al.*, 1996), i^6A_{37} (Motorin *et al.*, 1997) et t^6A_{34} (Morin *et al.*, 1998). Dans le cas de la formation des résidus I_{34} et i^6A_{37} , l'hypothèse a été émise que les enzymes reconnaîtraient l'ARNt en globalité et que leur fixation "tordrait" la structure tige-boucle de l'anticodon pour optimiser le contact avec le site actif de l'enzyme. En effet, il a été démontré qu'une flexibilité conformationnelle du bras anticodon est nécessaire à la synthèse de ces nucléotides modifiés, en plus de la séquence conservée $A_{36}A_{37}A_{38}$ dans le cas de la formation du résidu i^6A_{37} (Auxilien *et al.*, 1996; Motorin *et al.*, 1997).

En conclusion générale, il est à noter qu'il n'existe pas de cas type pour la reconnaissance des ARN cibles par les enzymes de modification. Cette reconnaissance peut dépendre de séquences primaires, de structures secondaires et/ou tertiaires, ainsi que de la présence des protéines associées à ces ARN. Nous verrons par la suite que ces modes de reconnaissance sont souvent corrélés à l'étape de maturation des ARN où la modification a lieu.

3.2. Les snoRNA guides nucléolaires des Ψ -synthèses modifiant les ARNr eucaryotes

Chez les eucaryotes, 3 des 4 ARNr matures cytoplasmiques (5,8 S, 18 S et 26/28 S) sont produits dans le nucléole à partir d'un transcrit unique (le pré-ARNr 35S chez la levure et 45S chez les mammifères). Ce précurseur est mûré par un processus complexe impliquant une cascade de clivages endonucléolytiques et de digestions par des exonucléases (pour revues, Morrissey et Tollervey, 1995; Venema et Tollervey, 1995; Tollervey, 1996). La formation des résidus Ψ et 2'-O-méthylés a lieu dans le pré-ARNr (Maden et Salim, 1974; Brand *et al.*, 1979). La synthèse de ces ARNr cytoplasmiques dans les cellules eucaryotes implique l'action d'une large population de snoRNA. Un petit nombre de ces ARN sont requis pour les étapes de clivage de l'ARNr précurseur, les autres sont maintenant connus pour définir les sites de formation des résidus Ψ et 2'-O-méthylés. Les données actuelles sont résumées dans plusieurs revues récentes (Bachellerie et Cavallé, 1997; 1998; Lafontaine et Tollervey, 1998a; Ofengand et Fournier, 1998). Certains snoRNA interviennent à la fois dans

A. Boîtes H/ACA



B. Boîtes C/D

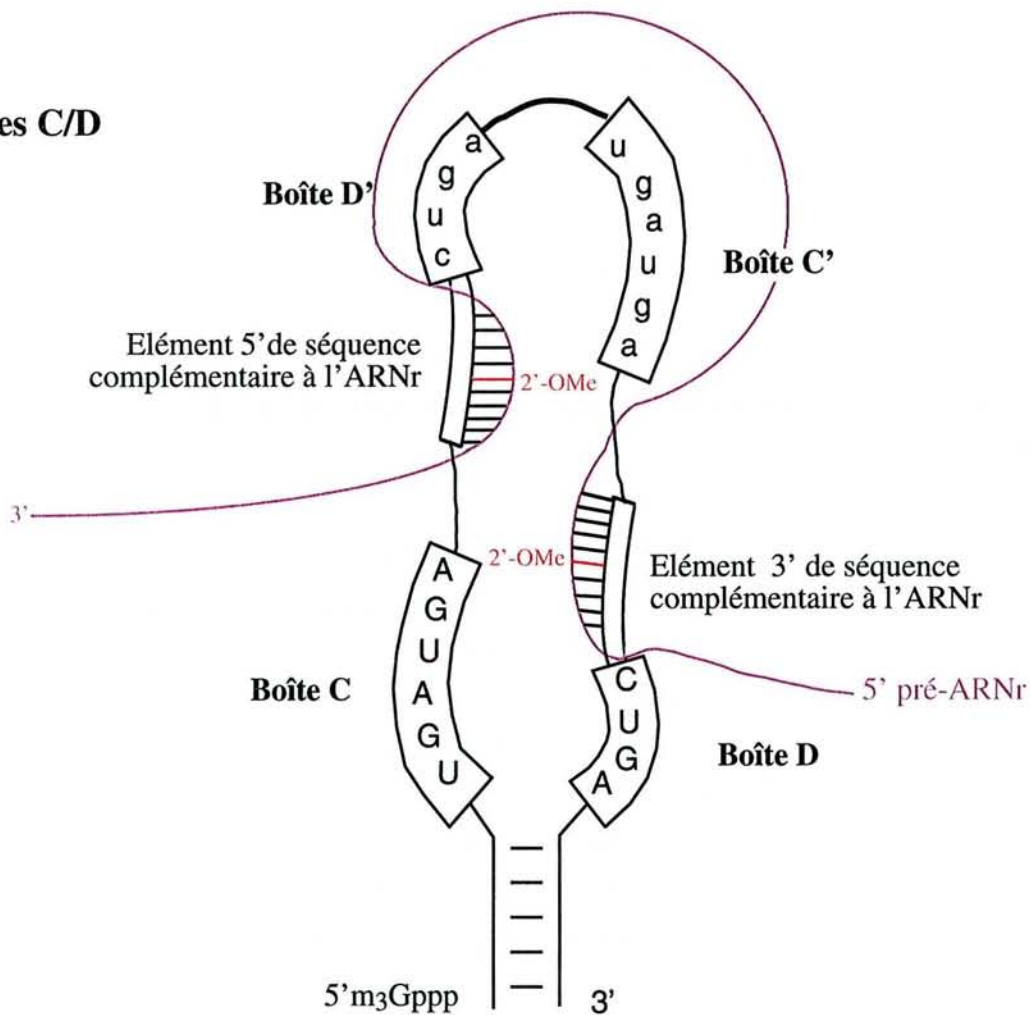


Figure 39: Structures schématiques des complexes formés entre le pré-ARNr et les snoRNA à boîtes H/ACA (A) et C/D (B).

Les séquences consensus des différentes boîtes conservées sont indiquées dans les cadres.

D'après la revue, Lafontaine et Tollervey, 1998 .

les étapes de clivage et de modification du pré-ARNr. C'est le cas pour les snoRNA de levure U14 et snR10, le premier étant un guide de méthylation, le second un guide de pseudouridylation (Li *et al.*, 1990; Liang et Fournier, 1995; Kiss-Laszlo *et al.*, 1996; Ni *et al.*, 1997).

On ne sait pas encore si les snoRNA guides sont impliqués dans la formation de tous les résidus Ψ et 2'-O-méthylés des ARNr des ribosomes cytoplasmiques eucaryotes. On connaît maintenant les snoRNA responsables de la formation de 51 méthylations sur le ribose sur les 55 présentes dans les ARNr de *S. cerevisiae* (Lowe et Eddy, 1999). En ce qui concerne les résidus Ψ , des snoRNA guides ont été identifiés pour 26 des 44 sites présents dans les ARNr de *S. cerevisiae* (Samarsky et Fournier, 1999).

Les snoRNA guides peuvent être subdivisés en deux classes, sur la base de leur conservation en structure primaire et secondaire: les snoRNA à boîtes C/D et les snoRNA à boîtes H/ACA (Fig. 39). Tous les snoRNA peuvent être inclus dans une des deux familles, même ceux ne participant pas à la formation des modifications post-transcriptionnelles, mais intervenant dans les clivages tels que les snoRNA U3 et U8 (snoRNA à boîtes C/D) et snR30 (snoRNA à boîtes H/ACA) (pour revue, Tollervey et Kiss, 1997). La seule exception mentionnée est l'ARN de la RNase MRP, une particule ribonucléoprotéique dont l'ARN porte l'activité catalytique (pour revues, Morrissey et Tollervey, 1995; Tollervey et Kiss, 1997). Il faut également noter que sur la base de leur biosynthèse, les snoRNA se répartissent aussi en deux classes qui ne recouvrent pas les deux précédentes: ceux produits à partir d'un gène classique et ceux produits par maturation d'introns contenus dans des ARN pré-messagers (pour revue, Maxwell et Fournier, 1995; Tollervey et Kiss, 1997).

3.2.1. Les snoRNP à boîtes H/ACA

a/ Structure et mode d'action des snoRNA H/ACA

Ce groupe de snoRNA partage une structure commune qui consiste en deux structures tige-boucles, séparées par une région simple brin (Fig. 39A) (Balakin *et al.*, 1996; Ganot *et al.*, 1997b). Cette partie simple brin, ainsi que l'extrémité 3' des snoRNA, contiennent des séquences conservées, qui sont respectivement la boîte H (consensus ANANNA) et la boîte ACA (ou dans certain cas AAA ou AUA) (Balakin *et al.*, 1996; Ganot *et al.*, 1997b).

Les boîtes H et ACA sont absolument requises pour la stabilité des snoRNA, que ce soit chez les mammifères ou chez la levure. Il en est de même pour les deux hélices à la base des deux structures tige-boucles (Balakin *et al.*, 1996; Ganot *et al.*, 1997b; Bortolin *et al.*, 1999). L'altération de la distance entre la boîte ACA et l'hélice terminale en 3' des snoRNA résulte en une diminution de la concentration cellulaire des snoRNA (Balakin *et al.*, 1996). De

même, la position de la boîte H par rapport à l'hélice 5' intervient dans la définition de l'extrémité 5' des snoRNA H/ACA (Bortolin *et al.*, 1999). Ces observations suggèrent que la boîte H avec l'hélice 5' et la boîte ACA avec l'hélice 3', constituent les signaux de reconnaissance des protéines des snoRNP. En s'associant, ces protéines doivent protéger les snoRNA contre l'action des endo et exonucléases et, de cette façon, elles doivent participer au contrôle de la formation des extrémités des snoRNA H/ACA (Bortolin *et al.*, 1999).

Ces snoRNA H/ACA sont nécessaires à la formation des résidus Ψ dans les ARNr eucaryotes (Ganot *et al.*, 1997a; Ni *et al.*, 1997). A l'intérieur de chaque hélice, une boucle interne appelée « poche de pseudouridylation » sélectionne la région de l'ARNr à modifier en formant, avec l'ARNr, deux petites hélices (3 à 10 pb) qui sont séparées par 2 nt non appariés de l'ARNr (Ganot *et al.*, 1997a; Ni *et al.*, 1997). Le premier nucléotide non apparié, en 5' de l'ARNr, est un résidu U qui est alors converti en résidu Ψ . Une poche de pseudouridylation peut être trouvée dans l'hélice 3' ou dans l'hélice 5' et beaucoup de snoRNA possèdent deux poches de pseudouridylation, une dans chaque hélice. Ils peuvent donc diriger la pseudouridylation de deux sites distincts de l'ARNr (Ganot *et al.*, 1997a). C'est le cas représenté dans la figure 39A.

Les boîtes H et ACA sont localisées, en moyenne, 14 nts en aval des centres catalytiques contenus dans les hélices 5' et 3', respectivement. Une variation de longueur de la région entre la boîte ACA et le centre catalytique 3' entraîne une mauvaise pseudouridylation dans la poche en 3' du snoRNA (Ni *et al.*, 1997). Ceci montre que la sélection du résidu à modifier dépend à la fois d'une bonne complémentarité entre snoRNA et pré-ARNr (Ganot *et al.*, 1997a; Ni *et al.*, 1997), mais également de la position de la boîte ACA par rapport au centre catalytique 3' du snoRNA. Une hypothèse est que la boîte H possède la même fonction que la boîte ACA, mais pour la sélection du résidu à modifier dans la poche de pseudouridylation 5' du snoRNA (Bortolin *et al.*, 1999). De plus, les domaines 3' et 5' semblent agir en coopération car la délétion de l'un des domaines empêche la pseudouridylation du pré-ARNr dans l'autre domaine (Bortolin *et al.*, 1999).

b/ Les protéines des snoRNP H/ACA

Les snoRNP H/ACA contiennent trois protéines communes: Gar1p, Nhp2p et Cbf5p (Balakin *et al.*, 1996; Ganot *et al.*, 1997b; Lafontaine *et al.*, 1998a; Watkins *et al.*, 1998a). Une étude en microscopie électronique des snoRNP snR30 et snR42 de levure purifiées, a révélé deux structures symétriques pour chaque snoRNP avec une masse moléculaire qui correspondrait à la présence d'un seul snoRNA mais à deux copies de chacune de ces protéines (Watkins *et al.*, 1998a). Une quatrième protéine, Nop10p, semble également être associée aux snoRNP H/ACA (Henras *et al.*, 1998).

Cbf5p porte l'activité catalytique Ψ -synthase (Lafontaine *et al.*, 1998a). Des homologues à cette enzyme ont été identifiés chez le rat: NAP57 (Meier et Blobel, 1994), la drosophile: Nop60p (Phillips *et al.*, 1998), *Kluyveromyces lactis* (Winkler *et al.*, 1998) et l'homme: DKC1 (Heiss *et al.*, 1998). Cbf5p est essentielle à la croissance cellulaire (Jiang *et al.*, 1993; Lafontaine *et al.*, 1998a). Nop60p, exprimée dans une souche de levure où le gène codant pour Cbf5p a été délété, peut restaurer la croissance, indiquant ainsi que cette enzyme est conservée au cours de l'évolution (Phillips *et al.*, 1998). Des mutations dans la protéine humaine DKC1 sont responsables d'une maladie génétique rare touchant la moelle osseuse, la dyskératose congénitale (Heiss *et al.*, 1998).

Nhp2p contient un domaine putatif de fixation à l'ARN trouvé également dans certaines protéines ribosomiques (Kolodrubetz et Burgum, 1991; Koonin *et al.*, 1994; Henras *et al.*, 1998).

Gar1p n'est pas nécessaire à la stabilité des snoRNA H/ACA, mais de nombreuses études ont montré que cette protéine leur était associée (Girard *et al.*, 1992; Balakin *et al.*, 1996; Ganot *et al.*, 1997b; Watkins *et al.*, 1998a) et qu'elle est nécessaire à leur activité de pseudouridylation (Bousquet-Antonelli *et al.*, 1997), ainsi qu'à la maturation du pré-ARNr (Girard *et al.*, 1992). A l'heure actuelle, on ne sait pas si Gar1p interagit directement ou indirectement avec les snoRNA H/ACA. Cette protéine contient deux domaines renfermant des répétitions de séquences glycine/arginine: domaines GAR, localisés à chacune de ses extrémités. Ces domaines sont retrouvés dans de nombreuses protéines nucléolaires et sont supposés permettre la fixation à l'ARN. Cependant, il a été reporté que Gar1p se fixerait aux snoRNA snR10 et snR30 par son domaine central conservé (Bagni et Lapeyre, 1998). Gar1p pourrait également aider à l'ouverture de la structure du pré-ARNr par l'intermédiaire d'une autre protéine: Rok1p, une hélicase ATP-dépendante putative, qui lui est génétiquement associée (Venema *et al.*, 1997).

Nop10p est une protéine essentielle à la croissance et nécessaire à l'activité de pseudouridylation des snoRNA H/ACA (Henras *et al.*, 1998).

Les 3 protéines Cbf5p, Nhp2p et Nop10p sont nécessaires à la stabilité des snoRNA H/ACA et de la protéine Gar1p, et sont également nécessaires à la maturation du pré-ARNr (Kolodrubetz et Burgum, 1991; Lafontaine *et al.*, 1998a; Watkins *et al.*, 1998a; Henras *et al.*, 1998). La délétion de ces 3 protéines entraîne donc une instabilité totale des snoRNA H/ACA indispensables à la maturation des pré-ARNr, comme snR30. Cette observation expliquerait le fait que ces 3 protéines soient essentielles à la viabilité cellulaire. Cette dépendance vis-à-vis

de Cbf5p pour la maturation des pré-ARNr rend très difficile l'évaluation d'un éventuel défaut de croissance dû à l'absence de formation des résidus Ψ dans le pré-ARNr.

3.2.2. Les snoRNP à boîtes C/D

a/ Structure et mode d'action des snoRNA C/D

Ces snoRNA se caractérisent par la présence de deux séquences conservées: les boîtes C et D dont les séquences consensus sont respectivement UGAUGA et CUGA (Fig. 39B). Ces boîtes sont nécessaires à la stabilité des snoRNA qui les portent (pour revues, Tollervey et Kiss, 1997; Bachellerie et Cavaillé, 1998) et à leur localisation nucléolaire (Lange *et al.*, 1998; Narayanan *et al.*, 1999). Les snoRNA à boîtes C et D portent parfois un second couple de boîtes C et D imparfaites, nommées C' et D' (pour revues, Tollervey et Kiss, 1997; Bachellerie et Cavaillé, 1998). Les snoRNA C/D contiennent une à deux régions complémentaires au pré-ARNr, la première est localisée immédiatement en 5' de la boîte D et est toujours observée dans les snoRNAs C/D, la deuxième peut être présente en 5' de la boîte D' (Kiss-Laszlo *et al.*, 1996; Nicoloso *et al.*, 1996; Cavaillé *et al.*, 1996). L'appariement du pré-ARNr au niveau de ces séquences complémentaires est le signal pour que le résidu U, en cinquième position de l'appariement, soit méthylé. La distance entre les boîtes D ou D' et le site à méthyler dans le pré-ARNr est invariant. En effet, l'ajout ou la délétion de nucléotides après la boîte D dans le snoRNA modifie la position du site méthylé dans le pré-ARNr (Kiss-Laszlo *et al.*, 1996; Cavaillé *et al.*, 1996). Les snoRNA C/D sont indispensables à la méthylation du pré-ARNr (Cavaillé *et al.*, 1996; Kiss-Laszlo *et al.*, 1996, Tycowski *et al.*, 1996) et leur fonction est sensible à des perturbations de l'appariement snoRNA/pré-ARNr (Cavaillé *et al.*, 1996; Cavaillé et Bachellerie, 1998).

La délétion des boîtes C' et D' ne modifie pas la méthylation médiée par le couple de boîtes C et D en 3' du snoRNA, mais inhibe complètement la méthylation médiée par ce couple C'/D' (Kiss-Laszlo *et al.*, 1998, Watkins *et al.*, 1998b). Les boîtes C et D, et C' et D', seraient des sites de fixation de protéines (Kiss-Laszlo *et al.*, 1998).

b/ Les protéines des snoRNP C/D

La protéine **Nop1p** chez la levure, ou fibrillarine chez l'homme, est associée aux snoRNA C/D par l'intermédiaire des boîtes C et D (Ochs *et al.*, 1985; Schimmang *et al.*, 1989, pour revues, Maxwell et Fournier, 1995; Tollervey et Kiss, 1997). Son homologue chez *X. laevis* requiert également des boîtes C et D intactes pour son association aux snoRNA (Caffarelli *et al.*, 1998). La délétion du gène codant Nop1p affecte la maturation du pré-ARNr (Tollervey *et al.*, 1991).

Nop56p et **Nop58p** sont des protéines essentielles à la croissance cellulaire, liées génétiquement et pouvant être co-purifiées avec Nop1p (Gautier *et al.*, 1997; Wu *et al.*, 1998). La délétion des gènes codant ces protéines entraîne aussi un défaut de maturation des pré-ARNr (Gautier *et al.*, 1997; Wu *et al.*, 1998). Nop58p (également appelée Nop5p) est associée aux snoRNA C/D et est nécessaire à leur stabilité (Wu *et al.*, 1998; Lafontaine et Tollervey, 1999).

La 2'-O-méthylase associée aux snoRNA C/D n'a pas encore été identifiée.

3.2.3. Y a-t-il des ARN guides pour la modification d'autres ARN que les ARNr eucaryotes?

Deux snoRNA à boîtes C/D sont requis pour la méthylation *in vivo* de certains nucléotides du snRNA U6 de vertébrés (Tycowski *et al.*, 1998): un pour méthyler la position 47 et l'autre pour méthyler la position 77. Ce deuxième snoRNA peut également s'apparier à l'ARN 28S et méthyler la position 2970 de cet ARN. Un troisième snoRNA C/D pourrait être impliqué dans la 2'-O-méthylation du résidu A₅₃ du snRNA U6 de vertébrés. Cette découverte était surprenante du fait que les UsnRNA n'étaient pas connus pour pouvoir transiter dans le nucléole. Mais il a été montré très récemment que le snRNA U6, qui ne va pas dans le cytoplasme, pourrait aller dans le nucléole où il serait 2'-O-méthylé et pseudouridylé (Ganot *et al.*, 1999).

Il est à noter que des ARN guides sont également utilisés dans le phénomène d'"editing" (pour revues, Hajduk et Sabatini, 1998; Rueter et Emeson, 1998). On peut citer comme exemple le cas de l'ARNm de la protéine GluR, une sous-unité d'un récepteur au glutamate. La conversion d'adénosine en inosine dans cet ARNm nécessite la présence d'un duplex d'ARN impliquant une séquence d'ARN intronique agissant en cis (Cattaneo, 1994). De même, l'insertion ou la délétion de nucléotides spécifiques dans les ARN mitochondriaux du trypanosome implique l'appariement au site à modifier d'un petit ARN guide agissant en trans (Benne, 1992; Seiwert et Stuart, 1994).

3.2.4. Certains snoRNA portent eux-mêmes des modifications post-transcriptionnelles

Certains snoRNA portent des nucléotides modifiés. Malheureusement, les données connues sont rares car elles ne sont dûes qu'à l'utilisation, il y a une vingtaine d'années, de techniques de séquençages chimiques directs des ARN. Les modifications post-transcriptionnelles ont été analysées pour quatre snoRNA. Les snoRNA U3 et U8 humains portent des résidus Ψ et 2'-O-méthylés (Reddy *et al.*, 1985, Kato et Harada, 1984). Wise *et al.*

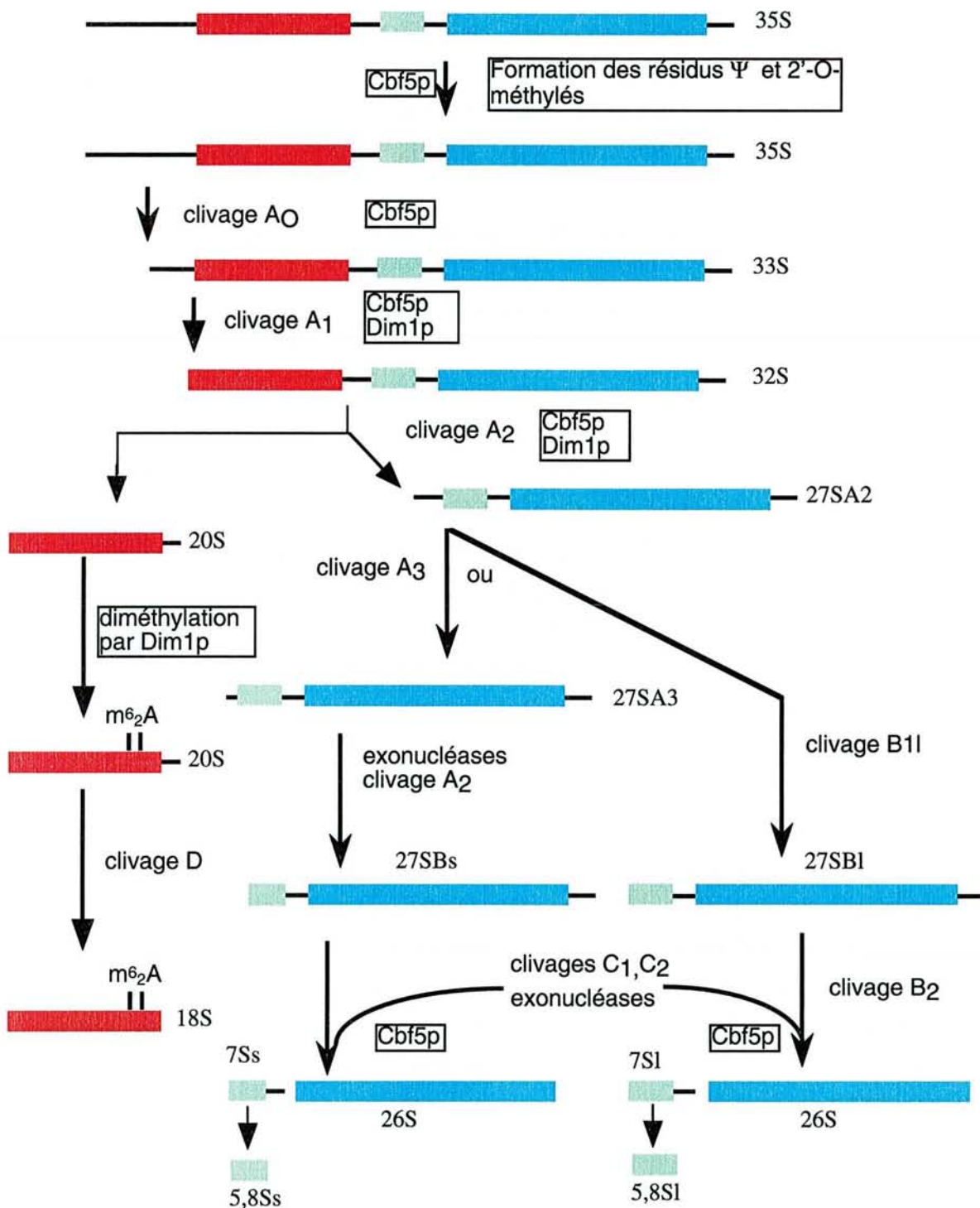


Figure 40: Les étapes de maturation du précurseur 35S des ARN ribosomiques de la levure *Saccharomyces cerevisiae*

D'après Lafontaine et Tollervey, 1998b.

Les étapes, pour lesquelles les protéines Cbf5p et/ou Dim1p sont requises, soit pour la maturation, soit pour la modification du pré-ARNr, sont indiquées.

(Wise *et al.*, 1983) ont aussi détecté des résidus Ψ et 2'-O-méthylés dans les snoRNA snR4 et snR8 de *S. cerevisiae*, mais leurs positions n'ont pas été identifiées.

Les snoRNA U3 et U8 sont des snoRNA à boîtes C/D impliqués dans les étapes de clivage des pré-ARNr, mais snR4 est un ARN guide de méthylation à boîtes C/D et snR8 est un snoRNA à boîtes H/ACA, guide de pseudouridylation (Ni *et al.*, 1997). Les ARN guides de modification peuvent donc eux-mêmes être modifiés.

D/ RELATIONS ENTRE MATURATION PAR DES RNases ET MODIFICATION DE NUCLÉOTIDES DANS LES ARN. LOCALISATION CELLULAIRE DES ENZYMES DE MODIFICATION EUKARYOTES

1. CAS DES ARNr

Peu de données sont disponibles sur la chronologie de la formation des modifications des nucléotides dans les ARNr procaryotes (pour revue, Maden, 1998). *In vitro*, seule la Ψ -synthase RsuA nécessite l'assemblage des protéines ribosomiques pour modifier l'ARNr 16S d'*E. coli* (Wrzesinski *et al.*, 1995b). Les autres Ψ -synthases identifiées peuvent modifier un ARNr 23S nu produit *in vitro* (voir §C.3.1.2) et semblent donc agir à un stade précoce de la formation du ribosome. Aucune relation n'a été établie entre étape de clivage et étape de modification des nucléotides.

Chez les eucaryotes, comme nous l'avons déjà mentionné, les résidus Ψ et 2'-O-méthylés sont formés dans le précurseur des ARNr cytoplasmiques (Fig. 40). Cette synthèse pourrait être concomitante avec la transcription du pré-ARNr (Maden et Salim, 1974; Brand *et al.*, 1979). Les méthylations sur les bases semblent, par contre, apparaître plus tard dans la maturation du pré-ARNr (Maden et Salim, 1974). En effet, il a été montré que la diméthylation m⁶₂A a lieu dans l'intermédiaire de maturation 20S (Fig. 40) (Lafontaine *et al.*, 1995). Cette modification, catalysée par l'enzyme Dim1p chez *S. cerevisiae*, est conservée dans les autres organismes (Lafontaine *et al.*, 1995). Elle a lieu aux positions 1779 et 1780 dans l'ARNr de la petite sous-unité ribosomique. Les modifications ayant lieu dans le nucléole, les enzymes correspondantes doivent être nucléolaires. Cbf5p et ses homologues chez le rat et la Drosophile ont effectivement été localisées dans le nucléole (Jiang *et al.*, 1993; Meier et Blobel, 1994; Phillips *et al.*, 1998).

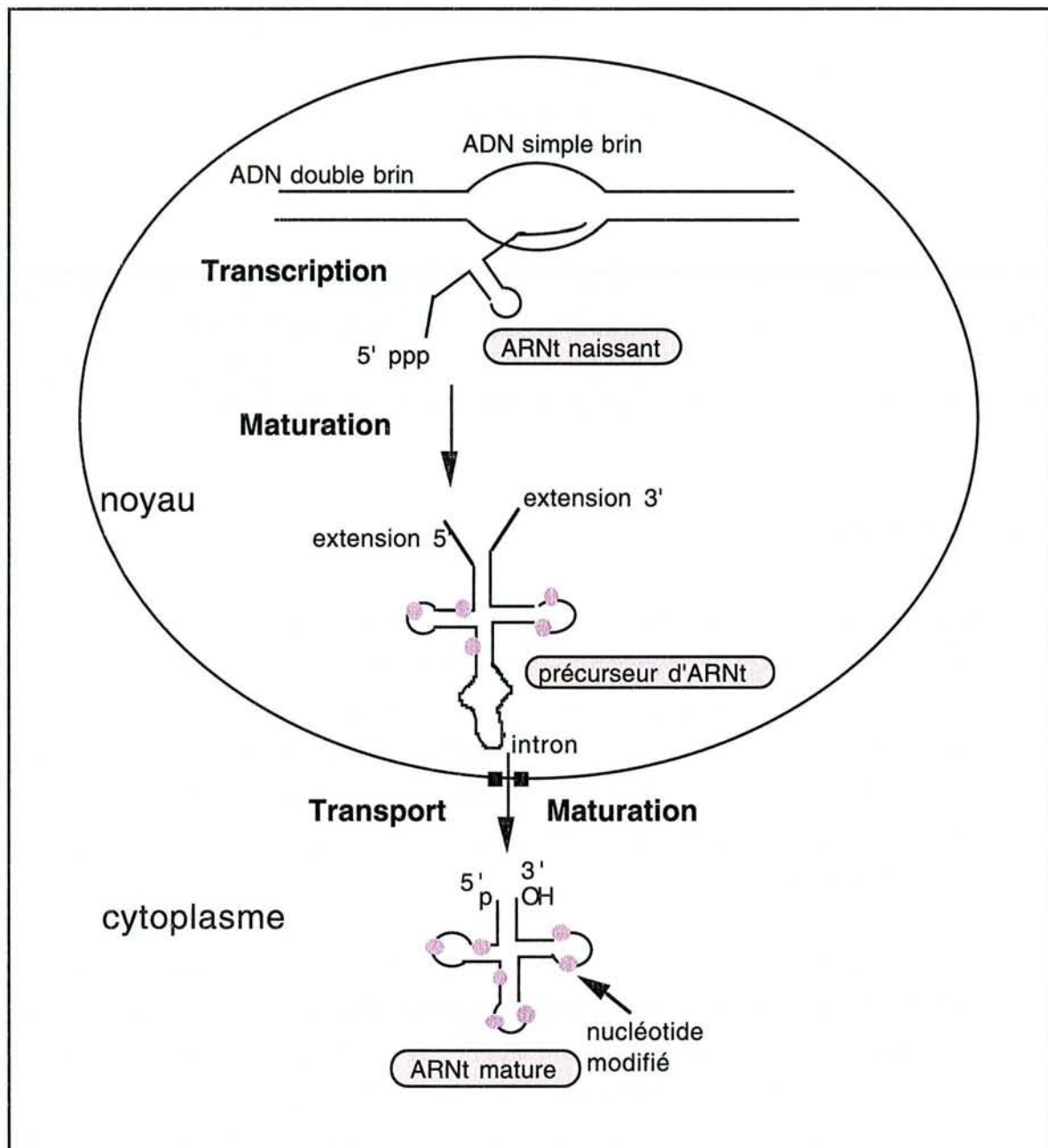


Figure 41: Principales étapes de la biosynthèse des ARNt dans une cellule eucaryote.

Une seule enzyme de modification des ARNr mitochondriaux est connue. Elle est codée par un gène nucléaire, *PET 56*. Il s'agit de la méthyltransférase requise pour la formation du résidu Gm₂₂₇₀ dans l'ARNr de la grande sous-unité ribosomique mitochondriale (Sirum-Connolly et Mason, 1993). On ne connaît pas à quelle étape de maturation des ARNr se font les modifications dans les mitochondries.

2. CAS DES ARNt

2.1. ARNt procaryotes et ARNt cytoplasmiques eucaryotes

Pour être biologiquement actifs, les ARNt de tous les organismes, procaryotes ou eucaryotes, sont issus d'un précurseur par des additions ou délétions de nucléotides aux extrémités 3' et 5' et par l'élimination d'introns. Dans le cas des ARNt cytoplasmiques eucaryotes, cette maturation se fait dans le noyau avant le transport de ces ARN vers le cytoplasme (Fig. 41). Plusieurs études, dans les années 70 et au début des années 80, avaient montré que la formation des nucléotides modifiés dans les ARNt procaryotes et eucaryotes se faisait tout au long de la maturation des ARNt et suivait donc un ordre temporel (Schaefer *et al.*, 1973; Ciampi *et al.*, 1977; Melton *et al.*, 1980; Nishikura et De Robertis, 1981; Koski et Clarkson, 1982). Par exemple, les nucléotides modifiés de la boucle TΨ (Ψ₅₅, m¹A₅₈ et m⁵U₅₄) se font, *in vivo*, dans l'ARNt précurseur, d'autres comme le résidu Ψ₃₉ sont synthétisés dans le pré-ARNt non épissé mais mûri à ses deux extrémités, et enfin, certaines modifications comme i⁶A₃₇ et Q₃₄ ne se font que dans l'ARNt mature.

Les modifications ont, par la suite, été classées en fonction de leur sensibilité à la présence d'un intron dans les ARNt (Jiang *et al.*, 1997, pour revue, Grosjean *et al.*, 1997). Certaines modifications (comme nous l'avons vu au §C.3.1.2) ne se forment que dans le pré-ARNt non épissé: m⁵C₃₄, Ψ₃₄, Ψ₃₅, Ψ₃₆ et m⁵C₄₀. Certaines requièrent l'absence d'intron pour leur formation dont, par exemple, les résidus Ψ₃₂, Gm₁₈, Q₃₄ et i⁶A₃₇. D'autres encore peuvent se former, *in vitro*, indépendamment de la présence ou non d'un intron. Il existe une corrélation entre l'étape où se font les modifications et les éléments structuraux dans l'ARNt requis pour leur synthèse. En effet, les modifications qui se forment tôt dans la maturation de l'ARNt, particulièrement celles du bras TΨ, sont celles qui sont le moins sensibles à la structure tertiaire de l'ARNt. A l'inverse, la plupart des modifications qui apparaissent plus tard, c'est-à-dire dans l'ARNt mature, sont celles qui sont très sensibles à la structure tertiaire de l'ARNt.

La majorité des nucléotides modifiés se faisant dans le pré-ARNt, les enzymes de modification des ARNt cytoplasmiques eucaryotes ont, le plus souvent, une localisation

nucléaire. En ce qui concerne les Ψ -synthases, les enzymes de levure Pus1p, Pus3p se trouvent effectivement dans le noyau (Simos *et al.*, 1996; Lecointe *et al.*, 1998). L'ARNt ligase, un composé de la machinerie d'épissage des ARNt, est localisée au niveau des pores nucléaires. L'épissage des ARNt est suivi du transport de ces ARN dans le cytoplasme. Donc, les quelques modifications se formant seulement après l'élimination de l'intron doivent se faire dans le cytoplasme, ou au moins pendant le transport des ARNt.

Le gène de la Ψ -synthase Pus1p de levure a été découvert car il est lié génétiquement aux gènes codant des protéines faisant partie des pores nucléaires, Los1p et Nsp1p (Simos *et al.*, 1996). Los1p est une protéine intervenant dans l'export des ARNt dans le cytoplasme (Hellmuth *et al.*, 1998). La délétion simultanée, mais pas individuelle, des 3 gènes *PUS1*, *LOS1* et *NSP1*, entraîne une létalité des cellules, indiquant un lien fonctionnel entre ces gènes (Simos *et al.*, 1996). Ces résultats suggèrent que l'enzyme Pus1p est importante pour le transport des ARNt vers le cytoplasme, soit parce que l'ARNt doit être modifié pour son transport, soit parce que Pus1p elle-même contribue au processus d'export des ARNt. La première hypothèse semble prévaloir car l'interaction entre des ARNt naturels, et donc modifiés, avec la protéine homologue à Los1p chez *X. laevis* (l'exportin-t) est plus stable d'un facteur 10, en moyenne, par rapport à l'interaction entre cette protéine et des ARNt produits par transcription *in vitro* (Kutay *et al.*, 1998; Lipowsky *et al.*, 1999).

2.2. ARNt mitochondriaux eucaryotes

Seules 5 enzymes de modification des ARNt mitochondriaux sont connues chez la levure *S. cerevisiae*. Les enzymes Trm1p et Mod5p sont codées par des gènes nucléaires et catalysent respectivement la formation des modifications m^2G_{26} et i^6A_{37} . Elles sont à la fois actives sur les ARNt cytoplasmiques et les ARNt mitochondriaux. Chacune de ces enzymes existe sous la forme de 2 isoformes, codés par un seul gène. L'isoforme contenant une extension en N-term est adressé vers la mitochondrie, l'autre, ne possédant pas cette extension, reste dans le cytoplasme (pour revues, Björk, 1995; Maden, 1998). Trois autres protéines modifient, à la fois, les ARNt mitochondriaux et cytoplasmiques: la protéine Trm2p, responsable de la formation du résidu m^5U_{54} , et les Ψ -synthases Pus3p et Pus4p (Hopper *et al.*, 1982; Becker *et al.*, 1997a; Lecointe *et al.*, 1998).

3. CAS DES UsnRNA

Le snRNA U6 semble 2'-O-méthylé et pseudouridylé dans le nucléole (Ganot *et al.*, 1999). On ne sait pas à quelle étape de la formation des UsnRNP se forment les modifications portées par les snRNA U1, U2, U4 et U5. Seuls quelques travaux, notamment de l'équipe de

J. Patton, apportent quelques éléments de réponse sur ce point (pour revue, Patton, 1994a). Chez l'homme, la formation des modifications portées par les snRNA U1, U2, U4 et U5 pourrait se produire à la fois dans le cytoplasme et dans le noyau, la formation de certains résidus Ψ nécessitant la présence des protéines Sm.

Pour le snRNA U2 de cellules HeLa, la fixation des protéines Sm n'est pas nécessaire à la formation de certains résidus Ψ (Kleinschmidt *et al.*, 1989). Certains nucléotides modifiés en 5' du snRNA U2 de *X. laevis* semblent nécessaires à la fixation des protéines spécifiques de la snRNP U2 (Yu *et al.*, 1998), et doivent donc se former à une étape antérieure, soit sur l'ARN nu, soit sur la pré-particule.

D/ ROLES DES RESIDUS PSEUDOURIDINES DANS LES ARN

Les rôles des nucléotides modifiés dans la fonction des ARNr et des UsnRNA sont encore peu connus. Certains rôles putatifs ont été émis, souvent sur la base de l'observation de la localisation des modifications dans des régions qui sont, en général, les régions fonctionnelles des ARN. En ce qui concerne les ARNr et les UsnRNA, l'idée générale qui semble émerger est que chaque modification post-transcriptionnelle améliorerait l'efficacité fonctionnelle de l'ARN, sans pour autant avoir un caractère indispensable.

À ce jour, les ARNt sont les seuls ARN où un lien direct a été démontré entre nucléotides modifiés et fonction de l'ARN. Ces modifications ont des effets sur les propriétés biochimiques des nucléotides. Elles affectent la structure et la stabilité des ARNt et également la fonction de ces ARN. Ainsi, les nucléotides modifiés augmentent l'efficacité et la spécificité de certaines reconnaissances codon/anticodon et ARNt/aminoacyl ARNt-synthétase. Ils interviennent également dans l'interaction entre les ARNt et les ARN génomiques de certains rétrovirus, comme c'est le cas entre l'ARNt^{lys}₃ humain et l'ARN d'HIV 1 (Isel *et al.*, 1993; pour revue, Marquet, 1998).

Nous présenterons, dans un premier temps, le rôle des résidus Ψ dans les ARNt et ensuite un état des lieux des connaissances actuelles des fonctions possibles de ce nucléotide modifié dans les ARNr et les UsnRNA.

1. ROLES DES RESIDUS Ψ DANS LES ARNt

Les résidus Ψ dans les ARNt, dont on a pu empêcher la formation dans les cellules, ne semblent pas jouer un rôle essentiel pour la survie de ces cellules. En effet, à part la délétion des gènes, codant les enzymes responsables de la formation des résidus Ψ en position 38, 39 et 40 chez *E. coli*, *S. typhimurium* et *S. cerevisiae* (*TRUA* et *PUS3*), qui induit une réduction de la vitesse de croissance cellulaire (Singer *et al.*, 1972, Carbone *et al.*, 1991, Tsui *et al.*, 1991, Lecointe *et al.*, 1998), la délétion des gènes codant les autres Ψ -synthases connues n'entraîne aucun phénotype de croissance particulier.

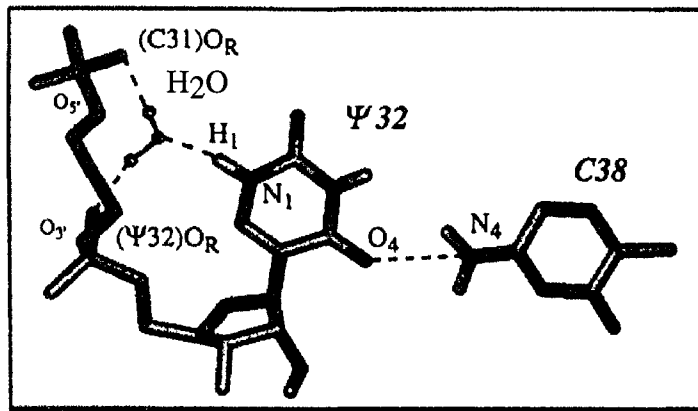
1.1. Rôles structuraux des résidus Ψ

La présence de l'ensemble des nucléotides modifiés dans un ARNt stabilise la structure tertiaire de cet ARN. Ainsi, des études de protection vis-à-vis de sondes enzymatiques et chimiques, réalisées respectivement sur l'ARNt^{Asp} de levure et l'ARNt^{Val} d'*E. coli*, ont montré que les interactions tertiaires entre les bras D et T sont partiellement déstabilisées dans des transcrits produits *in vitro*, par rapport à l'ARNt naturel (Perret *et al.*, 1990a; Derrick et Horowitz, 1993). Des études par RMN confortent ces résultats et indiquent que l'absence de certaines modifications post-transcriptionnelles peut conduire à des perturbations locales de la structure des ARNt (Hall *et al.*, 1989; Yue *et al.*, 1994; Kintanar *et al.*, 1994). Les modifications post-transcriptionnelles augmentent aussi la rigidité de la boucle anticodon et la rendent moins accessible aux nucléases (Przykorska, 1995). Les ARNt naturels présentent donc une stabilité thermique plus élevée que les transcrits correspondants (Sampson et Uhlenbeck, 1988; Perret *et al.*, 1990a; Derrick et Horowitz, 1993).

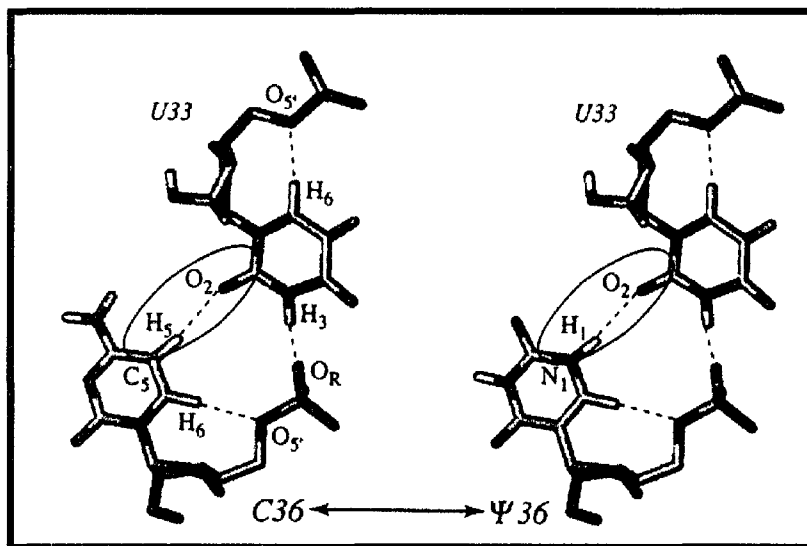
Ainsi, dans la plupart des cas, les nucléotides modifiés stabilisent la conformation active des ARNt, mais ne sont pas indispensables à l'établissement de cette structure. Un exemple est néanmoins connu où l'absence des nucléotides modifiés empêche la structuration de l'ARNt. Il s'agit de l'ARNt^{Lys} mitochondrial humain (Helm *et al.*, 1998). L'absence d'une seule méthylation sur le ribose du nucléotide en position 9 serait responsable de la formation de structures secondaires artéfactuelles.

Les résidus Ψ apportent leur contribution à la structuration de motifs locaux. Ainsi le résidu Ψ_{55} est impliqué dans la stabilisation de la boucle T et les résidus Ψ en position 32, 38 et 39 dans celle du bras de l'anticodon (pour revues, Agris, 1996; Davis, 1998; Auffinger et Westhof, 1998). Les résidus Ψ sont plus rigides que les résidus U. De ce fait, ils augmentent l'empilement des bases dans une hélice et rigidifient les boucles, stabilisant ces deux motifs structuraux (Davis, 1995; Davis *et al.*, 1998). D'autres nucléotides modifiés apportent également une stabilité locale, comme par exemple les résidus m⁵C, s²U ou les méthylations

A



B



C

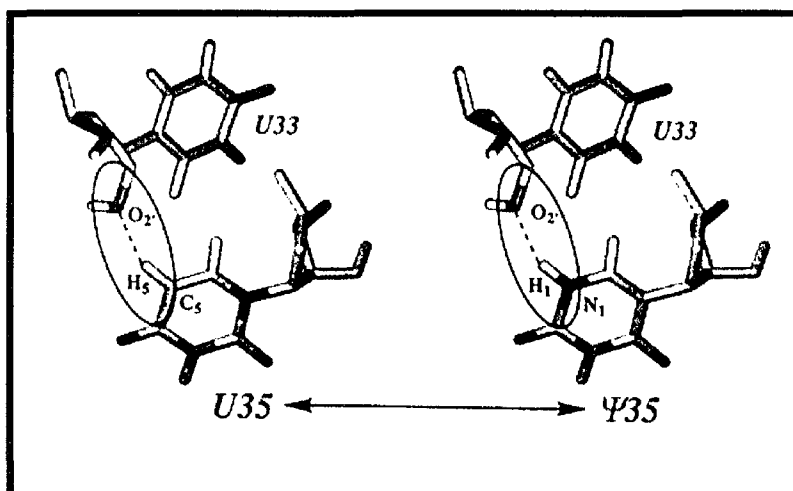


Figure 42: Les liaisons hydrogènes impliquant les résidus Ψ en position 32 (A), 36 (B) et 35 (C) des ARNt.

D'après Auffinger et Westhof, 1998.

sur le ribose. D'autres, comme la dihydrouridine, amènent une flexibilité de la région de l'ARN où elles sont localisées.

Des études par RMN avaient montré que la fonction amine supplémentaire N1-H, portée par les résidus Ψ , a un faible taux d'échange avec le solvant, ce qui démontrait que cette fonction était impliquée dans la formation d'une liaison hydrogène (Davis et Poulter, 1991; Hall et McLaughlin, 1991; 1992; Davis, 1995). Des simulations de dynamique moléculaire, ainsi que la cristallisation d'ARNt, indiquent que le mécanisme de stabilisation par les résidus Ψ se fait par un renforcement des interactions entre le résidu Ψ et la chaîne polynucléotidique par l'intermédiaire d'une molécule d'eau (Fig. 42A). En effet, la fonction amine N1-H participe à une liaison hydrogène du type $N_1-H(\Psi)\dots O$ (d'une molécule d'eau), stabilisant ainsi une molécule d'eau et liant les oxygènes anioniques portés par les phosphates du résidu Ψ et du résidu en 5' (Arnez et Steitz, 1994; Westhof *et al.*, 1988; Auffinger et Westhof, 1997) (Fig. 42A). Cette liaison réduit la mobilité conformationnelle du résidu Ψ et, par conséquent, celle des éléments structuraux proches du nucléotide modifié. Ce processus de stabilisation entre un résidu Ψ et le squelette nucléotidique, quand il existe, a pour conséquence une protection de la fonction amine N1-H qui ne peut plus alors participer à des interactions ARN-ARN ou ARN-protéine, sauf s'il y a dépliement local de la structure de l'ARN.

1.2. Rôles probables des résidus Ψ dans le bras anticodon

1.2.1. Les résidus Ψ en positions 32, 38 et 39

D'après les structures d'ARNt obtenues par cristallographie, la première et la dernière base de la boucle anticodon (position 32 et 38, respectivement) forment une pseudo-paire de bases avec une seule liaison hydrogène du type $(\Psi_{32})O_4\dots N_4(C_{38})$ pour l'ARNt^{Asp} de levure (Westhof *et al.*, 1985) (Fig. 42A), $(C_{m32})O_2\dots N_6(A_{38})$ pour l'ARNt^{Phe} de levure (Quigley et Rich, 1976) ou $(U_{m32})N_3\dots O_4(\Psi_{38})$ pour l'ARNt^{Gln} d'*E. coli* (Rould *et al.*, 1991). L'hydratation induite par la présence de résidus Ψ contribue à la stabilisation de cette interaction 32-38 (pour revue, Auffinger et Westhof, 1998). Dans toutes les structures d'ARNt connues, cette pseudo-paire de bases allonge l'hélice de la tige anticodon et prend part à la structure de cette région. Elle est, de ce fait, importante pour la fonction de l'ARNt. En effet, une mutation du résidu Ψ_{32} en C_{32} induit une perte de l'habilité de discrimination des codons par l'ARNt^{Gly} d'*E. coli* (Lustig *et al.*, 1993; Claesson *et al.*, 1995). Les résidus Ψ , en position 38 et également en position 39, interviennent également dans la lecture du codon (Bossi et Ruth, 1980; Palmer *et al.*, 1983; Hagervall *et al.*, 1990; Li *et al.*, 1997).

En effet, les modifications en position 39 contribuent à la stabilisation de l'appariement 32-38, en augmentant l'empilement des bases. Ainsi, la tige anticodon de l'ARNt^{Lys}₃ humain, contenant un résidu Ψ en position 39, est thermodynamiquement plus stable que la tige

correspondante non modifiée (Durant et Davis, 1997; 1999). Le pont d'eau, créé entre le résidu Ψ_{39} et le squelette ARN, diminue la mobilité conformationnelle de la paire de base A_{31} - Ψ_{39} et donc augmente la stabilité de cette paire de bases qui devient thermodynamiquement équivalente à celle des paires de bases G-C ou C-G (pour revue, Auffinger et Westhof, 1998).

Le rôle fonctionnel des résidus Ψ proches de l'anticodon est donc de stabiliser la structure de la boucle anticodon pour empêcher une mauvaise lecture du codon. Le fait qu'un grand nombre de nucléotides modifiés (Ψ mais également des 2'-O-méthylations) soit regroupés à la fin de la tige anticodon montre la nécessité de préserver la stabilité structurale de cette région.

1.2.2. Pas de résidus Ψ localisé en position 37

Des nucléotides modifiés sont souvent présents en positions 37 des ARNt (Fig. 31), position centrale entre l'hélice de la structure tige-boucle anticodon et l'hélice formée par l'interaction codon/anticodon. Ces nucléotides modifiés, tels que ms^2i_6A , t^6A , m^1G ou W , augmentent l'empilement des bases entre ces deux hélices et donc la stabilité de l'appariement entre la position 36 de l'anticodon et la position en 3' du codon (pour revues, Björk, 1995; Grosjean *et al.*, 1998). Ces nucléotides modifiés en position 37 peuvent également prévenir le décalage du cadre de lecture.

1.2.3. Rôles des résidus Ψ localisés dans l'anticodon (positions 34, 35 et 36)

a/ Rôles dans l'interaction ARNt/ aminoacyl-ARNt synthétase

La fidélité de la synthèse protéique repose sur la fidélité de cette interaction. Les nucléotides modifiés peuvent influencer directement cette reconnaissance entre ARNt et aminoacyl ARNt synthétase, en jouant le rôle de déterminants ou d'antidéterminants (pour revue, Björk, 1995). Les deux ARNt^{le} de *S. cerevisiae* en sont des exemples. L'ARNt^{le} majeur possède un anticodon de séquence IAU et l'ARNt mineur un anticodon $\Psi A \Psi$. Les résidus inosine et pseudouridine en position 34 sont des déterminants positifs pour l'isoleucyl-ARNt synthétase, mais c'est également le cas de l'uridine (Senger *et al.*, 1997). Il a été proposé que les groupements N-H et C=O ou C=N, présents dans ces nucléotides, soient à la base de la reconnaissance par l'enzyme. De la même façon, on peut citer le nucléotide $mnm^5s^2U_{34}$, porté par l'ARNt^{Glu} d'*E. coli*, qui joue le rôle de déterminant positif pour la glutamyl-ARNt synthétase (Sylvers *et al.*, 1993). La présence respective des résidus Ψ_{35} et Q_{34} , dans les ARNt^{Tyr} de levure et d'*E. coli*, augmente l'affinité des tyrosyl-ARNt synthétases pour ces ARNt (Noguchi *et al.*, 1978; Bare et Uhlenbeck, 1986) et l'absence des nucléotides

modifiés présents dans l'ARNt^{lle} mitochondrial humain diminue l'efficacité d'isoleucylation de cet ARN d'environ 50 fois (Degoul *et al.*, 1998). Par contre, le résidu lysidine (k²C) en position 34, dans l'ARNt^{lle} d'*E. coli*, est un antidéterminant pour la méthionyl-ARNt synthétase (Muramatsu *et al.*, 1988). De même, les nucléotides modifiés spécifiques de l'ARNt^{ASP} de levure empêchent son aminoacylation par l'arginyl-ARNt synthétase (Perret *et al.*, 1990b).

b/ Rôles dans l'interaction codon/anticodon

La position 34 de l'anticodon permet d'accommoder la dégénérescence du code génétique à la spécificité de la reconnaissance codon/anticodon, grâce au « Wobble », c'est-à-dire la reconnaissance de plusieurs codons par un même ARNt par l'établissement d'interactions non Watson et Crick entre la position 34 et le nucléotide en 3' du codon (Crick, 1966).

Les modifications de bases en position 34 peuvent faciliter ou restreindre la reconnaissance multiple entre codon et anticodon, en fonction de leur propriété biochimique. Ainsi, l'inosine, dérivée de l'adénine, peut s'apparier à l'uridine, la cytosine et l'adénine, augmentant de cette façon les possibilités d'interaction avec les codons (pour revue, Björk, 1995). La lysidine, dérivée de la cytosine, permet de stabiliser une interaction avec l'adénine au détriment de la guanine, ce qui change totalement la lecture du codon (Muramatsu *et al.*, 1988). Enfin, certains nucléotides modifiés, comme c'est le cas pour la grande majorité des dérivés de l'uridine, diminuent les possibilités d'appariements. Un exemple bien étudié est celui des 2-thiouridines (s²U), substituées en position 5, qui imposent la reconnaissance d'une adénine en troisième position du codon et excluent les codons portant une guanosine à cette position. Le thiol en position 2 introduit une rigidité du nucléotide et les substituants en position 5 empêchent la formation d'une liaison hydrogène avec l'azote 1 de la guanine (pour revue, Björk, 1995; Curran, 1998).

L'ARNt^{lle} mineur de levure possède un codon $\Psi A \Psi$ (Szweykowska-Kulinska *et al.*, 1994) qui se lie au codon AUA, mais pas au codon AUG. Il a été proposé que le résidu Ψ en position 34 inhibe, par sa rigidité, l'appariement avec la guanosine (Senger *et al.*, 1997; pour revues, Grosjean *et al.*, 1997; Curran, 1998), mais son implication stéréochimique reste encore à prouver. Le résidu Ψ_{36} peut servir à rigidifier la boucle anticodon en interagissant par une liaison du type $(\Psi_{36})N1-H \dots O_2(U_{33})$ avec la base généralement conservée en position 33 (Auffinger et Westhof, 1996; pour revue, Auffinger et Westhof, 1998) (Fig. 42B). Cette stabilisation additionnelle de la boucle anticodon semble augmenter l'efficacité de l'interaction codon/anticodon et prévenir la mauvaise lecture des codons contenant en 3' une guanosine (Senger *et al.*, 1997).

Le résidu Ψ est le seul nucléotide modifié localisé en position 35 (Fig. 31). Il est trouvé dans les ARNt^{Tyr} cytoplasmiques eucaryotes et les ARNt^{Asn} d'échinodermes et de plathelminthes. Il semble être également impliqué dans une interaction (Ψ_{35})N1-H...O₂(U₃₃), remplaçant ainsi une interaction plus faible (U₃₅)C₅-H...O₂(U₃₃) (pour revue, Auffinger et Westhof, 1998) (Fig. 42C). Ceci permettrait d'augmenter la stabilité de la boucle, stabilité requise pour la fidélité codon/anticodon. En effet, il a été démontré que la présence du résidu Ψ_{35} permet d'augmenter l'efficacité de suppression du codon stop UAA de l'ARNt^{Tyr} (U Ψ A) sup6 de *S. cerevisiae*, en augmentant la stabilité de l'interaction Ψ (anticodon):A (codon) (Johnson et Abelson, 1983). De même, le résidu Ψ_{35} des ARNt^{Tyr} (G Ψ A) de plantes permet la suppression des codons stop UAA et UAG (Zerfass et Beier, 1992).

1.3. Rôles probables du résidu Ψ en position 55

Un résidu Ψ est trouvé dans 90% des ARNt, en position 55 (Fig. 32). Il contribue à la structuration de la boucle T. En effet, la fonction amino N₃-H du résidu Ψ_{55} interagit avec un atome d'oxygène chargé négativement du phosphate 58 par l'intermédiaire d'une molécule d'eau (pour revue, Auffinger et Westhof, 1998). De plus, la fonction amino N₁-H lie le résidu 55 au résidu 54, également par une molécule d'eau (comme indiqué dans la Fig. 42A pour la Ψ_{32}) (Davis et Poulter, 1991; Westhof *et al.*, 1988). Ce résidu 55 intervient également dans la structuration tertiaire de l'ARNt, en favorisant l'interaction entre boucle D et boucle T. Ainsi, une liaison hydrogène s'établit entre l'O₄ de la Ψ_{55} et la guanine 18 de la boucle D (Westhof *et al.*, 1983; Romby *et al.*, 1987).

Les nucléotides modifiés jouent donc un rôle actif dans toutes les fonctions des ARNt, préservant ainsi la fidélité du processus de traduction. Ils peuvent intervenir de façon cruciale ou subtile, seuls ou de manière concertée, mais leur présence apporte, dans la plupart des cas, un avantage aux cellules contenant les ARNt qui les portent. Les résidus Ψ sont essentiellement trouvés dans des positions où ils contribuent, grâce à leur structure chimique particulière, à la stabilisation de motifs structuraux spécifiques. Il ne serait absolument pas étonnant que les résidus Ψ stabilisent les UsnRNA et les ARNr de la même façon. Il n'est également pas à exclure un rôle de ces nucléotides modifiés dans les interactions ARN/ARN ou ARN/protéine (comme c'est le cas de l'interaction ARNt/ aminoacyl-ARNt synthétase).

2. ROLES POSSIBLES DES RESIDUS Ψ DANS LES ARNr

Comme pour les ARNt, l'inactivation des gènes de la plupart des Ψ -synthases procaryotes modifiant les ARNr n'entraîne aucun phénotype de croissance. Seule l'insertion d'un transposon dans le gène codant une des ARNr: Ψ -synthase d'*E. coli* (RluD) entraîne un sévère défaut de croissance (Raychaudhuri *et al.*, 1998). De même, 16 des 30 résidus Ψ du grand ARN de la grande sous-unité ribosomique eucaryote ont été éliminés individuellement par délétion des gènes des snoRNA correspondants, sans perturber la croissance cellulaire (pour revue, Ofengand et Fournier, 1998). La délétion des gènes de levure, codant les protéines Cbf5p et Dim1p, entraîne une létalité des cellules (Lafontaine *et al.*, 1994; 1998a). Cette létalité ne serait pas imputée à un défaut de modification, mais à un défaut de maturation des pré-ARNr (Fig. 40) (Lafontaine *et al.*, 1995; 1998a). D'ailleurs, un mutant du gène *DIM1*, qui permet des clivages corrects du pré-ARNr, mais pas la formation des méthylations, a une croissance cellulaire normale (Lafontaine *et al.*, 1998b).

2.1. Rôles probables des résidus Ψ dans la fonction des ribosomes

2.1.1. Rôles des résidus Ψ localisés dans l'ARN de la petite sous-unité ribosomique

Un seul résidu Ψ est présent chez *E. coli*, mais il n'est pas conservé chez la levure ni chez *H. halobium* (Fig. 33A). La variation en nombre et en localisation des nucléotides modifiés dans l'ARN de la petite sous-unité ribosomique, entre les espèces, rend difficile l'établissement d'un rôle putatif pour chaque modification. Il est possible d'assembler, *in vitro*, une petite sous-unité ribosomique à partir d'un transcrit ARNr, produit *in vitro*, et les protéines ribosomiques. Cependant, son activité dans la synthèse protéique est réduite de moitié par rapport à une petite sous-unité assemblée avec un ARNr 16S naturel. Ce défaut a été attribué à l'absence des nucléotides modifiés (Cunningham *et al.*, 1991).

2.1.2. Rôles des résidus Ψ localisés dans le grand ARN de la grande sous-unité ribosomique

La majorité des nucléotides modifiés sont regroupés dans trois domaines du grand ARN de la grande sous-unité ribosomique, qui font partie de la région fonctionnelle de cette sous-unité (voir §B.2) et interagissent avec la petite sous-unité 30S chez *E. coli* (Mitchell *et al.*, 1992; Merryman *et al.*, 1999). Deux résidus Ψ sont conservés dans toutes les espèces étudiées (positions 1915 et 1917 chez *E. coli* ou 2257 et 2259 chez *S. cerevisiae*) et un troisième est semi-conservé (U_{2253} chez *S. cerevisiae* ou Ψ_{1911} chez *E. coli*) (Fig. 34B). Cette conservation laisse présager une importance fonctionnelle de ces trois résidus Ψ , ce qui est

également suggéré par le fort défaut de croissance qu'entraîne l'inactivation du gène *RLUD* d'*E. coli*, codant la Ψ -synthase responsable de la formation de ces trois résidus (Raychaudhuri *et al.*, 1998). Cependant, on ne peut pas exclure que l'enzyme RluD pourrait également avoir une autre fonction dans la cellule et ce défaut de croissance serait alors dû à l'absence de cette deuxième fonction et non à l'absence de la formation des trois résidus Ψ aux positions 1911, 1915 et 1917. Cette situation est déjà rencontrée pour la m^5U_{54} méthyltransférase d'*E. coli* (Persson *et al.*, 1992). Toutefois, ces trois résidus Ψ pourraient intervenir dans l'assemblage des sous-unités ribosomiques. En effet, il a été montré chez *E. coli* que les résidus 1912 à 1920 de l'ARN 23S sont protégés, vis-à-vis de sondes chimiques, après l'association des deux sous-unités (Merryman *et al.*, 1999) et ces mêmes résidus peuvent être pontés à l'ARNr 16S (Mitchell *et al.*, 1992). Ces trois résidus Ψ pourraient également intervenir dans le processus de décodage, puisque les résidus 1916 et 1918 peuvent être pontés à l'ARNt fixé au ribosome (Moazed et Noller, 1989) et que des mutations des résidus C_{1914} et A_{1916} affectent la fidélité de la reconnaissance codon/anticodon (O'Connor et Dahlberg, 1995).

Des sous-unités 50S chez *E. coli* ont été reconstituées, *in vitro*, en utilisant des fragments d'ARN couplés avec des fragments de transcrits, produits *in vitro*. Un segment d'ARN (position 2445 à 2523), contenant les modifications m^2G_{2445} , D_{2449} , Ψ_{2457} , Cm_{2498} , m^2A_{2503} et Ψ_{2504} , est requis pour la reconstitution de la particule, *in vitro*, et son activité peptidyl-transférase (Green et Noller, 1996). Ceci signifierait qu'un ou plusieurs de ces nucléotides modifiés soient importants pour l'activité du ribosome chez *E. coli*. Il est à noter que ce fragment d'ARN ne porte ni les résidus Ψ conservés, mentionnés ci-dessous, ni les deux résidus méthylés sur le ribose phylogénétiquement conservés dans les ARN 23/28S (Gm_{2251} et Um_{2552} selon la nomenclature de l'ARN 23S d'*E. coli*). Cela signifierait que ces modifications conservées, bien que pouvant intervenir dans la fonction du ribosome, ne sont pas essentielles à celle-ci.

L'hypothèse a été émise, que la fonction amino N1-H d'un résidu Ψ servirait d'intermédiaire accepteur de l'extrémité C-terminale de la chaîne peptidique en formation, avant sa réaction avec l'acide aminé suivant (Lane *et al.*, 1992; 1995). Cette hypothèse est basée sur le fait qu'un résidu Ψ peut être acétylé et peut transférer *in vitro* ce groupement acétylé (Wood *et al.*, 1995). Le fait qu'il n'existe pas de résidu Ψ complètement invariant au centre peptidyl-transférase, et que le caractère universel de la formation des protéines impliquerait, au moins, un résidu Ψ conservé directement impliqué dans cette fonction, est en défaveur de cette hypothèse.

3. ROLES PUTATIFS DES RESIDUS Ψ DANS LES UsnRNA

Revue

Posttranscriptional modifications in the Usmall nuclear RNAs.

S. Massenet, A. Mougin and C. Branlant

In H. Grosjean and R. Benne (eds), 1998. The modification and editing of RNA, ASM press, Washington DC, p201-228.

Revue

**Posttranscriptional modifications in the Usmall nuclear
RNAs**

S. Massenet, A. Mougin and C. Branlant.

The Modification and Editing of RNA

H. Grosjean and R. Benne (eds), ASM press, Washington DC, p 201-227

Chapter 11

Posttranscriptional Modifications in the U Small Nuclear RNAs

SÉVERINE MASSENET, ANNIE MOUGIN, AND CHRISTIANE BRANLANT

In eukaryotes, the nucleus is the site for RNA transcription and RNA maturation and most of the RNA maturation processes involve small nuclear ribonucleoprotein particles (snRNPs) containing metabolically stable small nuclear RNAs (snRNAs). This chapter deals with the posttranscriptional modifications of these snRNAs and the possible link to the function of these molecules.

In the nucleolus, the 18S, 5.8S, and 28S rRNAs are transcribed by RNA polymerase I as a long precursor, and this primary transcript is processed by an ordered series of RNase cleavages and posttranscriptional modifications (for reviews, see Maden, 1990; Henry et al., 1994; Eichler and Craig, 1994; Raué and Planta, 1995; Ofengand et al., 1995; Bachellerie et al., 1995) (see also Chapters 12 to 15). This complex maturation process involves the participation of a large series of small nucleolar ribonucleoproteins (snoRNPs) containing small nucleolar RNAs (snoRNAs) (for reviews, see Maxwell and Fournier, 1995; Tollervey and Kiss, 1997; Smith and Steitz, 1997). The snoRNAs can be divided into two classes: the snoRNAs required for the ordered cascade of RNase cleavage (U3, U8, U14, U22, and the MRP RNA) and the snoRNAs that serve as guides for the 2'-O-methylation and pseudouridine formation of rRNAs. U14 snoRNA is a special case because it seems to be required for both processes. The snoRNAs are either transcribed from their own genes by RNA polymerase II or RNA polymerase III or derived from the maturation of protein gene introns (for reviews, see Maxwell and Fournier, 1995; Bachellerie et al., 1995; Tollervey and Kiss, 1997; Smith and Steitz, 1997) (also see Chapter 13). The latter situation is more frequent for the snoRNAs that serve as a guide for posttranscriptional modifications. Depending on how they are produced, snoRNAs have different 5'-terminal structures: a trimethylated cap

structure for RNA polymerase II transcripts (for reviews, see Busch et al., 1982; Reddy et al., 1992); a γ -methyl triphosphate for RNA polymerase III transcripts (Singh and Reddy, 1989; for a review, see Reddy et al., 1992); and a phosphate when they are produced from introns (for review, Maxwell and Fournier, 1995). Because most of the U snoRNA primary structures have been established from DNA sequence analysis, except for the terminal structure, the information on snoRNA posttranscriptional modifications is very limited. Only the posttranscriptional modifications of the rat hepatoma (Reddy et al., 1985; Reddy et al., 1979) and the *Saccharomyces cerevisiae* (our unpublished results) U3 snoRNAs and of the mouse U8 snoRNA (Kato and Harada, 1984) have been investigated. The results obtained will be presented in this chapter, because they may be meaningful for understanding the relationships between U snoRNA function and posttranscriptional modifications.

In addition to rRNAs, the second main class of transcripts produced in the nucleus is pre-mRNAs. They are transcribed in the nucleoplasm by RNA polymerase II and are matured into messenger RNAs before their transport to the cytoplasm. Indeed, most of the nuclear genes coding for proteins contain introns that interrupt the coding sequences. They are eliminated in the pre-mRNAs by assembly onto the intron of a large macrocomplex, called the spliceosome, that catalyzes the two transesterification steps of the splicing reaction (Fig. 1) (for a review, see Moore et al., 1993). The first step is the nucleophilic attack of the phosphate at the 5' end of the intron by the 2'-OH of an adenosine residue from the intron branch-site sequence. For vertebrate introns, the branch-site sequence is located between positions 18 and 40 upstream of the 3'-splice site (Keller and Noon, 1984; Ruskin et al., 1984; Shapiro and Sen-

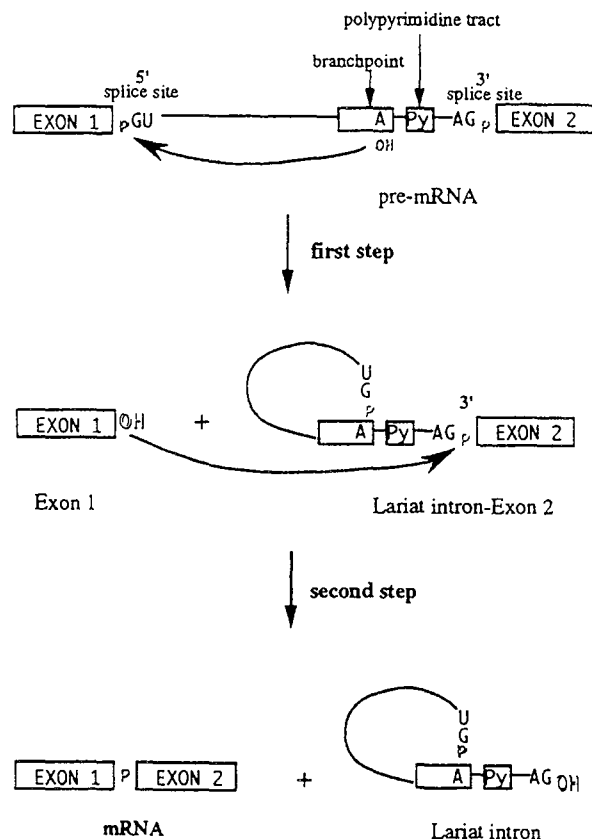


Figure 1. The two chemical steps of nuclear pre-mRNA splicing. The highly conserved nucleotides and structures of the intermediates and products are shown (for a review, see Moore et al., 1993).

apathy, 1987). In yeast, the situation is more complex: two classes of introns that differ by the distance between the branch-site sequence and the 3'-splice site were characterized (Parker and Patterson, 1987). As a result of the first nucleophilic attack, a lariat intermediate is formed and the first exon is released. The second step consists of the nucleophilic attack of the phosphate at the 3' end of the intron by the 3' hydroxyl of the free exon 1. The intron is liberated in a lariat form and the mature mRNA is produced (Ruskin et al., 1984; Padgett et al., 1984; Konarska et al., 1985). Catalysis requires a very precise assembly of the spliceosomal components: U1 and U2 snRNAs play a crucial role in intron recognition, U5 snRNA aligns the exon sequences for correct ligation, and U2 and U6 snRNAs probably play a direct role in catalysis (for reviews, Moore et al., 1993; Madhani and Guthrie, 1994; Umen and Guthrie, 1995).

Introns are defined by a limited number of conserved sequences. One essential sequence consists of the 5' extremity of the intron and the few nucleotides at the 3' end of the first exon; this segment displays complementarity with the 5' end of U1 snRNA (Ler-

ner et al., 1980; Rogers and Wall, 1980; Zhuang and Weiner, 1986; Séraphin et al., 1988; Siliciano and Guthrie, 1988; Séraphin and Rosbash, 1990) and is recognized by the U1 snRNP in an early step of spliceosome assembly (Mount et al., 1983; Black et al., 1985) (Fig. 2). The U1 snRNA interaction is later disrupted and replaced by an interaction with U6 snRNA (Sawa and Shimura, 1992; Lesser and Guthrie, 1993; Kandels-Lewis and Séraphin, 1993). The intron branch-site sequence, which displays a complementarity with U2 snRNA, is also essential. It is recognized by factor U2AF (Kohtz et al., 1994; Valcarcel et al., 1996) and then by U2 snRNA (Parker et al., 1987; Zhuang and Weiner, 1989; Wu and Manley, 1989) (Fig. 2). Finally, the third intron element required for spliceosome assembly and function is the AG 3'-terminal dinucleotide (Krainer et al., 1984; Parker and Siliciano, 1993). Hence, U1 and U2 snRNAs and a certain number of splicing factors play an essential role in the first steps of spliceosome assembly. First, the U1 snRNP is associated, then the U2 snRNP, and finally the U4-U6 and U5 snRNPs as a preformed (U4/U6,U5) tri-snRNP complex (Behrens and Lührmann, 1991) (Fig. 2). At this stage, important conformational changes occur: the U4-U6 snRNA interaction that is preserved in the tri-snRNP

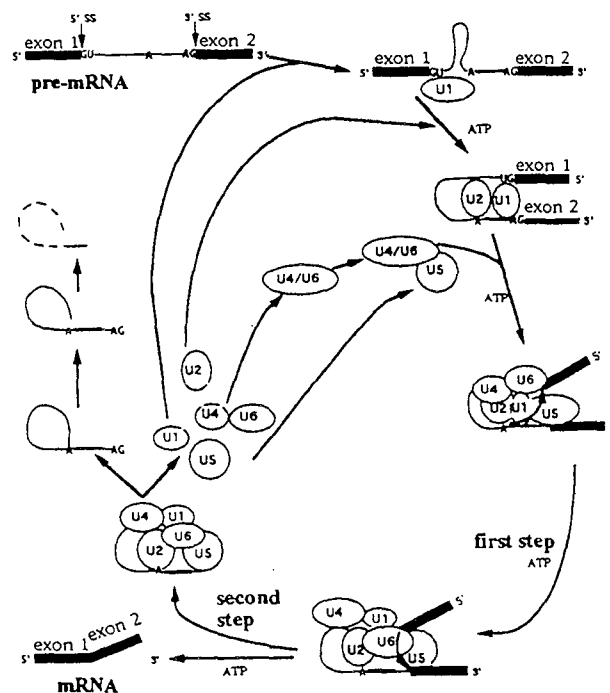


Figure 2. Steps of spliceosome assembly in the vertebrates. Substrate (pre-mRNA), intermediate (lariat-intron-exon2), and products (exon1, mRNA and lariat intron) of the reaction are indicated as well as the order of assembly of the snRNP components. First and second steps are the same as first and second steps in Fig. 1. (For references, see text.)

complex (our unpublished results) is dissociated and a U6–U2 interaction is formed (Hausner et al., 1990; Wu and Manley, 1991; Datta and Weiner, 1991; Madhani and Guthrie, 1992; Sun and Manley, 1995); U6 snRNA displaces the U1-intron interaction (Sawa and Abelson, 1992; Kandels-Lewis and Séraphin, 1993; Lesser and Guthrie, 1993; Sun and Manley, 1995); and U5 snRNA may participate to this transition and interacts with the two extremities of the exons (Newman and Norman, 1991 and 1992; Wyatt et al., 1992; Cortes et al., 1993; Sontheimer and Steitz, 1993; Newman et al., 1995). After splicing completion, U snRNAs are dissociated from the introns, the U2–U6 snRNA interaction is probably disrupted and the U4–U6 snRNA interaction reformed. This complete process implies the successive formation and dissociation of several intermolecular RNA–RNA interactions and as it is described later, the U snRNA sequences involved in these interactions are highly posttranscriptionally modified.

Indeed, posttranscriptional modifications in the spliceosomal U snRNAs have been studied long ago for vertebrate species, plants, and *Drosophila* (for reviews, see Jacob et al., 1984; Reddy, 1988; Myslinski et al., 1984; Solymosy and Pollák, 1993) and more recently for a few lower eukaryotes, *Physarum polycephalum* (Szkukálek et al., 1995), *Schizosaccharomyces pombe* (Gu et al., 1996) and *Saccharomyces cerevisiae* (our unpublished data). As mentioned above, U1, U2, U4, and U5 snRNAs that are transcribed by RNA polymerase II (Hellung-Larsen et al., 1980; Chadrasekharappa et al., 1983; for a review, see Dahlberg and Lund, 1988) have a trimethylated cap structure (Reddy et al., 1992), and this cap structure plays a very important role for the spliceosomal U snRNP biogenesis. U6 snRNA, which is transcribed by RNA polymerase III (Kunkel et al., 1986; Reddy et al., 1987; Krol et al., 1987), has a γ -methyl phosphate at the 5' end (Singh and Reddy, 1989; for a review, see Reddy et al., 1992), except in *Physarum*, where it bears a trimethyl G cap (Adams et al., 1987).

In addition to this major spliceosomal machinery that is universal in eukaryotes, a minor spliceosomal machinery was recently found in metazoan cells. It is required for the splicing of a minor class of introns, of which 17 members have been detected to date (Wu and Krainer, 1997). They are characterized by the replacement of the 5'-GU and 3'-AG dinucleotides, which are universally conserved in the major form of introns, by 5'-AU and 3'-AC dinucleotides, respectively (Jackson, 1991; Hall and Padgett, 1994). In addition, their 5'-terminal sequence is highly conserved and different from that of the classical introns. It is complementary to the 5'-terminal sequence of a minor nucleoplasmic snRNA denoted U11 (Kolosa

sova and Padgett, 1977). Their branch-site sequence is also different from that of classical introns, and it was found to be complementary to another minor snRNA called U12 (Tarn et al., 1995; Hall and Padgett, 1996). By depletion experiments, U11 and U12 snRNAs and U5 snRNA (which is used for assembly of the major form of spliceosome) were found to be required for excision of the AT–AC introns (AT–AC I introns) (Tarn and Steitz, 1996a). Finally, minor snRNAs U4atac and U6atac, which are the counterparts of the U4 and U6 snRNAs of the major form of spliceosome, have been identified (Tarn and Steitz, 1996b). Hence, this minor form of intron is spliced in a minor form of spliceosome that involves four U snRNAs distinct from those involved in the major form of spliceosome. However, the minor form of spliceosome seems to use the same U5 snRNA as the major form of spliceosome. Recently, a functionally distinct subclass of AT–AC introns (AT–AC II introns) was shown to be spliced by the major spliceosomal snRNAs (Wu and Krainer, 1997). Therefore, it is also possible that there are introns spliced by hybrid pathways, involving combinations of snRNAs or protein factors from the GT–AG and AT–AC I pathways. In both splicing machineries, the modified nucleotides may play a role. We started a study of the posttranscriptional modifications in the AT–AC specific spliceosomal snRNAs, and the results obtained will be described in this chapter. By several lines of evidence, there are similarities between the minor splicing machinery in vertebrates and the yeast splicing machinery, and this also will be discussed.

TECHNIQUES USED FOR DETECTION OF POSTTRANSCRIPTIONAL MODIFICATIONS IN U snRNAs AND U snoRNAs

The first identification of posttranscriptional modifications in U snRNAs were made long ago when RNA nucleotides sequences were determined directly using the chemical sequencing method of Peattie (1979). This method and two-dimensional, thin-layer chromatography were used for the identification of modified nucleotides. The 2'-O-methylations were easily detected after alkaline hydrolysis of end-labeled RNA because the phosphodiester bond in which a 2'-O-methyl nucleotide is implicated is resistant to hydrolysis.

One way to identify pseudouridine residues was the observation of an absence of reactivity of these residues to hydrazine in the chemical sequencing conditions described by Peattie. The presence of pseudouridine residues was verified by thin-layer chromatography (Harada et al., 1980; Reddy et al., 1981;

Krol et al., 1981a and 1981b, as examples). A rapid method for the detection of pseudouridine residues was recently described (Bakin and Ofengand, 1993) (also see Chapter 12) and used for the detection in the *S. pombe* (Gu et al., 1996) and *S. cerevisiae* snRNAs and in the human AT-AC snRNAs (our unpublished results).

Other base modifications in U6 snRNA (Gu et al., 1996; for reviews, see Reddy, 1988; Solymosy and Pollak, 1993) and U2 snRNA (Reddy et al., 1981; Westin et al., 1984; Gu et al., 1996) were identified by thin-layer chromatography.

POSTTRANSCRIPTIONAL MODIFICATIONS OF NUCLEOLAR snoRNAs

5'-Terminal Posttranscriptional Modifications

All the nucleolar snoRNAs that were characterized in vertebrates, insects, and yeast were found to be transcribed by RNA polymerase II, either from a classical gene or from an intron sequence (for reviews, see Maxwell and Fournier, 1995; Bachellerie et al., 1995; Tollervey and Kiss, 1997; Smith and Steitz, 1997); to date, the only exceptions to this rule are the plant and the *Chlamydomonas reinhardtii* U3 snoRNAs, which are transcribed by RNA polymerase III (Kiss et al., 1985; Kiss and Solymosy, 1990; Marshallsay et al., 1990; Leader et al., 1994; Antal et al., personal communication). Hence, depending on the phylum, a gene coding for a given RNA can be transcribed by different RNA polymerases and, as a consequence, the RNAs produced have different 5'-terminal structures.

All RNA polymerase II transcripts (as well the premessenger RNAs and the U snRNAs) undergo addition of a monomethylated guanosine residue to the first encoded nucleotide via a 5'-5' linkage to the 5'-terminal triphosphate (Fig. 3) (also see Chapter 10). This addition takes place shortly after the start of transcription (Shatkin, 1976; Salditt-Georgieff et al., 1980; Coppola et al., 1983). In addition, for all the characterized U snRNAs (the nucleolar and nucleoplasmic U snRNAs), this cap structure has the peculiarity to undergo two further methylations on position 2 of the terminal guanosine (Fig. 3) (Busch et al., 1982; Mattaj, 1986; for a review, see Reddy et al., 1992). In vertebrates, this is the case for the nucleolar U3, U8, and U13 RNAs (Tyc and Steitz, 1989; Reddy et al., 1979), the other snoRNAs being produced from introns. In yeast, a larger number of snoRNAs have a trimethylated cap structure, U3, snR3, R4, R5, R8, R9, R10, R11, R13, R30, R31, R32, R33, R34, R35, R37, and R189 (Tollervey, 1987; Riedel et al., 1986; Bally et al., 1988; Thomp-

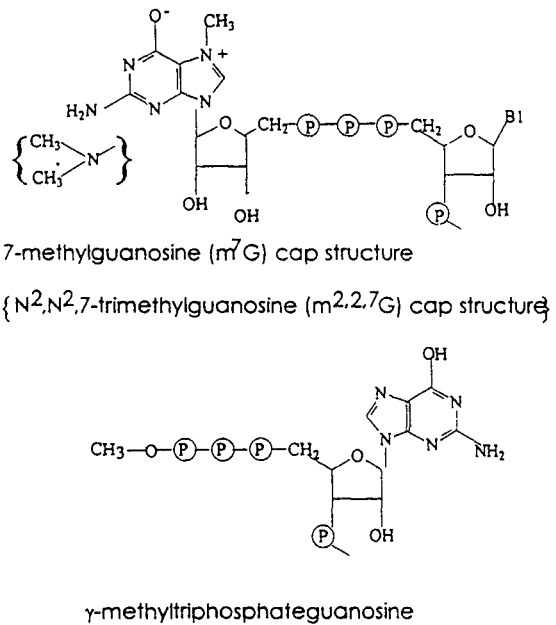


Figure 3. The various 5'-cap structures of U snoRNAs and spliceosomal U snRNAs. The m⁷G cap of the RNA polymerase II transcripts (Shatkin, 1976); N²,N²,7-trimethylguanosine cap structure (in square brackets) of the spliceosomal U1, U2, U4, U5, U11, U12, and U4atac snRNAs (Busch et al., 1982; Mattaj, 1986) and of the vertebrate U3, U8 and U13 snoRNAs (Tyc and Steitz, 1989; Reddy et al., 1979); and γ-methyl phosphate guanosine cap structure of U6 snRNA, U6atac snRNA (Singh and Reddy, 1989; Tarn and Steitz, 1996b), and plant U3 snoRNA (Marshallsay et al., 1990) are shown.

son et al., 1988; Balakin et al., 1993 and 1996; Samarsky et al., 1995).

According to the analysis of Zieve and Penman (1976), the idea was that the precursor for U3 snoRNA, like the precursors for the spliceosomal U1, U2, U4, and U5 snRNAs, is exported to the cytoplasm where it is matured (3'-terminal trimming and cap trimethylation), before being reimported into the nucleus. Based on *Xenopus laevis* oocyte microinjection assays, Baserga et al. (1992) also concluded that the biogenesis of the U3 snoRNP includes a cytoplasmic step. However, based on the same approach, a common maturation pathway was recently proposed for U3 and U8 snoRNAs that does not include a cytoplasmic step (Terns et al., 1995). Indeed, precursors of these snoRNAs were found to be retained and matured in the nucleus. Box D (see Fig. 4), a conserved element found in these and several of the snoRNAs (Hughes et al., 1987), plays a key role in their nuclear retention, probably by binding to specific proteins. Retention of U3 and U8 snoRNAs in the nucleus is saturable and relies on one or more common factors. This saturation property may explain the observation of U3 snoRNA in the cytoplasm after microinjection of a large quantity of U3

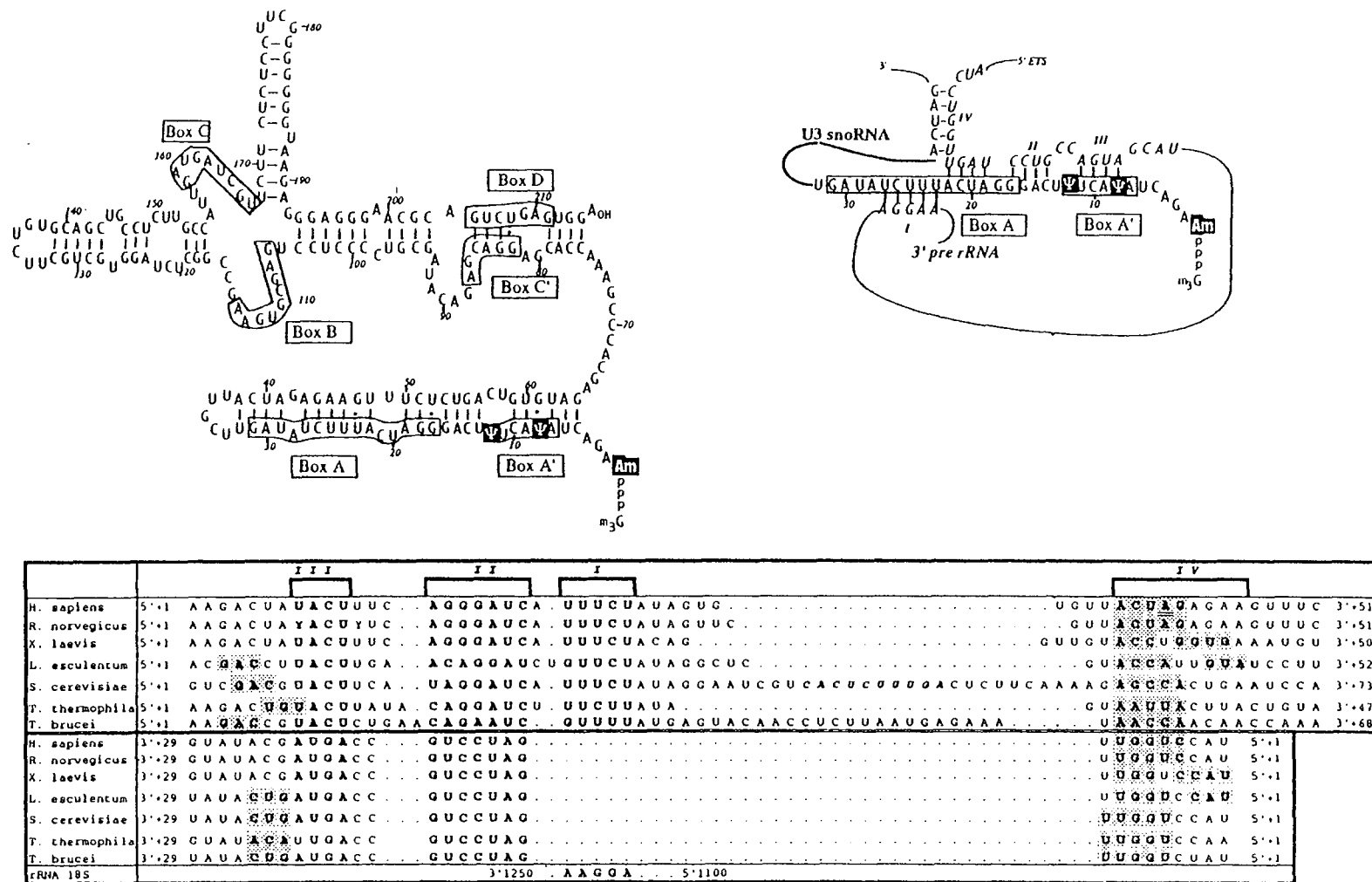


Figure 4. Rat U3B snoRNA. On the left, the possible secondary structure of the rat U3B snoRNA, based on the secondary structure experimentally established for the HeLa cell U3 snoRNA (Stroke and Weiner, 1985; Parker and Steitz, 1987). The primary sequence was from Reddy et al. (1985). The two pseudouridines identified in the rat hepatoma U3 snoRNA are shown (squared in black). The evolutionarily conserved sequences denoted Box A (Wise and Weiner, 1980), A' (Myslinski et al., 1990), B (Wise and Weiner, 1980), C (Wise and Weiner, 1980), C' (Tycowski et al., 1993) and D (Hughes et al., 1987) are indicated. On the right, interaction of U3 snoRNA with 18S pre-rRNA, as proposed by Hughes (1996) (helices I, II, and III) and Méreau et al. (1997), (helix IV). In the table at the bottom, phylogenetic conservation of the heterologous helices, I, II, III and IV between U3 snoRNA and 18S rRNA sequences according to Hughes (1996), and Méreau et al. (1997) is shown. The 5' domains of U3 snoRNAs from a wide range of organisms (top) are aligned to the complementary sequences in the respective SSU (small subunit) rRNA (bottom). The sequences are derived from the following GenEMBL database files, in order of listing in the table: U3 snoRNAs, M14061, K00780, X07318, X14411, M26649, X63788, and S74818; SSU rRNAs, X03205, V01270, XLRN01, X51576, J01353, M10932, and 152676. The putative extension of helix III in lower eukaryotes and the putative helix IV are in grey boxes.

snoRNA in the nucleus of *X. laevis* oocytes (Baserga et al., 1992). In agreement with the observation of a nuclear retention of U3 and U8 snoRNAs, Terns and Dahlberg (1994) observed the presence of a nuclear methylating activity that specifically converts the m⁷G cap of U3 snoRNA into an m^{2,2,7}G cap. Trimethylation of U3 snoRNA can also occur in the cytoplasm, but is much slower than in the nucleus (Baserga et al., 1992), suggesting that in vivo U3 snoRNA is mainly hypermethylated in the nucleus. The methylase activity is probably located within the nucleoplasm because hypermethylation can occur in nuclear extracts lacking nucleoli (Terns and Dahlberg, 1994). Hence, in vertebrates, after transcription by polymerase II, U3 snoRNA is retained within the nucleoplasm by a mechanism to be identified, and after maturation and assembly into a U3 snoRNP is transported to the nucleolus by a mechanism that involves RCC1, the abundant nuclear guanine nucleotide exchange factor for the guanosine triphosphatase Ran/TC4 (Cheng et al. 1995).

Internal Posttranscriptional Modifications

Studies on the internal posttranscriptional modifications of snoRNAs have been very limited. They were restricted to the rat hepatoma (Reddy et al., 1985) and *Saccharomyces cerevisiae* (our unpublished results) U3 snoRNAs and to the mouse U8 snoRNA (Kato and Harada, 1984) (Table 1).

In rat U3 snoRNA, three modified nucleotides were detected (Fig. 4): a 2'-O-methyladenosine residue on the 3' side of the cap triphosphate, and two pseudouridines that are located in the highly phylogenetically conserved box A' (Reddy et al., 1985; Myslinski et al., 1990) (see Fig. 4). U3 snoRNA is known to be required for the first nucleolytic cleavage at the 5' extremity of the 45S pre-rRNA precursor (Kass et al., 1990; Hughes and Ares, 1991; Mougey et al., 1993) and it was also shown to be required for further endonucleolytic steps that generate the mature 18S rRNA (Li et al., 1990; Hughes and Ares, 1991; Morrisey and Tollervey, 1993). In yeast, U3 snoRNA was first shown to base pair with a 10-nt segment of the 5'-external spacer (5'ETS) of the 45S pre-rRNA (Beltrame et al., 1994; Brulé et al., 1996) (Fig. 4). More recently, it was proposed to interact with 18S rRNA sequences from the precursor rRNA that are involved in the formation of a pseudoknot structure in mature 18S rRNA (Hughes, 1996; Méreau et al., 1997) (see Fig. 4). Using DMS as the probe, we got very strong arguments in favor of the formation of this base pair interaction in vivo (Méreau et al., 1997). Interestingly, Box A', which contains the two pseudouridines, is involved in one

of the proposed base pair interactions with 18S rRNA sequences (helix III [Fig. 4]). Based on our recent analysis, there is no pseudouridine in *Saccharomyces cerevisiae* U3 snoRNA. In this species, as well as in the other lower eukaryotes studied, the possible interaction between the 5'-terminal region of U3 snoRNA including box A' and the 18S rRNA sequence (helix III) is three base pairs longer than in vertebrates (Fig. 4) (Hughes, 1996). Substitution of uridines by pseudouridines is known to stabilize RNA-RNA interactions (for a review, see Agris, 1996) (also see Chapters 5 and 6). Hence, it might be that the presence of pseudouridines in the box A' of the vertebrate U3 snoRNAs compensates for the lower complementarity between box A' and the 18S rRNA.

The mouse U8 snoRNA was found to contain a 2'-O-methylated adenine and a 2'-O-methylated uridine on the 3' side of the triphosphate cap, and a pseudouridine residue at position 20 (Fig. 5) (Kato and Harada, 1984). U8 snoRNA was shown to be required for the formation of 5.8S and 28S rRNAs (Peculis and Steitz, 1993). The sequence of the 5' end of this snoRNA is able to base pair with the 5' end of 28S rRNA (Peculis, 1997). The role of this base pairing is to prevent the formation of premature 5.8S-28S or inappropriate 28S interactions during the pre-rRNA maturation process. It facilitates pre-rRNA processing and U8 snoRNA works as a "guide" to initiate the 5.8S-28S interaction essential for the completion of ribosome biogenesis (Peculis, 1997). None of the modified nucleotides is implicated in this U8-28S interaction (see Fig. 5). Hence, at the moment it is difficult to know whether there is a link between the presence of a pseudouridine at position 20 in U8 snoRNA and the mechanism of action of this molecule.

POSTTRANSCRIPTIONAL MODIFICATIONS OF THE SPLICEOSOMAL U snRNAs INVOLVED IN THE SPLICING OF THE GT-AG INTRONS

5'-Terminal Structure and Nuclear Export

Except for U6 snRNA, which is transcribed by RNA polymerase III (Kunkel et al., 1986; Reddy et al., 1987; Krol et al., 1987) (but not for the slime mold *P. polycephalum*: [see the introduction]), the other spliceosomal RNAs are transcribed by RNA polymerase II (for reviews, see Dahlberg and Lund, 1988; Reddy et al., 1992). U6 snRNA is retained in the nucleus probably because of the γ methylation of its 5'-terminal tri-phosphate (Fig. 3) (Singh and Reddy, 1989). Like the precursors of the vertebrate

Table 1. Distribution of intramolecular nucleotide modifications in analyzed spliceosomal U snRNAs and U snoRNAs

RNA ^a	Modifications ^b
U1	
Hs ^d	Am ₁ Um ₂ Ψ ₅ Ψ ₆ Am ₇₀
Rn ^d	Am ₁ Um ₂ Ψ ₅ Ψ ₆ Am ₇₀
Gg ^d	Am ₁ Um ₂ Ψ ₅ Ψ ₆ Am ₇₀
Mm ^d	Am ₁ Um ₂ Ψ ₅ Ψ ₆ Am ₇₀
Dm ^d	Am ₁ Um ₂ Ψ ₅ Ψ ₆ Am ₇₀
Cs ^b	Ψ _{m5} Um ₆
Sp ^c	Ψ ₃
Sc ^e	Ψ ₅ Ψ ₆
U2	
Hs ^d	Am ₁ Um ₂ Gm ₁₁ Gm ₁₂ Gm ₁₉ Gm ₂₅ Am ₃₀ Cm ₄₀ Ψ ₄₁ Cm ₆₁
Rn ^d	Am ₁ Um ₂ Ψ ₆ Ψ ₇ Gm ₁₁ Gm ₁₂ Ψ ₁₅ Gm ₁₀ Gm ₂₅ m ⁵ Am ₃₀ Ψ ₃₄ Ψ ₃₇ Ψ ₃₉ Cm ₄₀ Ψ ₄₁ Ψ ₄₃ Ψ ₄₄ Um ₄₇ Ψ ₅₄ Ψ ₅₈ (Ψ ₆₀) Cm ₆₁ (Ψ ₆₉)(Ψ ₇₂)(Ψ ₈₉) Ψ ₉₁
Vf ^d	Ψ ₆ Ψ _{m7} Ψ ₉ Ψ ₁₅ Ψ ₁₇ Gm ₁₉ Gm ₂₅ Cm ₂₈ Am ₃₀ Ψ ₃₄ Ψ ₃₇ Am ₃₈ Ψ ₃₉ Cm ₄₀ Ψ ₄₁ Ψ ₄₃ Ψ ₄₄ Ψ ₅₈ Ψ ₆₀
Sp ^{c,f}	Um ₇ Um ₁₁ Gm ₂₅ m ⁶ A ₃₀ Ψ ₃₄ Ψ ₃₇ Ψ ₃₉ Cm ₄₀ Ψ ₄₁ Ψ ₄₃ Ψ ₄₄ Ψ ₅₈
Sc ^e	Ψ ₃₅ Ψ ₄₂ Ψ ₄₄
U12	
Hs ^d	Ψ ₁₉ Ψ ₂₈
U4	
Hs ^b	Am ₁ Gm ₂ Ψ ₄ Cm ₈ Am ₆₅ Ψ ₇₂ Ψ ₇₉
Rn ^b	Am ₁ Gm ₂ Ψ ₄ Cm ₈ Am ₆₅ Ψ ₇₂ Ψ ₇₉
Gg ^b	Am ₁ Gm ₂ Ψ ₄ Cm ₈ Am ₆₅ Ψ ₇₂ Ψ ₇₉
Mm ^d	Am ₁ Gm ₂ Ψ ₄ Cm ₈ Am ₆₅ Ψ ₇₂ Ψ ₇₉
Dm ^d	Am ₁ Am ₆₅ Ψ ₁₃₁ Ψ ₁₃₂
Vf ⁱ	Ψ _{m5} Ψ _{m11} Ψ ₃₈ Ψ ₅₇ Am ₆₇ Ψ ₇₄
U4atac	
Hs ^g	Ψ ₁₂
U5	
Hs ^d	Am ₁ Um ₂ Gm ₃₇ Um ₄₁ Ψ ₄₃ Cm ₄₅ Ψ ₄₆ Ψ ₅₃
Rn ^d	Am ₁ Um ₂ Gm ₃₇ Um ₄₁ Ψ ₄₃ Cm ₄₅ Ψ ₄₆ Ψ ₅₃
Mm ^d	Am ₁ Um ₂ Gm ₃₇ Um ₄₁ Ψ ₄₃ Cm ₄₅ Ψ ₄₆ Ψ ₅₃
Cc ^d	m ⁶ A ₁ Um ₂ Cm ₃ Ψ ₁₀ Ψ ₂₉ Ψ ₃₂ Um ₃₃ Gm ₃₅ Um ₃₉ Ψ ₄₁ Cm ₄₃ Ψ ₄₄
Tt ^d	Am ₁ Um ₂ Ψ ₃₃ Gm ₃₆ Um ₄₀ Ψ ₄₂ Cm ₄₄ Ψ ₄₅ Ψ ₅₂
Ps ^d	Um ₃₀ Ψ ₃₃ Gm ₃₆ Ψ _{m40} Ψ ₄₂ Cm ₄₄ Ψ ₄₅ Ψ ₅₅
Pp ⁱ	Am ₁ Cm ₂ Ψ ₁₂ Ψ ₁₄ Ψ ₃₀ Ψ ₃₄ Um ₃₅ Gm ₃₇ Um ₄₁ Ψ ₄₃ Cm ₄₅ Ψ ₄₆
Dm ⁱ	Am ₁ Um ₂ Ψ ₃₂ Gm ₃₈ Um ₄₂ Ψ ₄₄ Cm ₄₆ Ψ ₄₇ Ψ ₅₄
Cr ⁱ	Am ₁ Um ₂ Gm ₂₄ Ψ ₃₁ Ψ ₃₆ Gm ₃₉ Um ₄₃ Ψ ₄₅ Cm ₄₇ Ψ ₄₈
Tp ⁱ	Am ₁ Um ₂ Ψ ₃₃ Gm ₃₆ Um ₄₀ Ψ ₄₂ Cm ₄₄ Ψ ₄₅ Ψ ₅₂
Sp ^c	Ψ ₃₈ Gm ₄₆ Um ₅₀ Ψ ₅₂ Cm ₅₄
Sc ^e	Ψ ₉₉
U6	
Rn ^d	Ψ ₃₁ Ψ ₄₀ m ⁶ A ₄₃ Am ₄₇ Am ₅₃ Gm ₅₄ Cm ₆₀ Cm ₆₂ Cm ₆₃ Am ₇₀ mG ₇₂ Cm ₇₇ Ψ ₈₆

Continued on following page

Table 1. Continued

RNA	Modifications ^a													
Mm ^d ...	Ψ ₁₁	Ψ ₄₀	m ⁶ A ₄₃	Am ₄₇	Am ₅₃	Gm ₅₄	Cm ₆₀	Cm ₆₂	Cm ₆₃	Am ₇₀	mG ₇₂	Cm ₇₇	Ψ ₈₆	
Vf ^k ...	Ψ ₂₆		m ⁶ A ₃₈	Am ₄₂	Am ₄₈			Cm ₅₇		Cm ₆₃	Am ₆₅	Gm ₇₅	Cm ₈₄	Ψ ₉₀
Sp ^c ...	Ψ ₂₆	Gm ₂₅	Ψ ₄₀	m ⁶ A ₃₇	Am ₄₁			Cm ₅₆	Cm ₅₇					
U3														
Rn ^f ...	Am ₁		Ψ ₈		Ψ ₁₂									
U8														
Mm ^m ...	Am ₁	Um ₂						Ψ ₂₀						

^a Reddy, 1988 (review).^b Kiss et al., 1988a.^c Gu et al., 1996.^d Westin et al., 1984.^e Reddy et al., 1981. The pseudouridines in parentheses are additional pseudouridines found in a study of posttranscriptional modifications in the 5' half of rat liver and brain snRNA (Branlant et al., 1982).^f For the *S. pombe* U2 snRNA, the presence of Ψ59 is uncertain, as it was shown in a figure but not cited in the text in a paper by Gu et al. (1996). No Ψ or 2'-O methylation was detected in *S. pombe* U4 snRNA.^g Massenet et al., unpublished results. Only the presence of pseudouridines was studied in the *S. cerevisiae* U1, U2, U4, U5, and U6 snRNAs and in the atac-specific U snRNAs from HeLa cells (U11, U12, U4atac, and U6atac). No pseudouridine was detected in the *S. cerevisiae* U4 or U6 snRNA or in the HeLa cell U11 or U6atac snRNA.^h Branlant et al., 1982.ⁱ Kiss et al., 1988b.^j Szukálek et al., 1995.^k Solymosy and Pollák, 1993 (review).^l Reddy et al., 1979.^m Kato and Harada, 1984.ⁿ Abbreviations: Hs, *Homo sapiens*; Rn, *Rattus norvegicus*; Gg, *Gallus gallus*; Mm, *Mus musculus*; Vf, *Vicia faba*; Ps, *Pisum sativum*; Dm, *Drosophila melanogaster*; Cs, *Chlorella saccharophila*; Cr, *Chlamydomonas reinhardtii*; Cc, *Cryptosporidium cohnii*; Tp, *Tetrahymena pyriformis*; Tt, *Tetrahymena thermophila*; Sp, *Schizosaccharomyces pombe*; Sc, *Saccharomyces cerevisiae*; Am, Gm, Um and Cm, 2'-O-methyl adenosine, 2'-O-methyl guanosine, 2'-O-methyl uridine, and 2'-O-methyl cytidine, respectively; m⁶A, N⁶-methyladenosine; Ψ, pseudouridine; mG, 2-methylguanosine. All these U snRNAs are capped by a 5'm₇G (N^{2,2,7}-trimethylguanosine), except the U6 and U6atac snRNAs, which are capped by a 5'γmG (γ-monomethyltriphosphate guanosine).

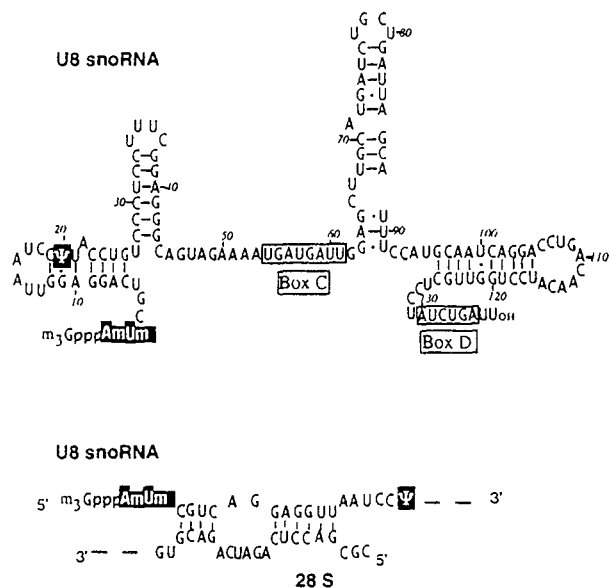


Figure 5. Secondary structure proposed for mouse U8 snoRNA. The sequence was from Kato and Harada (1984), and the structure was proposed by Tyc and Steitz (1989). The evolutionarily conserved sequences denoted Box C and D are indicated, as well as the identified posttranscriptional modifications. The interaction of U8 snoRNA with 28S rRNA, as proposed by Peculis (1997), is shown. The posttranscriptional modifications are squared in black.

U3 and U8 snoRNAs, after transcription by RNA polymerase II, the precursors of the U1, U2, U4, and U5 snRNAs carry a monomethyl-G-cap structure (for reviews, see Busch et al., 1982; Dahlberg and Lund, 1988). However, in contrast to the U snoRNA precursors, the U1, U2, U4, and U5 snRNA precursors are transported to the cytoplasm where they are matured (hypermethylation of the cap structure, 3' trimming and binding of Sm proteins) (see Fig. 6).

As a first step in nuclear export of U1, U2, U4, and U5 snRNA precursors, the monomethyl-G-cap structure is recognized by the nuclear cap-binding-protein complex (CBC), which was first identified by Ohno et al. (1990) (Fig. 6). This complex is composed of two proteins, CBP80 and CBP20 (Izaurralde et al., 1994) (Fig. 6), and it is also involved in the recognition of the 5'-cap structure of premessenger RNAs (Jarmolowski et al., 1994). Protein CBP20 belongs to the large family of proteins containing a conserved (RNP) motif, including the highly conserved RNP1 and RNP2 sequences. However, it does not exhibit specific cap-binding activity in the absence of protein CBP80 (Izaurralde et al., 1995). Demonstration of the role of protein CBP20 in pre-U1 and pre-U5 snRNA export was shown by microinjection of CBP20 antibodies in *X. laevis* oocytes (Izaurralde et al., 1995). The effect of cap structure elimination was found to be much more dramatic for U snRNA ex-

port than for mRNA export, probably because other proteins like hnRNP A1 are also involved in mRNA export (Piñol-Roma and Dreyfuss, 1992; Michael et al., 1995; Lee et al., 1996; Visa et al., 1996). However, based on microinjection of various U1 snRNA mutants, the m⁷G cap structure was proposed to be required but not sufficient for nuclear export of U1 snRNA. Both the 3' terminal stemloop structure and sequence in the 5' terminal 124-nt region were proposed to contribute to efficient export of this RNA (Terns et al., 1993). Nevertheless, in this kind of experiment, it is somehow difficult to make a distinction between the effect on export and the effect on RNA stability. By using U1 snRNA derivatives carrying different cap structures as competitors of export, it was demonstrated unequivocally that the export of U snRNAs requires saturable factors that bind to their 5' caps, CBP20 and CBP80 proteins (Jarmolowski et al., 1994). Their distribution in the nucleoplasm was shown to be relatively uniform (Izaurralde et al., 1994; Visa et al., 1996). The yeast counterpart of the vertebrate CBP20, the protein Mud13p, was recently isolated (Color et al., 1996) and it was shown to be involved in the formation of splicing commitment complexes. Protein CBP80 contains leucine-rich sequences that somehow resemble the hydrophobic nuclear export signal (NES). However, there is no evidence that these sequences are required for CBC function in export. CBC acts rather by the recruitment of a NES-containing polypeptide to the RNP (for a review, see Görlich and Mattaj, 1996). Recent results demonstrate that importin- α , a protein first characterized as a protein nuclear localization signal (NLS) receptor (Görlich et al., 1995a and 1995b), forms a trimeric complex with proteins CBP20 and CBP80 and this trimeric complex binds RNA in a cap-dependent fashion, which is consistent with the idea that the CBC RNA complex is exported from the nucleus as a complex with importin- α (Görlich et al., 1996a) (Fig. 6).

In addition to proteins CBP20, CBP80 and importin- α , the GTPase RAN/TC4 and its nucleotide exchange factor RCC1 were found to be required for nuclear export of spliceosomal RNAs (Kadowaki et al., 1992 and 1993; Belhumeur et al., 1993; Amberg et al., 1993; Schlenstedt et al., 1995; Cheng et al., 1995). The abundant RCC1 protein was found to stimulate GDP to GTP exchange at the catalytic center of the Ran/TC4 GTPase (Moore and Blobel, 1994; Bischoff et al., 1995). Ran is essentially present in a form associated with GDP in the cytoplasm and it plays a function in protein import (Görlich et al., 1996b). In contrast, due to the activity of RCC1 in the nucleus, Ran is essentially present in a form associated with GTP in this compartment (Ohtsubo et

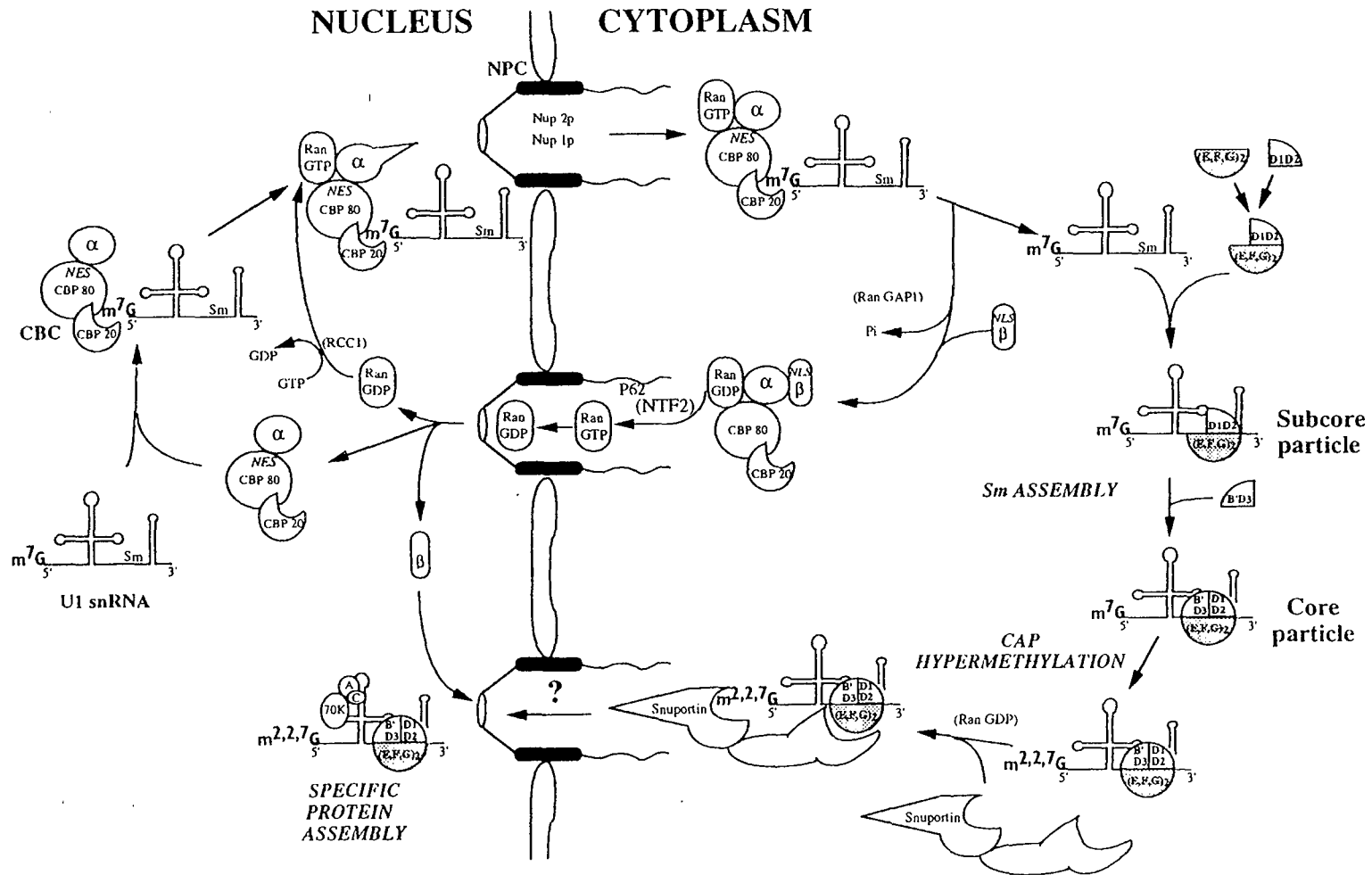


Figure 6. A model for the nuclear transport cycle of the spliceosomal U snRNPs. In the nucleus, newly transcribed snRNA (U1 as an example) binds the CBC-importin- α complex, and the Ran/TC4 guanosine triphosphatase in its GTP-associated form (Ran GTP). This complex is exported to the cytoplasm through the nuclear pore complex (NPC). Importin- α (α) interacts with the pore proteins Nup1p and Nup2p. In the cytoplasm, GTP is hydrolyzed in the presence of Ran GAP1, association of importin- β promotes the liberation of the m⁷G U snRNA, and the protein complex (Ran; importins α , β ; CBC) is allowed to move back to the nucleus. The m⁷G U snRNAs associate with the Sm proteins in a sequential pathway, and hypermethylation of the cap structure occurs. Both events provide the bipartite nuclear localization signal (NLS), recognized by a cytosolic factor (snuportin) which drives the core snRNP back to the nucleus, where specific proteins can associate. (For references, see the text.)

al., 1989; Bischoff and Ponstingl, 1991; Cheng et al., 1995). It associates to the CBC-importin- α -U snRNA complex to allow its translocation through the NPC where importin- α interacts with the NPC proteins Nup1p and Nup2p (Belanger et al., 1994) (Fig. 6).

Once in the cytoplasm, importin- α forms a heterodimer with importin- β (Görlich et al., 1996a) (Fig. 6), which results in the dissociation of the RNA from the RNA-CBC-importin- α complex delivered to the cytoplasm (Görlich et al., 1996c). Based on these data, it was proposed that the ability of importin- β to dissociate the export complex is blocked by Ran GTP in the nucleus (Görlich et al., 1996c). In the cytoplasm, the Ran-GAP1-GTPase activating protein activates the hydrolysis of GTP by the Ran protein, allowing the binding of importin- β . The importin- β has an NLS sequence that drives back the CBC-importin- α complex to the nucleus, to enter a new transport cycle (Görlich et al., 1996a). This implicates that the binding of Ntf2 (linked to the nucleoporin P62) to Ran-GDP and a GDP to GTP exchange, followed by GTP hydrolysis, allow translocation of the import complex across the central channel of the NPC (Wong et al., 1997; Panté and Aebi, 1996).

Hence, in the present stage of the knowledge, the presence of the m⁷G cap structure on U1, U2, U4, and U5 snRNA precursors allows recognition by the cap-binding protein complex (CBP) and associated proteins and this mediates nuclear export. Although bearing the same m⁷G cap structure, the nucleolar U3, U8, and U13 snoRNAs escape from this process by a retention mechanism involving the phylogenetically conserved Box D.

5'-Terminal Structure and Nuclear Import

After their liberation in the cytoplasm, the precursors of U1, U2, U4, and U5 snRNAs associate with a common set of eight proteins, called the Sm proteins (Lerner and Steitz, 1979; Liautard et al., 1982; Kastner et al., 1990) (Fig. 6). The Sm binding sites of the 4 snRNAs is a single-stranded segment with the consensus sequence PuA(U)_{3,6}Gpu, which is flanked by two stable helical structures (Branlant et al., 1982; Jarmolowski and Mattaj, 1993). Correct assembly of the eight Sm proteins is needed to get the conversion of the m⁷G-cap structure into a m^{2,2,7}G-cap structure (Mattaj, 1986; Plessel et al., 1994). Trimming also occurs at the 3' end of the precursor molecules, but the information on this step is very limited (Jacobson et al., 1993; Huang et al., 1997; Chanfreau et al., 1997). Formation of the U snRNP core complexes, upon association of U snRNAs with the Sm proteins, involves both

protein-protein and RNA-protein interactions and the assembly pathway of the U1 snRNP core complex has been established in vitro (Raker et al., 1996) (Fig. 6). The Sm proteins E, F, and G first form a hexamer containing two copies of each protein. In association with proteins D1 and D2, the E.F.G. hexamer can form a stable complex with U1 snRNA (U1 snRNP-subcore particle), which is converted into a U1 snRNP-core particle by association with a preformed B/B'.D3 complex. When microinjected into *X. laevis* oocyte cytoplasm, this reconstituted U1-core snRNP particle is the substrate of the trans-activesnRNA-(guanosine-N2)-methyl transferase leading to the trimethylated cap structure and is transported to the nucleus (Plessel et al., 1994).

Both binding of Sm proteins and hypermethylation of the cap provide the two parts of the bipartite NLS of the core-U snRNP particles (Mattaj and De Robertis, 1985; Fisher and Lührmann, 1990; Hamm et al., 1990). Only the presence of the core Sm protein complex seems to be completely indispensable for nuclear import of UsnRNAs (Fisher et al., 1993). The importance of the m₃G cap-structure for nuclear import is variable from one U snRNA to the others. It is more marked for U1 and U2 snRNAs than for U4 and U5 snRNAs. U snRNPs containing only the Sm-core RNP domain can enter the nucleus (Fisher et al., 1993), which suggests that the m₃G cap and the Sm-core-RNP domain may be recognized independently by the nuclear receptor(s) involved in the nuclear import of UsnRNPs, with the binding of the second element certainly strengthening the signal of the first one. In agreement with this hypothesis, a specific m₃G-cap-binding protein, called snuportin, was cross-linked to an RNA oligonucleotide corresponding to the 5' terminus of U1 snRNA (Marshall-say and Lührmann, 1994; Huber and Lührmann, personal communication). Once in the nucleus, completion of the snRNP is performed by the assembly of the specific proteins. Until now, nothing is known about the step(s) of U snRNP assembly where internal posttranscriptional modifications of U snRNA take place. Hence U1, U2, U4, and U5 snRNAs' biogenesis involves a bidirectional traffic of the U snRNAs throughout the NPC, and the posttranscriptional modifications at the 5' end of these RNAs plays an essential role in this traffic.

The Internal Posttranscriptional Modifications and Their Biogenesis

Internal posttranscriptional modifications in spliceosomal snRNAs were studied for four vertebrate species (chicken, mouse, and human [Harada et al., 1980; Kato and Harada, 1981; Krol et al., 1981a,

1981b; Branlant et al., 1982] and rat [for a review, see Reddy, 1988]), for the insect *Drosophila melanogaster* (Myslinski et al., 1984), for plants (Kiss et al., 1988a, 1988b; Solymosy and Pollák, 1993), for the fission yeast *S. pombe* (Gu et al., 1996) and for the budding yeast *S. cerevisiae* (our unpublished data). The study was extended to a larger number of species in the case of U5 snRNA (Fig. 7).

The identity and location of all the detected posttranscriptional modifications are shown in Table

1. In addition to 2'-O-methylations, they essentially consist of pseudouridines and a few methylations at position 6 of adenosines (Table 2). For most of the U snRNAs, modifications are clustered in specific areas of the molecules and they are highly conserved throughout evolution: the 5' terminal region of U1 snRNA, the 5' third of U2 snRNA, the 5'-terminal and central region of U4 snRNA, the terminal loop of stem-loop structure 1 of U5 snRNA, and the central region of U6 snRNA (Table 1; Fig. 7, 8, and 9A).

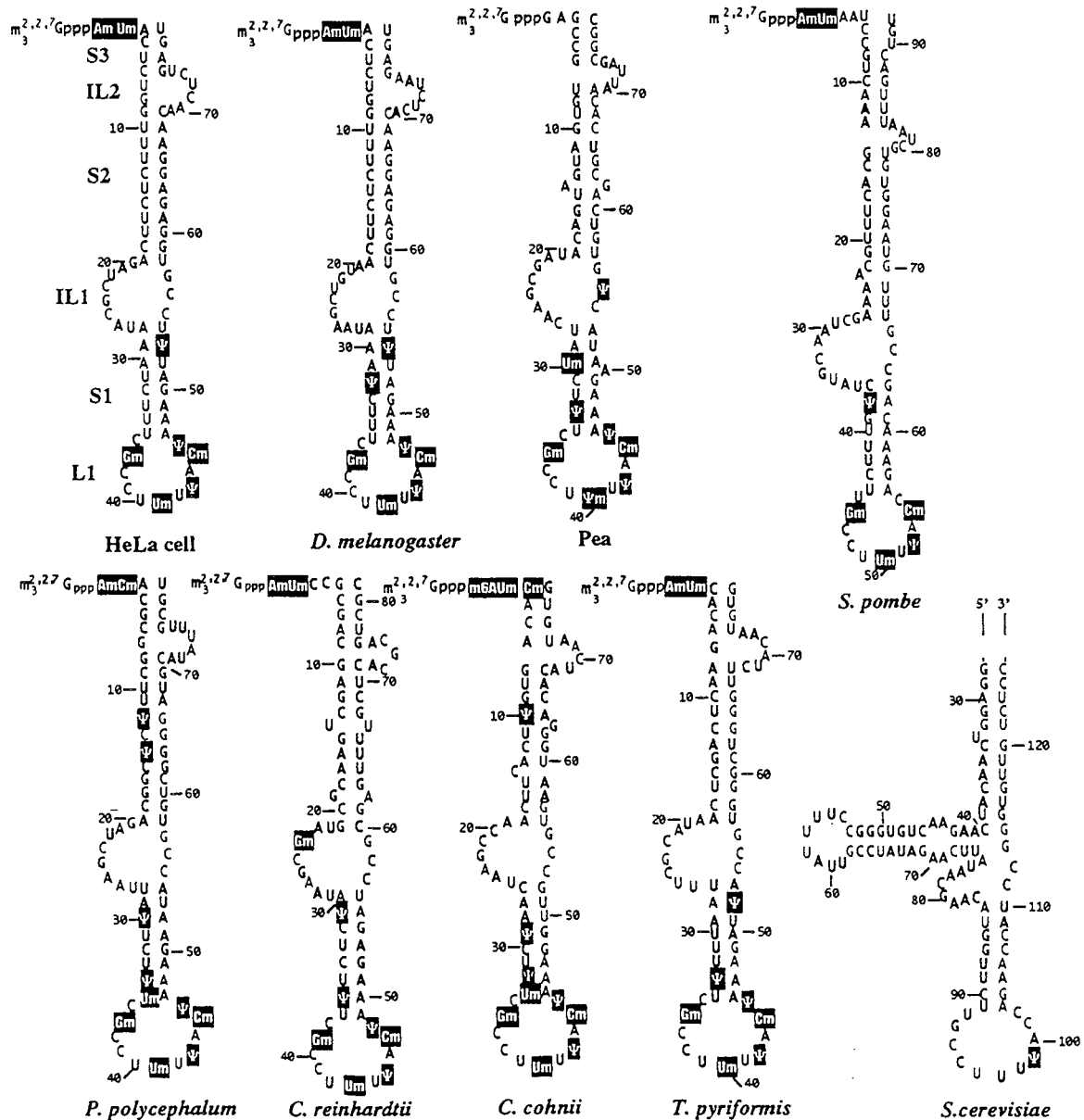


Figure 7. A comparison of modified nucleotides in U5 snRNAs from different species. Only the stem-loop structure I that contains modified nucleotides is shown. The nucleotide sequences are from HeLa cells (Reddy, 1988), *Drosophila melanogaster* (Myslinski et al., 1984), Pea (Krol et al., 1983), *Saccharomyces pombe* (Gu et al., 1996), *Physarum polycephalum* (Szkukalek et al., 1995), *Chlamydomonas reinhardtii* (Jakab et al., 1992), *Cryptocodium cohnii* (Reddy et al., 1983), *Tetrahymena pyriformis* (Branlant et al., 1983) and *Saccharomyces cerevisiae* (Massenet et al., unpublished results). The posttranscriptional modifications are squared in black.

Table 2. Internal posttranscriptional modifications in analyzed U snRNAs and spliceosomal U snRNAs^a

Common name	Symbol	Chemical Abstracts Index name
N ⁶ -Methyladenosine	m ⁶ A	Adenosine, N-methyl-
2'-O-Methyladenosine	Am	Adenosine, 2'-O-methyl-
N ⁶ ,2'-O-Methyladenosine	m ⁶ Am	Adenosine, N-methyl-2'-O-methyl-
2'-O-Methylcytidine	Cm	Cytidine, 2'-O-methyl-
N ² -Methylguanosine	m ² G	Guanosine, N-methyl-
2'-O-Methylguanosine	Gm	Guanosine, 2'-O-methyl-
Pseudouridine	Ψ	2,4(1 <i>H</i> ,3 <i>H</i>)-Pyrimidinedione, 5-β-D-ribofuranosyl-
2'-O-Methyluridine	Um	Uridine, 2'-O-methyl-
2'-O-Methylpseudouridine	Ψm	2,4(1 <i>H</i> ,3 <i>H</i>)-Pyrimidinedione, 5-(2-O-methyl-β-D-ribofuranosyl)-

^a For a review, see Solymosy and Pollák (1993). The chemical modification structures are shown in Appendix 1.

With respect to the established secondary structures of U1, U2, U4, U5, and U6 snRNAs (Branlant et al., 1981; Mount and Steitz, 1981; Krol et al. 1981a, 1981b; Branlant et al. 1982; Myslinski and Branlant, 1991; Zavanelli et al., 1994), a large number of the posttranscriptionally modified nucleotides are located in RNA segments predicted to be single stranded.

Both 2'-O-methylations and pseudouridines in *S. pombe* U snRNAs as well as pseudouridines in *S. cerevisiae* U snRNAs were studied (Table 1). Surprisingly, U snRNA posttranscriptional modifications are considerably less numerous in these two species than in all the other species studied, and this difference is more marked for *S. cerevisiae* than for *S. pombe*. We detected only six pseudouridines in the five spliceosomal U snRNAs of *S. cerevisiae*: two in U1 snRNA,

three in U2 snRNA (against 12 in rat hepatoma) and one in U5 snRNA.

Little is known on the biogenesis of the U snRNA internal modifications. Only the in vitro conversion of uridines into pseudouridines of U snRNAs produced by in vitro transcription was studied in HeLa cell nuclear extract or S100 extract (for a review, see Patton, 1994a). The strategy used to determine whether synthesis of pseudouridine moieties in U1, U2, and U5 snRNAs was catalyzed by the same or by different pseudouridine synthases was to test for the pseudouridine synthase activity of HeLa cell extracts, in the presence of U1, U2, or U5 snRNA containing 5-fluorouridine (5-FU). Indeed, RNAs containing 5-FU at positions where uridines are converted into pseudouridines were previously shown to

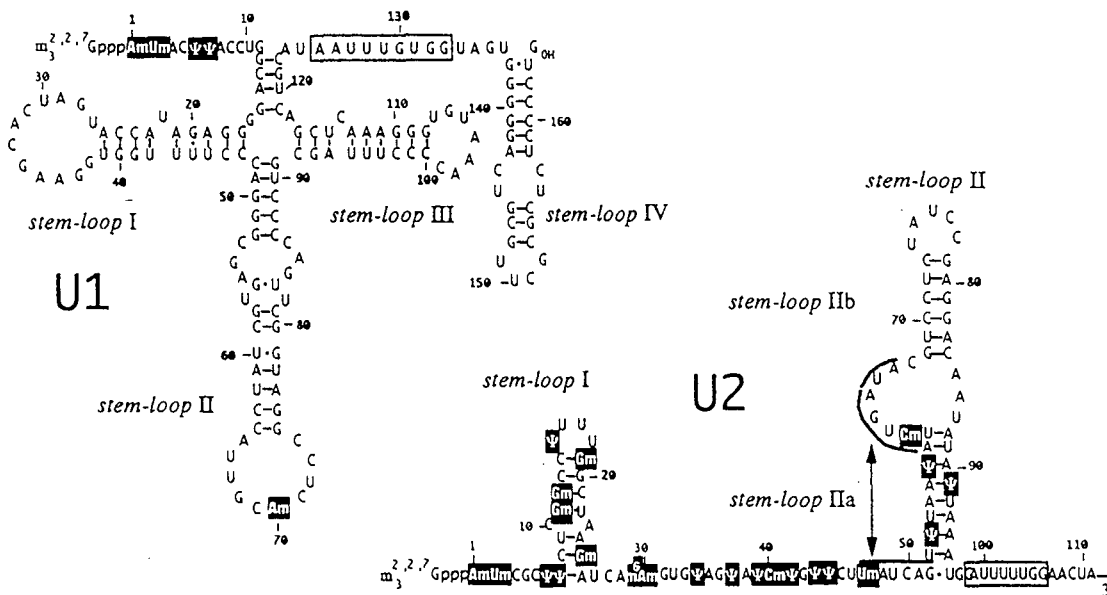
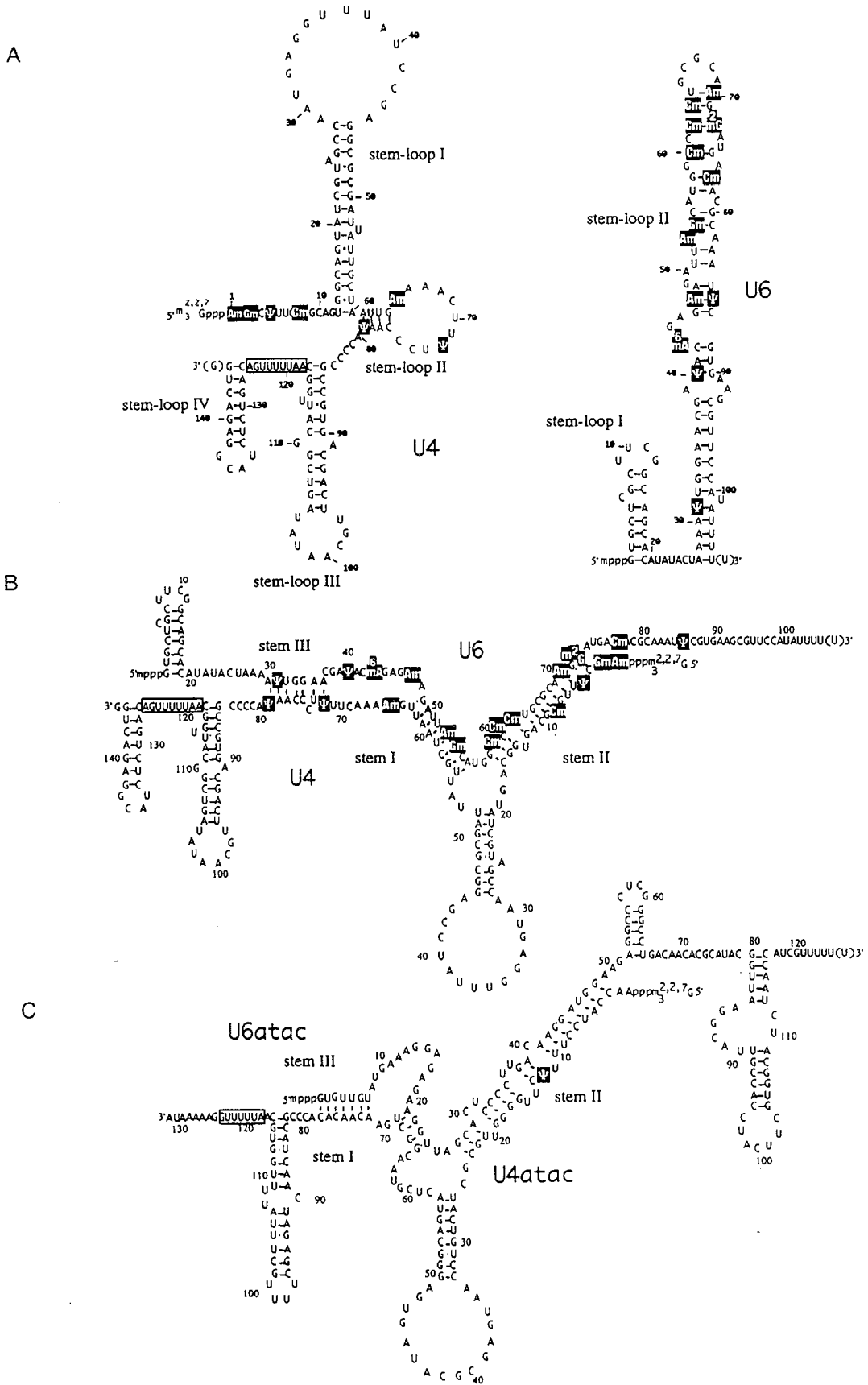


Figure 8. Secondary structures of rat spliceosomal U1 and U2 snRNAs. The secondary structures are from Branlant et al. (1981) and Mount and Steitz (1981) (U1) and Branlant et al. (1982) and Zavanelli et al. (1994) (U2). For the U2 snRNA structure, the sequences implicated in the proposed pseudoknot (Zavanelli et al., 1994) are overlined and joined by a double arrow. The Sm binding sites are boxed. The 3' extremity is not shown. The posttranscriptional modifications are squared in black.



block pseudouridine-synthase activity by formation of an irreversible complex (Samuelson, 1991). Using this strategy, U1 snRNA containing 5-FU was found to inhibit pseudouridine formation in U1 snRNA transcripts, but not in U2 or U5 snRNA transcripts (Patton, 1993). The same situation was found for U2 and U5 snRNAs. Hence, distinct pseudouridine synthases catalyze the pseudouridine formation in each of these three U snRNAs. A more detailed analysis was performed on pseudouridine formation in U2 and U5 snRNAs. For U2 snRNA, the inhibition experiments with 5-FU containing RNAs were extended to a large series of U2 snRNAs containing deletions and point mutations, and the data supported the hypothesis that there is no obligatory pathway for pseudouridine formation in U2 snRNA because formation at one site in U2 snRNA was not dependent on the formation of pseudouridine at any other site. As a whole, the data also suggest that there are probably multiple pseudouridine synthase activities that modify U2 snRNA (Patton et al., 1994).

Studies of pseudouridine formation in U5 snRNA found additional evidence for the existence of multiple pseudouridine synthase activities that modify a single U snRNA. Indeed, among the three pseudouridines that are present in the HeLa cell U5 snRNA (Kato and Harada, 1981; Branlant et al., 1983), formation of pseudouridine at positions 43 and 46 in the highly phylogenetically conserved terminal loop of the stem-loop structure I (Fig. 7) is obtained *in vitro* using an S100 extract or a nuclear extract (Patton, 1991, 1994b). Formation of pseudouridine 53 in the stem of this stem-loop structure is not obtained with an S100 extract. However, it can be obtained with a nuclear extract (Patton, 1991; Patton, 1994b). Hence, formation of pseudouridine 53, on the one hand, and formation of pseudouridines 43 and 46, on the other hand, are catalyzed by at least two distinct pseudouridine synthases. A combination of nuclear and S100 extracts gave the most efficient total formation of pseudouridines in the human U5 snRNA (Patton, 1994b).

Formation of pseudouridines in U5 snRNA at all three positions was found to be dependent on the integrity of the Sm binding site (Patton, 1991; Patton, 1994b). Such dependence on the Sm binding site was

not observed for U2 snRNA, although it cannot be ruled out that some of the pseudouridines are not formed in U2 snRNA upon deletion of the Sm binding site. Indeed, a detailed analysis of the pseudouridines formed in these conditions was not performed (Kleinschmidt et al. 1989). In fact, no systematic analysis of pseudouridines in the human U2 snRNA has ever been performed, and to interpret their results, Patton and coworkers made the hypothesis that the formation of pseudouridines in the human U2 snRNAs takes place at the same positions as in rat hepatoma cells.

The requirements for pseudouridine formation in U4 and U6 snRNAs in HeLa cell nuclear extract and S100 extract were also investigated *in vitro*. For U4 snRNA, optimal pseudouridine formation was observed when U4 RNA transcripts were first incubated in S100 extract, followed by addition of nuclear extract. The presence of an intact Sm binding site was also found to strongly increase the efficiency of pseudouridine formation (Zerby and Patton, 1996). *In vitro* formation of pseudouridines in U6 snRNA transcripts is optimal with incubation in nuclear extract, followed by the addition of S100 extract. This is surprising because U6 snRNA does not exit the nucleus *in vivo*. Interestingly, pseudouridine formation in both U4 and U6 snRNA transcripts was found to be highly increased on formation of a U4-U6 snRNA duplex (Fig. 9B) (Zerby and Patton, 1996, 1997). Two distinct U4 pseudouridine synthase activities seem to be required for U4 full modification (Zerby and Patton, 1997), and U4 snRNA with a truncated Sm binding site was found to be less efficient in increasing pseudouridine formation in U6 snRNA (Zerby and Patton, 1996). Among the three pseudouridines of the human U6 snRNA (positions 31, 40, and 86), only pseudouridine residues at positions 40 and 86 could be formed *in vitro* (Zerby and Patton, 1996). Interestingly, in addition to the heterologous stems I and II, proposed by Brow and Guthrie (1988), a third heterologous stem can be formed between U4 and U6 snRNAs, and this possibility is phylogenetically conserved (Fig. 9B) (Jakab et al., 1997). Residue 31 belongs to the U6 snRNA segment involved in the formation of this third putative heterologous stem, whereas residue 40 is located at one of its extremities.

Figure 9. Secondary structures of rat spliceosomal U4 and U6 snRNAs and U4-U6 and U4atac-U6atac snRNA interactions in vertebrates. (A) The secondary structures are from Krol et al. (1981a) and Myslinski and Branlant (1991) (U4) and Mouglin et al. (unpublished results) (U6). (B) U4-U6 snRNA duplex. The nucleotide sequences shown are from rat (Reddy, 1988). The heterologous stems I and II are as proposed by Brow and Guthrie (1988), and the additional stem III is as proposed by Jakab et al. (1997). (C) U4atac-U6atac snRNA duplex. The nucleotide sequences are from HeLa cells as determined by Tarn and Steitz (1996b). The heterologous stems I and II were proposed by Tarn and Steitz (1996b) based on the model proposed by Brow and Guthrie (1988). An additional potential stem III is shown as proposed by Jakab et al. (1997). Pseudouridines were investigated in both RNAs; only one residue (pseudouridine 14 in U4atac) was detected (Massenet et al., unpublished results). The posttranscriptional modifications are squared in black.

Residue 86 is located in a single-stranded region at the 3' extremity of the molecule (Fig. 9B). Hence, difference of conformation of the substrate-RNA segment may explain the different levels of modifications observed in vitro at positions 31, 40, and 86 of U6 snRNA. U4 snRNA is considered to be a modulator of U6 snRNA activity by formation of the U4-U6 RNA duplex that masks U6 snRNA activity. U4 snRNA may also play a positive role for U6 snRNA catalytic activity, by allowing pseudouridine formation in this molecule (Zerby and Patton, 1996). Hence, according to these data, after transcription, U4 snRNA exits the nucleus and is assembled into a core ribonucleoprotein particle with its 5' cap hypermethylated before returning to the nucleus (Fig. 6). Because base pair interaction between U4 and U6 snRNAs allows pseudouridine formation in both U4 and U6 snRNAs, it suggests that modifications in both snRNAs might occur simultaneously, probably in the nuclear compartment.

Nevertheless, based on the present knowledge, there is no clear idea of the compartment where internal posttranscriptional modifications of the spliceosomal U snRNAs take place. As suggested by the efficiency of S100 extracts for producing some of the pseudouridines, they are probably initiated in the cytoplasm after core-U snRNP formation and completed in the nucleus after the nuclear import of the core-U snRNP particles, U6 and U4 snRNAs being modified after their association.

Possible Implication of Posttranscriptional Modifications in Spliceosome Assembly and Function

Interestingly, the posttranscriptional modifications identified in spliceosomal RNAs are mainly concentrated in the U snRNA segments that were shown to be involved in intermolecular RNA-RNA interactions or in U snRNA segments showing two alternative secondary structures.

In U1 snRNA, except for the 2'-O-methyladenosine at position 70, all the other posttranscriptional modifications are located at the 5' end of the molecule (Fig. 8). The two pseudouridines at positions 5 and 6, which are conserved in vertebrates, *Drosophila*, and *S. cerevisiae*, are involved in the base pair interaction with the 5' extremity of the intron (Fig. 10A). Nucleotides at positions 5 and 6 are also modified in plants ($\Psi_{m_5}U_{m_6}$ in the alga *Chlorella saccharophila*) (Kiss et al., 1988a). In *S. pombe*, only one pseudouridine is detected at position 3 (Table 1). The 2'-O methylated residue 70 belongs to the binding site of protein U1A (Scherly et al., 1989; Bach et al., 1990; Oubridge et al., 1994) and it may play a role

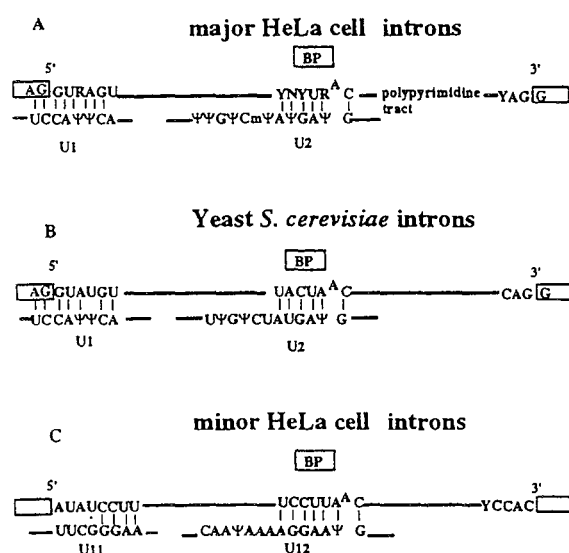


Figure 10. The 5' and 3' splice sites and branch-site consensus sequences of HeLa cells and *S. cerevisiae* introns and their interactions with the spliceosomal U snRNAs. (A) Consensus sequences of the major introns from vertebrates; (B) consensus sequences of the yeast introns; (C) consensus sequences of the minor introns from vertebrates. The complementary sequences of U1 and U11 for 5' splice sites and U2 and U12 for the branch-point sequences (BP) are shown.

in RNA-protein interaction. U1 snRNA transcripts produced in vitro, could complement a HeLa nuclear extract depleted in U1 snRNAs (Will et al., 1996). This means that 2'-O methylations are not essential for U1 snRNA function. No conclusion can be drawn on the role of pseudouridines in the human U1 snRNA function in splicing because no investigation on the in vitro formation of pseudouridines in U1 snRNA was performed.

A huge number of posttranscriptional modifications are gathered in the 5'-terminal third of U2 snRNA. This region was found to contain a pseudoknot structure (Zavanelli et al., 1994) (Fig. 8). Interestingly, the sequence involved in where this pseudoknot structure is located contains posttranscriptional modifications: two 2'-O methylations and three pseudouridines (Fig. 8). Here, posttranscriptional modifications may be required for stabilization of the pseudoknot structure.

Upon association with the intron, U2 snRNA interacts with the intron branch-site sequence. As found for the 5' segment of U1 snRNA, in vertebrates the segment of U2 snRNA that interacts with the branch-site sequence contains two pseudouridines (Fig. 10A). They are conserved in plants and in *S. pombe* (Table 1). However, only one of them is present in *S. cerevisiae* (Fig. 10B). Within the base pair interaction formed by U2 snRNA and the branch-site sequence, the adenosine residue that makes the first nucleo-

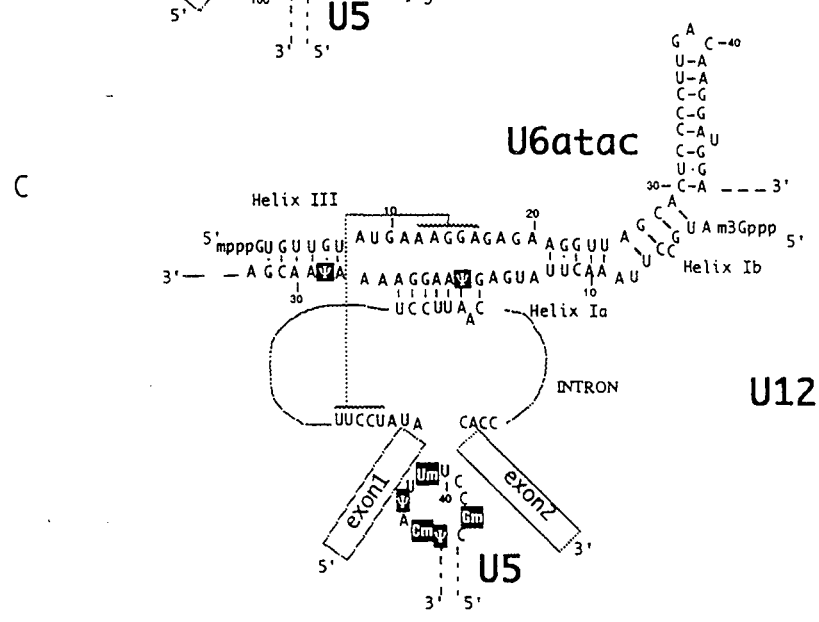
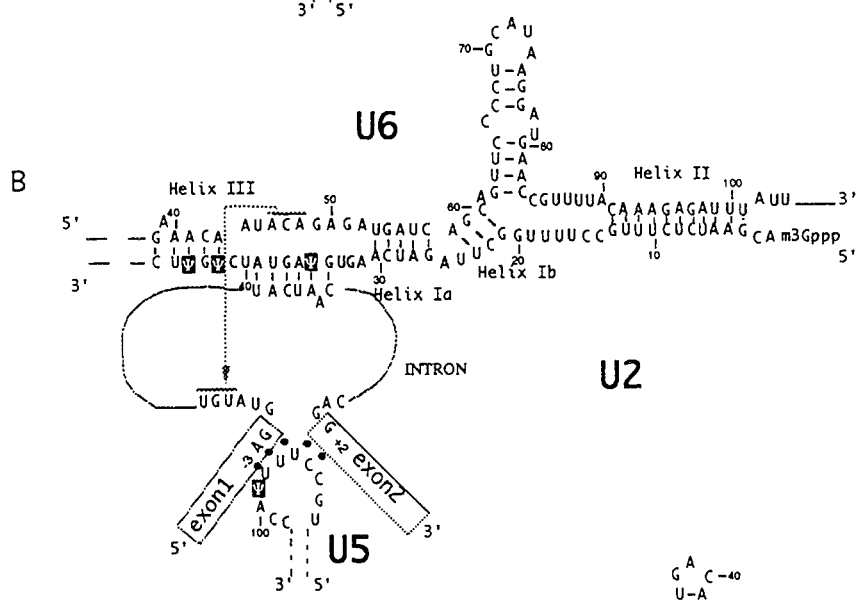
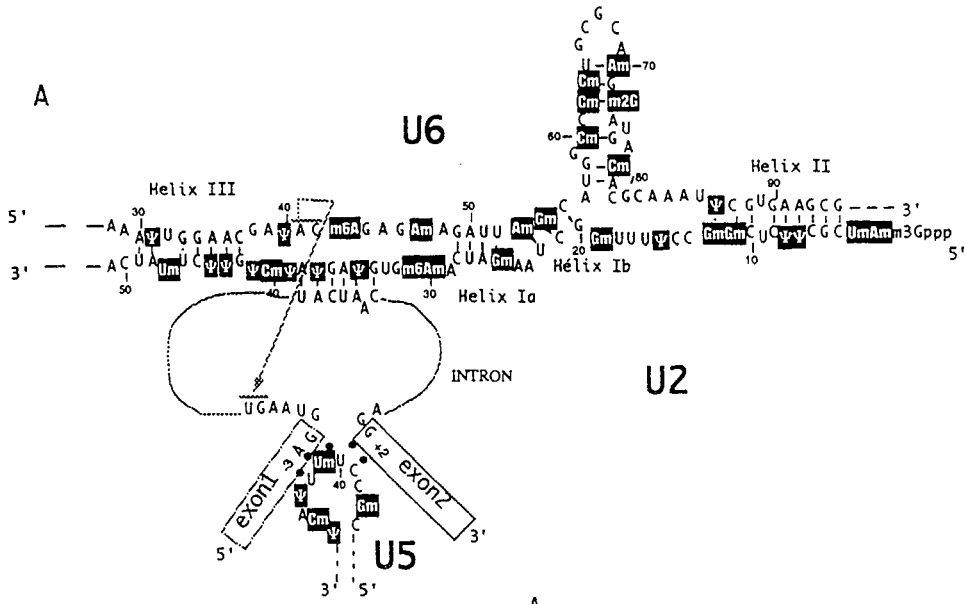
philic attack at the 5' end of the intron is a bulged nucleotide. Interestingly, the unique pseudouridine conserved in *S. cerevisiae* is involved in the formation of the base pair located on the 5' side of the bulged adenosine residue. It may be required for the stabilization of the interaction or for its recognition by proteins. The branch-site sequence is highly conserved in *S. cerevisiae*, in contrast to the situation in vertebrates (for reviews, see Jackson, 1991, and Raymond and Rosbash, 1992). Hence, the presence of a second pseudouridine residue in the vertebrate U2 snRNA sequence that base pairs with the branch-site sequence may be required to compensate for the lower level of complementarity with the introns.

After association of the (U4/U6, U5) tri-snRNP particle with prespliceosomal complexes, U2 snRNA undergoes a strong conformational change and forms several base pair interactions with U6 snRNA (Fig. 11). The ability of U2 and U6 snRNAs to form helices Ia, II, and III was found to be required for splicing in vertebrates (Fig. 11A). However, it is not known whether all of these intermolecular interactions take place together. Interestingly, the entire region involved in the formation of the four intermolecular interactions with U6 snRNA is highly posttranscriptionally modified in the vertebrate and plant U2 snRNAs (Fig. 11A; Table 1). We only detected three pseudouridines in the *S. cerevisiae* U2 snRNA: the one in the segment complementary to the branch-site sequence that we mentioned above and two others in the segment that is involved in helix III formation (Fig. 11B). No strict requirement for the formation of helix III was found in *S. cerevisiae* (Yan and Ares, 1996). However, the U2 snRNA region downstream from the branch-site complementary sequence was found to be the binding site of several splicing factors both in yeast (Ruby et al., 1993; Hodges and Beggs, 1994; Wells and Ares, 1994; Wells et al., 1996; Yan and Ares, 1996) and in vertebrates (Behrens et al., 1993). Hence, the two highly conserved pseudouridines in this area may either facilitate the formation of helix III or be involved in RNA-protein interaction. In yeast, mutation at position 35, corresponding to the pseudouridine residue in the sequence complementary to the branch site, strongly reduced splicing efficiency (McPheeters and Abelson, 1992). However, because this destroys the base pair, one cannot conclude whether the presence of a pseudouridine is absolutely required at this position. Mutation at position 42 had no phenotype (Yan and Ares, 1996), which indicates that this pseudouridine is dispensable in yeast. In contrast, mutation at position 44 strongly affects growth, and it was proposed that the presence of a uridine at this position (in fact, a pseudouridine according to our unpublished results) is required to

avoid the formation of an artifactual nonfunctional U2 snRNA secondary structure (Yan and Ares, 1996). This hypothesis is not exclusive of the involvement of pseudouridine 44 in RNA-protein interactions or helix III formation.

The importance of the posttranscriptional modifications of the human U2 snRNA for spliceosome activity was clearly demonstrated by reconstitution of the spliceosome with HeLa cell nuclear extracts depleted of U2 snRNA and complemented with in vitro transcribed U2 snRNA. Whereas active spliceosomes were obtained when the same kinds of experiments were performed with U5 snRNA transcripts, no active spliceosome could be obtained with U2 snRNA transcripts (Ségault et al., 1995). Active spliceosomes were obtained when the same kinds of experiments were performed in yeast (McPheeters et al., 1989). This confirmed the lower importance of the U snRNA posttranscriptional modifications for the yeast splicing machinery and this shows that some of the posttranscriptional modifications, which are not formed upon incubation of an in vitro transcribed human U2 snRNA in a HeLa nuclear extract, are required for the formation of efficient spliceosomes.

In vertebrates, the U6 snRNA regions involved in base pair interactions with U2 snRNA are also highly modified (Fig. 11A). One highly phylogenetically conserved U6 snRNA sequence (the ACAGAG sequence) is essential for the function of U6 snRNA in splicing (Fabrizio and Abelson, 1990 and 1992; Madhani et al., 1990; Madhani and Guthrie, 1992). Interestingly, this sequence contains an m⁶ residue that was found to be conserved in all the species studied: vertebrates, plants, and *S. pombe* (Table 1). It will be interesting to see whether it is also conserved in *S. cerevisiae*. As U6 snRNA is expected to play a direct role in catalysis with a possible implication of the ACAGAG sequence, the m⁶ residue of the ACAGAG sequence may be important for the splicing reaction. However, this residue is certainly not essential, because active spliceosomes could be obtained upon substitution for the authentic U6 snRNA by an in vitro produced transcript in the HeLa, *Xenopus*, and yeast systems (Fabrizio et al., 1989; Bindereif et al., 1990; Vankan et al., 1990; Wolff and Bindereif, 1995). In vitro, in the absence of S-adenosyl methionine (the expected methyl donor), no m⁶A formation is expected. It may be noticed that methylation at position 6 of adenine in U6 snRNA occurs in a sequence that fits the consensus sequence of messenger RNA m⁶A methylation sites (see Chapter 10). Another part of U6 snRNA, which was found to be very important for splicing, is the segment 57 to 78 (Wolff and Bindereif, 1993 and 1995; Fortner et al., 1994; Sun and Manley, 1995 and 1997; MCPheeters,



1996), that forms a stem-loop structure interrupting the U2–U6 interactions. Interestingly, this segment is also hypermodified in vertebrates (Fig. 11A), and mutation of some of the modified residues was found to impair splicing (Wolff and Bindereif, 1995; Sun and Manley, 1997). However, in the present stage of the study, there is no direct evidence that the effect is linked to the absence of posttranscriptional modifications. Before the interaction with U2 snRNA, the U6 snRNA segment 59 to 72 interacts with U4 snRNA and forms the heterologous stem II of the U4–U6 snRNA duplex (Brow and Guthrie, 1988) (Fig. 9B). As previously mentioned, in addition to the previously proposed heterologous stems I and II, a third intermolecular stem III can be formed (Jakab et al., 1997). U4 snRNA contains a low level of posttranscriptional modifications, even in vertebrates. Only three pseudouridines were identified. All three are contained in segments that can form intermolecular interactions with U6: stem II and stem III.

U5 snRNA carries two types of posttranscriptional modifications: species-specific pseudouridines in the stem of stem-loop structure I (Fig. 7), and highly phylogenetically conserved pseudouridines and 2'-O methylated residues in the terminal loop of stem-loop structure I (Fig. 7). In *S. pombe* and *S. cerevisiae*, this terminal loop I shows one difference as compared to other species: the pseudouridine at the 3' end of the terminal loop is replaced by a C residue (Fig. 7). In addition, no 2'-O methylation was detected in the terminal loop I of the *S. cerevisiae* U5 snRNA (our unpublished results); therefore, this loop contains a single modified residue. Hence, except for *S. cerevisiae*, the terminal loop of the U5 snRNA stem-loop structure I, which aligns the exon extremities for correct ligation (Newman and Norman, 1991 and 1992; Wyatt et al., 1992; Sontheimer and Steitz, 1993; Cortes et al., 1993; Newman et al., 1995) is highly posttranscriptionally modified, and modifications are highly conserved throughout evolution, probably because they ensure a precise conformation of this loop. Such peculiar structure may help U5 snRNA to bridge the exon extremities together with proteins. In yeast, both U5 snRNA and the internal structure of the intron may be involved

in bridging the exon extremities. Indeed, constraints on the intron secondary structure were specifically observed in *S. cerevisiae*, and it was proposed that the observed base pair interactions are required to limit the distance between the 5' and 3' splice sites (Goguel and Rosbash, 1993; Libri et al., 1995; Mouglin et al., 1996; Charpentier and Rosbash, 1996). This may explain a lower pressure for the conservation of posttranscriptional modifications in the *S. cerevisiae* U5 snRNA terminal loop I. However, yeast spliceosome reconstitution with a synthetic U5 snRNA bearing a deleted or a mutated terminal loop I was found to block the second step of splicing (O'Keefe et al., 1996), whereas splicing could be obtained when the same kind of experiment was performed in the HeLa cell system (Ségault et al., unpublished data). Based on these two observations, a conclusion in contradiction with the preceding one can be drawn: the yeast splicing reaction seems to be more dependent on the integrity of the U5 snRNA terminal loop I than on the vertebrate spliceosomes. However, it should be pointed out that the pre-messenger RNA used for the assay in the HeLa cell nuclear extract is very efficiently spliced *in vitro* and it might be that splicing dependence upon the structure of the U5 snRNA terminal loop I is greater for other pre-mRNAs. On the other hand, the terminal loop of U5 snRNA may have additional functions in the formation of active spliceosomes. Indeed, psoralen was recently found to produce a U5–U1 snRNA cross-link before the first step of splicing and, interestingly, the cross-linked residue of U5 snRNA was pseudouridine 43 (Ast and Weiner, 1997). Based on these data, U5 snRNA was proposed to be involved in the process of replacement of U1 snRNA by U6 snRNA in the interaction with the intron 5'-extremity. The facts that pseudouridines located in the stem I of the human U5 snRNA are species specific, whereas the pseudouridines located in the terminal loop I are phylogenetically conserved, are in accord with the observation that these two kinds of pseudouridines are formed by different pseudouridine synthases in human cells (Patton, 1991 and 1994b). Furthermore, the possible formation of these pseudouridines in nuclear cell extract may explain the functionality of in

Figure 11. U snRNA–U snRNA and U snRNA–pre-mRNA interactions at the catalytic center of the spliceosomes. (A) Major spliceosome of vertebrates; (B) yeast spliceosome. In panel A, the nucleotide sequences are from rat. The heterologous helices I, II, and III between U2 and U6 snRNAs are shown, as well as the base-pairing interaction between U2 snRNA and the branch-point sequence, and the interaction between U6 snRNA and the sequence close to the 5'-splice site. Interactions between the terminal loop of U5 snRNA and the exon extremities are also represented (for a review, see Moore et al., 1993, and text for references). (C) Minor spliceosome of vertebrates. The nucleotide sequences are from HeLa cells (Montzka and Steitz, 1988; Tarn and Steitz, 1996b). The predicted interactions (Tarn and Steitz, 1996b) between U12 and U6atac snRNAs are shown. Potential interactions deduced from comparisons with the major spliceosome are also shown: the potential interaction between U6atac snRNA and the sequence at the 5' end of the intron and the potential interaction of U5 snRNA with the exon extremities are shown. The posttranscriptional modifications are shown in black squares.

vitro transcribed U5 snRNA in an in vitro splicing reaction (Ségault et al., 1995).

In conclusion, as shown in Fig. 11, all the U snRNA segments that are expected to be at or near the spliceosome catalytic center are highly modified, and most of the modifications are conserved throughout evolution. These observations are strongly in favor of an important role of the internal posttranscriptional modifications of U snRNAs in splicing. In contrast, the 5'-terminal cap structure is probably not involved in the splicing reaction because in vitro transcribed U1 snRNA was found to be active in splicing when it carries an ApppAm cap structure instead of the m₃GpppAm cap structure (Will et al., 1996). A m⁷GpppAm cap structure was found to impair U1 snRNA function probably because it binds the CBP protein complex (Will et al., 1996).

POSTTRANSCRIPTIONAL MODIFICATIONS OF THE SPLICEOSOMAL U snRNAs INVOLVED IN THE SPLICING OF THE AT-AC INTRONS

Terminal and Internal Posttranscriptional Modifications

As previously mentioned, a new class of introns recently have been characterized, and they were found to be spliced in a spliceosome structure gathering around the intron, the U11, U12, U4atac, U6atac, and U5 snRNAs. U11 and U12 (Fig. 10) have been characterized previously (Montzka and Steitz, 1988) as minor nucleoplasmic U snRNAs. They were shown to bear a m^{2,2,7}G cap structure, together with an Sm-binding site (Montzka and Steitz, 1988), which suggests that they are processed in the cytoplasm as U1 and U2 snRNAs, their counterparts of the main splicing machinery.

U4atac and U6atac snRNAs (Fig. 9C) were recently characterized and, as was found for the U4 and U6 snRNAs of the main splicing machinery, U4atac snRNA has a m^{2,2,7}G cap structure and an Sm-binding site, whereas U6atac snRNA, which is transcribed by RNA polymerase III, has a γ -methyl triphosphate at its 5'-extremity. Hence, with respect to the 5'-terminal modifications, the situation is completely identical for the two classes of spliceosomal snRNAs.

We looked for the presence of pseudouridines in the AT-AC spliceosomal RNAs. The results obtained show some similarity with the *S. cerevisiae* GT-AG spliceosomal RNAs. U11 does not contain any pseudouridine; U12 contains only two pseudouridines (Fig. 10; Table 1). As found for the *S. cerevisiae* U2 snRNA, one pseudouridine residue is involved in the interaction with the branch-site sequence and is lo-

cated in the base pair on the 5' side of the bulged A residue. The second one is located in the segment upstream of the branch-site complementary sequence, which is involved in formation of the putative heterologous helix III of the U12-U6atac interaction (Tarn and Steitz, 1996b) (Fig. 11C). No pseudouridine residue was detected in the U6atac snRNA, and only one was found in the U4atac snRNA (Fig. 9C). It belongs to the helix II of the U4-U6atac RNA duplex. Analysis of the 2'-O methylations is underway in our laboratory.

Possible Implications of the AT-AC U snRNAs Posttranscriptional Modifications for Splicing of AT-AC Introns

The 5'-terminal modifications of the AT-AC U snRNAs probably play the same role in AT-AC U snRNP biogenesis as those of the GT-AG U snRNPs.

Concerning the internal posttranscriptional modifications, the low level of pseudouridine residues observed may be explained by the strong conservation of the 5'-terminal segment of the intron that base pairs with U11 snRNA and of the branch-site sequence that base-pairs with U12 snRNAs (Jackson, 1991; Hall and Padgett, 1994; Wu and Krainer, 1997). An optimal complementarity with U12 snRNA is found for all the introns detected in metazoan cells (Tarn et al., 1995). Hence, as it is the case for the *S. cerevisiae* GT-AG machinery, there is no need to reinforce the U snRNAs-pre-mRNA interactions by several posttranscriptional modifications. Only the pseudouridine residue that forms the base pair on the 5' side of the bulged branched site is conserved. This strongly suggests that this pseudouridine residue is essential in both splicing machineries, as is the case for the pseudouridine residue located a few nucleotides downstream from the branch-site complementary sequence. The ACm⁶AGAG sequence, that is essential for the GT-AG U6 snRNA function, is replaced by GGAGAG sequence in the U6atac snRNA. It will be interesting to test whether the first adenosine residue bears a methylation on position 6 in the U6atac snRNA. Several of the residues that carry a 2'-O methylation in the GT-AG U6 snRNA are conserved in the AT-AC U6 snRNA; it will be interesting to see whether they are also 2'-O methylated. Experiments are under way in our laboratory to test this possibility.

CONCLUSIONS

It has now been clearly demonstrated that 5'-terminal posttranscriptional modifications play an es-

essential role for the biogenesis of the spliceosomal U snRNPs in eukaryotes, and this is probably also the case for the U snRNAs involved in the recently discovered minor splicing machinery of vertebrates. The role of the internal modifications of the spliceosomal U snRNAs from the major splicing machinery is not yet clear. A trivial explanation could be that the U snRNA sequences involved in intermolecular RNA-RNA base pair interactions are free of protein interactions and freely accessible for the numerous posttranscriptional modifying enzymes present in the cytoplasm and the nucleus. If this was the only explanation, the strongly observed conservation of several of the modifications, throughout evolution, would not have been expected. Furthermore, the atac U snRNAs would carry a similar number of posttranscriptional modifications. We showed that this is not the case.

Another trivial explanation could be that posttranscriptional modifications of U snRNA segments involved in intermolecular base pair interactions are needed to confer a resistance of these RNA segments to the nucleases of the nuclear compartment. Although this explanation cannot be completely ruled out, on one hand, the nucleus is not known to contain high levels of nucleases; on the other hand, atac U snRNAs would also need to be protected by posttranscriptional modifications. Finally, the various recent experiments that we previously summarized are strongly in favor of a direct functional role of U snRNA posttranscriptional modifications in spliceosome assembly and splicing.

Although clear proof is still lacking, the essential role of human U2 snRNA internal posttranscriptional modifications in splicing was already shown (Ségault et al., 1995). It should be pointed out that, taken individually, the posttranscriptional modifications of vertebrate U snRNAs are perhaps not essential for splicing, but together they may have a positive effect on splicing, especially for introns that deviate from the consensus sequences.

The current view is that the pre-mRNA splicing machinery has evolved from a common ancestor to the autocatalytic introns of group II (Sharp, 1985; Cech, 1986). According to this theory, some of the DNA segments coding for autocatalytic intron sequences would have dissociated to generate the spliceosomal U snRNA genes (Sharp, 1991). Arguments in favor of this hypothesis are the following. First, splicing of the autocatalytic group II introns and of the pre-mRNA introns proceeds through the same two chemical steps (for a review, see Moore et al., 1993). Second, some of the secondary structure elements formed by U snRNAs and the pre-mRNA at the catalytic center of the spliceosome share strong

homology with secondary structure elements of introns of group II (Wu and Manley, 1989; Jacquier, 1990; Newman and Norman, 1992; for a review, see Madhani and Guthrie, 1994). Based on its high dependence on the pre-mRNA secondary structure, the nuclear pre-mRNA splicing machinery of *S. cerevisiae* seems to be closer to the autocatalytic intron ancestor of the spliceosome than the nuclear pre-mRNA splicing machinery of vertebrates. Introns have probably played a very important role in the evolution of the genomes of pluricellular eukaryotes that developed complex cell differentiation processes. It seems likely that for an easier evolution of the genome, the dependence on intron primary and secondary structure, which probably existed in the ancestral pre-mRNA introns and whose vestige may be the yeast introns, has been decreased in the course of evolution. One may imagine that the multiplication of internal posttranscriptional modifications in some of the U snRNAs participated in this evolution of introns in pluricellular eukaryotes. The participation of proteins, in particular the SR proteins (for a review, see Manley and Tacke, 1996), may also have contributed. Introns are rare in yeast. They probably did not contribute strongly in the evolution of the yeast genomes, and this may explain the presence of pre-messenger introns that may be more similar to the ancestral ones than those in vertebrates, with as corollary a more limited number of U snRNA internal modifications in yeasts.

The recent discovery of a new category of intron in vertebrates, spliced by a minor form of spliceosome, opens new questions concerning the phylogeny of pre-mRNA introns. Interestingly, the pattern of pseudouridine formations in the atac U snRNAs has similarity to that of *S. cerevisiae*. However, the functional sequences of these introns that are recognized by the atac splicing machinery are very different from the ones recognized by the main splicing machinery of yeast and vertebrates. Hence, the similarity of posttranscriptional modifications in the yeast *S. cerevisiae* U snRNAs and in the atac U snRNAs may result from convergent evolution rather than from the existence of a common ancestor.

Little is known about the internal posttranscriptional modifications in the U snoRNAs. However, clearly for U3 snoRNA, as is the case for spliceosomal U snRNAs, posttranscriptional modifications are found in a segment involved in an intermolecular RNA-RNA interactions.

Acknowledgments. Work in the authors' laboratory is supported by the Centre National de la Recherche Scientifique and the French Ministère de la Recherche et de l'Enseignement Supérieur. S. Massener is a fellow of the French Ministère de la Recherche et de l'Enseignement Supérieur. We thank H. Grosjean, R. Benne, J.-P.

Bachelier, and S. Jacquenet for critical reading of the manuscript and useful comments.

REFERENCES

- Adams, D. S., D. Noonan, T. C. Burn, and H. B. Skinner. 1987. A library of trimethylguanosine-capped small RNAs in *Physarum polycephalum*. *Gene* 54:93-103.
- Agris, P. F. 1996. The importance of being modified: roles of modified nucleosides and Mg²⁺ in RNA structure and function. *Prog. Nucleic Acid Res. Mol. Biol.* 53:79-129.
- Amberg, D. C., M. Fleischmann, I. Stagljar, C. N. Cole, and M. Aebi. 1993. Nuclear PRP20 protein is required for mRNA export. *EMBO J.* 12:233-241.
- Antal et al. Personal communication.
- Ast, G., and A. M. Weiner. 1997. A novel U1/U5 interaction indicates proximity between U1 and U5 snRNAs during an early step on mRNA splicing. *RNA* 3:371-381.
- Bach, M., A. Krol, and R. Lührmann. 1990. Structure-probing of U1 snRNPs gradually depleted of the U1-specific proteins A, C, and 70k. Evidence that A interacts differentially with developmentally regulated mouse U1 snRNA variants. *Nucleic Acids Res.* 18:449-457.
- Bachelier, J. P., B. Michot, M. Nicoloso, A. Balakin, J. Ni, and M. J. Fournier. 1995. Antisense snoRNAs: a family of nucleolar RNAs with long complementarities to rRNA. *Trends Biochem. Sci.* 20:261-264.
- Bakin, A., and J. Ofengand. 1993. Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique. *Biochemistry* 32:9754-9762.
- Balakin, A. G., G. S. Schneider, M. S. Corbett, J. Ni, and M. J. Fournier. 1993. SnR31, snR32, and snR33: three novel, non-essential snRNAs from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 21:5391-5397.
- Balakin, A. G., L. Smith, and M. J. Fournier. 1996. The RNA world of the nucleolus: two major families of small RNAs defined by different elements with related functions. *Cell* 86:823-834.
- Bally, M., J. Hughes, and G. Cesareni. 1988. SnR30: a new, essential small nuclear RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 16:5291-5303.
- Baserga, S. J., M. Gilmore-Hebert, and X. W. Yang. 1992. Distinct molecular signals for nuclear import of the nucleolar snRNA U3. *Genes Dev.* 6:1120-1130.
- Behrens, S. E., and R. Lührmann. 1991. Immunoaffinity purification of a [U4/U6.U5] tri-snRNP from human cells. *Genes Dev.* 5:1439-1452.
- Behrens, S. E., K. Tyc, B. Kastner, J. Reichelt, and R. Lührmann. 1993. Small nuclear ribonucleoprotein (RNP) U2 contains numerous additional proteins and has a bipartite RNP structure under splicing conditions. *Mol. Cell. Biol.* 13:307-319.
- Belanger, K. D., M. A. Kenna, S. Wei, and L. I. Davis. 1994. Genetic and physical interactions between Srp1p and nuclear pore complex proteins Nup1p and Nup2p. *J. Cell Biol.* 126:619-630.
- Belhumeur, P., A. Lee, R. Tam, T. Di Paolo, N. Fortin, and M. W. Clark. 1993. GSP1 and GSP2, genetic suppressors of the prp20-1 mutant in *Saccharomyces cerevisiae*: GTP-binding proteins involved in the maintenance of nuclear organization. *Mol. Cell. Biol.* 13:2152-2161.
- Beltrame, M., Y. Henry, and D. Tollervy. 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.* 22:4057-4065. (Corrected and republished in *Nucleic Acids Res.*, 22:5139-5147.)
- Bindereif, A., T. Wolff, and M. R. Green. 1990. Discrete domains of human U6 snRNA required for the assembly of U4/U6 snRNP and splicing complexes. *EMBO J.* 9:251-255.
- Bischoff, F. R., and H. Ponstingl. 1991. Mitotic regulator protein RCC1 is complexed with a nuclear Ras-related polypeptide. *Proc. Natl. Acad. Sci. USA* 88:10830-10834.
- Bischoff, F. R., H. Krebber, T. Kempf, I. Hermes, and H. Ponstingl. 1995. Human RanGTPase-activating protein RanGAP1 is a homolog of yeast RNA1p involved in messenger RNA processing and transport. *Proc. Natl. Acad. Sci. USA* 92:1749-1753.
- Black, D. L., B. Chabot, and J. A. Steitz. 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. *Cell* 42:737-750.
- Branlant, C., A. Krol, J. P. Ebel, H. Gallinaro, E. Lazar, and M. Jacob. 1981. The conformation of chicken, rat and human U1A RNAs in solution. *Nucleic Acids Res.* 9:841-858.
- Branlant, C., A. Krol, J. P. Ebel, E. Lazar, B. Haendler, and M. Jacob. 1982. U2 RNA shares a structural domain with U1, U4, and U5 RNAs. *EMBO J.* 1:1259-1265.
- Branlant, C., A. Krol, E. Lazar, B. Haendler, M. Jacob, L. Galegoudias, and C. Pousada. 1983. High evolutionary conservation of the secondary structure and of certain nucleotide sequences of U5 RNA. *Nucleic Acids Res.* 11:8359-8367.
- Brow, D. A., and C. Guthrie. 1988. Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature* 334:213-218.
- Brulé, F., J. Venema, V. Ségault, D. Tollervy, and C. Branlant. 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* 2:183-197.
- Busch, H., R. Reddy, L. Rothblum, and Y. C. Choi. 1982. SnRNAs, SnRNPs, and RNA processing. *Annu. Rev. Biochem.* 51:617-654.
- Cech, T. R. 1986. The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell* 44:207-210.
- Chandrasekharappa, S. C., J. H. Smith, and G. L. Eliceiri. 1983. Biosynthesis of small nuclear RNAs in human cells. *J. Cell. Physiol.* 117:169-174.
- Chanfreau, G., S. A. Elela, M. Ares, Jr., and C. Guthrie. 1997. Alternative 3'-end processing of U5 snRNA by RNase III. *Genes Dev.* 11:2741-2751.
- Charpentier, B., and M. Rosbash. 1996. Intramolecular structure in yeast introns aids the early steps in vitro spliceosome assembly. *RNA* 2:509-522.
- Cheng, Y., J. E. Dahlberg, and E. Lund. 1995. Diverse effects of the guanine nucleotide exchange factor RCC1 on RNA transport. *Science* 267:1807-1810.
- Colot, H. V., F. Stutz, and M. Rosbash. 1996. The yeast splicing factor Mud13 is a commitment complex component and corresponds to CBP20, the small subunit of the nuclear cap-binding complex. *Genes Dev.* 10:1699-1708.
- Coppola, J. A., A. S. Feild, and D. S. Luse. 1983. Promoter-proximal pausing by RNA polymerase 2 in vitro: transcripts shorter than 20 nucleotides are not capped. *Proc. Natl. Acad. Sci. USA* 80:1251-1255.
- Cortes, J. J., E. J. Sontheimer, S. D. Seiwert, and J. A. Steitz. 1993. Mutations in the conserved loop of human U5 snRNA generate use of novel cryptic 5' splice sites in vivo. *EMBO J.* 12:5181-5189.
- Dahlberg, J. E., and E. Lund. 1988. The genes and transcription of the major small nuclear RNAs, p. 100-114. In M. L. Birnstiel (ed.), *Structure and Function of the Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer Verlag, Berlin, Germany.
- Datta, B., and A. M. Weiner. 1991. Genetic evidence for base-pairing between U2 and U6 snRNA in mammalian mRNA splicing. *Nature* 352:821-824.

- Eichler, D. C., and N. Craig. 1994. Processing of eukaryotic ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* 49:197-239.
- Fabrizio, P., D. S. McPheeters, and J. Abelson. 1989. *In vitro* assembly of yeast U6 snRNP: a functional assay. *Genes Dev.* 3: 2137-2150.
- Fabrizio, P., and J. Abelson. 1990. Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. *Science* 250:404-409.
- Fabrizio, P., and J. Abelson. 1992. Thiophosphates in yeast U6 snRNA specifically affect pre-mRNA splicing *in vitro*. *Nucleic Acids Res.* 20:3659-3664.
- Fisher, U., and R. Lührmann. 1990. An essential signaling role for the m₃G cap in the transport of U1 snRNP to the nucleus. *Science* 249:786-790.
- Fisher, U., V. Sumpster, M. Sekine, T. Satoh, and R. Lührmann. 1993. Nucleocytoplasmic transport of UsnRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m₃G cap. *EMBO J.* 12: 573-583.
- Fortner, D. M., R. G. Troy, and D. A. Brow. 1994. A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. *Genes Dev.* 8:221-223.
- Goguel, V., and M. Rosbash. 1993. Splice site choice and splicing efficiency are positively influenced by pre-mRNA intramolecular base pairing in yeast. *Cell* 72:893-901.
- Görlich, D., S. Kostka, R. Kraft, C. Dingwall, R. A. Laskey, E. Hartmann, and S. Prehn. 1995a. Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. *Curr. Opin. Cell Biol.* 5: 383-392.
- Görlich, D., F. Vogel, A. D. Mills, E. Hartmann, and R. A. Laskey. 1995b. Distinct functions for the two importin subunits in nuclear protein import. *Nature* 377:246-248.
- Görlich, D., and I. W. Mattaj. 1996. Nucleocytoplasmic transport. *Science* 271:1513-1518.
- Görlich, D., P. Henklein, E. Hartmann, and R. A. Laskey. 1996a. A 41 amino acid motif in importin- α confers binding to importin- β and hence transit into the nucleus. *EMBO J.* 15: 1810-1817.
- Görlich, D., N. Panté, U. Kutay, U. Aebi, and F. R. Bischoff. 1996b. Identification of different roles for RanGDP and RanGTP in nuclear protein import. *EMBO J.* 15:5584-5594.
- Görlich, D., R. Kraft, S. Kostka, F. Vogel, E. Hartmann, R. A. Laskey, I. W. Mattaj, and E. Izaurralde. 1996c. Importin provides a link between nuclear protein import and UsnRNA export. *Cell* 87:21-32.
- Gu, J., J. R. Patton, S. Shimba, and R. Reddy. 1996. Localisation of modified nucleotides in *Saccharomyces pombe* spliceosomal small nuclear RNAs: modified nucleotides are clustered in functionally important regions. *RNA* 2:909-918.
- Hall, S. L., and R. A. Padgett. 1994. Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. *J. Mol. Biol.* 239:357-365.
- Hall, S. L., and R. A. Padgett. 1996. Requirement of U12 snRNA for the *in vivo* splicing of a minor class of eukaryotic nuclear pre-mRNA introns. *Science* 271:1716-1718.
- Hamm, J., E. Darzynkiewicz, S. M. Tahara, and I. W. Mattaj. 1990. The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. *Cell* 62: 569-577.
- Harada, F., N. Kato, and S. Nishimura. 1980. The nucleotide sequence of nuclear 4.8S RNA of mouse cells. *Biochem. Biophys. Res. Comm.* 95:1332-1340.
- Hausner, T. P., L. M. Giglio, and A. M. Weiner. 1990. Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles. *Genes Dev.* 4:2146-2156.
- Hellung-Larsen, P., I. Kulamowicz, and S. Frederiksen. 1980. Synthesis of low molecular weight RNA components in cells with a temperature-sensitive polymerase II. *Biochem. Biophys. Acta* 609:201-204.
- Henry, Y., H. Wood, J. P. Morrissey, E. Petfalski, S. Kearsey, and D. Tollervey. 1994. The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. *EMBO J.* 13: 2452-2463.
- Hodges, P. E., and J. D. Beggs. 1994. RNA splicing. U2 fulfils a commitment. *Curr. Opin. Cell Biol.* 4:264-267.
- Huang, Q., M. R. Jacobson, and T. Pederson. 1997. 3' processing of human pre-U2 small nuclear RNA: a base-pairing interaction between the 3' extension of the precursor and an internal region. *Mol. Cell. Biol.* 17:7178-7185.
- Huber. Personal communication.
- Hughes, J. M. X., D. A. Konings, and G. Cesareni. 1987. The yeast homologue of U3 snRNA. *EMBO J.* 6:2145-2155.
- Hughes, J. M. X., and M. Ares. 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.* 10:4231-4239.
- Hughes, J. M. X. 1996. Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit. *J. Mol. Biol.* 259:645-654.
- Izaurralde, E., J. Lewis, C. McGuigan, M. Jankowska, E. Darzynkiewicz, and I. W. Mattaj. 1994. A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* 78:657-668.
- Izaurralde, E., J. Lewis, C. Gamberi, A. Jarmolowski, C. McGuigan, and I. W. Mattaj. 1995. A cap-binding protein complex mediating UsnRNA export. *Nature* 376:709-712.
- Jackson, I. J. 1991. A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Res.* 19:3794-3798.
- Jacob, M., E. Lazar, B. Haendler, H. Gallinaro, A. Krol, and C. Branlant. 1984. A family of small nucleoplasmic RNAs with common structural features. *Biol. Cell.* 51:1-9.
- Jacobson, M. R., M. Rhoadhouse, and T. Pederson. 1993. U2 small nuclear RNA 3' end formation is directed by a critical internal structure distinct from the processing site. *Mol. Cell. Biol.* 13:1119-1129.
- Jacquier, A. 1990. Self-splicing group II and nuclear pre-mRNA introns: how similar are they? *Trends Biochem. Sci.* 15:351-354.
- Jakab, G., A. Mougin, M. Kis, T. Pollák, M. Antal, C. Branlant, and F. Solymosy. 1997. *Chlamydomonas* U2, U4 and U6 snRNAs. An evolutionarily conserved putative third interaction between U4 and U6 snRNAs which has a counterpart in the U4atac-U6atac snRNA duplex. *Biochimie* 79:387-395.
- Jakab, G., M. Kis, and F. Solymosy. 1992. Nucleotide sequence of U5 RNA from a green alga, *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* 20:5224.
- Jarmolowski, A., and I. W. Mattaj. 1993. The determinants for Sm protein binding to *Xenopus* U1 and U5 snRNAs are complex and non-identical. *EMBO J.* 12:223-232.
- Jarmolowski, A., W. C. Boelens, E. Izaurralde, and I. W. Mattaj. 1994. Nuclear export of different classes of RNA is mediated by specific factors. *J. Cell Biol.* 124:627-635.
- Kadowaki, T., Y. Zhao, and A. Tartakoff. 1992. Mtr1, a conditional yeast mutant deficient in mRNA transport from nucleus to cytoplasm. *Proc. Natl. Acad. Sci. USA* 89:2312-2316.
- Kadowaki, T., D. Goldfarb, L. M. Spitz, A. M. Tartakoff, and M. Ohno. 1993. Regulation of RNA processing and transport by a nuclear guanine nucleotide release protein and members of the Ras superfamily. *EMBO J.* 12:2929-2937.
- Kandels-Lewis, S., and B. Séraphin. 1993. Role of U6 snRNA in 5' splice site selection. *Science* 262:2035-2039.
- Kass, S., K. Tyc, J. A. Steitz, and B. Sollner-Webb. 1990. The U3 small nucleolar ribonucleoprotein functions in the first step of pre-ribosomal RNA processing. *Cell* 60:897-908.

- Kastner, B., M. Bach, and R. Lührmann. 1990. Electron microscopy of small nuclear ribonucleoprotein (snRNP) particles U2 and U5: evidence for a common structure-determining principle in the major U snRNP family. *Proc. Natl. Acad. Sci. USA* 87: 1710-1714.
- Kato, N., and F. Harada. 1981. Nucleotide sequence of nuclear 5S of mouse cells. *Biochem. Biophys. Res. Commun.* 99: 1468-1476.
- Kato, N., and F. Harada. 1984. Nucleotide sequence of nuclear 5.4S RNA of mouse cells. *Biochem. Biophys. Acta* 782:127-131.
- Keller, E. B., and W. A. Noon. 1984. Intron splicing: a conserved internal signal in introns of animal pre-mRNAs. *Proc. Natl. Acad. Sci. USA* 81:7417-7420.
- Kiss, T., M. Toth, and F. Solymosy. 1985. Plant small nuclear RNAs. Nucleolar U3 snRNA is present in plants: partial characterization. *Eur. J. Biochem.* 152:259-266.
- Kiss, T., M. Antal, H. Hegyi, and F. Solymosy. 1988a. Plant small nuclear RNAs. IV. The structure of U1 RNA from *Chlorella saccharophila*: a phylogenetic support, in terms of RNA structure, for the probable interaction between U1 and U2 RNPs during the splicing of pre-mRNA. *Nucleic Acids Res.* 16:2734.
- Kiss, T., G. Jakab, M. Antal, Z. Palfi, H. Hegyi, M. Kis, and F. Solymosy. 1988b. Plant small nuclear RNAs. V. U4 RNA is present in broad bean plants in the form of sequence variants and is base-paired with U6 RNA. *Nucleic Acids Res.* 16:5407-5426.
- Kiss, T., and F. Solymosy. 1990. Molecular analysis of a U3 RNA gene locus in tomato: transcription signals, the coding region, expression in transgenic tobacco plants and tandemly repeated pseudogenes. *Nucleic Acids Res.* 18:1941-1949.
- Kleinschmidt, A. M., J. R. Patton, and T. Pederson. 1989. U2 small nuclear RNP assembly *in vitro*. *Nucleic Acids Res.* 17: 4817-4828.
- Kohtz, J. D., S. F. Jamison, C. L. Will, P. Zuo, R. Lührmann, M. A. Garcia-Blanco, and J. L. Manley. 1994. Protein-protein interactions and 5' splice site recognition in mammalian mRNA precursors. *Nature* 368:119-124.
- Kolosova, I., and R. A. Padgett. 1997. U11 snRNA interacts *in vivo* with the 5' splice site of U12-dependent (AU-AC) pre-mRNA introns. *RNA* 3:227-233.
- Konarska, M. M., R. A. Padgett, and P. A. Sharp. 1985. Trans-splicing of mRNA precursors *in vitro*. *Cell* 42:165-171.
- Krainer, A. R., T. Maniatis, B. Ruskin, and M. R. Green. 1984. Normal and mutant human beta-globin pre-mRNAs are faithfully and efficiently spliced *in vitro*. *Cell* 36:993-1005.
- Krol, A., C. Branlant, E. Lazar, H. Gallinaro, and M. Jacob. 1981a. Primary and secondary structures of chicken, rat and man nuclear U4 RNAs. Homologies with U1 and U5 RNAs. *Nucleic Acids Res.* 9:2699-2716.
- Krol, A., H. Gallinaro, E. Lazar, M. Jacob, and C. Branlant. 1981b. The nuclear 5S RNAs from chicken, rat and man. U5 RNAs are encoded by multiple genes. *Nucleic Acids Res.* 9: 769-787.
- Krol, A., J. P. Ebel, J. Rinke, and R. Lührmann. 1983. U1, U2 and U5 small nuclear RNAs are found in plant cells. Complete nucleotide sequence of the U5 RNA family from pea nuclei. *Nucleic Acids Res.* 11:8583-8594.
- Krol, A., P. Carbon, J. P. Ebel, and B. Appel. 1987. *Xenopus tropicalis* U6 snRNA genes transcribed by Pol III contain the upstream promoter elements used by Pol II dependent UsnRNA genes. *Nucleic Acids Res.* 15:2463-2478.
- Kunkel, G. R., R. L. Maser, J. P. Calvet, and T. Pederson. 1986. U6 small nuclear RNA is transcribed by RNA polymerase III. *Proc. Natl. Acad. Sci. USA* 83:8575-8579.
- Leader, D. J., S. Connely, W. Filipowicz, and J. W. Brown. 1994. Characterization and expression of a maize U3 snRNA gene. *Biochim. Biophys. Acta* 1219:145-147.
- Lee, M. S., M. Henry, and S. A. Silver. 1996. A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export. *Genes Dev.* 10:1233-1246.
- Lerner, M. R., and J. A. Steitz. 1979. Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proc. Natl. Acad. Sci. USA* 76: 5495-5499.
- Lerner, M. R., J. A. Boyle, S. M. Mount, S. L. Wolin, and J. A. Steitz. 1980. Are snRNPs involved in splicing? *Nature* 283: 220-224.
- Lesser, C. F., and C. Guthrie. 1993. Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science* 262:1982-1988.
- Li, H. V., J. Zagorski, and M. J. Fournier. 1990. Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 10:1145-1152.
- Liautard, J. P., J. Sri-Widada, C. Brunel, and P. Jeanteur. 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.* 162:623-643.
- Libri, D., F. Stutz, T. McCarthy, and M. Rosbash. 1995. RNA structural patterns and splicing: molecular basis for an RNA-based enhancer. *RNA* 1:425-436.
- Maden, B. E. 1990. The numerous modified nucleotides in eukaryotic ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* 39: 241-303.
- Madhani, H. D., R. Bordonne, and C. Guthrie. 1990. Multiple roles for U6 snRNA in the splicing pathway. *Genes Dev.* 4: 2264-2277.
- Madhani, H. D., and C. Guthrie. 1992. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* 71:803-817.
- Madhani, H. D., and C. Guthrie. 1994. Dynamic RNA-RNA interactions in the spliceosome. *Annu. Rev. Genet.* 28:1-26.
- Manley, J. L., and R. Tacke. 1996. SR proteins and splicing control. *Genes Dev.* 10:1569-1579.
- Marshallsay, C., T. Kiss, and W. Filipowicz. 1990. Amplification of plant U3 and U6 snRNA gene sequences using primers specific for an upstream promoter element and conserved intragenic regions. *Nucleic Acids Res.* 18:3459-3466.
- Marshallsay, C., and R. Lührmann. 1994. *In vitro* nuclear import of snRNPs: cytosolic factors mediate m⁷G-cap dependence of U1 and U2 snRNPs transport. *EMBO J* 13:222-231.
- Massenet S., et al. Unpublished results.
- Mattaj, I. W., and E. M. De Robertis. 1985. Nuclear segregation of U2 snRNA requires binding of specific snRNP proteins. *Cell* 40:111-118.
- Mattaj, I. W. 1986. Cap trimethylation of UsnRNA is cytoplasmic and dependent on UsnRNP protein binding. *Cell* 46:905-911.
- Maxwell, E. S., and M. J. Fournier. 1995. The small nucleolar RNAs. *Annu. Rev. Biochem.* 67:6523-6527.
- McPheeters, D. S., P. Fabrizio, and J. Abelson. 1989. *In vitro* reconstitution of functional yeast U2 snRNPs. *Genes Dev.* 3: 2124-2136.
- McPheeters, D. S., and J. Abelson. 1992. Mutational analysis of the yeast U2 snRNA suggests a structural similarity to the catalytic core of group I introns. *Cell* 71:819-831.
- McPheeters, D. S. 1996. Interactions of the yeast U6 RNA with the pre-mRNA branch site. *RNA* 2:1110-1123.
- Méreau, A., R. Fournier, A. Grégoire, A. Mougou, P. Fabrizio, R. Lührmann, and C. Branlant. 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.* 273:552-571.

- Michael, W. M., M. Choi, and G. Dreyfuss. 1995. A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. *Cell* 83:415-422.
- Montzka, K. A., and J. A. Steitz. 1988. Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. *Proc. Natl. Acad. Sci. USA* 85:8885-8889.
- Moore, M. J., C. C. Query, and P. A. Sharp. 1993. Splicing of precursors to messenger RNAs by the spliceosome, p. 303-358. In R. Gesteland and J. Atkins (eds.), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Moore, M. S., and G. Blobel. 1994. A G protein involved in nucleocytoplasmic transport: the role of Ran. *Trends Biochem. Sci.* 19:211-216.
- Morrissey, J. P., and D. Tollervy. 1993. Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol. Cell. Biol.* 13:2469-2477.
- Mougey, E. B., L. K. Pape, and B. Sollner-Webb. 1993. A U3 small nuclear ribonucleoprotein-requiring processing event in the 5' external transcribed spacer of *Xenopus* precursor rRNA. *Mol. Cell. Biol.* 13:5990-5998.
- Mougin, A., et al. Unpublished results.
- Mougin, A., A. Grégoire, J. Banroques, V. Ségault, R. Fournier, F. Brulé, M. Chevrier-Miller, and C. Branlant. 1996. Secondary structure of the yeast *Saccharomyces cerevisiae* pre-U3A snoRNA and its implication for splicing efficiency. *RNA* 2:1079-1093.
- Mount, S. M., and J. A. Steitz. 1981. Sequence of U1 RNA from *Drosophila melanogaster*: implications for U1 secondary structure and possible involvement in splicing. *Nucleic Acids Res.* 9:6351-6368.
- Mount, S. M., I. Pettersson, M. Hinterberger, A. Karmas, and J. A. Steitz. 1983. The U1 small nuclear RNA-protein complex selectively binds a 5' splice site *in vitro*. *Cell* 33:509-518.
- Myslinski, E., C. Branlant, E. D. Wieben, and T. Pederson. 1984. The small nuclear RNAs of *Drosophila*. *J. Mol. Biol.* 180:927-945.
- Myslinski, E., V. Ségault and C. Branlant. 1990. An intron in the genes for U3 small nucleolar RNAs of the yeast *Saccharomyces cerevisiae*. *Science* 247:1213-1216.
- Myslinski, E., and C. Branlant. 1991. A phylogenetic study of U4 snRNA reveals the existence of an evolutionarily conserved secondary structure corresponding to "free" U4 snRNA. *Biochimie.* 73:17-28.
- Newman, A. J., and C. Norman. 1991. Mutations in yeast U5 snRNA alter the specificity of 5' splice site cleavage. *Cell* 65:115-123.
- Newman, A. J., and C. Norman. 1992. U5 snRNA interacts with exon sequences at 5' and 3' splice sites. *Cell* 68:743-754.
- Newman, A. J., S. Teigelkamp, and J. D. Beggs. 1995. SnRNA interactions at 5' and 3' splice sites monitored by photoactivated crosslinking in yeast spliceosomes. *RNA* 1:968-980.
- Ofengand, J., A. Bakin, J. Wrzesinski, K. Nurse, and B. G. Lane. 1995. The pseudouridine residues of ribosomal RNA. *Biochem. Cell Biol.* 73:915-924.
- Ohno, M., N. Kataoka, and Y. Shimura. 1990. A nuclear cap binding protein from HeLa cells. *Nucleic Acids Res.* 18:6989-6995.
- Ohtsubo, M., H. Okasaki, and T. Nishimoto. 1989. The RCC1 protein, a regulator for the onset of chromosome condensation, locates in the nucleus and binds to DNA. *J. Cell. Biol.* 109:1389-1397.
- O'Keefe, R. T., C. Norman, and A. J. Newman. 1996. The invariant U5 snRNA loop I sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. *Cell* 86:679-689.
- Oubridge, C., N. Ito, P. R. Evans, C. H. Teo, and K. Nagai. 1994. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* 372:432-438.
- Padgett, R. A., M. M. Konarska, P. J. Grabowski, S. F. Hardy, and P. A. Sharp. 1984. Lariat RNAs as intermediates and products in the splicing of messenger RNA precursors. *Science* 225:898-903.
- Panté, N., and U. Aebi. 1996. Toward the molecular dissection of protein import into nuclei. *Curr. Opin. Cell Biol.* 8:397-406.
- Parker, R., P. G. Siliciano, and C. Guthrie. 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* 49:229-239.
- Parker, R., and B. Patterson. 1987. Architecture of fungal introns: implications for spliceosome assembly, p. 133-148. In M. Inouye and B. Dubock (eds.), *Molecular Biology of RNA: New Perspective*. Academic Press, New York, N.Y.
- Parker, K. A., and J. A. Steitz. 1987. Structural analysis of the human U3 ribonucleoprotein particle reveal a conserved sequence available for base pairing with pre-rRNA. *Mol. Cell. Biol.* 7:2899-2913.
- Parker, R., and P. G. Siliciano. 1993. Evidence for an essential non-Watson-Crick interaction between the first and last nucleotides of a nuclear pre-mRNA intron. *Nature* 361:660-662.
- Patton, J. R. 1991. Pseudouridine modification of U5 RNA in ribonucleoprotein particles assembled *in vitro*. *Mol. Cell. Biol.* 11:5998-6006. (Erratum, 12:904, 1992.)
- Patton, J. R. 1993. Multiple pseudouridine synthase activities for small nuclear RNAs. *Biochem. J.* 290:595-600.
- Patton, J. R. 1994a. Pseudouridine formation in small nuclear RNAs. *Biochimie* 76:1129-1132.
- Patton, J. R., M. R. Jacobson, and T. Pederson. 1994. Pseudouridine formation in U2 small nuclear RNA. *Proc. Natl. Acad. Sci. USA* 91:3324-3328.
- Patton, J. R. 1994b. Formation of pseudouridine in U5 small nuclear RNA. *Biochemistry* 33:10423-10427.
- Peattie, D. A. 1979. Direct chemical method for sequencing RNA. *Proc. Natl. Acad. Sci. USA* 76:1760-1764.
- Peculis, B. A., and J. A. Steitz. 1993. Disruption of U8 nucleolar snRNA inhibits 5.8S and 28S rRNA processing in the *Xenopus* oocyte. *Cell* 73:1233-1245.
- Peculis, B. A. 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.* 17:3702-3713.
- Piñol-Roma, S., and G. Dreyfuss. 1992. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature* 355:730-732.
- Plessel, G., U. Fisher, and R. Lührmann. 1994. m₁G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) *in vitro*: evidence that the U1 small nuclear RNA-(guanosine-N2)-methyl-transferase is a non-snRNP cytoplasmic protein that requires a binding site on the Sm core domain. *Mol. Cell. Biol.* 14:4160-4172.
- Raker, V. A., G. Plessel, and R. Lührmann. 1996. The snRNP core assembly pathway: identification of stable core protein heterodimeric complexes and an snRNP subcore particle *in vitro*. *EMBO J.* 15:2256-2269.
- Raué, H. A., and R. J. Planta. 1995. The pathway to maturity: processing of ribosomal RNA in *Saccharomyces cerevisiae*. *Gene Expression* 5:71-77.
- Reddy, R., D. Henning, and H. Busch. 1979. Nucleotide sequence of nucleolar U3B RNA. *J. Biol. Chem.* 254:11097-11105.
- Reddy, R., D. Henning, P. Epstein, and H. Busch. 1981. Primary and secondary structure of U2 snRNA. *Nucleic Acids Res.* 9:5645-5658.
- Reddy, R., D. Spector, D. Henning, M. H. Liu, and H. Busch. 1983. Isolation and partial characterization of dinoflagellate U1-U6 small RNAs homologous to rat U small nuclear RNAs. *J. Biol. Chem.* 258:13965-13969.

- Reddy, R., D. Henning, S. Chirala, L. Rothblum, D. Wright, and H. Busch. 1985. Isolation and characterization of three rat U3 RNA pseudogenes colinear with U3 RNA. *J. Biol. Chem.* 260: 5715-5719.
- Reddy, R., D. Henning, G. Das, M. Harless, and D. Wright. 1987. The capped U6 small nuclear RNA is transcribed by RNA polymerase III. *J. Biol. Chem.* 262:75-81.
- Reddy, R. 1988. Compilation of small RNA sequences. *Nucleic Acids Res.* 16:71-85.
- Reddy, R., and H. Busch. 1988. Small nuclear RNAs: RNA sequences, structures and modifications. In M. L. Birnstiel (ed.), *Structure and Function of Major Minor Small Nuclear Ribonucleoprotein Particles*. Springer-Verlag, Berlin, Germany.
- Reddy, R., R. Singh, and S. Shimba. 1992. Methylated cap structures in eukaryotic RNAs: structure, synthesis and functions. *Pharmacol. Ther.* 54:249-267.
- Riedel, N., J. A. Wise, H. Swerdlow, A. Mak, and C. Guthrie. 1986. Small nuclear RNAs from *Saccharomyces cerevisiae*: unexpected diversity in abundance, size and molecular complexity. *Proc. Natl. Acad. Sci. USA* 83:8097-8101.
- Rogers, J., and R. Wall. 1980. A mechanism for RNA splicing. *Proc. Natl. Acad. Sci. USA* 77:1877-1879.
- Ruby, S. W., T. H. Chang, and J. Abelson. 1993. Four yeast spliceosomal proteins (PRP5, PRP9, PRP11, and PRP21) interact to promote U2 snRNP binding to pre-mRNA. *Genes Dev.* 7: 1909-1925.
- Ruskin, B., A. R. Krainer, T. Maniatis, and M. R. Green. 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing *in vitro*. *Cell* 38:317-331.
- Rymond, B. C., and M. Rosbash. 1992. Yeast pre-mRNA splicing, p. 143-192. In E. W. Jones, J. R. Pringle, and J. R. Broach (ed.), *Yeast Pre-mRNA Splicing*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Salditt-Georgieff, M., M. Harpold, S. Chen-Kiang, and J. J. Darnell. 1980. The addition of 5' cap structures occurs early in hnRNA synthesis and prematurely terminated molecules are capped. *Cell* 19:69-78.
- Samarsky, D. A., A. G. Balakin, and M. J. Fournier. 1995. Characterization of three new snRNAs from *Saccharomyces cerevisiae*: snR34, snR35, and snR36. *Nucleic Acids Res.* 23: 2548-2554.
- Samuelson, T. 1991. Interactions of tRNA pseudouridine synthases with RNAs substituted with fluorouracil. *Nucleic Acids Res.* 19: 6139-6144.
- Sawa, H., and Y. Shimura. 1992. Association of U6 snRNA with the 5' splice site region of pre-mRNA in the spliceosome. *Genes Dev.* 6:244-254.
- Sawa, H., and J. Abelson. 1992. Evidence for a base-pairing interaction between U6 small nuclear RNA and 5' splice site during the splicing reaction in yeast. *Proc. Natl. Acad. Sci. USA* 89: 11269-11273.
- Scherly, D., W. Boelens, W. J. Van Venrooij, N. A. Dathan, J. Hamm, and I. W. Mattaj. 1989. Identification of the RNA binding segment of human U1 A protein and definition of its binding site on U1 snRNA. *EMBO J.* 8:4163-4170.
- Schlenstedt, G., D. H. Wong, D. M. Koepp, and P. A. Silver. 1995. Mutants in a yeast Ran binding protein are defective in nuclear transport. *EMBO J.* 14:5367-5378.
- Ségault, V., et al. Unpublished data.
- Ségault, V., C. L. Will, B. S. Sproat, and R. Lührmann. 1995. *In vitro* reconstitution of mammalian U2 and U5 snRNPs active in splicing: Sm proteins are functionally interchangeable and are essential for the formation of functional U2 and U5 snRNPs. *EMBO J.* 14:4010-4021.
- Séraphin, B., L. Kretzner, and M. Rosbash. 1988. A U1 snRNA: pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. *EMBO J.* 7:2533-2538.
- Séraphin, B., and M. Rosbash. 1990. Exon mutations uncouple 5' splice site selection from U1 snRNA pairing. *Cell* 63:619-629.
- Shapiro, M. B., and P. Senapathy. 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res.* 15: 7155-7174.
- Sharp, P. A. 1985. On the origin of RNA splicing and introns. *Cell* 42:397-400.
- Sharp, P. A. 1991. Five easy pieces. *Science* 254:663.
- Shatkin, A. 1976. Capping of eukaryotic mRNAs. *Cell* 9:645-653.
- Siliciano, P. G., and C. Guthrie. 1988. 5' splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. *Genes Dev.* 2:1258-1267.
- Singh, R., and R. Reddy. 1989. Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proc. Natl. Acad. Sci. USA* 86:8280-8283.
- Smith, C., and J. A. Steitz. 1997. Sno storm in the nucleolus: new roles for myriad small RNPs. *Cell* 89:669-672.
- Solyomosi, F., and T. Pollák. 1993. Uridylate-rich small nuclear RNAs (UsnRNAs), their genes and pseudogenes, and UsnRNPs in plants: structure and function. A comparative approach. *Crit. Rev. Plant Sci.* 12:275-369.
- Sontheimer, E. J., and J. A. Steitz. 1993. The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science* 262:1989-1996.
- Stroke, I. L., and A. M. Weiner. 1985. Genes and pseudogenes for rat U3A and U3B small nuclear RNAs. *J. Mol. Biol.* 184: 183-193.
- Sun, J. S., and J. L. Manley. 1995. A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes Dev.* 9: 843-854.
- Sun, J. S., and J. M. Manley. 1997. The human U6 snRNA intramolecular helix: structural constraints and lack of sequence specificity. *RNA* 3:514-526.
- Szukálek, A., E. Mysliński, A. Mouglin, R. Lührmann, and C. Branlant. 1995. Phylogenetic conservation of modified nucleotides in the terminal loop 1 of the spliceosomal U5 snRNA. *Biochimie* 77:16-21.
- Tarn, W., T. R. Yario, and J. A. Steitz. 1995. U12 snRNA in vertebrates: Evolutionary conservation of 5' sequences implicated in splicing of pre-mRNAs containing a minor class of introns. *RNA* 1:644-656.
- Tarn, W. Y., and J. A. Steitz. 1996a. A novel spliceosome containing U11, U12 and U5 snRNPs excises a minor class (AT-AC) intron *in vitro*. *Cell* 84:1-20.
- Tarn, W. Y., and J. A. Steitz. 1996b. Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science* 273:1824-1832.
- Terns, M. P., J. E. Dahlberg, and E. Lund. 1993. Multiple cis-acting signals for export of pre-U1 snRNA from the nucleus. *Genes Dev.* 7:1898-1908.
- Terns, M. P., and J. E. Dahlberg. 1994. Retention and 5' cap trimethylation of U3 snRNA in the nucleus. *Science* 264:1: 959-961.
- Terns, M. P., C. Grimm, E. Lund, and J. E. Dahlberg. 1995. A common maturation pathway for small nucleolar RNAs. *EMBO J.* 14:4860-4871.
- Thompson, J. R., J. Zagorski, J. L. Woolford, and M. J. Fournier. 1988. Sequence and genetic analysis of a dispensible 189 nucleotides snRNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 16:5587-5601.
- Tollervey, D. 1987. A yeast small nuclear RNA is required for normal processing of preribosomal RNA. *EMBO J.* 6: 4169-4175.

- Tollervey, D., and T. Kiss. 1997. Function and synthesis of small nucleolar RNAs. *Curr. Opin. Cell Biol.* 9:337-342.
- Tyc, K., and J. A. Steitz. 1989. U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus. *EMBO J.* 8:3113-3119.
- Tycowski, K. T., M. D. Shu, and J. A. Steitz. 1993. A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. *Genes Dev.* 7:1176-1190.
- Umen, J. G., and C. Guthrie. 1995. The second catalytic step of pre-mRNA splicing. *RNA* 1:869-885.
- Valcarcel, J., R. K. Gaur, R. Singh, and M. R. Green. 1996. Interaction of U2AF⁶⁵RS region with pre-mRNA of branch point and promotion base-pairing with U2 snRNA. *Science* 273:1706-1709.
- Vankan, P., C. McGuigan, and I. W. Mattaj. 1990. Domains of U4 and U6 snRNAs required for snRNP assembly and splicing complementation in *Xenopus* oocytes. *EMBO J.* 9:3397-3404.
- Visa, N., A. T. Alzhanova-Ericsson, X. Sun, E. Kiselava, B. Björkroth, T. Wurtz, and B. Daneholt. 1996. A pre-mRNA binding protein accompanies the RNA from the gene through the nuclear pores and into polysomes. *Cell* 84:253-264.
- Wells, S. E., and M. Ares. 1994. Interactions between highly conserved U2 small nuclear RNA structures and Prp5p, Prp9p, Prp11p, and Prp21p proteins are required to ensure integrity of the U2 small nuclear ribonucleoprotein in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 14:6337-6349.
- Wells, S. E., M. Neville, M. Haynes, J. Wang, H. Igel, and M. Ares. 1996. CUS1, a suppressor of cold-sensitive U2 snRNA mutations, is a novel yeast splicing factor homologous to human SAP 145. *Genes Dev.* 10:220-232.
- Westin, G., E. Lund, J. T. Murphy, U. Pettersson, and J. E. Dahlberg. 1984. Human U2 and U1 RNA genes use similar transcription signals. *EMBO J.* 3:3295-3301.
- Will, C. L., S. Rümpler, J. K. Gunnewiek, W. J. VanVenrooij, and R. Lührmann. 1996. *In vitro* reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E) spliceosomal complexes. *Nucleic Acids Res.* 24:4614-4623.
- Wise, J. A., and A. M. Weiner. 1980. *Dictyostelium* small nuclear RNA D2 is homologous to rat nucleolar RNA U3 and is encoded by a dispersed multigene family. *Cell* 22:109-118.
- Wolff, T., and A. Bindereif. 1993. Conformational changes of U6 RNA during the spliceosome cycle: an intramolecular helix is essential both for initiating the U4-U6 interaction and for the first step of splicing. *Genes Dev.* 7:1377-1389.
- Wolff, T., and A. Bindereif. 1995. Mutational analysis of human U6 RNA: stabilizing the intramolecular helix blocks the spliceosomal assembly pathway. *Biochim. Biophys. Acta* 1263:39-44.
- Wong, N., A. H. Corbett, H. M. Kent, M. Steward, and P. A. Silver. 1997. Interaction between the small GTPase Ran/Gsp1p and Ntf2p is required for nuclear transport. *Mol. Cell. Biol.* 17:3755-3767.
- Wu, J., and J. L. Manley. 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base-pairing. *Genes Dev.* 3:1553-1561.
- Wu, J. A., and J. L. Manley. 1991. Base-pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. *Nature* 352:818-821.
- Wu, Q., and A. R. Krainer. 1997. Splicing of a divergent subclass of AT-AC introns requires the major spliceosomal snRNAs. *RNA* 3:586-601.
- Wyatt, J. R., E. J. Sontheimer, and J. A. Steitz. 1992. Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. *Genes Dev.* 6:2542-2553.
- Yan, D., and M. J. Ares. 1996. Invariant U2 RNA sequences bordering the branchpoint recognition region are essential for interaction with yeast SF3A and SF3B subunits. *Mol. Cell. Biol.* 16:818-828.
- Zavanelli, M. I., J. S. Britton, A. H. Igel, and M. Ares. 1994. Mutations in an essential U2 small nuclear RNA structure cause cold-sensitive U2 small nuclear ribonucleoprotein function by favoring competing alternative U2 RNA structures. *Mol. Cell. Biol.* 14:1689-1697.
- Zerby, D. B., and J. R. Patton. 1997. The modification of U4 RNA requires U6 RNA and multiple pseudouridine synthases. *Nucleic Acids Res.* 25:4808-4815.
- Zerby, D. B., and J. R. Patton. 1996. Metabolism of pre-messenger RNA splicing cofactors: modification of U6 RNA is dependent on its interaction with U4 RNA. *Nucleic Acids Res.* 24:3583-3589.
- Zhuang, Y., and A. M. Weiner. 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell* 46:827-835.
- Zhuang, Y., and A. M. Weiner. 1989. A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes Dev.* 3:1545-1552.
- Zieve, G., and S. Penman. 1976. Small RNA species of the HeLa cell: metabolism and subcellular localization. *Cell* 8:19-31.

Résultats
Discussion

*PARTIE 1: LOCALISATION DES RESIDUS Ψ DANS LA GRANDE
SOUS-UNITÉ RIBOSOMIQUE D'UNE Archaea
HYPERTHERMOPHILE: *S. acidocaldarius**

A/ INTRODUCTION ET RESULTATS

La découverte de snoRNA guidant les enzymes de modification des ARNr chez les eucaryotes soulève la question de l'origine évolutive de ces ARN. La présence de modifications post-transcriptionnelles a été étudiée depuis longtemps dans les ARNr d'*E. coli* (Branlant *et al.*, 1981b) et, plus récemment, les enzymes générant ces modifications ont commencé à être étudiées. Peu de résidus Ψ sont présents dans les ARNr d'*E. coli* (Branlant *et al.*, 1981b; Bakin et Ofengand, 1993; Bakin *et al.*, 1994). Il en est d'ailleurs de même pour les autres bactéries étudiées (Ofengand et Bakin, 1997). Clairement, la formation de ces quelques résidus Ψ est catalysée par des enzymes ne renfermant pas d'ARN guides (pour revue, Ofengand et Fournier, 1998; voir Chap. C de la deuxième partie de l'introduction). En admettant qu'*E. coli* reflète une image du système le plus ancestral de formation des résidus Ψ et 2'-O-méthylés, comment un système impliquant une seule enzyme et plusieurs ARN guides a-t-il pu naître d'un système utilisant plusieurs enzymes avec des spécificités différentes? Lafontaine et Tollervey (1998a) ont tenté de répondre à cette question et proposent un modèle évolutif basé sur le fait que, chez *S. cerevisiae*, Pus4p (responsable de la formation du résidu Ψ_{55} dans les ARNt) et Cbf5p présentent toutes deux de fortes homologues de séquence avec la Ψ -synthase d'*E. coli* TruB, elle-même responsable de la formation de la Ψ_{55} dans les ARNt (Koonin, 1996; Becker *et al.*, 1997a; Lafontaine *et al.*, 1998a). Ces auteurs ont proposé que la duplication du gène codant la protéine TruB ait donné naissance aux ancêtres de Cbf5p et Pus4p. Cette dernière aurait gardé la fonction de modification des ARNt et Cbf5p aurait acquis la possibilité de reconnaître un duplex d'ARN formé entre le pré-ARN ribosomique et les ancêtres des ARN guides actuels. Selon leur modèle, les ARN guides de modification auraient pu être générés par duplication des gènes codant les petits ARN nucléolaires impliqués dans les étapes de clivage du pré-ARNr. Cette proposition est basée sur le fait que les ARN guides de méthylation et les ARN guides de pseudouridylation partagent des séquences communes avec certains des ARN impliqués dans les étapes de clivage.

Si une telle duplication de gène a eu lieu, à quel moment de l'évolution s'est-elle déroulée? Les génomes complètement séquencés des *Archaea Methanobacterium thermoautotrophicum* (Smith *et al.*, 1997), *Archaeoglobus fulgidus* (Klenk *et al.*, 1997) et *Methanococcus jannaschii* (Bult *et al.*, 1996) ne présentent qu'une phase ouverte de lecture homologue à celle du gène *truB* (Lafontaine et Tollervey, 1998a). Cependant, l'analyse plus approfondie de cette ORF chez *M. jannaschii* montre que la protéine codée présente 62% de similarité avec Cbf5p et seulement 29% avec TruB. Les auteurs en concluent que la protéine des *Archaea* serait l'homologue de Cbf5p et non celle de TruB (Watkins *et al.*, 1998a).

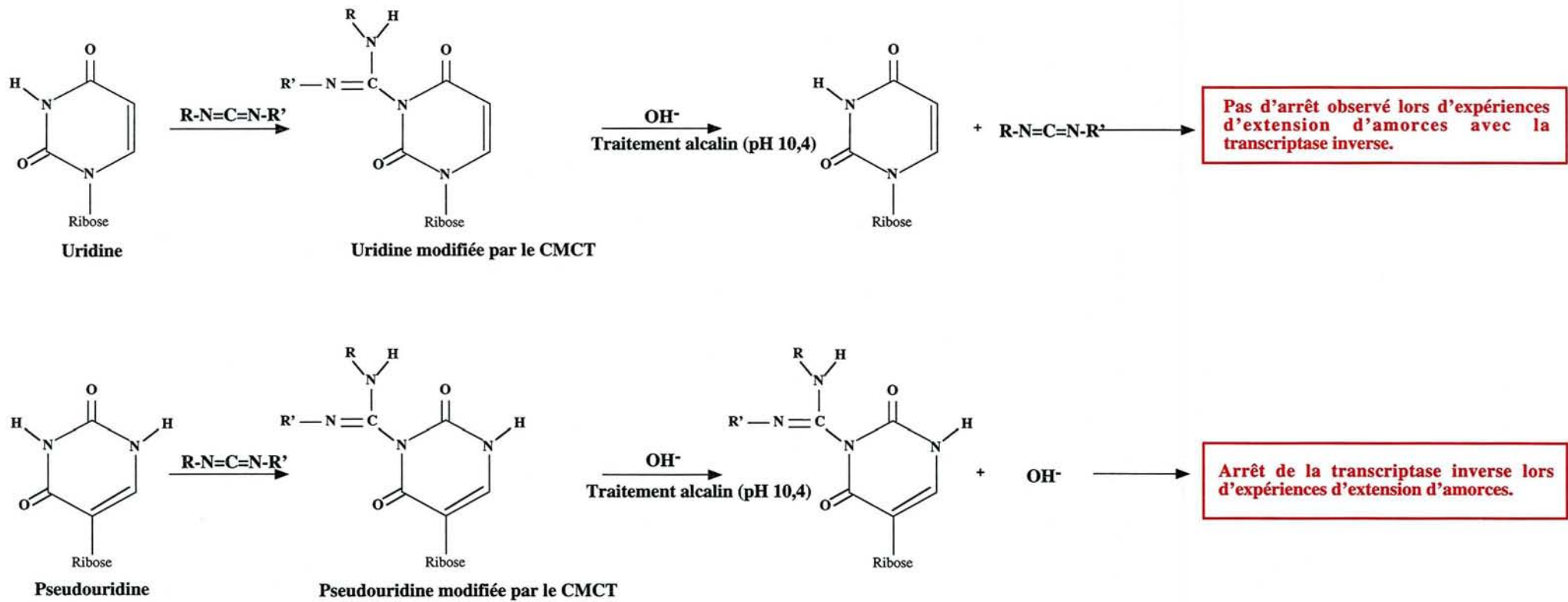


Figure 43: Principe de détection des résidus Ψ par modification chimique des ARN par le CMCT.

Si l'enzyme des *Archaea* est plus proche de Cbf5p, existe-t-il des ARN guides de pseudouridylation chez les *Archaea*? En défaveur de cette hypothèse, aucune protéine homologue à Gar1p n'a été jusqu'à maintenant détectée dans les génomes séquencés d'*Archaea* (Lafontaine et Tollervey, 1998a). Cette protéine Gar1p est associée aux snoRNA H/ACA chez les eucaryotes et est nécessaire à l'activité de pseudouridylation. Néanmoins, on pourrait imaginer qu'une autre protéine soit associée aux ARN guides de pseudouridylation des *Archaea*.

Une étape préliminaire et essentielle, pour savoir s'il y a vraiment des ARN guides de pseudouridylation chez les *Archaea*, est de localiser les résidus Ψ dans les ARNr d'*Archaea*. La cartographie des résidus Ψ contenus dans le grand ARNr de la grande sous-unité ribosomique avait, en effet, seulement été réalisée chez une *Archaea* halophile, *H. halobium* (Ofengand et Bakin, 1997). Seuls 4 résidus Ψ avaient été détectés dans cet ARN, 2 dans le domaine IV et 2 dans le domaine V (Fig. 34 de l'introduction). Cette observation suggérerait donc que les ARNr d'*Archaea* comportent un nombre de résidus Ψ voisin de celui des bactéries et très inférieur à celui trouvé pour les eucaryotes (9 Ψ dans le grand ARN de la grande sous-unité ribosomique d'*E. coli* contre 55 chez l'homme, pour revue, Ofengand et Fournier, 1998). Il était important d'élargir ce type d'étude à une autre *Archaea*. Nous avons choisi d'entreprendre l'identification des résidus Ψ présents dans l'ARNr 23S d'une *Archaea* hyperthermophile acidophile, *Sulfolobus acidocaldarius* (croissance à 85°C).

Comme nous l'avons vu dans l'introduction de ce manuscrit, trois domaines du grand ARNr, les domaines II, IV et V, sont connus pour renfermer un grand nombre de résidus modifiés chez les eucaryotes et pour être les seuls à porter des modifications post-transcriptionnelles chez les bactéries. L'identification des résidus Ψ , présents dans ces trois domaines de l'ARNr 23S de *S. acidocaldarius*, a été réalisée à l'aide de la méthode de modification chimique par le CMCT (Bakin et Ofengand, 1993) (Fig. 43). J'ai utilisé cette méthode de manière récurrente tout au long de mon travail de thèse. Nous en résumerons ici son principe. Cette technique repose sur l'addition d'un groupement carbodiimide sur les résidus uridyliques. Une addition du groupement CMCT a lieu en position 3 des résidus U et des résidus Ψ et, à un moindre degré, sur l'azote en position 1 des résidus G. Les groupements CMCT liés aux résidus U et G sont facilement éliminés par traitement alcalin, contrairement à ceux liés aux résidus Ψ . Après traitement alcalin à pH 10,4, la présence de modifications peut être détectée par la méthode d'extension d'amorces. Le principe de cette analyse repose sur: i) l'arrêt de l'élongation d'amorce au niveau des résidus U et Ψ modifiés après traitement par le CMCT, sans traitement alcalin, ii) l'élimination de l'arrêt au niveau des résidus U, mais pas des résidus Ψ après traitement alcalin. Il est à noter que l'arrêt d'élongation de la transcriptase inverse est observé au niveau du nucléotide suivant la position modifiée. L'utilisation de la

transcriptase inverse pour l'analyse avait, par ailleurs, l'avantage de permettre l'analyse de régions données du grand ARN de la grande sous-unité ribosomique.

La publication n°1 présente les résultats obtenus lors de la localisation des résidus Ψ dans le grand ARN de la grande sous-unité ribosomique de *S. acidocaldarius*. Ces résultats sont discutés sur la base d'une comparaison des positions modifiées dans différentes espèces, ainsi que sur la base des connaissances actuelles sur les Ψ -synthases agissant sur les ARNr.

Publication n°1:

The first determination of pseudouridine residues in 23S ribosomal RNA from hyperthermophilic *Archaea Sulfolobus acidocaldarius*.

S. Massenet, I. Ansmant, Y. Motorin and C. Branlant.

FEBS letters 462, 94-100, 1999.

Publication n°1

**The first determination of pseudouridine residues in 23S
ribosomal RNA from hyperthermophilic *Archaea*
*Sulfolobus acidocaldarius***

S. Massenet, I. Ansmant, Y. Motorin and C. Branlant.

FEBS letters 462, 94-100, 1999

The first determination of pseudouridine residues in 23S ribosomal RNA from hyperthermophilic *Archaea Sulfolobus acidocaldarius*

S everine Massenet¹, Isabelle Ansmant¹, Yuri Motorin, Christiane Branlant*

Laboratoire de Maturation des ARN et Enzymologie Mol culaire, UMR 7567 CNRS-UHP Nancy I, Facult  des Sciences, P.O. Box 239, 54506 Vandoeuvre-les-Nancy Cedex, France

Received 16 October 1999

Abstract We describe the first identification of pseudouridine (Ψ) residues in ribosomal RNA (23S rRNA) of an hyperthermophilic *Archaea Sulfolobus acidocaldarius*. In contrast to *Eucarya* rRNA, only six Ψ residues were detected, which is rather close to the situation in *Bacteria*. However, three modified positions (Ψ_{2479} , Ψ_{2535} and Ψ_{2550}) are unique for *S. acidocaldarius*. Two Ψ residues at positions 2060 and 2594 are universally conserved, while one other Ψ (position 2066) is also common to *Eucarya*. Taken together the results argue against the conservation of Ψ -synthases between *Archaea* and *Bacteria* and provide a basis for the search of snoRNA-like guides for Ψ formation in *Archaea*.

  1999 Federation of European Biochemical Societies.

Key words: Pseudouridine; rRNA; SnoRNA; Pseudouridine-synthase; *Archaea*; *Sulfolobus acidocaldarius*

1. Introduction

The rRNAs and tRNAs of all organisms and the *Eucarya* UsnRNAs examined so far undergo extensive post-transcriptional modification (for review see [1–4]). The most frequent chemical modifications in rRNAs are methylations of the 2'-hydroxyl of the ribose moieties (2'-O-methylation) and isomerisations of uracil residues into pseudouridines (Ψ). The Ψ residues in rRNA of the large ribosomal subunit (LSU) have been mapped for several of species from the *Eucarya* and *Bacteria* domains (*Homo sapiens*, *Mus musculus*, *Saccharomyces cerevisiae*, *Escherichia coli*, etc.), for one *Archaea* (*Halobacterium halobium*), and also for mitochondria and chloroplasts of several *Eucarya* cells [5–10]. The number of Ψ residues per rRNA molecule varies widely from *Bacteria* and *Archaea* to *Eucarya* (from nine Ψ residues in *E. coli* LSU rRNA to 55 Ψ residues in its human counterpart). However, all detected Ψ residues are concentrated in three domains of the LSU rRNA (domains II, IV and V). These domains are located at or are functionally linked to the peptidyl transferase centre (PTC) ([5,6,11]; for review see [2]). Despite the clustering of Ψ residues in the functional site of the ribosome, the exact importance of these Ψ residues in ribosomal assembly and peptidyl transferase reaction remains unknown. Several Ψ

sites present in LSU rRNAs are evolutionarily conserved. The situation is quite different for the Ψ residues present in the rRNA of the small ribosomal subunit (SSU). In this rRNA, the Ψ residues are mostly present at sites specific for each organism and no apparent clustering of these modified nucleotides in defined areas of SSU rRNA was observed (for review see [2]).

Synthesis of Ψ residues in RNA molecules is catalysed by a large family of enzymes, the RNA: Ψ -synthases. Recent progress in studies of Ψ residue formation demonstrated that *Bacteria* and *Eucarya* cells do not share the same process to modify U residues in rRNAs. Indeed, in *Bacteria*, Ψ formation is ensured by site- or multisite-specific rRNA: Ψ -synthases. Several such enzymes have recently been characterised in *E. coli* and *Bacillus subtilis* [12–18]. Two multisite-specific *E. coli* enzymes, RluC and RluD, catalyse Ψ formation at six distinct sites in 23S rRNA (positions 955, 2504 and 2580, and 1911, 1915 and 1917, respectively) [14–16]. Three other Ψ -synthases, RluB in *B. subtilis* and RsuA and RluE in *E. coli*, each modify a unique U residue into a Ψ residue [13,18,19]. The first discovered rRNA: Ψ -synthase in *E. coli*, RluA, displays a dual substrate specificity, since it is implicated in the formation of Ψ_{746} in 23S rRNA and of Ψ_{32} in tRNAs [12,17].

In *Eucarya*, pre-rRNA maturation involves many different small nucleolar RNAs (snoRNAs) of two different classes: the H/ACA and the C/D snoRNAs. Most of these H/ACA and C/D snoRNAs are implicated in selection of sites of Ψ formation and 2'-O-methylations, respectively, by site-specific base pairing with the pre-rRNA (for review see [2,4]). Specific H/ACA snoRNA guides have been found for at least 23 pseudouridylation sites out of 30 in yeast 26S rRNAs [20] and, in most cases, the direct implication of the corresponding snoRNA in modification at a given site was proved experimentally [21,22]. In yeast, this snoRNA-guided modification system comprises several protein components and a unique Ψ -synthase, Cbf5p [23,24]. *S. cerevisiae* Cbf5p and the tRNA: Ψ_{55} -synthase Pus4p are both homologous to the *E. coli* tRNA: Ψ_{55} -synthase TruB [23,25]. Cbf5p counterparts have been previously reported for a variety of *Eucarya*, such as *Kluyveromyces lactis* [26], *Drosophila melanogaster* (Nop60p) [27], rat (NAP 57) [28] and human (DKC1) [29].

Although the *Archaea* and *Bacteria* appear rather similar in terms of general genome organisation, some of the *Archaea* genes show greater sequence similarity to *Eucarya* than to *Bacteria* counterparts (for review see [30]). Up to now, it is not known whether the *Archaea* domain also possesses RNA-guided modification systems. Several observations argue for the existence of only C/D-type RNPs, but not for H/ACA-type RNPs in *Archaea* [31]. Indeed, the *S. cerevisiae* C/D snoRNAs are associated with proteins Nop1p (fibrillar in

*Corresponding author. Fax: (33) 3.83.91.20.93.
E-mail: cbranlant@maem.uhp-nancy.fr

¹ These authors contributed equally to the work.

Abbreviations: Ψ , pseudouridine; nt, nucleotides; RT, reverse transcriptase; CMCT, *N*-cyclohexyl-*N'*-(2-morpholinoethyl)-carbodiimide metho-*p*-toluolsulfonate; PTC, peptidyl transfer centre; LSU, large ribosomal subunit; SSU, small ribosomal subunit

human), Nop56p and Nop58p [32–35]. A gene homologous to the yeast *NOPI* gene was detected in three methanogenic *Archaea* [36,37], and a homologue of the yeast *NOP58* gene was found in the *Archaea Archaeoglobus fulgidus*, *Methanobacterium thermoautotrophicum* and *Methanococcus jannaschii* [31,37–39]. The estimated number of 2'-*O*-methylations in *Sulfolobus solfataricus* LSU rRNA (about 44 sites) [40] is much higher than that found in *Bacteria* (four sites in *E. coli* LSU rRNA), and quite similar to that observed in *Eucarya* (55 sites in yeast LSU rRNA, for review see [4]). This huge number of 2'-*O*-methylation sites in *Archaea* rRNAs suggests that this type of post-transcriptional modification in *Archaea* may be RNA-guided like in *Eucarya*. As far as rRNA pseudouridylation is concerned, the situation with *Archaea* seems to be different. H/ACA snoRNA guides directing pseudouridylation of rRNA in *S. cerevisiae* are found to be associated with the Gar1p protein [24,41–43] and no homologous protein has been reported up to now in *Archaea* [31]. In fact, the only observation in favour of an RNA-guided pseudouridylation system in *Archaea* is the finding of an open reading frame, conserved in all studied *Archaea* genomes, that shows 62% sequence similarity with Cbf5p and only 29% with TruB [24]. The authors concluded that the *Archaea* protein is most likely a Cbf5p and not a TruB homologue.

Thus, the question remains open about the existence of a RNA-guided system for Ψ synthesis in *Archaea* rRNAs. To search for the presence of genes encoding H/ACA-like guide RNAs in the *Archaea* genomes, one has first to identify the positions of Ψ residues in rRNAs from these organisms. Up to now, the location of Ψ residues in LSU rRNAs was determined for only one halophilic *Archaea* (*H. halobium*) [10]. In this paper we report the mapping of Ψ residues in the LSU rRNA of a hyperthermophile *Archaea*, *Sulfolobus acidocaldarius*. Six Ψ residues were detected in the studied rRNA regions,

which is slightly higher than the 3.7 ± 0.1 mol/mol of Ψ residues detected by chromatography/mass spectrometry in the *S. sulfolobus* LSU rRNA [40] and the four Ψ residues localised in the *H. halobium* LSU rRNA. Only a few modification positions are conserved between *Bacteria* and *Archaea*, while the others seem to be specific for *Archaea* or common to *Archaea* and *Eucarya*. These observations will be discussed in terms of the evolution of Ψ -synthase activities.

2. Materials and methods

2.1. RNA preparation

S. acidocaldarius cells were a generous gift of P. Forterre (Orsay, France). The method of RNA preparation was adapted from Aiba et al. [44]. 0.3–0.6 g of cells were washed with 10 ml of 50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl pH 8.2 buffer and centrifuged for 10 min at 4000 rpm. Cells were resuspended in 300 μ l of 1 mM EDTA, 0.5% SDS, 20 mM sodium acetate pH 5.5 buffer, treated with 300 μ l of phenol, incubated for 10 min at 65°C and centrifuged. The extraction was done twice, recovered RNAs were treated with equal volumes of chloroform/isoamyl alcohol (24:1) and ethanol precipitated in the presence of 0.3 M NaCl.

2.2. Mapping of Ψ residues in the *S. acidocaldarius* LSU rRNA

The *N*-cyclohexyl-*N'*-(2-morpholinoethyl)-carbodiimide metho-*p*-toluolsulfonate (CMCT) modification protocol was adapted from Bakin and Ofengand [8], with the modification described previously [45]. Modification by CMCT was performed with 10 μ g of *S. acidocaldarius* total RNA.

Positions of CMCT modifications were identified by primer extension analysis, using the AMV RT (Life Science, USA), in the conditions described by Mougin et al. [46]. The oligonucleotides complementary to the following regions of the *S. acidocaldarius* LSU rRNAs were used as primers: for domain II, nucleotides (nt) 1121–1139 (primer 1), for domain IV, nt 2158–2178 (primer 2) and nt 2072–2091 (primer 3) and for domain V, nt 2575–2592 (primer 4), nt 2638–2656 (primer 5) and nt 2736–2751 (primer 6). The hybridisation conditions were optimised to improve the yield of reverse transcrip-

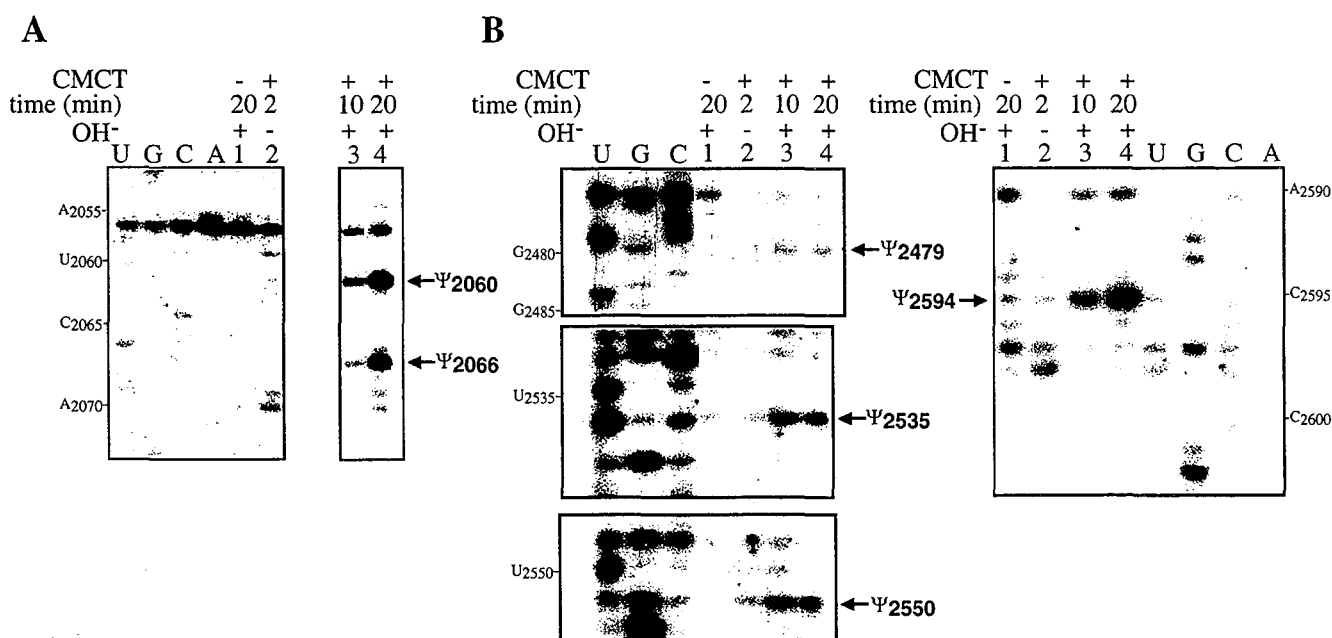


Fig. 1. Primer extension analysis of Ψ residues in domains IV (A) and V (B) of the *S. acidocaldarius* LSU rRNA. Modification of total *S. acidocaldarius* RNA by CMCT was done for 2, 10 and 20 min (lanes 2, 3 and 4, respectively). Experimental conditions were as described in Section 2. In lanes 3 and 4, the CMCT-modified RNA was subjected to alkaline treatment at pH 10.4. A control extension experiment was made without CMCT treatment (lane 1). Lanes U, G, C and A correspond to the RNA sequencing ladder. The reverse transcription stops, in lanes 3 and 4, indicate the presence of Ψ residues at these positions and are shown by arrows.

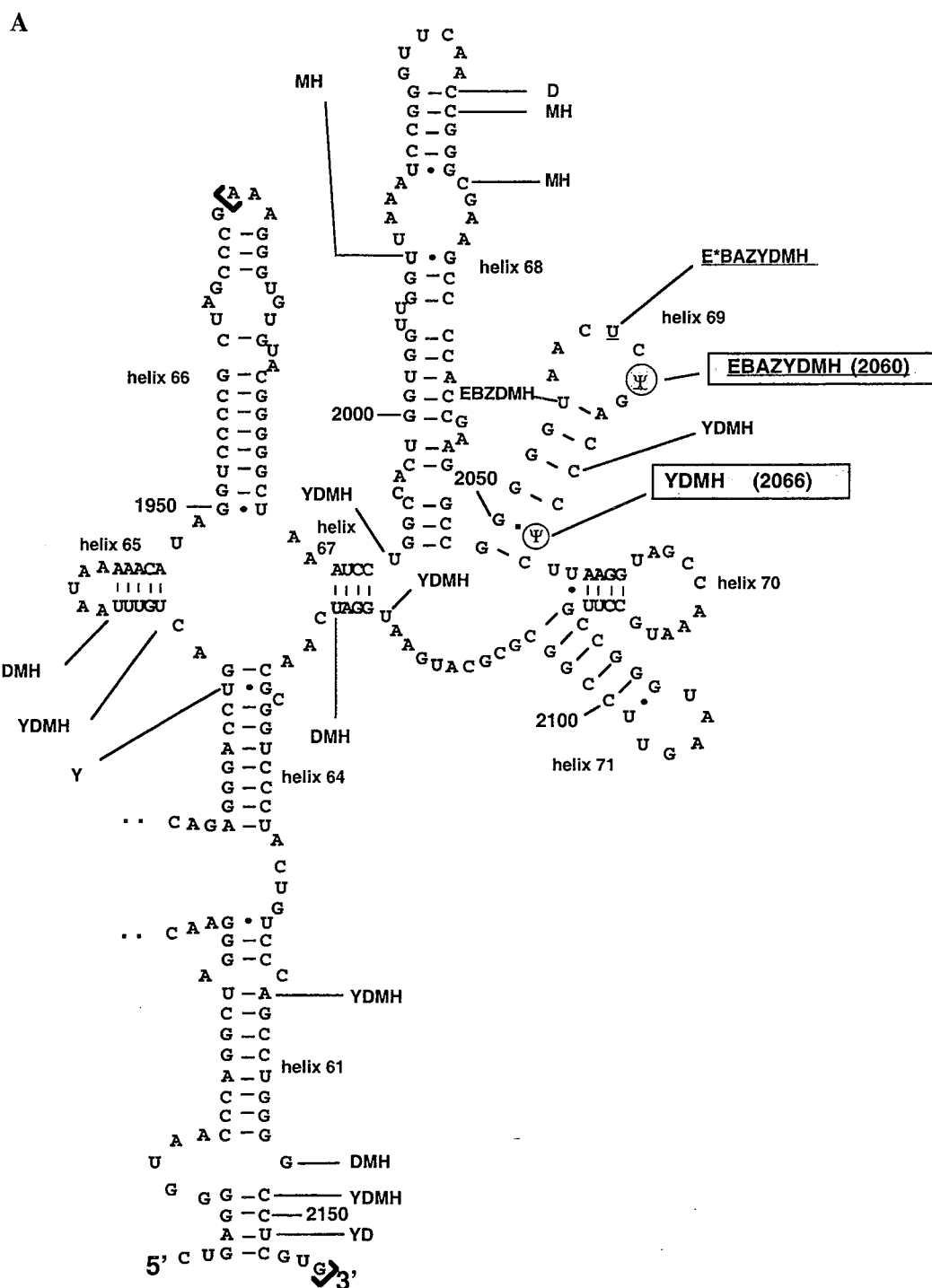


Fig. 2. Location of Ψ residues in domains IV (A) and V (B) of the *S. acidocaldarius* LSU rRNA and comparison with the locations of Ψ residues in LSU rRNAs of other species. The sequence shown is that of *S. acidocaldarius* [59]. The six Ψ residues detected in this study are shown in grey. The positions of the Ψ residues found in other organisms are indicated by: E, *E. coli*; B, *B. subtilis*; A, *H. halobium*; Y, *S. cerevisiae*; D, *D. melanogaster*; M, *M. musculus*; H, *H. sapiens*; Z, *Zea mays* chloroplast; Ym, Mm, Tm and Hm, *S. cerevisiae*, *M. musculus*, *Trypanosoma brucei* and *H. sapiens* mitochondria, respectively (for review see [2]). The asterisk (*) indicates an N^3 -methyl Ψ residue [48]. Nucleotide positions, starting from the 5'-terminal nucleotide of *S. acidocaldarius* LSU rRNA, are indicated. The regions analysed by primer extension analysis are between brackets.

tion. Primers 1, 2 and 4 were annealed for 10 min at 65°C followed by a 10 min incubation in ice, while oligonucleotides 3, 5 and 6 were hybridised for 30 s at 94°C followed by incubations for 5 min at 65°C and 5 min at 55°C. Primers were 5'-end labeled using [γ - 32 P]ATP (3000 Ci/mmol) and T4 polynucleotide kinase. RNA sequencing was done with 40 μ g of total *S. acidocaldarius* RNA.

3. Results and discussion

3.1. General strategy

Mapping Ψ residues in the *S. acidocaldarius* LSU rRNA was done on a total RNA fraction using the chemical method

B

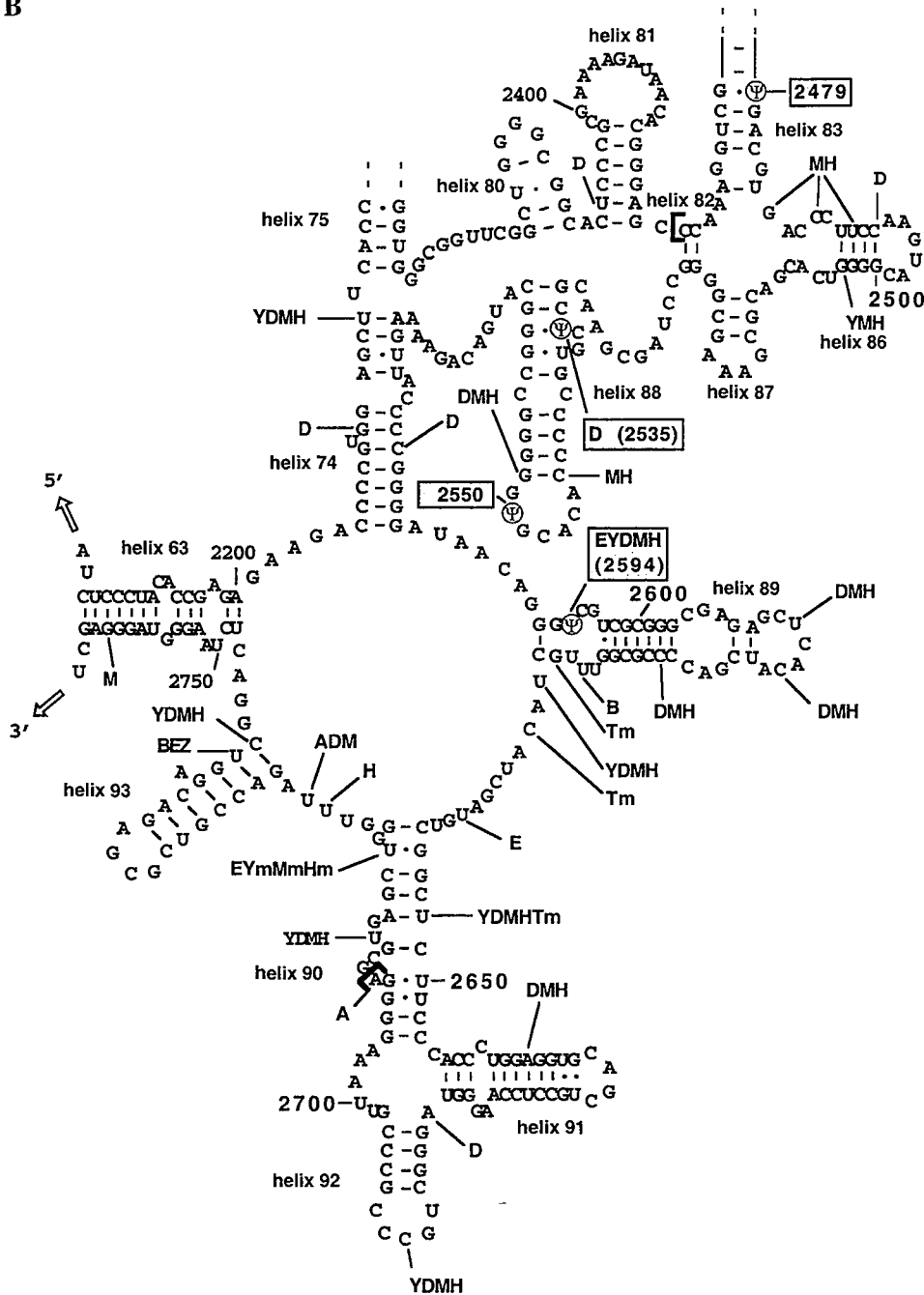


Fig. 2 (continued).

based on the alkaline-resistant modification of Ψ residues by CMCT, followed by their detection by primer extension analysis [8]. Due to the high GC content of *S. acidocaldarius* LSU rRNA, the specific conditions of hybridisation had to be determined for each oligonucleotide primer (see Section 2). The Ψ mapping in *S. acidocaldarius* LSU rRNA was done for the LSU rRNA regions that were found to contain most of the Ψ residues in the other species studied ([5,6,8–10]; for review see [2]). Thus, three regions were analysed by CMCT treatment followed by primer extension analysis: nt 1000–1120 in domain II, 1966–2155 in domain IV and 2419–2708 in domain V.

The results of primer extension analysis show that the re-

gion 1000–1120 in domain II of *S. acidocaldarius* 23S rRNA does not contain Ψ residues. It is noteworthy that *E. coli* is the only prokaryote found to have two Ψ residues in this domain [5,10]. Domains II of *B. subtilis* and *H. halobium* LSU rRNA were found to be devoid of Ψ residues [10].

3.2. Two Ψ residues are found in *S. acidocaldarius* domain IV 23S rRNA

We detected two Ψ residues in domain IV of the *S. acidocaldarius* 23S rRNA, at positions 2060 and 2066, respectively (Fig. 1A). Fig. 2A shows the location of these two Ψ residues in *S. acidocaldarius* domain IV as compared to locations of Ψ residues found in other species. All the pseudouridylation sites

detected in domain IV of the LSU rRNAs from *S. acidocaldarius*, *H. halobium* and *Bacteria* are clustered in a single stem-loop structure (numbered 69 according to the nomenclature of [47]). In contrast, in the cytoplasmic LSU rRNA from *Eucarya*, Ψ residues are more widely distributed in domain IV (Fig. 2A). The two Ψ residues underlined in Fig. 2A were found in all studied LSU rRNAs, except the mitochondrial ones, whose domain IV does not contain Ψ residues [10] and, as shown by the present data, a U residue is present at position 2058 in *S. acidocaldarius* LSU rRNA. The activity required to convert U₂₀₅₈ into a Ψ residue is probably absent in *S. acidocaldarius*. In *E. coli*, the Ψ residue corresponding to the *S. acidocaldarius* U₂₀₅₈ is a hypermodified N³-methyl Ψ residue [48]. The presence of this residue generates a strong RT stop upon primer extension analysis, even in the absence of CMCT treatment [8]. This is not the case for position 2058 in the *S. acidocaldarius* LSU rRNA. CMCT clearly reacts with uridine residue at this position (Fig. 1A, lane 2) and the CMCT modification is alkaline-labile (Fig. 1A, lanes 3 and 4). Interestingly, the second Ψ residue detected in stem-loop structure 69 of the *S. acidocaldarius* LSU rRNA was found only in *Eucarya* species (*S. cerevisiae*, *D. melanogaster*, *M. musculus* and *H. sapiens*, Fig. 2A). The presence of a Ψ residue at this position was never detected in LSU rRNAs from *Bacteria* or *Archaea*. Finally, the third Ψ residue present in stem-loop structure 69 of *E. coli* and *B. subtilis*, which is rather well conserved in *Eucarya* LSU rRNAs (position 1911 in *E. coli*), is not detected in *S. acidocaldarius* (position 2054). Since unmodified U residue is present at this position in *S. acidocaldarius* rRNA this means that the corresponding enzymatic activity is absent. Clearly, stem-loop structure 69 is the target for Ψ -synthases in all living organisms. However, positions of modifications vary between *Bacteria*, *Archaea* and *Eucarya*. They vary even from one *Archaea* to the other. It should be pointed out that in *E. coli*, stem-loop structure 69 of domain IV was shown to be located at the active site of the 50S subunit. It is protected against nuclease degradation upon association of the 30S and 50S subunits [49] and, as shown by cross-linking of the *E. coli* 23S rRNA segment from position 1912 to 1920 (corresponding to nt 2055–2063 in *S. acidocaldarius*) to the decoding centre of 16S rRNA [50], it is juxtaposed to the decoding centre of the 30S subunit. These regions of 23S and 16S rRNAs are not only close in space, but also functionally related, since fidelity of codon recognition is disturbed upon mutation at the positions in *E. coli* 23S rRNA corresponding to positions 2057 and 2059 in the *S. acidocaldarius* 23S rRNA (Fig. 2A) [51]. In addition, chemical footprinting of 23S rRNA revealed a protection of the bases from stem-loop structure 69 by the tRNA located at the P site [52]. The strong phylogenetic conservation of one of the Ψ residues in the terminal loop of stem-loop structure 69 suggests a functional importance of this modification for ribosome function.

3.3. Four Ψ residues are found in *S. acidocaldarius* domain V 23S rRNA

The analysis of a large region of the *S. acidocaldarius* domain V reveals the presence of four pseudouridylation sites at positions 2479, 2535, 2550 and 2594 (Fig. 1B). The strong RT stops observed at the positions 2535, 2550 and 2594 suggest a complete conversion of U into Ψ . In contrast, based on the low intensity of the RT stop at position 2479 (Fig. 1B), the pseudouridylation at this position may be partial. As found

for domain IV, pseudouridylation sites in domain V are tightly clustered in *Bacterial* LSU rRNAs, whereas the numerous Ψ residues found in domain V of the *Eucarya* LSU rRNAs are rather dispersed in the entire domain (Fig. 2B) ([10], for review see [2]). Four Ψ residues were detected in domain V of the *E. coli* LSU rRNA. They are clustered in the area expected to play an important role for the peptidyl transferase activity [8]. Only one of these four Ψ residues is present in domain V of the *S. acidocaldarius* LSU rRNA (position 2594). Modification at this position is highly conserved in *Bacteria*, *Archaea* and *Eucarya*. This conserved Ψ residue is located at or very close to the peptidyl transferase centre. Indeed, it is found in stem-loop structure 89, proposed to be implicated in ribosomal subunit interaction, decoding function and peptidyl transferase activity ([51,53]; for review see [54,55]). It is possible that this conserved Ψ residue is implicated in one of these ribosomal functions.

The three other Ψ residues that we detected in the *S. acidocaldarius* LSU rRNA are located in another area of domain V (positions 2479–2550) (Figs. 1B and 2B). Based on studies in *E. coli*, this part of domain V is likely to be functionally important. It should be in close contact with the 3' end of the tRNA at the A-site, since the terminal loop of stem-loop structure 80 (Fig. 2B), containing a Gm residue in *E. coli*, was proposed to base pair with the CCA at the 3' extremity of tRNAs [56]. Two out of the three Ψ residues detected in this area of domain V are specific for the *S. acidocaldarius* LSU rRNA (positions 2479 and 2550) and one was found only in *D. melanogaster* LSU rRNA (position 2535) (see Fig. 2B).

3.4. Which kind of Ψ -synthases can form the detected Ψ residues?

The total number of Ψ residues detected in *S. acidocaldarius* LSU rRNA is quite similar to that found for *Bacteria* like *E. coli* or *B. subtilis*. However, their location shows differences as compared to that in *Bacteria*. Among the Ψ residues that we detected in the *S. acidocaldarius* LSU rRNA, two are highly conserved in *Bacteria* and *Eucarya* (positions 2594 and 2060), one of them was also detected in *H. halobium* (Ψ_{2060}) (for review see [2]). The third Ψ residue detected in the *S. acidocaldarius* LSU rRNA (Ψ_{2066}) is highly conserved in *Eucarya* but absent in *Bacteria*. Finally, among the three remaining ones, two have never been detected in any other species (Ψ_{2479} and Ψ_{2550}) and one (Ψ_{2535}) was detected only in *D. melanogaster*. In spite of this variability of modified positions, all the Ψ residues detected in *S. acidocaldarius* LSU rRNA are concentrated in the parts of the LSU rRNA domains IV and V, where Ψ residues were detected in the *Bacteria* and *Eucarya* LSU rRNAs. Hence, in all *Bacteria* and *Archaea* species studied, the same regions are the target of Ψ -synthase activities. However, the locations of modified positions are variable. How can this situation be explained in relation to Ψ -synthase specificities?

In *E. coli*, the three Ψ residues in loop 69 and its border (positions equivalent to positions 2054, 2058 and 2060 in *S. acidocaldarius*) are synthesised by the same Ψ -synthase RluD [14,16]. Since only one or two of these Ψ residues are found in *S. acidocaldarius* and *H. halobium*, respectively, the corresponding *Archaea* Ψ -synthases should differ in their substrate specificity as compared to the *E. coli* RluD enzyme. In *E. coli*, deletion of the *RLUD* gene was found to result in a severe

growth inhibition [16]. Based on our observation of the strong conservation of the modification at position 2060 (*S. acidocaldarius* numbering), the growth defect observed in the *E. coli* *rluD*-deleted strain may be due to the absence of Ψ formation at the position corresponding to Ψ_{2060} . However, one cannot exclude the possibility that RluD protein catalyses the formation of other essential Ψ residues or performs some other critically important functions unrelated to its Ψ -synthase activity.

In human, formation of Ψ residues in loop 69 is guided by the H/ACA snoRNA U19. This single RNA is predicted to direct Ψ formation at two sites equivalent to *S. acidocaldarius* U₂₀₅₈ and Ψ_{2060} [57]. Thus, if the formation of this residue in *S. acidocaldarius* is RNA-guided like in *Eucarya*, this RNA should lack the ability to direct Ψ synthesis at position 2058.

The formation of Ψ residue at position 2594 in *E. coli* LSU rRNA is predicted to be catalysed by the Ψ -synthase RluE [19] and is guided by the H/ACA snoRNAs snR34 in yeast or U65 in human. These two guide RNAs are also involved in the formation of a second Ψ residue at the position equivalent to U₂₆₄₈ in *S. acidocaldarius* [22]. Thus, again, if *S. acidocaldarius* Ψ formation in LSU rRNA was RNA-guided, the guide RNA would differ from that in *Eucarya*, directing Ψ formation only at one position. In spite of the strong conservation of the modification at position 2594, it should be noted that deletion of the snR34 gene has no effect on growth [58]. Taking individually post-transcriptional modifications of rRNAs were not found to be essential. However, they may bring selective advantages explaining their conservation.

The comparison of the Ψ sites common for *E. coli* and *S. acidocaldarius* shows that only one Ψ -synthase (RluE) may be conserved between *Bacteria* and *Archaea*. However, the homology search, based on the motif II sequence conserved in all the known Ψ -synthases [19], detected only two putative Ψ -synthase genes in *Archaea* genomes, in contrast to 10 enzymes in *E. coli*. In addition, the two *Archaea* proteins found in this way are members of the TruA or TruB families, acting on tRNAs at positions 38–40 and 55 respectively. No genes related to the RluA family, to which RluE belongs, were detected. In other words, based on the present data and current knowledge, the Ψ -synthase activities responsible for the Ψ formation in *Archaea* LSU rRNA seem to be different from those characterised in *Bacteria*. Either these activities are carried by purely protein enzymes that differ in specificity from those found in *Bacteria* or the specificity of the *Archaea* enzyme depends upon RNA guides. If such RNAs are present they must be quite less numerous than those found in *Eucarya*. Furthermore, they should be different among *Archaea*, since *H. halobium* and *S. acidocaldarius* share only one common modified position. Using the characteristic features of *Eucarya* snoRNA guides, one may now search for putative RNA guides for Ψ synthesis in the genomes of *S. acidocaldarius* and *H. halobium*.

Acknowledgements: This work was supported by laboratory funds from the Ministère de la Recherche et de l'Enseignement supérieur and the French Centre National de la Recherche Scientifique. S. Massenet and I. Ansmant are pre-doctoral fellows supported by a fellowship from the Ministère de la Recherche et de l'Enseignement supérieur. Y. Motorin is supported by an Associated Researcher position at the CNRS ('Poste Rouge'). P. Forterre is acknowledged for providing us with *S. acidocaldarius* cells.

References

- [1] Grosjean, H., Sprinzl, M. and Steinberg, S. (1995) *Biochimie* 77, 139–141.
- [2] Ofengand, J. and Fournier, M.J. (1998) in: *The Modification and Editing of RNA* (Grosjean, H. and Benne, R., Eds.), pp. 229–254. ASM Press, Washington, DC.
- [3] Massenet, S., Mougou, A. and Branlant, C. (1998) in: *The Modification and Editing of RNA* (Grosjean, H. and Benne, R., Eds.), pp. 201–228. ASM Press, Washington, DC.
- [4] Bachelier, J.P. and Cavallé, J. (1998) in: *The Modification and Editing of RNA* (Grosjean, H. and Benne, R., Eds.), pp. 255–272. ASM Press, Washington, DC.
- [5] Branlant, C., Krol, A., Machatt, M.A., Pouyet, J., Ebel, J.P., Edwards, K. and Kossel, H. (1981) *Nucleic Acids Res.* 9, 4303–4324.
- [6] Veldman, G.M., Klootwijk, J., de Regt, V.C., Planta, R.J., Branlant, C., Krol, A. and Ebel, J.P. (1981) *Nucleic Acids Res.* 9, 6935–6952.
- [7] Maden, B.E. (1990) *Prog. Nucleic Acid Res. Mol. Biol.* 39, 241–303.
- [8] Bakin, A. and Ofengand, J. (1993) *Biochemistry* 32, 9754–9762.
- [9] Bakin, A., Lane, B.G. and Ofengand, J. (1994) *Biochemistry* 33, 13475–13483.
- [10] Ofengand, J. and Bakin, A. (1997) *J. Mol. Biol.* 266, 246–268.
- [11] Brimacombe, R., Mitchell, P., Osswald, M., Stade, K. and Bochkarov, D. (1993) *FASEB J.* 7, 161–167.
- [12] Wrzesinski, J., Bakin, A., Nurse, K., Lane, B.G. and Ofengand, J. (1995) *Biochemistry* 34, 8904–8913.
- [13] Wrzesinski, J., Nurse, K., Bakin, A., Lane, B.G. and Ofengand, J. (1995) *RNA* 1, 437–448.
- [14] Huang, L., Ku, J., Pookanjanatavip, M., Gu, X., Wang, D., Greene, P.J. and Santi, D.V. (1998) *Biochemistry* 37, 15951–15957.
- [15] Conrad, J., Sun, D., Englund, N. and Ofengand, J. (1998) *J. Biol. Chem.* 273, 18562–18566.
- [16] Raychaudhuri, S., Conrad, J., Hall, B.G. and Ofengand, J. (1998) *RNA* 4, 1407–1417.
- [17] Raychaudhuri, S., Niu, L., Conrad, J., Lane, B.G. and Ofengand, J. (1999) *J. Biol. Chem.* 274, 18880–18886.
- [18] Niu, L. and Ofengand, J. (1999) *Biochemistry* 38, 629–635.
- [19] Conrad, J., Niu, L., Rudd, K., Lane, B.G. and Ofengand, J. (1999) *RNA* 5, 751–763.
- [20] Samarsky, D.A. and Fournier, M.J. (1999) *Nucleic Acids Res.* 27, 161–164.
- [21] Ni, J., Tien, A.L. and Fournier, M.J. (1997) *Cell* 89, 565–573.
- [22] Ganot, P., Bortolin, M.L. and Kiss, T. (1997) *Cell* 89, 799–809.
- [23] Lafontaine, D.L.J., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M. and Tollervey, D. (1998) *Genes Dev.* 12, 527–537.
- [24] Watkins, N.J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M. and Lüthmann, R. (1998) *RNA* 4, 1549–1568.
- [25] Becker, H.F., Motorin, Y., Planta, R.J. and Grosjean, H. (1997) *Nucleic Acids Res.* 25, 4493–4499.
- [26] Winkler, A.A., Bobok, A., Zonneveld, B.J., Steensma, H.Y. and Hooykaas, P.J. (1998) *Yeast* 14, 37–48.
- [27] Phillips, B., Billin, A.N., Cadwell, C., Buchholz, R., Erickson, C., Merriam, J.R., Carbon, J. and Poole, S.J. (1998) *Mol. Gen. Evol.* 260, 20–29.
- [28] Meier, U.T. and Blobel, G. (1994) *J. Cell. Biol.* 127, 1505–1514.
- [29] Heiss, N.S., Knight, S.W., Vulliamy, T.J., Klauck, S.M., Wiemann, S., Mason, P.J., Poustka, A. and Dokal, I. (1998) *Nature Genet.* 19, 32–38.
- [30] Brown, J.R. and Doolittle, W.F. (1997) *Microbiol. Mol. Biol. Rev.* 61, 456–502.
- [31] Lafontaine, D.L. and Tollervey, D. (1998) *Trends Biochem. Sci.* 23, 383–388.
- [32] Ochs, R.L., Lischwe, M.A., Spohn, W.H. and Busch, H. (1985) *Biol. Cell* 54, 123–133.
- [33] Schimmang, T., Tollervey, D., Kern, H., Frank, R. and Hurt, E.C. (1989) *EMBO J.* 8, 4015–4024.
- [34] Wu, P., Brockenbrough, J.S., Metcalfe, A.C., Chen, S. and Aris, J.P. (1998) *J. Biol. Chem.* 273, 16453–16463.

- [35] Lafontaine, D.L. and Tollervey, D. (1999) *RNA* 5, 455–467.
- [36] Amiri, K.A. (1994) *J. Bacteriol.* 176, 2124–2127.
- [37] Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A. and Gocayne, J.D. et al. (1996) *Science* 273, 1058–1073.
- [38] Klenk, H.P., Clayton, R.A., Tomb, J.F., White, O., Nelson, K.E., Ketchum, K.A., Dodson, R.J., Gwinn, M., Hickey, E.K. and Peterson, J.D. et al. (1997) *Nature* 390, 364–370.
- [39] Smith, D.R., Doucette-Stamm, L.A., Deloughery, C., Lee, H., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R. and Gilbert, K. et al. (1997) *J. Bacteriol.* 179, 7135–7155.
- [40] Noon, K.R., Bruenger, E. and McCloskey, J.A. (1998) *J. Bacteriol.* 180, 2883–2888.
- [41] Girard, J.P., Lehtonen, H., Caizergues-Ferrer, M., Amalric, F., Tollervey, D. and Lapeyre, B. (1992) *EMBO J.* 11, 673–682.
- [42] Balakin, A.G., Smith, L. and Fournier, M.J. (1996) *Cell* 86, 823–834.
- [43] Ganot, P., Caizergues-Ferrer, M. and Kiss, T. (1997) *Genes Dev.* 11, 941–956.
- [44] Aiba, H., Adhya, S. and de Crombrughe, B. (1981) *J. Biol. Chem.* 256, 11905–11910.
- [45] Massenet, S., Motorin, Y., Lafontaine, D.L., Hurt, E.C., Grosjean, H. and Branlant, C. (1999) *Mol. Cell. Biol.* 19, 2142–2154.
- [46] Mougin, A., Grégoire, A., Banroques, J., Ségault, V., Fournier, R., Brulé, F., Chevrier-Miller, M. and Branlant, C. (1996) *RNA* 2, 1079–1093.
- [47] Leffers, H., Kjems, J., Ostergaard, L., Larsen, N. and Garrett, R.A. (1987) *J. Mol. Biol.* 195, 43–61.
- [48] Kowalak, J.A., Bruenger, E. and McCloskey, J.A. (1995) *J. Biol. Chem.* 270, 17758–17764.
- [49] Merryman, C., Moazed, D., Daubresse, G. and Noller, H.F. (1999) *J. Mol. Biol.* 285, 107–113.
- [50] Mitchell, P., Osswald, M. and Brimacombe, R. (1992) *Biochemistry* 31, 3004–3011.
- [51] O'Connor, M. and Dahlberg, A.E. (1995) *J. Mol. Biol.* 254, 838–847.
- [52] Moazed, D. and Noller, H.F. (1989) *Cell* 57, 585–597.
- [53] Porse, B.T. and Garrett, R.A. (1995) *J. Mol. Biol.* 249, 1–10.
- [54] Green, R. and Noller, H.F. (1997) *Annu. Rev. Biochem.* 66, 679–716.
- [55] Noller, H.F. (1998) in: *RNA Structure and Function* (Simos, R.W. and Grunberg-Manago, M., Eds.), pp. 253–278, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [56] Samaha, R.R., Green, R. and Noller, H.F. (1995) *Nature* 377, 309–314.
- [57] Bortolin, M.L. and Kiss, T. (1998) *RNA* 4, 445–454.
- [58] Samarsky, D.A., Balakin, A.G. and Fournier, M.J. (1995) *Nucleic Acids Res.* 23, 2548–2554.
- [59] Durovic, P. and Dennis, P.P. (1994) *Mol. Microbiol.* 13, 229–242.

B/ DISCUSSION ET PERSPECTIVES

S.C.D. - I.I.H.P. NANCY 1
BIBLIOTHÈQUE DES SCIENCES
Rue du Jardin Botanique
54600 VILLERS-LES-NANCY

1. DES snoRNA GUIDES CHEZ LES *Archaea*?

Nous avons identifié 6 résidus Ψ dans les régions étudiées, des domaines IV et V de l'ARNr 23S de *S. acidocaldarius*. Ce nombre est comparable à ceux obtenus après identification des résidus Ψ , localisés dans les mêmes régions du grand ARNr d'une autre *Archaea*, *H. halobium* (3), et d'*E. coli* (5) (Ofengand et Bakin, 1997; pour revue, Ofengand et Fournier, 1998). Par contre, les régions correspondantes des ARNr 26-28S contiennent beaucoup plus de résidus Ψ chez les eucaryotes, à savoir 14 résidus chez *S. cerevisiae*, 24 chez l'homme ou encore, 25 chez *M. musculus* (pour revue, Ofengand et Fournier, 1998). Les nombres de résidus Ψ que portent respectivement les ARNr 16S et 23S de *S. solfataricus*, estimés par des méthodes combinées de chromatographie et de spectrométrie de masse, sont également comparables à ceux trouvés pour les bactéries (Noon *et al.*, 1998). Ces observations et, comme nous l'avons dit dans l'introduction de ce chapitre, l'absence pour le moment de détection de protéines homologues à Gar1p chez les *Archaea*, ne sont pas en faveur de l'existence d'un système d'ARN guides de pseudouridylation des ARNr d'*Archaea*, comparable à celui des eucaryotes. Cependant, l'hypothèse de l'existence de ce type d'ARN guides ne peut pas être complètement rejetée sur la base du nombre de modifications détectées. En effet, on ne peut pas exclure l'existence d'un système qui comporterait peu d'ARN guides, associés à des protéines différentes de celles trouvées chez les eucaryotes. L'existence, chez les *Archaea*, d'une Ψ -synthase homologue à Cbf5p est, en effet, en faveur de cette hypothèse. Si l'homologue de Cbf5p chez les *Archaea* est responsable de la formation des résidus Ψ dans les ARNr, est-ce que cette enzyme catalyse également la formation du résidu Ψ_{55} dans les ARNt? Si ce n'est pas le cas, il faut alors imaginer que ce résidu Ψ est formé chez les *Archaea* par une autre Ψ -synthase n'appartenant pas au groupe TruB/Pus4p/Cbf5p.

Contrairement à ce qui se passe pour les résidus Ψ , le nombre de 2'-O-méthylations dans les ARNr de *S. solfataricus*, est élevé et voisin de celui trouvé chez les eucaryotes. En effet, ce nombre a été estimé à 44 (Noon *et al.*, 1998) (contre 4 chez *E. coli* et 55 chez *S. cerevisiae*, pour revue, Bachellerie et Cavallé, 1998). L'existence d'ARN guides est donc plus probable dans ce cas. La découverte de gènes chez les *Archaea*, codant des ARN présentant de fortes homologues avec les snoRNA C/D des eucaryotes, a d'ailleurs été rapportée très récemment, lors du congrès RNA 99 à Edinbourg (Lowe et Eddy, résultats non publiés).

De plus, les snoRNP C/D des eucaryotes contiennent la protéine Nop1p (chez la levure) ou fibrillarine (chez l'homme). Des homologues de cette protéine ont été identifiés chez des *Archaea* méthanogènes thermophiles (Amiri, 1994; Bult *et al.*, 1996; Smith *et al.*, 1997). Des homologues du gène codant la protéine Nop58p de *S. cerevisiae*, également associée aux snoRNP C/D, ont aussi été détectés dans le génome de plusieurs *Archaea* méthanogènes thermophiles et dans une *Archaea* hyperthermophile, *Archeoglobus fulgidus* (Bult *et al.*, 1996; Klenk *et al.*, 1997; Smith *et al.*, 1997).

Actuellement, les données suggérant l'existence d'un système de méthylation des ARNr, basé sur l'utilisation d'ARN guides, ne sont disponibles que pour des *Archaea* thermophiles et hyperthermophiles. Cependant, les espèces concernées appartiennent aux deux grands règnes des *Archaea*: les *euryarchaeota* et les *crenarchaeota*. Des études devront être menées chez les *Archaea* halophiles ou mésophiles, afin de savoir si la présence d'ARN guides de méthylation peut être générale chez les *Archaea*. En effet, comme nous allons en discuter, la 2'-O-méthylation des riboses présente un intérêt particulier pour les organismes se développant à température élevée.

2. LA METHYLATION SUR LE RIBOSE, UNE MODIFICATION POST-TRANSCRIPTIONNELLE SELECTIONNEE PAR LES ORGANISMES SE DEVELOPPANT A TEMPERATURE ELEVEE?

La question qui se pose, en effet, est de savoir pourquoi les *Archaea* auraient développé un système, basé sur des ARN guides, pour méthyler leurs ARNr. La réponse la plus simple est que, chez ces organismes, la présence de nombreuses 2'-O-méthylations apporte un avantage sélectif aux cellules. Ce type de modification pourrait être une stratégie pour stabiliser les ARN d'organismes se développant à haute température, au même titre que les acétylations en position 4 des bases ou l'addition d'un soufre en position 2 (Edmonds *et al.*, 1991; Kowalak *et al.*, 1994; pour revue, Davis, 1998). Les modifications dans les ARNr d'*Archaea* augmentent la température de demi-dénaturation (T_m) de ces ARN d'environ 14 à 18°C (Kowalak *et al.*, 1994), alors que les modifications portées par les ARNr d'*E. coli* ou *S. cerevisiae* conduisent seulement à une augmentation du T_m de 4 à 5°C (Sampson et Uhlenbeck, 1988; Perret *et al.*, 1990a; Derrick et Horowitz, 1993). Ceci pourrait être, en partie, dû à la présence de certains résidus hypermodifiés, à la fois sur la base et le ribose (les résidus m^2_2Gm , m^5Cm , m^1Im ou ac^4Cm , par exemple), comme cela a été observé dans les ARNr d'*Archaea* thermophiles et hyperthermophiles (Edmonds *et al.*, 1991). Le taux de certains de ces résidus augmente même en fonction de l'augmentation de la température de croissance (Kowalak *et al.*, 1994), ce qui est également le cas pour le taux de 2' O-

méthylations dans les ARNr 16S et 23S de *S. solfataricus* (Noon *et al.*, 1998). Un effet similaire a aussi été observé pour les méthylations sur le ribose et les résidus m⁵s²U portés respectivement par les ARNt de bactéries thermophiles, *B. stearothermophilus* (Agris *et al.*, 1973) et *T. thermophilus* (Watanabe *et al.*, 1976).

La présence de méthylations sur les riboses augmente la rigidité des hélices d'ARN. C'est également le cas pour d'autres modifications post-transcriptionnelles comme les résidus Ψ (pour revue, Davis, 1998). Mais, en plus, la présence d'un grand nombre de méthylations sur les riboses, chez des organismes thermophiles et hyperthermophiles, pourrait apporter une stabilisation chimique. En effet, les 2'-O-méthylations limitent l'hydrolyse des liaisons phosphodiesters à haute température, qui requiert un 2'OH libre.

On peut donc se demander si un système d'ARN guides, permettant de méthyler de nombreuses positions avec une seule enzyme, n'est pas apparu tôt au cours de l'évolution chez des organismes se développant à température élevée. Ce système pourrait très bien être ancien et avoir été perdu par les bactéries se développant à température modérée.

*PARTIE 2: LOCALISATION DES RESIDUS Ψ DANS LES UsnRNA DE LA LEVURE *S. cerevisiae* ET DANS LES UsnRNA DES SPLICEOSOMES U12-DEPENDANTS DE CELLULES HeLa. IDENTIFICATION CHEZ *S. cerevisiae* D'UNE PSEUDOURIDINE-SYNTHASE PRESENTANT UNE DOUBLE SPECIFICITE DE SUBSTRATS.*

A/ INTRODUCTION ET RESULTATS

1. IDENTIFICATION DES RESIDUS Ψ PORTES PAR LES UsnRNA DE *S. cerevisiae* ET DE LA PREMIERE Ψ -SYNTHASE AGISSANT SUR LES UsnRNA

1.1. Localisation des résidus Ψ dans les UsnRNA de *S. cerevisiae*.

Comme nous l'avons mentionné dans l'introduction bibliographique, lorsque nous avons commencé ce travail, seules étaient connues les modifications post-transcriptionnelles portées par les UsnRNA de quelques espèces de vertébrés, de plantes et, dans le cas du snRNA U5, de certains eucaryotes inférieurs. Aucune étude n'avait été menée chez les levures, et en particulier chez *S. cerevisiae*, alors que cet organisme est très utilisé dans l'étude de la structure et du mécanisme d'action du spliceosome. Par ailleurs, comme nous l'avons décrit dans la revue insérée dans l'introduction de ce travail, les nucléotides modifiés des UsnRNA sont, pour la plupart, regroupés dans des régions fonctionnelles, en particulier la boucle terminale du snRNA U5 qui aligne les bordures exoniques, les régions du snRNA U2 qui s'apparient à la boîte de branchement des ARN pré-messagers ou au snRNA U6, la région 5'-terminale du snRNA U1 qui s'apparie au site 5' d'épissage et les régions du snRNA U6 formant des appariements de bases avec le snRNA U2. Il nous a donc paru important d'étudier les modifications post-transcriptionnelles portées par les UsnRNA de la levure *S. cerevisiae*.

L'absence d'études antérieures sur les nucléotides modifiés de cet organisme pouvait s'expliquer par les difficultés expérimentales que présentait cette analyse. Une première difficulté était liée au fait que les UsnRNA sont présents en très faible quantité chez les levures, ceci du fait que très peu de gènes nucléaires contiennent des introns. Une deuxième difficulté était la taille exceptionnellement grande des snRNA U1 et U2 chez *S. cerevisiae*. La méthode d'identification des résidus Ψ par le CMCT, que l'équipe de J. Ofengand venait de mettre au point (Bakin et Ofengand, 1993), pouvait être appliquée en dépit de ces difficultés. Nous avons décrit le principe de cette méthode dans la partie précédente des résultats. L'avantage de cette technique, basée sur une analyse par extension d'amorce, est que l'on peut travailler sur un mélange d'ARN et analyser des régions définies, au sein d'un grand ARN, par sélection d'amorces appropriées.

Afin de confirmer les résultats obtenus par cette approche, j'ai utilisé une deuxième technique, basée sur l'action différentielle de l'hydrazine sur les résidus U et Ψ . En effet, l'hydrazine s'additionne sur la double liaison carbone-carbone en 5-6 des uridines. Ceci

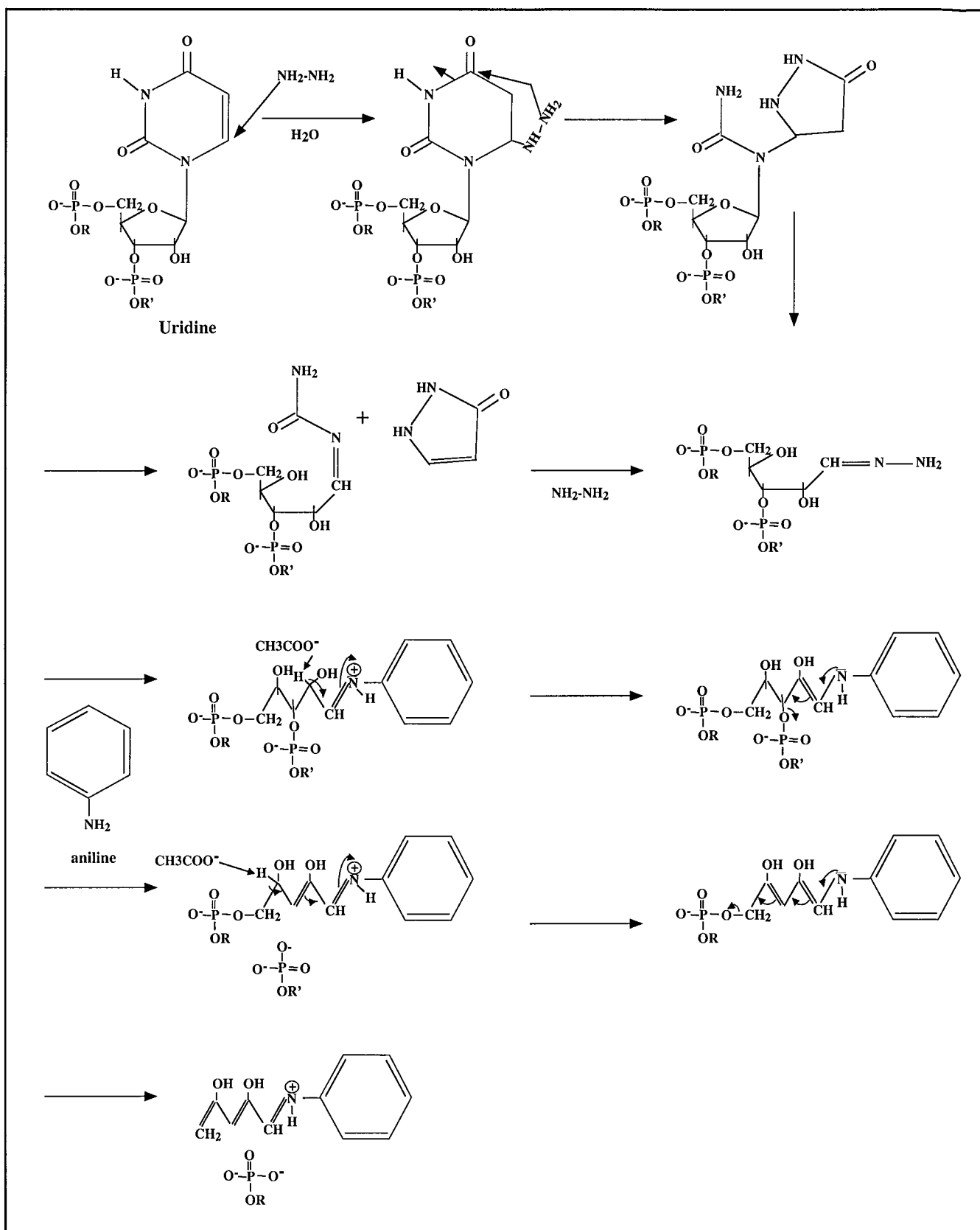


Figure 44: Mécanisme d'action de l'hydrazine/aniline en milieu aqueux.

conduit à une ouverture de la base qui est éliminée. L'aniline peut ensuite induire une β -élimination, avec rupture des liaisons phosphodiester en 3' et en 5' du nucléotide modifié (Fig. 44). Le fait que le carbone en position 5 du résidu Ψ soit lié au ribose limite la réactivité de la double liaison 5-6 vis-à-vis de l'hydrazine, et donc n'entraîne pas une élimination du nucléotide modifié par β -élimination. Cette absence de coupure de l'ARN au niveau des résidus Ψ , observée après traitement successif à l'hydrazine et à l'aniline d'un ARN marqué à son extrémité 3', avait déjà été utilisée par C. Branlant lors de l'étude de la séquence et des modifications post-transcriptionnelles de l'ARNr 23S d'*E. coli* (Branlant *et al.*, 1981b). Un marquage 3'-terminal des ARN nécessitait de travailler sur des ARN de petite taille et purifiés à homogénéité. L'application était donc difficile dans le cas des UsnRNA de *S. cerevisiae*. J'ai donc essayé de mettre en évidence les coupures produites par le traitement hydrazine/aniline par la méthode d'extension d'amorce par la transcriptase inverse. En fait, nous avons constaté que la modification à l'hydrazine altère assez fortement la base pour conduire à un arrêt de la transcriptase inverse, ce qui permet de distinguer les résidus Ψ des résidus U, car aucun arrêt n'est observé au niveau des résidus Ψ non modifiés par l'hydrazine. Il faut noter qu'une absence de réactivité à l'hydrazine est aussi observée pour les résidus m^5U (Bakin et Ofengand, 1993). La méthode basée sur l'utilisation de l'hydrazine, qui est plus difficile à mettre en œuvre et ne donne pas toujours des résultats d'excellente qualité, doit être considérée comme un complément de la méthode au CMCT. Il n'est, en effet, pas toujours facile d'obtenir une modification statistique des ARN, tous les résidus U n'étant pas modifié par l'hydrazine avec la même efficacité. Il est également parfois difficile, en utilisant cette approche, d'obtenir une analyse correcte par la transcriptase inverse. La technique basée sur l'action de l'hydrazine est donc une méthode complémentaire pour confirmer la présence de résidus Ψ dans les ARN et je l'ai utilisée comme telle.

Nous avons réalisé une analyse des résidus Ψ sur la quasi totalité des snRNA U4, U5 et U6 de *S. cerevisiae*, à l'exception d'une trentaine de nucléotides à leur extrémité 3'-terminale, région permettant l'hybridation de l'amorce de la transcriptase inverse. Les régions 3'-terminales ne contiennent pas de résidus Ψ dans les UsnRNA de vertébrés et de plantes. Nous avons donc postulé qu'il pouvait en être de même chez *S. cerevisiae*. En ce qui concerne les snRNA U1 et U2, vu leur grande taille (568 et 1175 nts, respectivement), nous avons restreint l'étude à leur extrémité 5'-terminale. De nouveau, nous avons formulé l'hypothèse que les modifications post-transcriptionnelles dans ces deux UsnRNA de *S. cerevisiae* devaient être localisées dans les régions où des modifications ont été mises en évidence dans les UsnRNA de vertébrés et plantes. Dans le cas des snRNA U2 de vertébrés, tous les résidus Ψ sont localisés dans les 91 premiers nucléotides en 5' de cet ARN. Les snRNA U1 de vertébrés portent seulement deux résidus Ψ aux positions 5 et 6. Nous avons donc analysé les

régions correspondantes des snRNA U1 et U2 de *S. cerevisiae*, à savoir les 50 nucléotides en 5' dans le cas du snRNA U1 et les 100 nucléotides en 5' dans le cas du snRNA U2 (Fig. 1 et 2 de la publication n°2). Ces régions avaient d'ailleurs été identifiées comme les régions fonctionnelles de ces ARN chez *S. cerevisiae*.

Notre étude a permis de montrer que le nombre de résidus Ψ est beaucoup plus faible dans les UsnRNA de *S. cerevisiae* que dans les UsnRNA de plantes et de vertébrés et même de certains eucaryotes unicellulaires tels que *T. pyriformis*. Les résidus Ψ présents sont localisés dans des séquences de haute importance fonctionnelle. En effet, deux résidus Ψ sont présents dans la région 5'-terminale du snRNA U1 qui s'apparie au site 5' des introns et trois dans la région du snRNA U2 qui s'apparie avec la boîte de branchement et la région immédiatement en aval qui fixe des facteurs d'épissage. Enfin, un résidu Ψ est localisé dans la boucle terminale du snRNA U5 de *S. cerevisiae* qui sert à aligner les exons. Notre observation d'un nombre faible de résidus Ψ dans les UsnRNA de *S. cerevisiae*, est en accord avec l'observation d'un nombre faible de résidus Ψ dans les UsnRNA d'une autre levure, *S. pombe*. Pour cette levure, les résidus modifiés ont été analysés par l'équipe de Reddy, en parallèle à notre étude (Gu *et al.*, 1996).

1.2. Pus1p: une Ψ -synthase à double spécificité de substrats.

Ayant identifié les résidus Ψ dans les UsnRNA de *S. cerevisiae*, nous étions alors en mesure de rechercher les Ψ -synthases responsables de la formation de ces modifications. Cette identification présentait un intérêt double. D'une part, aucune Ψ -synthase spécifique des UsnRNA n'avait été caractérisée, d'autre part, une approche possible, pour étudier le rôle des nucléotides modifiés dans l'épissage, est d'empêcher la formation de ces résidus de manière individuelle et d'observer le phénotype qui en résulte. Si la formation de chacun des résidus Ψ dans les UsnRNA est catalysée par une Ψ -synthase spécifique, un moyen d'atteindre ce but est de détruire spécifiquement les gènes codant chacune de ces enzymes. La levure est parfaitement adaptée au développement de ce type d'approche puisque, dans cet organisme, il est possible de réaliser des délétions de gènes de manière rapide et simple.

Comme nous l'avons déjà mentionné, 5 Ψ -synthases avaient été caractérisées chez *S. cerevisiae* (Tableau 10 de l'introduction). Trois d'entre elles sont spécifiques des ARNt (Pus1p, Pus3p et Pus4p) (Simos *et al.*, 1996; Becker *et al.*, 1997a; Lecointe *et al.*, 1998; Motorin *et al.*, 1998), Cbf5p modifie les ARNr (Lafontaine *et al.*, 1998a) et le (ou les) substrat(s) de la cinquième Ψ -synthase putative, Pus2p, ne sont pas encore identifié(s) (Simos *et al.*, 1996). Des données de la littérature montraient, sur la base d'expériences réalisées *in vitro*, que deux enzymes de modification des ARN présentent probablement une double

spécificité de substrats. En effet, la Ψ -synthase RluA d'*E. coli* modifie *in vitro* la position 32 dans les ARNt et la position 746 dans l'ARNr 23S (Wrzesinski *et al.*, 1995a) et la m⁵U méthyltransférase d'*E. coli* (RUMT) catalyse *in vitro* la formation des nucléotides modifiés m⁵U₅₄ dans les ARNt et m⁵U₇₈₈ dans l'ARNr 16S (Gu *et al.*, 1994). Cependant, aucune démonstration expérimentale de l'existence de cette double spécificité de substrats *in vivo* n'était disponible. Néanmoins, ces observations ouvraient la possibilité que certaines enzymes de modification des ARN puissent avoir une multispécificité de substrats. Nous avons donc, dans un premier temps, testé si des Ψ -synthases déjà caractérisées chez *S. cerevisiae*, et modifiant soit les ARNt soit les ARNr, pouvaient également agir sur les UsnRNA.

Pour cela, nous avons analysé le contenu en résidus Ψ des UsnRNA issus de différentes souches de *S. cerevisiae*. Quatre de ces souches présentaient une délétion spécifique d'un gène codant une des Ψ -synthases non essentielles à la croissance cellulaire, Pus1p, Pus2p, Pus3p et Pus4p. Dans la cinquième souche utilisée, le gène codant la Ψ -synthase Cbf5p, essentielle à la croissance, avait été placé sous contrôle du promoteur Gal afin de réguler son expression. Ce travail a été réalisé en collaboration avec plusieurs équipes. L'équipe de H. Hurt (Heidelberg, Allemagne) nous a fourni les souches de *S. cerevisiae* présentant une délétion des gènes *PUS1*, *PUS2* et *PUS3* et celle de R. Planta (Amsterdam, Pays-Bas), une souche présentant une délétion du gène *PUS4*. L'équipe de D. Tollervey (Edinbourg, Ecosse) (D. Lafontaine) nous a fourni les ARN issus de la souche où l'expression du gène *CBF5* avait été bloquée par croissance sur glucose. Enfin, l'équipe d'H. Grosjean (Gif-sur-Yvette) nous a fourni l'enzyme Pus1p purifiée, après production dans *E. coli*.

Nous avons ainsi pu montrer que, parmi les Ψ -synthases connues de *S. cerevisiae*, seule la Ψ -synthase Pus1p est impliquée dans la modification des UsnRNA. Elle modifie, aussi bien *in vivo* qu'*in vitro*, le snRNA U2 et ceci en plus de son action sur les ARNt. C'était la première mise en évidence, *in vivo*, d'une double spécificité de substrats d'une enzyme de modification des ARN. *In vitro*, l'enzyme Pus1p purifiée par Y. Motorin ne forme qu'un seul résidu Ψ dans la région du snRNA U2 que nous avons analysée, et ce résidu est le même que celui qui est modifié *in vivo* par cette enzyme. L'expérience a été réalisée sur un ARN correspondant au snRNA U2 entier (1175 nts), que nous avons produit, par transcription *in vitro*, par l'ARN T7 polymérase, d'un plasmide que nous avait fourni P. Fabrizio (Marburg, Allemagne). L'enzyme Pus1p avait été produite en grande quantité chez *E. coli* et contenait une étiquette poly-histidine, ce qui avait permis sa purification sur colonne d'agarose Ni-NTA (Arluisson *et al.*, 1998). Le mélange protéique ainsi obtenu avait ensuite été fractionné sur une colonne échangeuse d'anions, afin de permettre l'élimination d'éventuels ARN fixés à l'enzyme. Les conditions d'utilisation de l'enzyme *in vitro* avaient été établies par l'équipe d'H. Grosjean (Motorin *et al.*, 1998).

L'ensemble des résultats sur la localisation des résidus Ψ dans les UsnRNA de *S. cerevisiae*, ainsi que la mise en évidence de la double spécificité de substrats de la Ψ -synthase Pus1p, sont présentés dans la publication n°2.

Publication n°2

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.

Massenet S., Motorin Y., Lafontaine D.L., Hurt E.C., Grosjean H. and Branlant C.
Mol. Cell. Biol. 19, 2142-2154, 1999.

Publication n°2

Pseudouridine mapping in the *Saccharomyces cerevisiae* spliceosomal U small nuclear RNA (snRNAs) reveals that pseudouridine synthase Pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA.

S. Massenet, Y. Motorin, D. L. J. Lafontaine, E. C. Hurt, H. Grosjean and C. Branlant.

Mol. Cell. Biol. 19, 2142-2154, 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

SÉVERINE MASSENET,¹ YURI MOTORIN,² DENIS L. J. LAFONTAINE,³ EDUARD C. HURT,⁴
HENRI GROSJEAN,² AND CHRISTIANE BRANLANT^{1*}

Laboratoire de Maturation des ARN et Enzymologie Moléculaire, UMR7567 CNRS-UHP, Faculté des Sciences, 54506 Vandoeuvre-les-Nancy Cédex,¹ and Laboratoire d'Enzymologie et Biochimie Structurales, UPR CNRS, 91198 Gif-sur-Yvette,² France; Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh EH9 3JR, United Kingdom³; and University of Heidelberg, 69120 Heidelberg, Germany⁴

Received 24 August 1998/Returned for modification 5 October 1998/Accepted 30 November 1998

Pseudouridine (Ψ) residues were localized in the *Saccharomyces cerevisiae* spliceosomal U small nuclear RNAs (UsnRNAs) by using the chemical mapping method. In contrast to vertebrate UsnRNAs, *S. cerevisiae* UsnRNAs contain only a few Ψ residues, which are located in segments involved in intermolecular RNA-RNA or RNA-protein interactions. At these positions, UsnRNAs are universally modified. When yeast mutants disrupted for one of the several pseudouridine synthase genes (*PUS1*, *PUS2*, *PUS3*, and *PUS4*) or depleted in rRNA-pseudouridine synthase Cbf5p were tested for UsnRNA Ψ content, only the loss of the Pus1p activity was found to affect Ψ formation in spliceosomal UsnRNAs. Indeed, Ψ_{44} formation in U2 snRNA was abolished. By using purified Pus1p enzyme and in vitro-produced U2 snRNA, Pus1p is shown here to catalyze Ψ_{44} formation in the *S. cerevisiae* U2 snRNA. Thus, Pus1p is the first UsnRNA pseudouridine synthase characterized so far which exhibits a dual substrate specificity, acting on both tRNAs and U2 snRNA. As depletion of rRNA-pseudouridine synthase Cbf5p had no effect on UsnRNA Ψ content, formation of Ψ residues in *S. cerevisiae* UsnRNAs is not dependent on the Cbf5p-snoRNA guided mechanism.

Introns are universally present in the nuclear genes transcribed by RNA polymerase II. Introns with GU and AG terminal dinucleotides and some introns with AU and AC terminal dinucleotides are removed by spliceosomal complexes containing the U1, U2, U4, U5, and U6 small nuclear RNAs (UsnRNAs) (for reviews, see references 54 and 59), the remaining part of introns with AU and AC terminal dinucleotides being excised by complexes containing the U11, U12, U4atac, U5, and U6atac UsnRNAs (32, 79, 106, 105). In yeast cells, only introns with GU and AG borders have been detected, and their excision is catalyzed by ribonucleoprotein complexes containing UsnRNAs homologous to the vertebrate U1, U2, U4, U5, and U6 snRNAs (for a review, see reference 31). However, compared to their counterparts in other eukaryotes, the *Saccharomyces cerevisiae* spliceosomal UsnRNAs differ by their larger size. For example, U2 snRNA is 1,175 nucleotides (nt) long in *S. cerevisiae* versus 187 nt in humans, and U1 snRNA is 568 nt long in *S. cerevisiae* versus 164 nt in humans (for a review, see reference 31).

In spite of this difference, the splicing machineries for the elimination of the GU-AG type of introns, in both vertebrates and *S. cerevisiae*, share several common properties. In particular, UsnRNPs are assembled in the same sequential order (17, 27; for a review, see reference 59), and the same kinds of bi- and multimolecular RNA-RNA interactions are implicated. The picture that now emerges from a large body of experi-

ments in several laboratories is rather complex. First, upon U1 snRNP association, the 5' extremity of U1 forms a base-pair interaction with the intron 5' extremity (62, 94, 95, 96, 123). Then, the U2 snRNP is associated and a base-pair interaction is formed between U2 and the intron branch-point sequence (71, 72, 116, 124). A U4/U6 RNA duplex is present in the U4/U6 snRNP (34, 86) and a tri-snRNP is generated by association of the U4/U6 snRNP with the U5 snRNP (11). When this tri-snRNP particle joins the prespliceosomal complex, several conformational changes take place and the U1 snRNA interaction at the 5' end of the intron is replaced by a U6 snRNA interaction (41, 52, 91). The U4/U6 RNA duplex is disrupted and replaced by a U2/U6 RNA duplex (20, 35, 53, 102, 117), and the terminal loop I of U5 contacts the 3' extremity of the upstream exon (19, 65, 66, 67, 99, 118). These structural rearrangements are required for the first *trans*-esterification step to occur. Following this first step, other structural rearrangements take place that reveal the catalytic activity for the second step of the reaction. In particular, the highly conserved terminal loop I of U5 then interacts with the two exon extremities for proper alignment and ligation (65-67, 99). Although several proteins play an essential role in spliceosome assembly and function (for reviews, see references 45, 109, and 114), the general idea is that some of the UsnRNAs, in particular U2 and U6, may be directly involved in catalysis (for reviews, see references 18, 30, 40, 101).

The first determinations of UsnRNA sequences were made at the RNA level, and several posttranscriptional modifications were identified. This analysis was done on U1, U2, U4, U5, and U6 snRNAs from HeLa, chicken, mouse (13, 14, 33, 42, 47, 48), rat hepatoma (for a review, see reference 81), *Drosophila melanogaster* (63), and plant (for a review, reference 98) cells.

* Corresponding author. Mailing address: Laboratoire de Maturation des ARN et Enzymologie Moléculaire, UMR7567 CNRS-UHP Nancy I, Faculté des Sciences, BP 239, 54506 Vandoeuvre-les-Nancy Cédex, France. Phone: 33-3-83-91-20-92. Fax: 33-3-83-91-20-93. E-mail: cbranlant@scbim.u-nancy.fr.

Then, for a long period of time, UsnRNA sequences were deduced essentially from the corresponding gene sequences so that posttranscriptional modifications were only investigated for a limited number of UsnRNAs: the *Physarum polycephalum* U1 and U5 snRNAs (63, 103) and, more recently, the *Schizosaccharomyces pombe* U1, U2, U4, U5, and U6 snRNAs (28).

Except for the cap and cap-related modifications, the two most frequently found posttranscriptional modifications at internal positions of UsnRNAs are methylation at the 2'-*O* position of ribose and isomerization of uridine to pseudouridine (Ψ). Only a few base methylations (m^3C , m^6A , and m^2G) were detected (for a review, see reference 55). Since Ψ residues and 2'-*O*-methylated residues stabilize RNA double helices (for reviews, see references 2, 6, and 21), the functional importance of these modified nucleotides in UsnRNAs may be, at least in part, linked to the necessity to form base-pair interactions between RNA molecules at one or the other steps of the splicing process. Furthermore, in the case of Ψ residues, the presence of an additional free NH group at position 3 of the ring, as compared to uridine, generates the possibility to form an additional hydrogen bond with RNA or proteins. It is noteworthy that the modified residues of spliceosomal UsnRNAs are clustered in the segments involved in intermolecular interactions, and several of these modifications are conserved in all of the species that were studied so far (103; for a review, see reference 55). This is the case for one of the two Ψ residues located in the U1 region that base pairs with the intron, for the two Ψ residues present in the U2 region, which interacts with the branch site, and for the numerous posttranscriptional modifications of the U5 snRNA terminal loop I (28, 103).

The number of posttranscriptionally modified nucleotides present in the interacting regions of U6 and U2 snRNAs, supposed to form at least a part of the spliceosome active site, is rather impressive in vertebrate species. All of these observations strongly suggest an important role of UsnRNA posttranscriptional modifications in spliceosome assembly and function. An experimental evidence for a role of posttranscriptional modifications in splicing was obtained for the human U2 snRNA (93, 120). Indeed, whereas fully active U2 snRNPs were obtained upon *in vitro* reconstitution with HeLa cell U2 snRNP proteins and the authentic human U2 snRNA, the *in vitro*-produced human U2 snRNA lacking all of the posttranscriptional modifications failed to form functional splicing complexes (93). Replacement of the authentic *Xenopus laevis* U2 snRNA by chimeric U2 snRNAs, in which some sequences are from cellular-derived U2 and others are from *in vitro*-transcribed U2, demonstrated that the essential posttranscriptional modifications of the vertebrate U2 snRNA are restricted to the 27-nt 5' terminus (120).

The biogenesis of Ψ in U1, U2, U4, U5, and U6 snRNAs from HeLa cells have been the subject of several reports (121; reference 122 and references therein; for a review, see reference 76). Using *in vitro*-transcribed snRNAs and nuclear or S100 extracts from HeLa cells, the existence of multiple RNA-pseudouridine synthase activities that specifically recognize U1, U2, and U5 snRNAs was demonstrated (73–75, 77). However, to date, none of the implicated RNA-pseudouridine synthases has been identified, nor was their precise specificity elucidated.

The splicing machineries of HeLa cells and *S. cerevisiae* are the two most extensively studied, with respect to understanding how the spliceosome is assembled and how it functions. However, in the case of *S. cerevisiae* snRNAs, due to the low amounts of spliceosomal UsnRNAs and their unusual lengths, no internal nucleotide modifications have been identified so far. We started such a study by investigating the presence of Ψ

residues in the *S. cerevisiae* full-length U4, U5, and U6 snRNAs and in the regions of the *S. cerevisiae* U1 and U2 snRNAs that have a counterpart with modified residues in the vertebrate U1 and U2 snRNAs. We also looked for the pseudouridine synthases that may catalyze the formation of the identified residues. Based on *in vitro* experiments, a case of dual-specificity was already described for an *E. coli* RNA-pseudouridine synthase (the RluAp enzyme that catalyzes the site-specific formation of Ψ at position 32 of tRNA anticodon and at position 746 of 23 S rRNA) (115). We therefore asked whether some of the already characterized yeast pseudouridine synthases, which act on tRNAs or rRNAs, can also modify UsnRNAs. To this end, we performed the mapping of the Ψ residues in UsnRNAs extracted from yeast strains carrying separate disruptions of the genes coding for the already-characterized RNA-pseudouridine synthases Pus1p, Pus3p, and Pus4p acting on tRNAs (10, 51, 60, 97); the putative Pus2p enzyme, whose substrate has not been identified so far (97); and Cbf5p acting on rRNA complexed with snoRNA guides (50).

In this study, we report the mapping of Ψ residues in the *S. cerevisiae* spliceosomal UsnRNAs and demonstrate that one of the previously identified tRNA-pseudouridine synthases (Pus1p) is directly implicated in the pseudouridylation of U2 snRNA *in vivo*.

MATERIALS AND METHODS

Yeast strains and growth conditions. The following *S. cerevisiae* strains were used in this study: *S. cerevisiae* FL100 (ATCC 28383) (49); *S. cerevisiae* strains carrying disruptions in the *PUS1*, *PUS2*, and *PUS3* genes that were described previously (51, 97); an *S. cerevisiae* strain with a disrupted *PUS4* gene that was kindly provided by R. Planta (University of Amsterdam) (10); and an *S. cerevisiae* strain carrying a deletion in the *PUS1* gene transformed with a plasmid expressing an active recombinant protein ProtA-Pus1p (pUN100-*PUS1*) (97). All of these *S. cerevisiae* strains were grown at 30°C on YPD liquid medium (1% [wt/vol] yeast extract, 1% [wt/vol] Bacto Peptone, and 2% [wt/vol] glucose). The essential gene *CBF5* was fused on the chromosome to a *GAL*-repressible promoter (50). Transcription driven from *GAL*-regulated promoters is strongly repressed when strains are grown on glucose medium, allowing the effects of depletion of essential proteins to be monitored. For depletion of Cbf5p, cells growing exponentially in permissive conditions (2% galactose, 2% sucrose, and 2% raffinose minimal medium) at 30°C were harvested by centrifugation, washed, and resuspended in 2% glucose minimal medium. During growth, cells were diluted with prewarmed medium and constantly maintained in exponential phase. For the experiment described in Fig. 4, two independently isolated *GAL::cbf5* strains (YDL521-1 and YDL521-3) (50) were used. RNA was extracted from these strains after transfer to nonpermissive conditions for up to 70 h. At this time point of transfer, Ψ formation in rRNA is strongly inhibited and no H+ACA snoRNA were detected (50).

Preparation of RNA from *S. cerevisiae*. The soluble RNA fraction (containing mainly tRNAs and snRNAs) from wild-type, disrupted yeast strains and the *GAL::cbf5* strain was prepared as follow. Cells grown on YPD liquid medium until the late stationary phase were harvested by centrifugation and resuspended in twice their volume of lysis buffer containing 50 mM Tris-HCl (pH 7.5), 10 mM MgCl₂, 100 mM KCl, 0.1 mM EDTA, 10% glycerol, and 10 mM β -mercaptoethanol. The suspension was frozen in dry ice and passed through a French Press at about 3,000 lb/in². The homogenate obtained was centrifuged for 10 min at 10,000 \times g at 4°C. The supernatant was further centrifuged in a high-speed centrifuge TL100 (Beckman) for 2 h at 80,000 rpm and 4°C. The resulting S100 extract was successively treated by equal volumes of phenol, phenol-chloroform (1:1), and chloroform-isoamyl alcohol (24:1). The extracted RNA was ethanol precipitated and used for further analysis by reverse transcription or Northern hybridization.

***In vitro* transcription of snRNA genes.** Plasmids containing the *S. cerevisiae* U1, U2, U4, U5, and U6 snRNA coding sequences under the control of a T7 promoter were kindly provided by P. Fabrizio. Synthetic U1, U2, U4, U5, and U6 snRNAs were prepared by *in vitro* transcription of *Pvu*II-linearized pT7U1 plasmid, *Xho*I-linearized pT7U2 plasmid, *Syl*I-linearized pT7U4 plasmid, *Dra*I-linearized pT7U5 plasmid, and *Dra*I-linearized pT7U6 plasmid, respectively (23, 57).

Synthesis of transcripts was carried out in 30 μ l of buffer containing 40 mM Tris-HCl (pH 8.1), 20 mM MgCl₂, 5 mM dithiothreitol (DTT), 1 mM spermidine, 0.01% Triton X-100, 80 mg of polyethylene glycol 8000 per ml, 2 μ g of the linearized plasmid, 4 mM concentrations of each ribonucleoside triphosphate, 19 U of RNase Guard (Pharmacia), and 138 U of T7 RNA polymerase (Pharmacia). After 1 h of incubation at 37°C, the template DNA was digested by using 7.5 U

of RNase-free DNase I (Pharmacia) for 30 min at 37°C. After phenol extraction and ethanol precipitation, the RNA was dissolved in 100 μ l of sterile water.

Localization of Ψ residues in *S. cerevisiae* UsnRNAs and 26S rRNA by primer extension analysis. Total RNA (10 μ g) from the wild-type or mutant *S. cerevisiae* strains was used for reverse transcription. The CMCT [*N*-cyclohexyl-*N'*-(2-morpholinoethyl)-carbodiimid metho-*p*-toluolsulfonate] modification protocol was adapted from Bakin and Ofengand (8) with the following modifications: the CMCT treatment was performed for 2, 10, or 20 min; the treatment in bicarbonate buffer at pH 10.4 was done for 3 h; and all precipitations were done by using 0.3 M sodium acetate buffer (pH 5.3). The hydrazine reaction was performed essentially as described by Peattie (78).

Positions of CMCT and hydrazine modifications were identified by primer extension analysis with AMV reverse transcriptase (RT; Life Sciences) as described by Mougouin et al. (61). The oligonucleotides complementary to the following regions of UsnRNAs and 26S rRNA were used as primers for reverse transcription: U1 (nt 57 to 72), U2 (nt 104 to 126), U4 (nt 68 to 90 and nt 134 to 160), U5 (nt 159 to 182), U6 (nt 93 to 112), and 26S rRNA (nt 1144 to 1163). Oligonucleotides were 5' end labeled with [γ -³²P]ATP (3,000 Ci/mmol) and T4 polynucleotide kinase (90).

In vivo analysis of rp51A pre-mRNA and pre-U3 snoRNA splicing. The yield of rp51A pre-mRNA in the absence of an active *PUS1* gene was evaluated by primer extension analysis. Primer extension analysis on the U1 snRNA was used as a control. The oligonucleotide primer complementary to the rp51A pre-mRNA and its two splicing products were described by Teem and Rosbash (107). The oligonucleotide primer used for U1 snRNA was that described above for Ψ residue identification. Primer extension analysis was done in the conditions described by Mougouin et al. (61). For quantitative cDNA synthesis, the labeled primers were always added in excess, 8 ng per assay. At the end of the incubation period, the elongation mixtures were treated with 20 μ g of RNase A per ml for 30 min at 37°C and analyzed by electrophoresis on a 5% sequencing gel. To verify that the amount of synthesized cDNAs was proportional to the amount of rp51A mRNAs in the total RNA extract, the experiments on the *pus1*⁻ strain were made in triplicate by using 10, 20, and 50 μ g of total RNA. The linearity of the ³²P amount in cDNA bands versus the total RNA amount used as the template was verified by PhosphorImager measurement. For the wild-type strain, 10 and 50 μ g of total RNA were used for rp51A mRNA quantification, and for the U1 snRNA control assays, 5 and 15 μ g of total RNA were used. The amounts of synthesized U1 and rp51A cDNAs, in each assay, were estimated by PhosphorImager measurement. Based on the values obtained, the relative yields of rp51A mRNAs versus U1 snRNA were established for the wild-type and *pus1*⁻ strains.

The efficiency of pre-U3 snoRNA splicing in the absence or the presence of an active *PUS1* gene was analyzed by Northern blot analysis. A 5'-end ³²P-labeled oligonucleotide complementary to the *S. cerevisiae* U3 snoRNA (nt 1 to 16) was used as the probe. Total RNA (10 μ g) from the wild-type and the *pus1*⁻ disrupted strains (prepared as described above) was fractionated on a 5% polyacrylamide gel. In vitro transcripts (10 ng) corresponding to the spliced and unspliced pre-U3 snoRNAs were used as controls. The plasmid pV551:snR17A (92) was used to produce the U3A snoRNA transcript, while the pre-U3 snoRNA transcript was obtained by using the construct described by Mougouin et al. (61). Total RNA (10 μ g) from the JH84 *S. cerevisiae* strain (37) transformed by the pU3U14ds5' plasmid (61) was also applied on the gel. The total RNA fractions from the JH84 strain were prepared as described by Mèreau et al. (58). Plasmid pU3U14ds5' contains a U3 snoRNA gene carrying mutations that lead to an accumulation of unspliced pre-U3 snoRNA and of a degraded form of the pre-U3 snoRNA in vivo (24).

In vitro pseudouridine formation in yeast U2 snRNA transcripts. Two micrograms of *S. cerevisiae* U2 transcript dissolved in 9 μ l of buffer (100 mM Tris-HCl buffer [pH 8.0] containing 100 mM ammonium acetate, 5 mM MgCl₂, 2 mM DTT, and 0.1 mM EDTA) was heated for 3 min at 80°C and cooled down to 37°C. The purified Pus1p enzyme (2.5 μ g), prepared from the recombinant *E. coli* strain (97) as described by Motorin et al. (60), was added, and the reaction was performed for 30 min at 37°C. After incubation, the modified transcript was phenol extracted and ethanol precipitated. The presence of pseudouridine residues was analyzed by the CMCT-RT technique as described above.

RESULTS

Identification of Ψ residues in *S. cerevisiae* UsnRNAs. Two independent approaches were used for mapping Ψ residues in the *S. cerevisiae* spliceosomal UsnRNAs. The first approach was based on the CMCT-RT technique developed by Bakin and Ofengand (8). This method depends on the efficient chemical reaction of U and Ψ residues in RNA with CMCT (very strong for U residues, less strong for Ψ residues). Upon alkaline treatment with bicarbonate buffer at pH 10.4, the bulky CMC group linked to the U residues can be selectively hydrolyzed, while it remains bound to the Ψ residues. Stops of reverse transcription at CMCT-modified Ψ allow localization of Ψ residues by primer extension analysis. Guanosine residues

also react with CMCT but less efficiently than do U and Ψ . Moreover, the CMC groups bound to G residues are easily removed by the alkaline treatment.

The second complementary approach was based on the observation that Ψ (and also m⁵U) residues are resistant to hydrazine treatment under conditions developed for chemical sequencing of RNA (15, 78). Thus, with this analysis, the absence of reactivity indicates the presence of one of these two posttranscriptional modifications. The absence of hydrazine reactivity can be detected directly by observing the cleavage pattern of purified 3'-end-labeled RNA after aniline treatment or indirectly, as described above, by primer extension analysis with RT.

In both methods, primer extension was performed on total, unfractionated yeast RNA prepared as described in Materials and Methods. Oligonucleotide primers allowed to explore different regions of the yeast UsnRNA molecules. For the long U1 (568 nt in length) and U2 snRNAs (1,175 nt), only the 5'-terminal regions (positions 1 to 50 of U1 and positions 1 to 100 of U2) were analyzed. These regions were chosen taking into account the already localized modified nucleotides in the U1 and U2 snRNAs of vertebrates, plants (reviewed in references 81 and 98), and *S. pombe* (28) (Fig. 1A and 2A). Indeed, vertebrate U1 snRNA contains five modified nucleotides, four of which are found, respectively, at positions 1 (Am), 2 (Um), 5 (Ψ), and 6 (Ψ), and the fifth one, located at position 70, is a 2'-*O*-methylated adenosine residue (for a review, see reference 81). Likewise, the 5'-terminal 100-nt region of rat hepatoma U2 snRNA contains all of the posttranscriptional modifications that were found in this snRNA (12 Ψ , 9 2'-*O*-methylated, and 1 m⁶Am residues) (84). For the shorter U4 (160 nt), U5S (179 nt), U5L (214 nt), and U6 (112 nt) snRNAs, only a short 3'-terminal region was not explored due to the constraint of primer extension analysis: the 30 nt at the 3' end of the U4 snRNA, the 61 and 26 nt at the 3' extremity of the U5L and U5S, respectively, and the 23 nt at the 3' end of the U6 snRNA.

Only a few Ψ residues are found in the *S. cerevisiae* UsnRNAs. The primer extension analysis of U1, U2, U4, U5, and U6 snRNAs modified by CMCT was used to map the pseudouridine residues. Incubation of the RNAs with CMCT was performed in each case for 2, 10, or 20 min, followed or not (control experiment) by alkaline treatment at pH 10.4. The cleavage patterns upon hydrazine treatment were also analyzed in each case. Both approaches gave essentially the same conclusions. Representative examples of primer extension patterns are shown.

Figure 1B illustrates the analysis of the *S. cerevisiae* U1 snRNA modified by hydrazine or CMCT. As described for vertebrate UsnRNAs (Fig. 1A and 3B) (12, 83), two Ψ residues are found at positions 5 and 6 (Fig. 1B and C and 3A). Only one of these two Ψ residues was detected in *S. pombe* (28). Analysis of the 5'-terminal region of *S. cerevisiae* U2 snRNA by CMCT (Fig. 2B) and hydrazine modification (data not shown) revealed the presence of three Ψ residues, respectively, at positions 35, 42, and 44 (Fig. 2C and 3A). Three Ψ residues were found at the same positions in the rat hepatoma (Fig. 2A and 3B) and the *S. pombe* U2 snRNAs, but additional Ψ residues were also detected in these two RNAs (Fig. 3B) (28, 84). Only one Ψ residue was detected in the *S. cerevisiae* U5 snRNAs (Fig. 4C). It corresponds to one of the two phylogenetically highly conserved Ψ residue found in the U5 snRNA terminal loop I (Fig. 3B) (103). At the other conserved pseudouridylation site of this terminal loop, the uridine residue is replaced by a cytidine residue in *S. cerevisiae* (Fig. 3A). The same situation was found for *S. pombe* (28). As found for

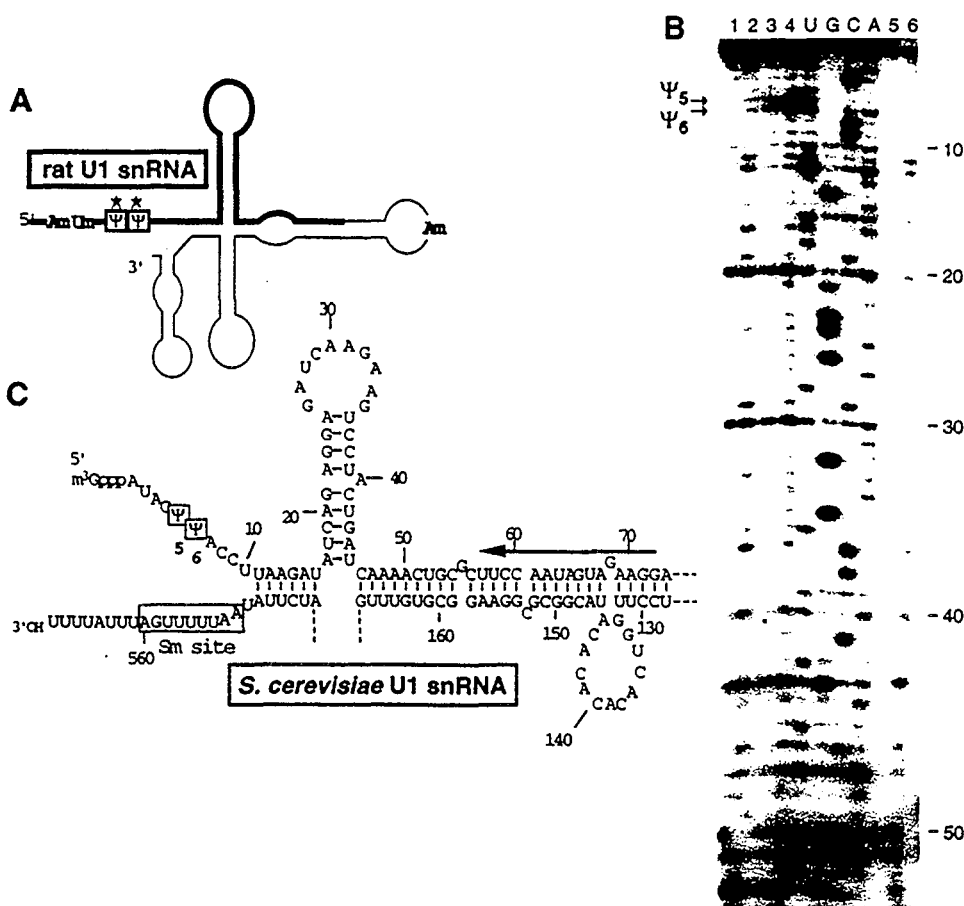


FIG. 1. Localization of Ψ residues in *S. cerevisiae* U1 snRNA. (A) Schematic representation of the secondary structure of vertebrate U1 snRNA (12, 62). The Ψ residues are boxed, and the 2'-O-methylated residues are also indicated (for a review, see reference 55). The thick line shows the region of rat hepatoma U1 snRNA that corresponds to the analyzed region of *S. cerevisiae* U1 snRNA. The Ψ residues that were conserved in *S. cerevisiae* U1 snRNA are indicated by stars. (B) Primer extension analysis of the *S. cerevisiae* U1 snRNA modified by CMCT in a total RNA fraction for 2, 10, and 20 min (lanes 2, 3, and 4, respectively). Experimental conditions were as described in Materials and Methods. In lanes 3 and 4, the CMCT-modified RNA was subjected to an alkaline treatment at pH 10.4 as described in Materials and Methods. A control extension experiment was made without CMCT treatment (lane 1). The two reverse-transcription stops, corresponding to residues Ψ_5 and Ψ_6 , are indicated by arrows on the left of panel B. U1 snRNA in a total RNA mixture was also treated by hydrazine under the conditions described in Materials and Methods (lane 5). The absence of hydrazine reactivity at positions 5 and 6 indicates the presence of a Ψ residue at these positions. Lanes U, G, C, and A correspond to the RNA sequencing ladder. Nucleotide positions, starting from the 5'-terminal nucleotide bound to the cap structure, are indicated on the right. (C) Nucleotide sequence of the analyzed region of the *S. cerevisiae* U1 snRNA. The oligonucleotide used as the primer for reverse transcription is indicated by a horizontal arrow. The two detected Ψ residues are boxed. The secondary structure was taken from Kretzner et al. (46).

U4 and U6 snRNAs from *S. pombe* (28), no Ψ residues were detected in the examined regions of U4 and U6 snRNAs (data not shown), while in the vertebrates, three Ψ residues were detected in both U4 (47, 82) and U6 snRNAs (Fig. 3B) (33). Hence, altogether only 6 Ψ residues were detected in the *S. cerevisiae* spliceosomal UsnRNAs (Fig. 3A), compared to 9 in the UsnRNAs from *S. pombe* (28) and 23 in the UsnRNAs from rat hepatoma (for a review, see reference 81).

Disruption of the *PUS1* gene results in the absence of Ψ_{44} in U2 snRNA. Based on sequence homology with known *E. coli* tRNA-pseudouridine synthases, several potential RNA-pseudouridine synthases were identified in the *S. cerevisiae* genome (44). Three genes (*PUS1*, *PUS3*, and *PUS4*) were shown to code for tRNA-pseudouridine synthases (respectively, Pus1p, Pus3p, and Pus4p [10, 51, 60, 97]), and a fourth one (*CBF5*) was shown to be involved in pseudouridine formation in rRNA (protein Cbf5p [50]). The target RNA for a putative pseudouridine synthase Pus2p has not yet been determined (97).

Total RNA was extracted from *S. cerevisiae* strains disrupted for one of the nonessential *PUS1*, *PUS2*, *PUS3*, or *PUS4* genes. Depletion of the Cbf5p enzyme was achieved in a *GAL::cbf5* conditional strain after transfer to nonpermissive conditions for 70 h (50; see Materials and Methods). The pseudouridine residues in UsnRNAs were mapped by the CMCT-RT method. Since the previous results showed that only U1, U2, and U5 snRNAs contain Ψ residues in *S. cerevisiae*, the analysis was limited to these three snRNAs.

The results of the mapping analysis indicate that only the disruption of the *PUS1* gene affects Ψ formation in the spliceosomal UsnRNAs (Fig. 4). Conversion of U into Ψ residues in U1, U2, and U5 UsnRNAs was unaffected by Cbf5p depletion under conditions where Ψ formation in pre-rRNA was strongly inhibited. This is illustrated by the analysis of the 26S rRNA segment (positions 1010 to 1140) that we performed together with the UsnRNA analysis (Fig. 4D) (50). Compared to a wild-type strain (WT), the *pus1*⁻ strain differed by the disappearance of Ψ_{44} in U2 snRNA (Fig. 4A, compare lanes 3

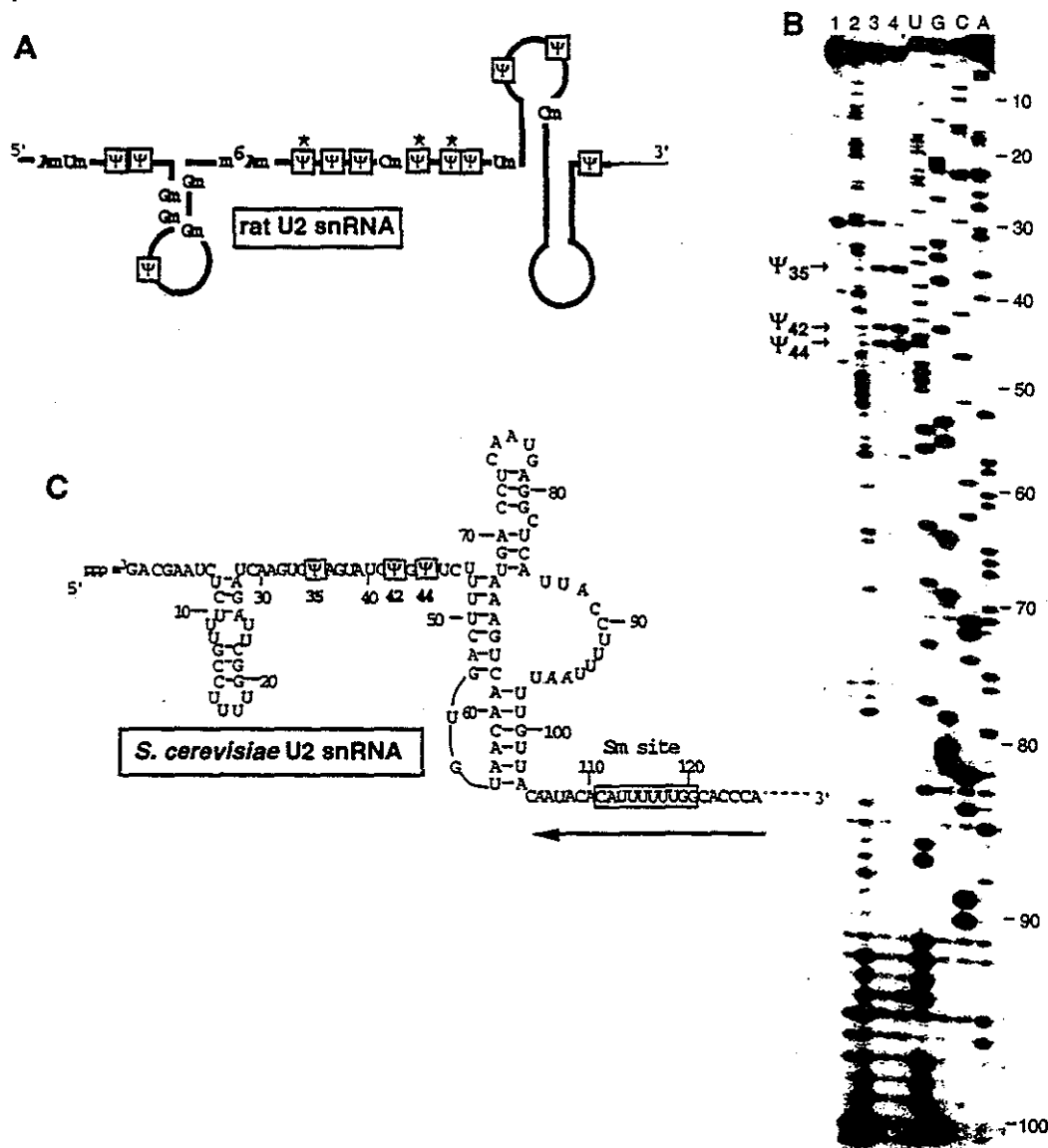


FIG. 2. Localization of Ψ residues in *S. cerevisiae* U2 snRNA. (A) Rat hepatoma U2 snRNA, folded according to the U2 snRNA secondary structure model of Ares and Igel (4). The Ψ residues are boxed, the 2'-O-methylated residues and base-methylated residues are indicated (for a review, see reference 81). The thick line shows the region of rat hepatoma U2 snRNA that corresponds to the analyzed region of *S. cerevisiae* U2 snRNA. The Ψ residues that are conserved in the *S. cerevisiae* U2 snRNA are marked by stars. (B) Primer extension analysis of the *S. cerevisiae* U2 snRNA modified by CMCT (lanes 2, 3, 4, and 1 [for the control]). The conditions are the same as those for Fig. 1. The alkaline-resistant RT stops in lanes 3 and 4 revealed the presence of Ψ residues at positions 35, 42, and 44. The nucleotide sequence (3) and the secondary structure (4, 38) of the analyzed region of *S. cerevisiae* U2 snRNA are shown in panel C. The position of the primer oligonucleotide is shown by the horizontal arrow. The identified Ψ residues are boxed.

and 4 [WT] to lanes 3 and 4 [*pus1*⁻]). No apparent change at the other positions (35 and 42) in U2 snRNA (Fig. 4A), as well as at positions 5 and 6 of U1 snRNA (Fig. 4B) and position 99 in U5 snRNA (Fig. 4C), was detected.

Pus1p catalyzes the pseudouridylation of U₄₄ in U2 snRNA in vivo and in vitro. To verify that Pus1p directly catalyzes Ψ formation at position 44 of U2 snRNA, the *pus1*⁻ disrupted strain was transformed with plasmid pUN100-*PUS1* (97) containing a wild-type *PUS1* gene. As shown in Fig. 4A (lanes 2 and 3 *pus1*⁻ + pUN100-*PUS1*), Ψ formation at position 44 of U2 snRNA was completely restored in the transformed yeast cells.

The capacity of the yeast Pus1p enzyme to catalyze U to Ψ conversion at position 44 in U2 snRNA was also tested in vitro.

To this end, we used a recombinant Pus1p enzyme purified from *E. coli* (60) and an in vitro-produced T7 transcript of the *S. cerevisiae* U2 snRNA. The results of CMCT mapping demonstrate the efficient formation of Ψ_{44} in the U2 snRNA transcript and show that only this residue is modified in vitro (Fig. 5).

The absence of Ψ_{44} in U2 snRNA does not influence pre-mRNA and pre-U3 snoRNA splicing. Although no apparent growth defect was found for the *PUS1* gene deletion (97), it was interesting to test whether the absence of a Ψ at position 44 in U2 snRNA could affect the efficiency of pre-mRNA splicing. We tested this possibility on the *rp51A* pre-mRNA, the precursor for the mRNA of the *S. cerevisiae* *rp51A* ribosomal protein (87). To this end, using primer-extension anal-

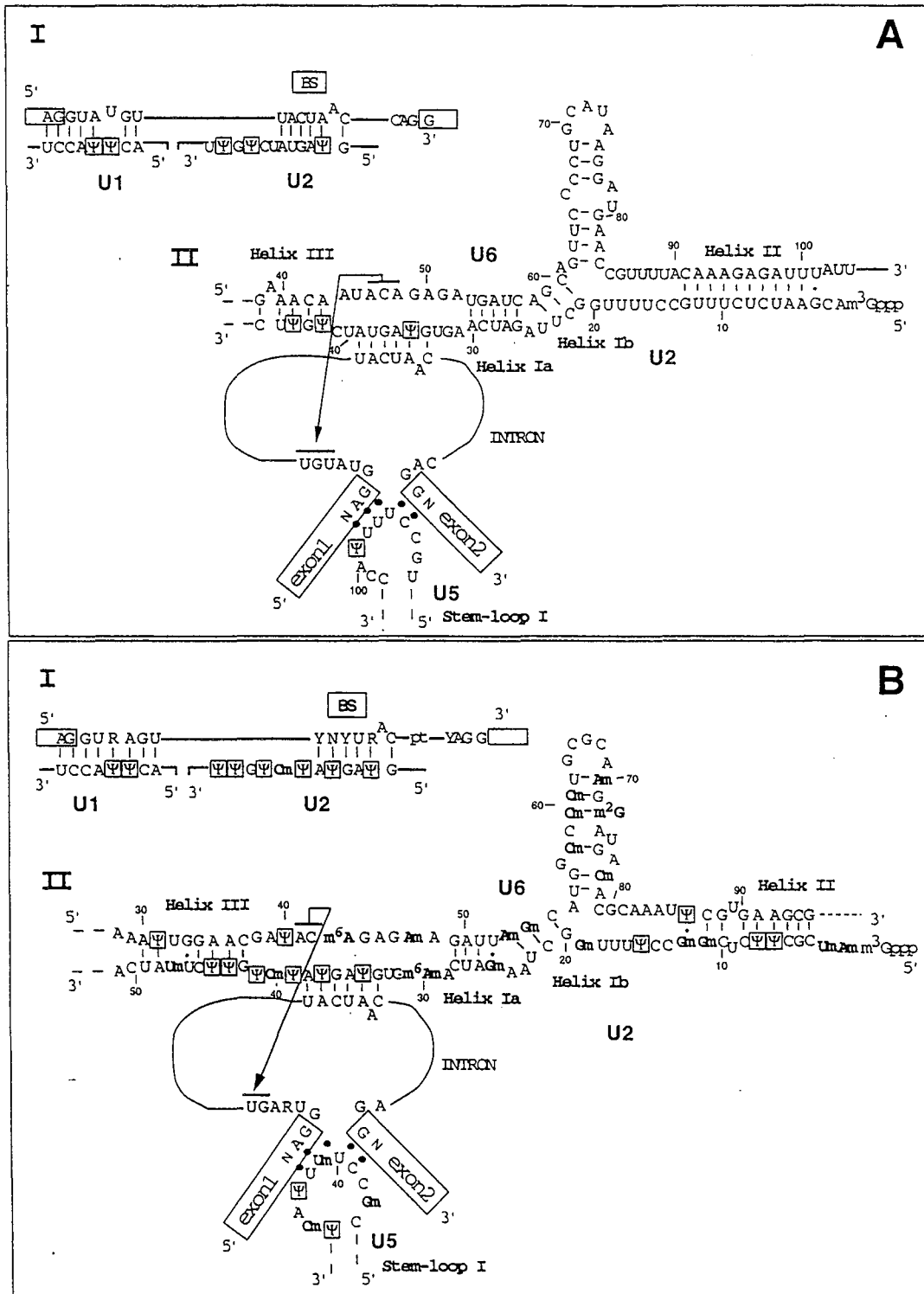


FIG. 3. RNA-RNA interactions in the *S. cerevisiae* spliceosomes (A) and in the GU-AG spliceosomes of higher eukaryotes (B). The interaction between the 5' and 3' splice sites and branch-site (BS) consensus sequences with U1 and U2 snRNAs are shown in scheme I of panels A and B. UsnRNA-UsnRNA and UsnRNA-pre-mRNA interactions at the catalytic center of the spliceosome are shown in scheme II of panels A and B. Helices Ia, Ib, II, and III between U2 and U6 snRNAs are represented, as well as the base-pairing interaction between U2 snRNA and the branch-site sequence. The interaction between U6 snRNA and the UGU trinucleotide, close to the 5' splice site, is indicated by overlined residues joined by an arrow; the interaction between the terminal loop I of U5 snRNA and the exon extremities is also shown (for a review of all of these interactions, see reference 54). The Ψ residues are boxed, and the base and ribose methylations are indicated in boldface (this study; for a review, see reference 55). Nucleotide positions, starting from the 5' extremity of each UsnRNA, are indicated.

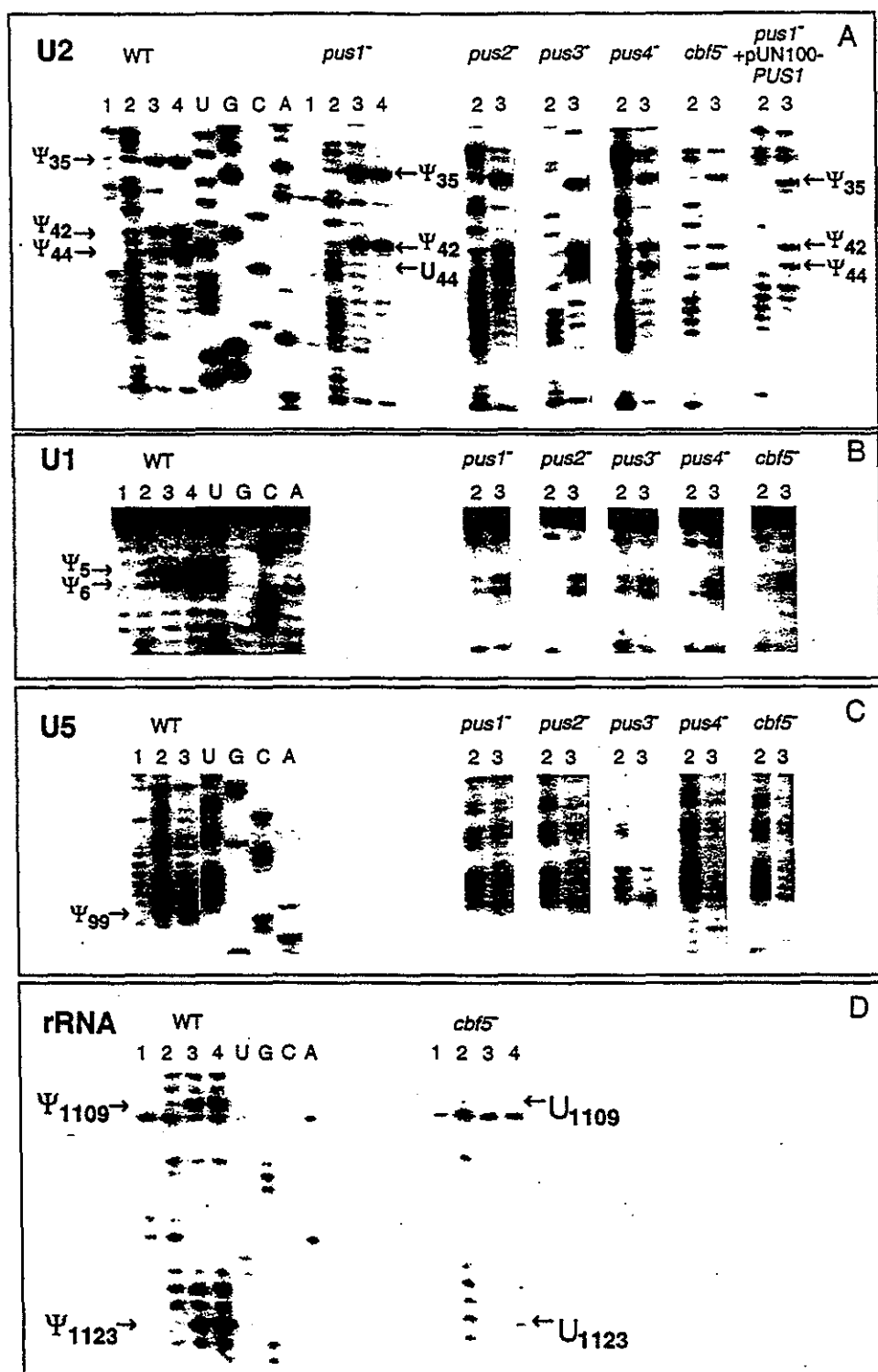


FIG. 4. UsnRNA and rRNA pseudouridylation pattern in yeast mutant strains. Mapping of Ψ residues in U2 (A), U1 (B), and U5 (C) snRNAs by the CMCT-RT method, was performed as indicated in Fig. 1. The total RNAs used were extracted from the wild-type *S. cerevisiae* (WT) and from strains carrying a disruption in one of the *PUS1*, *PUS2*, *PUS3*, or *PUS4* genes or from *GAL::cbf5* strain (*cbf5*⁻) grown as described in Materials and Methods. In the case of U2 snRNA (last row of panel A), an additional analysis was performed with total RNA extracted from the *pus1*⁻ strain transformed with plasmid pUN100-*PUS1* containing a wild-type *PUS1* gene. Lanes 1, 2, 3, and 4 correspond to the CMCT analysis according to the legend of Fig. 1. U, G, C, and A is the sequencing ladder. To minimize the figure size, only lanes 2 and 3 are shown for strains that show no variation as compared to wild type. The detected Ψ residues are indicated by arrows. The control experiment showing the absence of Ψ residues in the 26S rRNA domain II (positions 1010 to 1140) from the *GAL::cbf5* strain is shown in panel D. For the wild-type and the *GAL::cbf5* (*cbf5*⁻) strains, Ψ residues in this 26S region were mapped by CMCT as described in Materials and Methods. To minimize the figure size, only the portion of 26S rRNA containing the two Ψ residues 1109 and 1123 in the wild-type strain are shown.

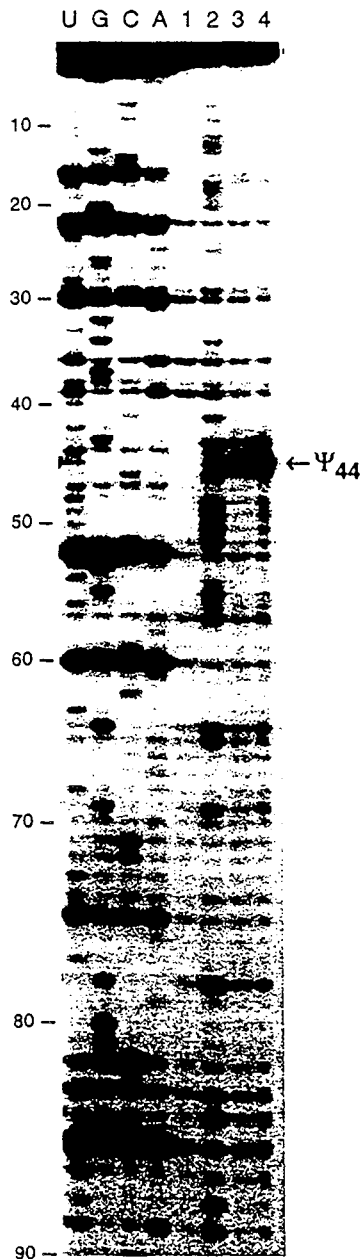


FIG. 5. In vitro pseudouridylation of U2 snRNA transcript by using recombinant Pus1p pseudouridine synthase. After incubation of the transcript with the purified Pus1p enzyme (as described in Materials and Methods), the phenol-extracted RNA was analyzed by the CMCT-RT method (lanes 1, 2, 3, and 4) as described in the legend to Fig. 1. The strong alkaline-resistant RT stop at position 44 is indicated by an arrow. Lanes U, G, C, and A correspond to the sequencing ladder.

ysis in conditions allowing quantification of the RNA template, we compared the relative yields of rp51A mRNAs and pre-mRNA in total RNA extracted from the wild-type and the *pus1*⁻ strains grown until mid-exponential phase. U1 snRNA, which was not spliced in *S. cerevisiae*, was used as an internal control. In both strains, no rp51A pre-mRNA was accumulated (not shown). Only the two cDNA products corresponding to the two rp51A mRNAs previously described (1, 107) were detected (Fig. 6A). As shown in Fig. 6, only a very slight decrease of the rp51A mRNA concentrations in total RNA

was observed in the absence of residue Ψ_{44} in U2 snRNA. Based on the rp51A mRNA/U1 snRNA ratio, estimated by PhosphorImager measurement, the rp51A mRNA concentration in total RNA was only reduced by 10% in the absence of residue Ψ_{44} in U2 snRNA. The observed difference is at the limit of the accuracy of mRNA quantification by primer extension analysis.

Since the *S. cerevisiae* U3 snoRNA is produced from a precursor RNA that is spliced in a spliceosome and since the branch-point sequence of the U3 snRNA intron shows a substitution at the 5' extremity compared to the consensus branch-site sequence (64), we tested whether splicing of this substrate may be more dramatically altered by the absence of a Ψ residue at position 44 of U2 snRNA. For this purpose, Northern blot analysis was performed with total RNA extracted from the wild-type and the *pus1*⁻ strains (Fig. 6B). An *S. cerevisiae* strain carrying a mutant U3 snoRNA gene that produces a pre-U3 snoRNA with a splicing defect (plasmid pU3U14ds5') (61) was used as a control. In this strain, the pre-U3 snoRNA and its main degradation product accumulate (24). As shown in Fig. 6B, the level of mature U3 snoRNA was very slightly decreased in the absence of Ψ_{44} in U2 snRNA. However, here again it was at the limit of the accuracy of this kind of experiment. Furthermore, no accumulation of the U3 snoRNA precursor was detected. Hence, at least for the two pre-RNAs tested, the presence of a Ψ residue at position 44 of U2 snRNA did not show a strong effect on splicing efficiency. As no precursor accumulation was detected, the slight differences observed were perhaps not due to a splicing defect but to differences in RNA stability in the *pus1*⁻ strain.

DISCUSSION

S. cerevisiae spliceosomal UsnRNAs contain a few Ψ residues located at sites of RNA-RNA or RNA-protein interactions. We detected only six Ψ residues in the studied regions of *S. cerevisiae* UsnRNAs. This is significantly lower than the 23 and 20 residues found in the corresponding parts of rat hepatoma and plant UsnRNAs, respectively (for a review, see reference 55). A low number of Ψ residues in RNAs seems to be a general feature of yeasts, as only nine Ψ residues were detected in the *S. pombe* spliceosomal UsnRNAs (28), and a low level of Ψ residues was also observed for rRNAs and tRNAs in yeasts (7, 9, 16, 100).

Despite the absence of phenotype when there was a lack of Ψ formation at position 44 of the *S. cerevisiae* U2 snRNA and the very low effect that we detected on the level of rp51A mRNAs and U3 snoRNA, which are among the very few *S. cerevisiae* RNAs processed in a spliceosome, it is noteworthy that the six Ψ residues that we detected in *S. cerevisiae* UsnRNAs are evolutionarily conserved. Furthermore, they are all located in or very close to the UsnRNA segments involved in intermolecular RNA-RNA interactions within the spliceosomes (Fig. 3A). Altogether, this suggests that at least some of them may have a functional importance. In U1 snRNA, the two Ψ residues detected are located in the segment that base pairs with the 5' extremity of introns. Posttranscriptional modifications at position 5 and 6 of the U1 snRNA are nearly universal. Two Ψ residues are found at these positions in vertebrates and insects (Fig. 3B), in plants a Ψ_m residue is present at position 5 and a Ψ_n residue is found at position 6. Surprisingly, only position 5 shows a posttranscriptional modification in *S. pombe* (28). In the *S. cerevisiae* U1 snRNA-intron interaction, residue Ψ_3 of U1 snRNA faces residue U_4 of the intron (Fig. 3A). The presence of a U residue at position 4 of yeast introns is required for the interaction with U6 snRNA in the active con-

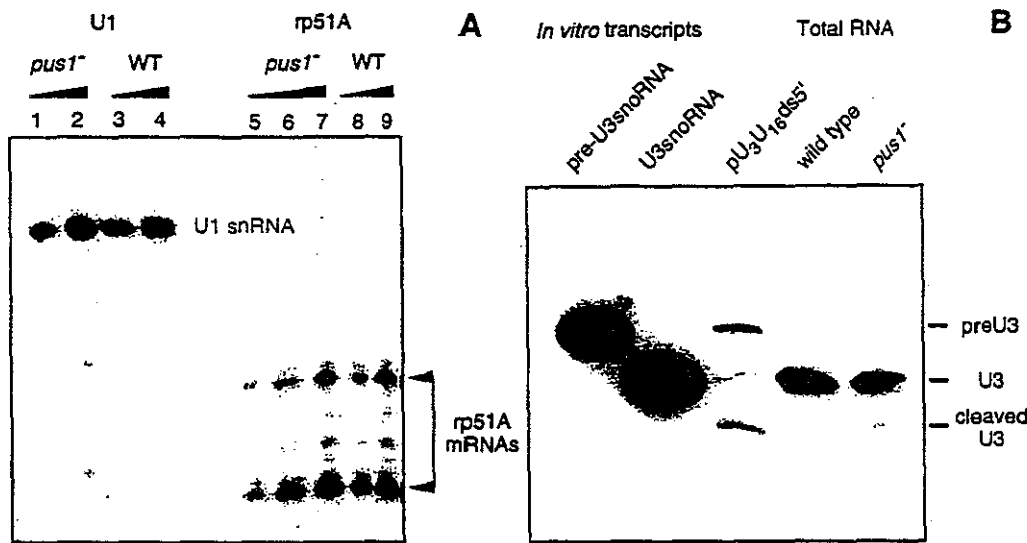


FIG. 6. No marked effect of the *PUS1* gene deletion on rp51A pre-mRNA and pre-U3 snoRNA splicing. In panel A, total RNA extracted from wild-type (WT) and the *pus1*⁻ strains were analyzed by primer extension analysis under the conditions described in Materials and Methods that allow rp51A pre-mRNA and mRNA quantification. The *pus1*⁻ strain primer extension analyses with the rp51A primer were performed on 10 μg (lane 5), 20 μg (lane 6), or 50 μg (lane 7) of total RNA, and for the wild-type strain the experiments were performed with 10 (lane 8) or 50 μg (lane 9) of total RNA, respectively. Control extension experiments with the U1 primer were performed with 5 μg (lanes 1 and 3) or 15 μg (lanes 2 and 4) of total RNA. The cDNA fractionation was performed on a 3% sequencing gel; only the portions of the gel corresponding to the two expected rp51A mRNA extension products are shown. The cDNA amounts in the U1 and rp51A bands of gel were estimated by PhosphorImager measurement. In panel B, total RNA extracted from the wild-type and the disrupted *pus1*⁻ strains was analyzed by Southern blot hybridization by using a labeled oligonucleotide complementary to the yeast U3 snoRNA from position 1 to 16 (92). In vitro transcripts corresponding to the pre-U3 snoRNA and the U3 snoRNA, respectively, were used as controls (the two lanes on the left), as well as the total RNA extracted from a pU3U16ds5' strain that shows the accumulation of unspliced and degraded pre-U3 snoRNA (24). The ³²P amounts in the U3 signals obtained for the wild-type and *pus1*⁻ strains were compared by PhosphorImager analysis.

formation of the spliceosome (Fig. 3A) (41, 52, 91), and the conversion of the Ψ₃-U₄ pair into a canonical Watson-Crick pair in the U1 snRNA-intron interaction results in a 50% increase of the cell doubling time (96). In this context, Ψ formation at position 5 of U1 snRNA may be of functional importance.

Among the 12 Ψ residues detected in the rat hepatoma U2 snRNA (Fig. 2A and 3B), only 3 are conserved in the *S. cerevisiae* U2 snRNA (Fig. 2C and 3A). None are found in the 27-nt 5'-terminal segment, whose modification at the posttranscriptional level was found to be essential for vertebrate U2 snRNA function (120). However, residue Ψ₃₅ may be of high functional importance, since it is involved in one of the two canonical base pairs that bulge out the A residue responsible for the nucleophilic attack in the first step of the splicing reaction. Helix stabilization by the presence of a Ψ-A pair (reviewed in references 2, 6, and 21) may be the reason for the universal conservation of a Ψ residue at this position. In addition, since it was shown for vertebrates that the branch-point bulged structure is recognized by proteins (26, 80; for a review, see reference 85), residue Ψ₃₅ may also be involved in this recognition. The presence of a second Ψ-A pair in the vertebrate U2 snRNA-branch site interaction (Fig. 3B) may be needed to compensate for the high degree of degeneracy of the vertebrate branch-site sequence (for reviews, see references 39 and 89).

Using in vitro splicing assays, the effect of base substitutions on splicing efficiency was tested for almost all residues of the 5'-terminal region of the *S. cerevisiae* U2 snRNA (56). Mutations at position 35 resulted in a strong decrease in the splicing efficiency. Pascolo and Séraphin (72) tested the effect of compensatory mutations in the branch-site sequence and its U2 snRNA recognition element in vivo. However, the data ob-

tained for the U2 Ψ₃₅-branch site A₅ pair are not sufficient to determine whether the presence of a Ψ residue at position 35 is required for high splicing efficiency.

The *S. cerevisiae* Prp5, Prp9, Prp11, Prp21, and Cus1 proteins, which are homologous to proteins from the human splicing factors SF3a and SF3b, associate with the *S. cerevisiae* U2 snRNA region from position 40 to 87 (36, 88, 110, 111, 119). This region contains the Ψ residues 42 and 44. In addition, the segments from positions 42 to 46 in the *S. cerevisiae* U2 snRNA and from positions 42 to 49 in the vertebrate U2 snRNA were found to form a base-pair interaction with U6 snRNA (Fig. 3). Such interaction delineates helix III, which is needed for active spliceosome formation in HeLa cells (102). In *S. cerevisiae*, no growth defect was detected for base substitution at positions 42 of U2 snRNA (119). This argues against an essential role of the U-to-Ψ conversion at position 42. Mutations at position 44 strongly affected growth (119). However, based on the absence of phenotype after deletion of the *PUS1* gene (reference 97 and this study), the defect was not related to the absence of a U-to-Ψ conversion at this position. Indeed, a proposed explanation of the defect was that the replacement of U₄₄ (in fact, Ψ₄₄) by an A or a G residue favors an alternative conformation of U2 snRNA which impairs active spliceosome formation (119).

The Ψ residue at position 99 in the 5'-terminal loop of U5 snRNA is at the border of the segment that interacts with the exon extremities. Crystallographic and nuclear magnetic resonance studies of the yeast tRNA^{Asp} and tRNA^{Gln} showed an essential role for Ψ₃₂ and Ψ₃₈ residues in the anticodon loop for maintaining the conformation required for proper anticodon recognition (5, 21, 22, 112, 113). As referred to this model, it may be that the U5 snRNA Ψ₉₉ favors a conformation of the U5 snRNA 5'-terminal loop, which facilitates the correct align-

ment of exons. In connection with this hypothesis, it should be noticed that mutations affecting the 5'-terminal loop of *S. cerevisiae* U5 snRNA block splicing after the first step of the reaction (68).

In conclusion, based upon available genetic data and the present work, four of the six Ψ residues detected in the *S. cerevisiae* spliceosomal UsnRNAs belong to RNA segments involved in intermolecular interactions and may be of high functional importance. The other two are not essential taken individually. However, their conservation from yeasts to humans suggests that they may provide some selective advantage, which is difficult to test in laboratory conditions.

Among several characterized yeast pseudouridine synthases, only the Pus1p tRNA-pseudouridine synthase acts on spliceosomal UsnRNAs. In yeast tRNAs, Ψ residues are found at 15 different locations, and at least five or six distinct enzymes are required for complete pseudouridylation of tRNAs. Only three of them, Pus1p (60), Pus3p (51), and Pus4p (10) have been characterized. The *S. cerevisiae* 17S and 26S rRNAs contain, respectively, 13 and 30 Ψ residues (16, 69, 70). The recent discovery of the snoRNA-guided mechanism of eukaryotic rRNA pseudouridylation has allowed to assign the Cbf5p (in yeast) and NAP57 (in higher eukaryotic) enzymes as the ones which act on the rRNA-snoRNP complex (25, 50). The *S. cerevisiae* pseudouridine synthases responsible for U isomerization in UsnRNAs have not been investigated so far. Studies of UsnRNA-specific pseudouridine synthases have been restricted to human cells. The results obtained revealed the presence of several distinct pseudouridine synthases; however, none of them was identified or cloned (73–75, 77, 121, 122).

Since no decrease of the level of Ψ residues was observed in UsnRNAs upon Cbf5p protein depletion, whereas Ψ synthesis in rRNA was strongly affected, our results strongly suggest that the Cbf5p enzyme is not involved in the pseudouridylation of UsnRNAs in *S. cerevisiae*. This implies that the U-to- Ψ conversion in the *S. cerevisiae* spliceosomal UsnRNAs is not based on the snoRNA guide mechanism involving the Cbf5p enzyme (50). This is in contrast to the recently reported mechanism of 2'-O-methylation of the vertebrate U6 snRNA (43, 108). Our results, obtained for yeasts, do not rule out the possibility that some pseudouridylation in the vertebrate U6 snRNA is generated by the snoRNA-guided mechanism. The nonimplication of Cbf5p in *S. cerevisiae* spliceosomal UsnRNA pseudouridylation may simply be due to the absence in *S. cerevisiae* U6 snRNA of an appropriate U residue that can be modified by the Cbf5p-snoRNA system. A larger number of Ψ residues are present in the vertebrate U6 snRNA (33), and some of them may be generated by a snoRNA-guided mechanism.

Two of the characterized *S. cerevisiae* pseudouridine synthases were found to act on tRNAs: Pus3p catalyzes the U isomerization at positions 38 and 39 in tRNAs (51), and Pus4p catalyzes the isomerization at position 55 (10). Our data show that these enzymes are not involved in the modification of UsnRNAs. Both of them are highly dependent on the global tRNA three-dimensional structure (10, 51), which probably explains their absence of reactivity on UsnRNAs. In contrast, Pus1p catalyzes Ψ formation at eight different sites in various tRNAs (positions 26 to 28, 34 to 36, 65, and 67) (60). Pseudouridine formation at the anticodon positions 34, 35, and 36 depends on the presence of an intron in tRNA, but it does not require the intact three-dimensional tRNA architecture (97, 104). These results obtained with tRNA substrates allowed the authors of these studies to conclude that Pus1p recognizes only a limited structural domain in RNA. The results presented here demonstrate that Pus1p is also capable of modifying, both in vivo and in vitro, one

of the uridine residues (U_{44}) in U2 snRNA. Evidently, this enzyme shows a dual specificity for both tRNA and snRNA.

The only well-documented case of a pseudouridine synthase with substrate dual specificity concerns the *E. coli* pseudouridine synthase RluAp. This enzyme catalyzes U-to- Ψ conversion at position 32 in several tRNAs and at position 746 in 23S rRNA (115). This dual specificity towards tRNA and rRNA was observed in vitro with synthetic RNA transcripts. In this case, the dual specificity was attributed to the recognition of a consensus sequence (UUNAAAA, where N is any of the four nucleotides) that is present in both the tRNA anticodon loop and in the loop of 23S rRNA bearing residue U_{746} . Similarly, the *E. coli* tRNA- U_{54} -methyltransferase catalyzes the in vitro formation of m^5U_{54} in tRNA and m^5U_{788} in a 16S rRNA fragment produced by in vitro transcription (29). However, no m^5U residue has been found in naturally occurring *E. coli* rRNA, attesting that in vitro assays do not necessarily reflect the in vivo situation.

How does Pus1p act on substrates with different architectures? From the present results, we can conclude that Pus1p does not require an external guide RNA for the formation of Ψ_{44} in U2 snRNA, since we obtained Ψ_{44} formation in vitro by using an RNA transcript and the purified enzyme. Previous results on tRNA modification suggest that Pus1p requires a double-stranded RNA portion for binding, while the target uridine should be present in a rather flexible RNA structure (internal loop or even single-strand) (60). The target nucleotide in U2 snRNA is in a single-stranded region, which was confirmed by secondary-structure probing on the U2 snRNA transcript (data not shown), and no common sequence element was found for Pus1p target sites in tRNAs and in U2 snRNA. To identify the determinants required for U-to- Ψ conversion at position 44 by the Pus1p enzyme, a mutational analysis of the U2 snRNA target site is underway.

The presence of numerous Ψ and 2'-O-methylated residues in eukaryotic tRNAs, rRNAs, and UsnRNAs raises the important question of how many enzymes are required to account for all of these posttranscriptional modifications. To minimize the number of required enzymes, the cell seems to use two main strategies: (i) the involvement of guide RNAs conferring different specificities to a unique enzymatic machinery and (ii) the utilization of a unique multisite-specific enzyme that allows the formation of posttranscriptional modifications at different locations in a given type of RNA, as well as in different RNA substrates, and this is exemplified by our observation with the Pus1p enzyme.

However, the observation that Pus1p catalyzes only one of the Ψ residues detected in the *S. cerevisiae* UsnRNAs is in agreement with the previous observation for vertebrate UsnRNAs showing the involvement of several pseudouridine synthases (73–75). Other putative pseudouridine synthase genes are present in the *S. cerevisiae* genome, and we have initiated systematic gene disruption experiments in order to identify the enzymes responsible for the formation of the five other Ψ residues that were detected in this study.

ACKNOWLEDGMENTS

This work was supported by laboratory funds from the Ministère de la Recherche et de l'Enseignement Supérieur, the CNRS, specific research grants from the CNRS (Programme Physique et Chimie du Vivant 1997) and the Action de Recherche sur le Cancer. S. Massenet is a predoctoral fellow supported by a fellowship from the Ministère de la Recherche et de l'Enseignement Supérieur. Y. Motorin is the recipient of an Associated Researcher position at the CNRS (Poste Rouge). D. L. J. Lafontaine was the recipient of a fellowship from the European Commission.

The plasmids containing yeast UsnRNA genes were kindly provided by P. Fabrizio (Institut für Molekularbiologie und Tumorforschung, Marburg, Germany). We also thank R. Planta (University of Amsterdam, Amsterdam, The Netherlands) for providing the yeast strain with a disrupted *PUS4* gene and G. Simos for providing the *PUS1*, *PUS2*, and *PUS3*-disrupted strains. D. Tollervey (University of Edinburgh, Edinburgh, United Kingdom) is thanked for the *GAL::cbf5* RNA preparation, prepared in his laboratory. J. Ugolini is thanked for technical assistance. A. Mougin is acknowledged for useful advice at the beginning of this work. We also thank J.-P. Waller (CNRS, Gif-sur-Yvette, France) for critical reading of the manuscript and for useful comments.

REFERENCES

- Abovich, N., and M. Rosbash. 1984. Two genes for ribosomal protein 51 of *Saccharomyces cerevisiae* complement and contribute to the ribosomes. *Mol. Cell. Biol.* 4:1871-1879.
- Agris, P. F. 1996. The importance of being modified: roles of modified nucleosides and Mg²⁺ in RNA structure and function. *Prog. Nucleic Acid Res. Mol. Biol.* 53:79-129.
- Ares, M. 1986. U2 RNA from yeast is unexpectedly large and contains homology to vertebrate U4, U5, and U6 small nuclear RNAs. *Cell* 47:49-59.
- Ares, M., and A. H. Igel. 1990. Lethal and temperature-sensitive mutations and their suppressors identify an essential structural element in U2 small nuclear RNA. *Genes Dev.* 4:2132-2145.
- Arnez, J. G., and T. A. Steitz. 1994. Crystal structure of unmodified tRNA^{Gln} complexed with glutamyl-tRNA synthetase and ATP suggests a possible role for pseudouridines in stabilization of RNA structure. *Biochemistry* 33:7560-7567.
- Auffinger, P., and E. Westhof. 1998. Effects of pseudouridylation on tRNA hydration and dynamics: a theoretical approach, p. 103-111. In H. Grosjean and R. Benne (ed.), *The modification and editing of RNA*. ASM Press, Washington, D.C.
- Bakin, A., B. G. Lane, and J. Ofengand. 1994. Clustering of pseudouridine residues around the peptidyltransferase center of yeast cytoplasmic and mitochondrial ribosomes. *Biochemistry* 33:13475-13483.
- Bakin, A., and J. Ofengand. 1993. Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique. *Biochemistry* 32:9754-9762.
- Bakin, A., and J. Ofengand. 1995. Mapping of the 13 pseudouridine residues in *Saccharomyces cerevisiae* small subunit ribosomal RNA to nucleotide resolution. *Nucleic Acids Res.* 23:3290-3294.
- Becker, H. F., Y. Motorin, R. J. Planta, and H. Grosjean. 1997. The yeast gene YNL292w encodes a pseudouridine synthase (Pus4) catalyzing the formation of Ψ55 in both mitochondrial and cytoplasmic tRNAs. *Nucleic Acids Res.* 25:4493-4499.
- Behrens, S. E., and R. Lührmann. 1991. Immunoaffinity purification of a [U4/U6.U5] tri-snRNP from human cells. *Genes Dev.* 5:1439-1452.
- Branlant, C., A. Krol, J. P. Ebel, H. Gallinaro, E. Lazar, and M. Jacob. 1981. The conformation of chicken, rat and human U1A RNAs in solution. *Nucleic Acids Res.* 9:841-858.
- Branlant, C., A. Krol, J. P. Ebel, E. Lazar, H. Gallinaro, M. Jacob, J. Sri-Widada, and P. Jeanteur. 1980. Nucleotide sequences of nuclear U1A RNAs from chicken, rat, and man. *Nucleic Acids Res.* 8:4143-4154.
- Branlant, C., A. Krol, J. P. Ebel, E. Lazar, B. Haendler, and M. Jacob. 1982. U2 RNA shares a structural domain with U1, U4, and U5 RNAs. *EMBO J.* 1:1259-1265.
- Branlant, C., A. Krol, M. A. Machatt, and J. P. Ebel. 1979. Structural study of ribosomal 23S RNA from *Escherichia coli*. *FEBS Lett.* 107:177-181.
- Branlant, C., A. Krol, G. M. Veldman, J. Klootwijk, V. Regt, R. J. Planta, and J. Ebel. 1981. The primary and secondary structure of yeast 26S rRNA. *Nucleic Acids Res.* 9:6935-6952.
- Brody, E., and J. Abelson. 1985. The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* 228:963-967.
- Cech, T. R. 1986. The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell* 44:207-210.
- Cortes, J. J., E. J. Sontheimer, S. D. Seiwert, and J. A. Steitz. 1993. Mutations in the conserved loop of human U5 snRNA generate use of novel cryptic 5' splice sites *in vivo*. *EMBO J.* 12:5181-5189.
- Datta, B., and A. M. Weiner. 1991. Genetic evidence for base-pairing between U2 and U6 snRNA in mammalian mRNA splicing. *Nature* 352:821-824.
- Davis, D. R. 1998. Biophysical and conformational properties of modified nucleosides in RNA (nuclear magnetic resonance studies), p. 85-102. In H. Grosjean and R. Benne (ed.), *The modification and editing of RNA*. ASM Press, Washington, D.C.
- Davis, D. R., and C. D. Poulter. 1991. 1H-15N NMR studies of *Escherichia coli* tRNA (Phe) from HisT mutants: a structural role for pseudouridine. *Biochemistry* 17:4223-4231.
- Fabrizio, P., D. S. McPheeters, and J. Abelson. 1989. *In vitro* assembly of yeast U6 snRNP: a functional assay. *Genes Dev.* 3:2137-2150.
- Fournier, R., et al. Unpublished data.
- Ganot, P., M. L. Bortolin, and T. Kiss. 1997. Site-specific pseudouridine formation in pre-ribosomal RNA is guided by small nucleolar RNAs. *Cell* 89:799-809.
- Gaur, R. K., J. Valcárcel, and M. R. Green. 1995. Sequential recognition of the pre-mRNA branch point by U2AF⁶⁵ and a novel spliceosome-associated 28-kDa protein. *RNA* 1:407-417.
- Grabowski, P. J., S. R. Seiler, and P. A. Sharp. 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. *Cell* 42:345-353.
- Gu, J., J. R. Patton, S. Shimba, and R. Reddy. 1996. Localisation of modified nucleotides in *Saccharomyces pombe* spliceosomal small nuclear RNAs: modified nucleotides are clustered in functionally important regions. *RNA* 2:909-918.
- Gu, X. G., J. Ofengand, and D. V. Santi. 1994. *In vitro* methylation of *Escherichia coli* 16S rRNA by tRNA (m⁵U54)-methyltransferase. *Biochemistry* 33:2255-2261.
- Guthrie, C. 1991. Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. *Science* 253:157-163.
- Guthrie, C., and B. Patterson. 1988. Spliceosomal snRNAs. *Annu. Rev. Genet.* 22:387-419.
- Hall, S. L., and R. A. Padgett. 1994. Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. *J. Mol. Biol.* 239:357-365.
- Harada, F., N. Kato, and S. Nishimura. 1980. The nucleotide sequence of nuclear 4.8S RNA of mouse cells. *Biochem. Biophys. Res. Commun.* 95:1332-1340.
- Hashimoto, C., and J. A. Steitz. 1984. U4 and U6 RNAs coexist in a single small nuclear ribonucleoprotein particle. *Nucleic Acids Res.* 12:3283-3293.
- Hausner, T. P., L. M. Giglio, and A. M. Weiner. 1990. Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles. *Genes Dev.* 4:2146-2156.
- Hodges, P. E., and J. D. Beggs. 1994. RNA splicing. U2 fulfills a commitment. *Curr. Biol.* 4:264-267.
- Hughes, J. M. X., and M. Ares. 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.* 10:4231-4239.
- Igel, A. H., and M. Ares. 1988. Internal sequences that distinguish yeast from metazoan U2 snRNA are unnecessary for pre-mRNA splicing. *Nature* 334:450-453.
- Jackson, I. J. 1991. A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Res.* 19:3794-3798.
- Jacquier, A. 1990. Self-splicing group II and nuclear pre-mRNA introns: how similar are they? *Trends Biochem. Sci.* 15:351-354.
- Kandels-Lewis, S., and B. Séraphin. 1993. Role of U6 snRNA in 5' splice site selection. *Science* 262:2035-2039.
- Kato, N., and F. Harada. 1981. Nucleotide sequence of nuclear 5S of mouse cells. *Biochem. Biophys. Res. Commun.* 99:1468-1476.
- Kiss, László, Z., Y. Henry, and T. Kiss. 1998. Sequence and structural elements of methylation guide snRNAs essential for site-specific ribose methylation of pre-rRNA. *EMBO J.* 17:797-807.
- Koonin, E. V. 1996. Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases. *Nucleic Acids Res.* 24:2411-2415.
- Krämer, A. 1996. The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu. Rev. Biochem.* 65:367-409.
- Kretzner, L., A. Krol, and M. Rosbash. 1990. *Saccharomyces cerevisiae* U1 small nuclear RNA secondary structure contains both universal and yeast-specific domains. *Proc. Natl. Acad. Sci. USA* 87:851-855.
- Krol, A., C. Branlant, E. Lazar, H. Gallinaro, and M. Jacob. 1981. Primary and secondary structures of chicken, rat and man nuclear U4 RNAs. Homologies with U1 and U5 RNAs. *Nucleic Acids Res.* 9:2699-2716.
- Krol, A., H. Gallinaro, E. Lazar, M. Jacob, and C. Branlant. 1981. The nuclear 5S RNAs from chicken, rat and man. U5 RNAs are encoded by multiple genes. *Nucleic Acids Res.* 9:769-787.
- Lacroute, F. 1968. Regulation of pyrimidine biosynthesis in *Saccharomyces cerevisiae*. *J. Bacteriol.* 95:824-832.
- Lafontaine, D. L. J., C. Bousquet-Antonelli, Y. Henry, M. Caisergues-Ferrer, and D. Tollervey. 1998. The box H+ACA snRNAs carry the rRNA pseudouridine synthase. *Genes Dev.* 12:527-537.
- Lecointe, F., G. Simos, A. Sauer, E. C. Hurt, Y. Motorin, and H. Grosjean. 1998. Characterization of yeast protein Deg 1 as pseudouridine synthase (Pus 3) catalysing the formation of Ψ₃₈ and Ψ₃₉ in tRNA anticodon loop. *J. Biol. Chem.* 273:1316-1323.
- Lesser, C. F., and C. Guthrie. 1993. Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science* 262:1982-1988.
- Madhani, H. D., and C. Guthrie. 1992. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic acti-

- vation of the spliceosome. *Cell* 71:803-817.
54. Madhani, H. D., and C. Guthrie. 1994. Dynamic RNA-RNA interactions in the spliceosome. *Annu. Rev. Genet.* 28:1-26.
 55. Massenet, S., A. Mougín, and C. Branlant. 1998. Posttranscriptional modifications in the U small nuclear RNAs, p. 201-227. In H. Grosjean and R. Benne (ed.), *The modification and editing of RNA*. ASM Press, Washington, D.C.
 56. McPheeters, D. S., and J. Abelson. 1992. Mutational analysis of the yeast U2 snRNA suggests a structural similarity to the catalytic core of group I introns. *Cell* 71:819-831.
 57. McPheeters, D. S., P. Fabrizio, and J. Abelson. 1989. *In vitro* reconstitution of functional yeast U2 snRNPs. *Genes Dev.* 3:2124-2136.
 58. Méreau, A., R. Fournier, A. Grégoire, A. Mougín, P. Fabrizio, R. Lührmann, and C. Branlant. 1997. An *in vivo* and *in vitro* structure-function analysis of the *Saccharomyces cerevisiae* U3A snoRNP: protein-RNA contacts and base-pair interactions with the pre-ribosomal RNA. *J. Mol. Biol.* 273:552-571.
 59. Moore, M. J., C. C. Query, and P. A. Sharp. 1993. Splicing of precursors to messenger RNAs by the spliceosome, p. 303-358. In R. Gesteland and J. Athis (ed.), *The RNA world*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 60. Motorin, Y., G. Keith, C. Simon, D. Foiret, G. Simos, E. Hurt, and H. Grosjean. 1998. The yeast tRNA:pseudouridine synthase Pus1p displays a multisite substrate specificity. *RNA* 4:856-869.
 61. Mougín, A., A. Grégoire, J. Banroques, V. Ségault, R. Fournier, F. Brulé, M. Chevrier-Miller, and C. Branlant. 1996. Secondary structure of the yeast *Saccharomyces cerevisiae* pre-U3A snoRNA and its implication for splicing efficiency. *RNA* 2:1079-1093.
 62. Mount, S. M., and J. A. Steitz. 1981. Sequence of U1 RNA from *Drosophila melanogaster*: implications for U1 secondary structure and possible involvement in splicing. *Nucleic Acids Res.* 9:6351-6368.
 63. Mysliński, E., C. Branlant, E. D. Wieben, and T. Pederson. 1984. The small nuclear RNAs of *Drosophila*. *J. Mol. Biol.* 180:927-945.
 64. Mysliński, E., V. Ségault, and C. Branlant. 1990. An intron in the genes for U3 small nucleolar RNAs of the yeast *Saccharomyces cerevisiae*. *Science* 247:1213-1216.
 65. Newman, A. J., and C. Norman. 1991. Mutations in yeast U5 snRNA alter the specificity of 5' splice site cleavage. *Cell* 65:115-123.
 66. Newman, A. J., and C. Norman. 1992. U5 snRNA interacts with exon sequences at 5' and 3' splice sites. *Cell* 68:743-754.
 67. Newman, A. J., S. Teigelkamp, and J. D. Beggs. 1995. SnRNA interactions at 5' and 3' splice sites monitored by photoactivated crosslinking in yeast spliceosomes. *RNA* 1:968-980.
 68. O'Keefe, R. T., and A. J. Newmann. 1998. Functional analysis of the U5 snRNA loop 1 in the second catalytic step of yeast pre-mRNA splicing. *EMBO J.* 17:565-574.
 69. Ofengand, J., and A. Bakin. 1997. Mapping to nucleotide resolution of pseudouridine residues in a large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaeobacteria, mitochondria and chloroplasts. *J. Mol. Biol.* 17:1993-2005.
 70. Ofengand, J., A. Bakin, J. Wrzesinski, K. Nurse, and B. G. Lane. 1995. The pseudouridine residues of ribosomal RNA. *Biochem. Cell Biol.* 73:915-924.
 71. Parker, R., P. G. Siliciano, and C. Guthrie. 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* 49:229-239.
 72. Pascolo, E., and B. Séraphin. 1997. The branchpoint residue is recognized during commitment complex formation before being bulged out of the U2 snRNA-pre-mRNA duplex. *Mol. Cell Biol.* 17:3469-3476.
 73. Patton, J. R. 1991. Pseudouridine modification of U5 RNA in ribonucleoprotein particles assembled *in vitro*. *Mol. Cell Biol.* 11:5998-6006.
 74. Patton, J. R. 1993. Multiple pseudouridine synthase activities for small nuclear RNAs. *Biochem. J.* 290:595-600.
 75. Patton, J. R. 1994. Formation of pseudouridine in U5 small nuclear RNA. *Biochemistry* 33:10423-10427.
 76. Patton, J. R. 1994. Pseudouridine formation in small nuclear RNAs. *Biochimie* 76:1129-1132.
 77. Patton, J. R., M. R. Jacobson, and T. Pederson. 1994. Pseudouridine formation in U2 small nuclear RNA. *Proc. Natl. Acad. Sci. USA* 91:3324-3328.
 78. Peattie, D. A. 1979. Direct chemical method for sequencing RNA. *Proc. Natl. Acad. Sci. USA* 76:1760-1764.
 79. Qiang, W., and A. R. Krainer. 1997. Splicing of a divergent subclass of AT-AC introns requires the major spliceosomal snRNAs. *RNA* 3:586-601.
 80. Query, C. C., S. A. Strobel, and P. A. Sharp. 1996. Three recognition events at the branch-site adenine. *EMBO J.* 15:1392-1402.
 81. Reddy, R. 1988. Compilation of small RNA sequences. *Nucleic Acids Res.* 16(Suppl.):71-85.
 82. Reddy, R., D. Henning, and H. Busch. 1981. The primary nucleotide sequence of U4 RNA. *J. Biol. Chem.* 256:3532-3538.
 83. Reddy, R., D. Henning, and H. Busch. 1981. Pseudouridine residues in the 5'-terminus of uridine-rich nuclear RNA I (U1 RNA). *Biochem. Biophys. Res. Commun.* 98:1076-1083.
 84. Reddy, R., D. Henning, P. Epstein, and H. Busch. 1981. Primary and secondary structure of U2 snRNA. *Nucleic Acids Res.* 9:5645-5658.
 85. Reed, R. 1996. Initial splice-site recognition and pairing during pre-mRNA splicing. *Curr. Opin. Gen. Dev.* 6:215-220.
 86. Rinke, J., B. Appel, M. Digweed, and R. Lührmann. 1985. Localization of a base-paired interaction between small nuclear RNAs U4 and U6 in intact U4/U6 ribonucleoprotein particles by psoralen cross-linking. *J. Mol. Biol.* 185:721-731.
 87. Rosbash, M., P. K. W. Harris, J. L. Woolford, and J. L. Teem. 1981. The effect of temperature-sensitive mutants on the transcription products from cloned ribosomal protein genes of yeast. *Cell* 24:679-686.
 88. Ruby, S. W., T. H. Chang, and J. Abelson. 1993. Four yeast spliceosomal proteins (PRP5, PRP9, PRP11, and PRP21) interact to promote U2 snRNP binding to pre-mRNA. *Genes Dev.* 7:1909-1925.
 89. Rymond, B. C., and M. Rosbash. 1992. Yeast pre-mRNA splicing, p. 143-192. In E. W. Jones, J. R. Pringle, and J. R. Broach (ed.), *The molecular and cellular biology of the yeast Saccharomyces*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 90. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 91. Sawa, H., and Y. Shimura. 1992. Association of U6 snRNA with the 5' splice site region of pre-mRNA in the spliceosome. *Genes Dev.* 6:244-254.
 92. Ségault, V., A. Mougín, A. Grégoire, J. Banroques, and C. Branlant. 1992. An experimental study of *Saccharomyces cerevisiae* U3 snRNA conformation in solution. *Nucleic Acids Res.* 20:3443-3451.
 93. Ségault, V., C. L. Will, B. S. Sproat, and R. Lührmann. 1995. *In vitro* reconstitution of mammalian U2 and U5 snRNPs active in splicing: Sm proteins are functionally interchangeable and are essential for the formation of functional U2 and U5 snRNPs. *EMBO J.* 14:4010-4021.
 94. Séraphin, B., L. Kretzner, and M. Rosbash. 1988. A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. *EMBO J.* 7:2533-2538.
 95. Séraphin, B., and M. Rosbash. 1990. Exon mutations uncouple 5' splice site selection from U1 snRNA pairing. *Cell* 63:619-629.
 96. Siliciano, P. G., and C. Guthrie. 1988. 5' splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. *Genes Dev.* 2:1258-1267.
 97. Simos, G., H. Tekotte, H. Grosjean, A. Segref, K. Sharma, D. Tollervey, and E. C. Hurt. 1996. Nuclear pore proteins are involved in the biogenesis of functional tRNA. *EMBO J.* 15:2270-2284.
 98. Solymosy, F., and T. Pollák. 1993. Uridylate-rich small nuclear RNAs (UsnRNAs), their genes and pseudogenes, and UsnRNPs in plants: structure and function. A comparative approach. *Crit. Rev. Plant Sci.* 12:275-369.
 99. Sontheimer, E. J., and J. A. Steitz. 1993. The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science* 262:1989-1996.
 100. Sprinzl, S., C. Horn, M. Brown, A. Ioudovitch, and S. Steinberg. 1998. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* 26:148-153.
 101. Steitz, T. A., and J. A. Steitz. 1993. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. USA* 90:6498-6502.
 102. Sun, J. S., and J. L. Manley. 1995. A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes Dev.* 9:843-854.
 103. Szukálek, A., E. Mysliński, A. Mougín, R. Lührmann, and C. Branlant. 1995. Phylogenetic conservation of modified nucleotides in the terminal loop 1 of the spliceosomal U5 snRNA. *Biochimie* 77:16-21.
 104. Szweykowska-Kulinska, Z., B. Senger, G. Keith, F. Fasiolo, and H. Grosjean. 1994. Intron-dependent formation of pseudouridines in the anticodon of *Saccharomyces cerevisiae* minor tRNA^{Ile}. *EMBO J.* 13:4636-4644.
 105. Tarrn, W. Y., and J. A. Steitz. 1996. Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science* 273:1824-1832.
 106. Tarrn, W. Y., and J. A. Steitz. 1996. A novel spliceosome containing U11, U12 and U5 snRNPs excises a minor class (AT-AC) intron *in vitro*. *Cell* 84:1-20.
 107. Teem, J. L., and M. Rosbash. 1983. Expression of a β -galactosidase gene containing the ribosomal protein 51 intron is sensitive to the *rna2* mutation in yeast. *Proc. Natl. Acad. Sci. USA* 80:4403-4407.
 108. Tycowski, K. T., Z. H. You, P. J. Graham, and J. A. Steitz. Personal communication.
 109. Valcárcel, J., and M. R. Green. 1996. The SR protein family: pleiotropic functions in pre-mRNA splicing. *Trends Biochem. Sci.* 21:296-301.
 110. Wells, S. E., and M. Ares. 1994. Interactions between highly conserved U2 small nuclear RNA structures and Prp5p, Prp9p, Prp11p, and Prp21p proteins are required to ensure integrity of the U2 small nuclear ribonucleoprotein in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 14:6337-6349.
 111. Wells, S. E., M. Neville, M. Haynes, J. Wang, H. Igel, and M. Ares. 1996. CUS1, a suppressor of cold-sensitive U2 snRNA mutations, is a novel yeast splicing factor homologous to human SAP 145. *Genes Dev.* 10:220-232.
 112. Westhof, E., P. Dumas, and D. Moras. 1985. Crystallographic refinement of

- yeast aspartic acid transfer RNA. *J. Mol. Biol.* 184:119-145.
113. Westhof, E., P. Dumas, and D. Moras. 1988. Hydration of transfer RNA molecules: a crystallographic study. *Biochimie* 70:145-165.
 114. Will, C. L., and R. Lührmann. 1997. Protein functions in pre-mRNA splicing. *Curr. Opin. Cell Biol.* 9:320-328.
 115. Wrzesinski, J., K. Nurse, A. Bakin, B. G. Lane, and J. Ofengand. 1995. A dual-specificity pseudouridine synthase: an *Escherichia coli* synthase purified and cloned on the basis of its specificity for Psi 746 in 23S RNA is also specific for Psi 32 in tRNA^{Phe}. *RNA* 1:437-448.
 116. Wu, J., and J. L. Manley. 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base-pairing. *Genes Dev.* 3:1553-1561.
 117. Wu, J., and J. L. Manley. 1991. Base-pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. *Nature* 352:818-821.
 118. Wyatt, J. R., E. J. Sontheimer, and J. A. Steitz. 1992. Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. *Genes Dev.* 6:2542-2553.
 119. Yan, D., and M. J. Ares. 1996. Invariant U2 RNA sequences bordering the branchpoint recognition region are essential for interaction with yeast SF3A and SF3B subunits. *Mol. Cell. Biol.* 16:818-828.
 120. Yu, Y., M. Shu, and J. A. Steitz. 1998. Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J.* 17:5783-5795.
 121. Zerby, D. B., and J. R. Patton. 1996. Metabolism of pre-messenger RNA splicing cofactors: modification of U6 RNA is dependent on its interaction with U4 RNA. *Nucleic Acids Res.* 24:3583-3589.
 122. Zerby, D. B., and J. R. Patton. 1997. The modification of U4 RNA requires U6 RNA and multiple pseudouridine synthases. *RNA* 25:4808-4815.
 123. Zhuang, Y., and A. M. Weiner. 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell* 46:827-835.
 124. Zhuang, Y., and A. M. Weiner. 1989. A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes Dev.* 3:1545-1552.

1.3. Recherche des résidus 2'-O-méthylés portés par le snRNA U5 de *S. cerevisiae*

1.3.1. Intérêt de l'étude et problèmes techniques rencontrés

Comme nous l'avons vu dans la revue insérée dans le chapitre introduction, les deux résidus modifiés les plus fréquents dans les UsnRNA sont les résidus Ψ et 2'-O-méthylés. Il était donc intéressant d'entreprendre également une étude des 2'-O-méthylations portées par les UsnRNA de *S. cerevisiae* et, pour cela, plusieurs techniques de détection de ces modifications ont été testées.

Tout d'abord, nous avons testé l'approche proposée par Maden *et al.* (1995), qui est basée sur le fait que des pauses de la transcriptase inverse peuvent être observées au niveau des résidus 2'-O-méthylés, lorsque la concentration en ddNTP, utilisée pour l'élongation, n'est pas optimale. Les tests que nous avons réalisés sur les UsnRNA de cellules HeLa, pour lesquels les positions des 2'-O-méthylations étaient connues, ont montré que cette technique ne pouvait pas être employée pour une recherche systématique et exhaustive des résidus 2'-O-méthylés dans un ARN. D'autres auteurs avaient fait indépendamment la même constatation. En effet, certains nucléotides ne portant pas de méthylation sur le ribose peuvent conduire à des pauses de la transcriptase inverse lorsque la concentration de ddNTP est faible et, au contraire, certains résidus 2'-O-méthylés ne sont pas détectés par cette méthode.

La méthode utilisée dans le passé par l'équipe consistait à détecter l'absence de rupture des liaisons phosphodiester en 3' des résidus 2'-O-méthylés, lors d'une hydrolyse thermique de l'ARN à 96°C dans l'eau (Krol *et al.*, 1981a; 1981b). Une possibilité était d'utiliser cette méthode en réalisant l'analyse par la transcriptase inverse. Cependant, les nombreuses pauses que cette enzyme marque lors de la synthèse de l'ADNc ne permettent pas d'obtenir un profil clair, avec une absence totale de bandes d'ADNc au niveau des résidus 2'-O-méthylés.

Ainsi, aucune méthode simple et fiable, basée sur une analyse par extension d'amorces, n'était disponible pour l'analyse des 2'-O-méthylations. La seule méthode sûre était, donc, celle précédemment utilisée par notre équipe et qui imposait de marquer les UsnRNA à leur extrémité 3' et de les purifier à homogénéité, avant de les soumettre à une hydrolyse à 96°C, les liaisons phosphodiester non clivées étant ensuite identifiées par électrophorèse en gel de polyacrylamide. Cette approche ne pouvait pas s'appliquer aux snRNA U1 et U2 dans leur totalité, du fait de leur grande taille. Elle était également très difficilement applicable au snRNA U6 qui porte un phosphate 2'-3' cyclique à son extrémité 3' (Lund et Dahlberg, 1992), ce qui limite fortement l'efficacité de marquage en 3'. Dans le cas

des snRNA U1 et U2, une possibilité était d'essayer de les fragmenter, puis de purifier et de marquer en 3' les fragments obtenus, avant de réaliser l'hydrolyse thermique dans l'eau. Nous avons tenté d'obtenir des clivages spécifiques par la RNase H, en utilisant des oligonucléotides ADN complémentaires des régions des UsnRNA à cliver. Cependant, les rendements de purification des fragments ont été trop faibles pour pouvoir mener à bien une analyse des 2'-O-méthylations.

Le seul UsnRNA de *S. cerevisiae*, pour lequel les rendements de marquage et de purification ont été suffisants pour réaliser une telle analyse a été le snRNA U5. Ayant de ce fait développé une méthode permettant de purifier les snRNA U5S et U5L, nous en avons tiré parti pour rechercher si le résidu Ψ_{99} , précédemment mis en évidence dans le snRNA U5 à partir d'un mélange d'ARN totaux (voir publication n°2), était bien présent dans les deux formes du snRNA U5.

1.3.2. La boucle terminale I du snRNA U5 de *S. cerevisiae* ne contient vraisemblablement pas de résidus 2'-O-méthylés. Le résidu Ψ_{99} est présent à la fois dans le snRNA U5S et le snRNA U5L.

a) Développement de la méthode d'analyse

Une simple purification par électrophorèse des snRNA U5 contenus dans un extrait total d'ARN de levure, après marquage en 3', ne s'est pas avérée suffisante pour obtenir un degré de pureté satisfaisant. Nous avons donc réalisé une première étape de purification des snRNA U5 de *S. cerevisiae*, à partir des ARN totaux, par hybridation d'un oligonucléotide biotinyllé. Cet oligonucléotide était complémentaire aux nts 159 à 182 du snRNA U5 et permettait ainsi la sélection des deux formes de cet ARN, la forme longue (U5L) et la forme courte (U5S) (Fig. 45A). La sélection a été réalisée sur des billes de streptavidine-agarose (comme décrit dans la publication n°3). Les ARN ont ensuite été décrochés des billes par hydrolyse à la protéinase K, suivie d'une extraction au phénol. Ces ARN ont ensuite été marqués en 3' par du pCp [$5'$ - 32 P] et de l'ARN T4 ligase, puis fractionnés par électrophorèse en gel de polyacrylamide dénaturant. Environ 5000 à 10000 cpm de chaque ARN purifié ont ensuite été incubés 1h30 à 96°C. Un séquençage chimique par l'hydrazine et le DEPC (Peattie, 1979) a été réalisé en parallèle, afin de repérer, respectivement, les résidus U et A.

b) Examen des résultats obtenus

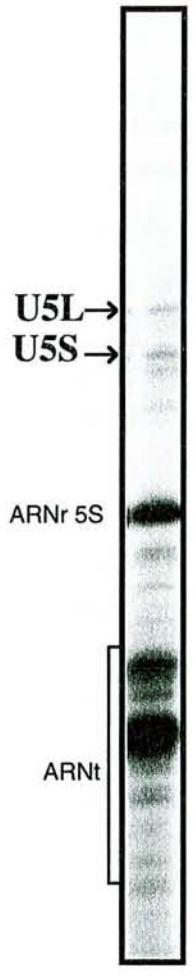
Dans toutes les espèces précédemment étudiées, trois résidus 2'-O-méthylés avaient été retrouvés dans la boucle terminale I du snRNA U5 (Fig. 7 de la revue et Fig. 45C) (Szkukalek *et al.*, 1995) et ceci même chez la levure *S. pombe* (Gu *et al.*, 1996). Or, sur la base de l'analyse des produits d'hydrolyse à 96°C du snRNA U5S de *S. cerevisiae*, qui est

illustrée dans la Figure 45B, nous n'avons pas observé d'absence de coupures de liaisons phosphodiester dans la boucle terminale I. En effet, une bande nette est observée aux 3 positions où une absence de clivage serait attendue si les 2'-O-méthylations présentes dans les autres espèces, Gm₃₇, Um₄₁ et Cm₄₅ (suivant la numérotation du snRNA U5 humain, Fig. 45C), existaient chez *S. cerevisiae*. Seule la bande observée au niveau du résidu G est d'intensité un peu plus faible. Cependant, dans l'ensemble de l'analyse, les clivages au niveau des résidus G ont tous conduit à des bandes d'intensité plus faible. Il semble donc qu'il n'y ait pas de résidus 2'-O-méthylés dans la boucle terminale I du snRNA U5S. Nous avons obtenu le même résultat en analysant le snRNA U5L. La levure *S. cerevisiae* représente donc une exception parmi tous les organismes étudiés. On peut imaginer que des différences subtiles dans les mécanismes fins de reconnaissance des extrémités exoniques par les complexes spliceosomaux existent entre *S. cerevisiae* et les autres organismes.

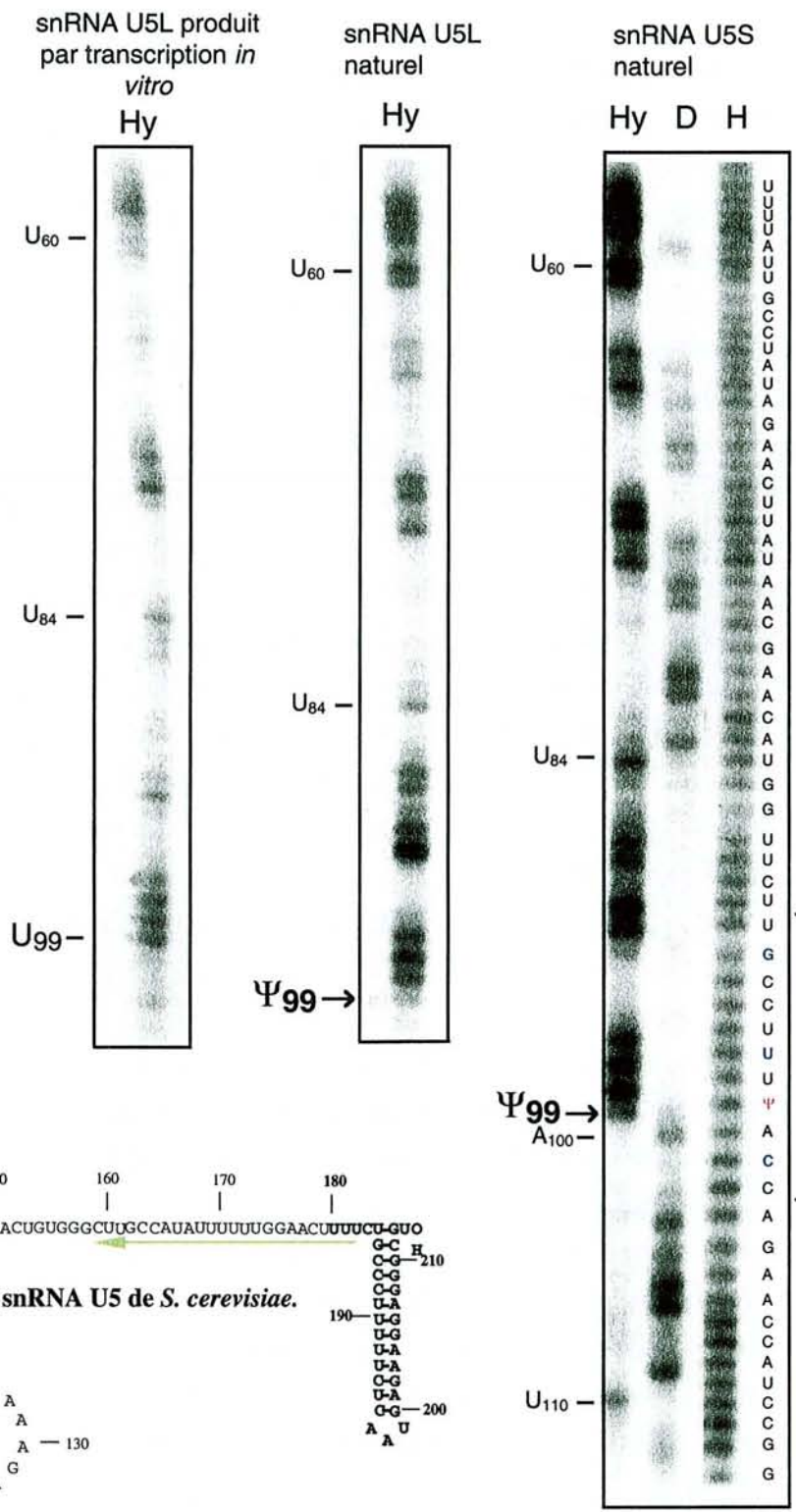
L'analyse illustrée dans la Figure 45B montre aussi qu'il ne semble pas y avoir de 2'-O-méthylations sur le ribose à d'autres positions au sein du snRNA U5, du moins à un taux détectable par le type d'analyse utilisé. Il est à noter, qu'à l'exception des 2'-O-méthylations des résidus jouxtant la structure coiffe, les seuls résidus 2'-O-méthylés mis en évidence dans les snRNA U5 de toutes les espèces étudiées sont les trois contenus dans la boucle terminale I. Vu la taille des snRNA U5 de *S. cerevisiae*, même pour le snRNA U5S, il a été difficile, à partir d'un marquage en 3', d'étudier les 2'-O-méthylations à l'extrémité 5'-terminale au niveau de la coiffe. Nous n'avons donc pas d'information sur cette région.

Au cours de cette analyse, nous avons pu observer, par ailleurs, que le résidu Ψ en position 99, est présent aussi bien dans le snRNA U5L que dans le snRNA U5S. Le taux de modification par l'hydrazine à cette position était plus faible qu'au niveau des résidus U. Néanmoins, l'absence de modification par l'hydrazine n'est pas complète pour ces deux ARN. Par quantification au Phospho Imager des différentes bandes du gel, le taux de conversion du résidu U₉₉ en résidu Ψ pourrait être compris entre 50% et 60%. Comme nous l'avons déjà mentionné, ce type d'analyse par l'hydrazine ne permet pas une quantification précise. Cependant, deux points semblent clairs: d'une part, la conversion ne se fait pas à 100%, d'autre part, elle a lieu aussi bien dans la forme longue que dans la forme courte du snRNA U5. Seuls quelques exemples de mise en évidence de modifications partielles à une position donnée d'un ARN, ont été décrits dans la littérature. C'est le cas pour la conversion du résidu U₆₅ en résidu Ψ de l'ARNt^{Trp} cytoplasmique (Keith *et al.*, 1971) et du résidu Ψ_{28} de l'ARNt^{Ser}_(UGA) mitochondrial de *S. cerevisiae* (Martin *et al.*, 1982). C'est également le cas pour plusieurs types de modifications dans les ARNt^{Lys}, ARNt^{Leu}_(UAA) et ARNt^{Pro} mitochondriaux humains (Brulé *et al.*, 1998; Helm *et al.*, 1999).

A/



B/



C/

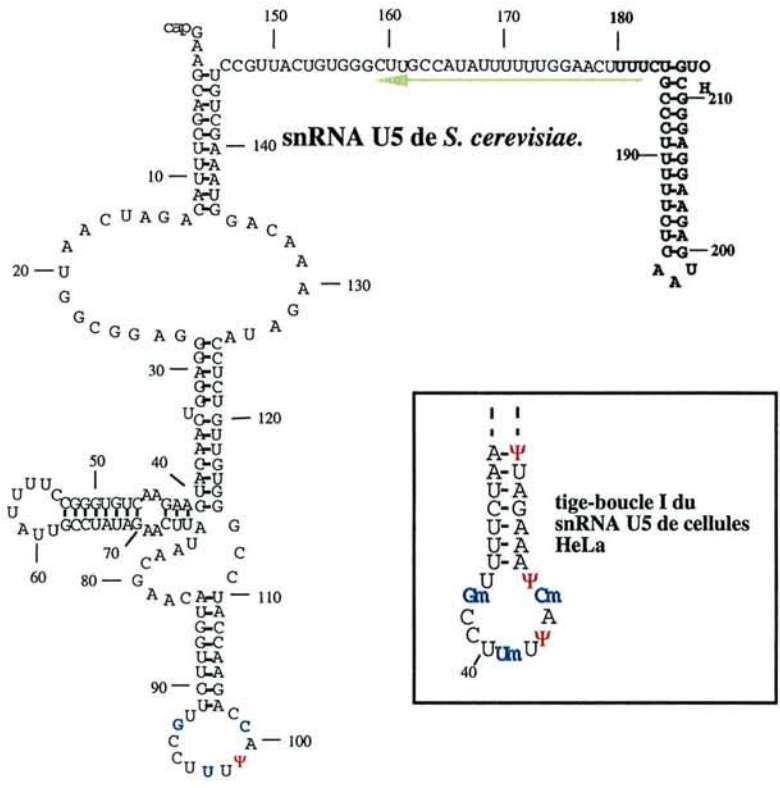


Figure 45: Analyse des résidus 2'-O-méthylés dans la boucle terminale I des snRNA U5 de *S. cerevisiae*.

Panneau A: Fractionnement par électrophorèse en gel de polyacrylamide dénaturant à 15% des snRNA U5S et U5L de *S. cerevisiae*, marqués à leur extrémité 3'. Au préalable, ces ARN avaient été partiellement purifiés, à partir d'un mélange d'ARN totaux, par hybridation d'un oligonucléotide biotinylé qui leur était complémentaire et par sélection de l'hybride obtenu sur billes de streptavidine-agarose. Les bandes correspondant aux ARN U5S et U5L sont indiquées. Après élution, ces deux ARN ont été traités à 96°C, afin d'identifier les résidus 2'-O-méthylés. Il faut noter que le taux de marquage en 3' des différents ARN, présents dans le mélange, n'est pas proportionnel à la quantité de chaque ARN. Les ARNt, dont l'extrémité 3' est très accessible, sont de meilleurs substrats de l'ARN T4 ligase et sont donc très fortement marqués.

Panneau B: Les analyses ont été réalisées sur des transcrits, produits *in vitro*, correspondants au snRNA U5L et sur les snRNA U5S et U5L naturels purifiés. Les échantillons ont été soit soumis à un traitement thermique à 96°C pendant 1h30 (H), soit traités par l'hydrazine (Hy) ou par le DEPC (D). Dans ces deux derniers cas, un traitement à l'aniline a été réalisé pour cliver les résidus modifiés. Les produits de coupure, générés par ces différentes réactions chimiques, ont ensuite été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 15%.

L'analyse montre un défaut de modification par l'hydrazine en position 99, aussi bien dans le snRNA U5S que dans le snRNA U5L, ce qui démontre la présence du résidu Ψ_{99} . La boucle terminale I est indiquée pour le snRNA U5S. Comme on peut le constater, après hydrolyse à 96°C, une bande est obtenue à chaque position correspondant à un résidu de la boucle, ce qui suggère fortement l'absence de 2'-O-méthylations. Les nucléotides de la boucle qui sont 2'-O-méthylés dans les autres espèces sont indiqués en bleu.

Panneau C: Structure du snRNA U5 de *S. cerevisiae* (Guthrie et Patterson, 1988). Les nucléotides excédentaires dans la forme longue de cet ARN sont indiqués en gras. La région complémentaire à l'oligonucléotide, utilisé pour la purification des deux formes du snRNA U5, est indiquée, ainsi que le résidu Ψ_{99} (en rouge). Les nucléotides modifiés portés par la structure tige-boucle terminale I du snRNA U5 humain sont indiqués dans l'encart à droite du panneau.

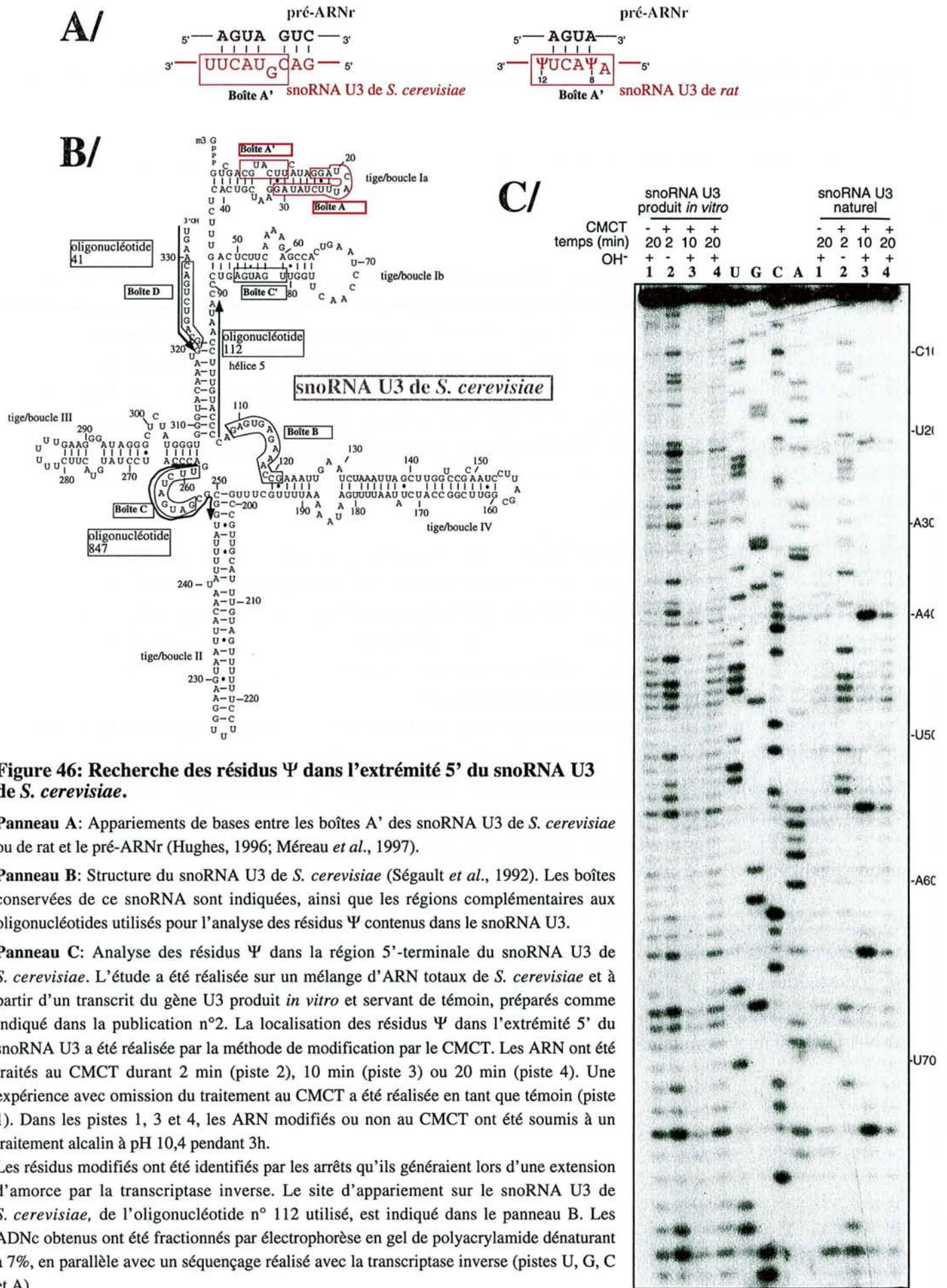


Figure 46: Recherche des résidus Ψ dans l'extrémité 5' du snoRNA U3 de *S. cerevisiae*.

Panneau A: Appariements de bases entre les boîtes A' des snoRNA U3 de *S. cerevisiae* ou de rat et le pré-ARNr (Hughes, 1996; Méreau *et al.*, 1997).

Panneau B: Structure du snoRNA U3 de *S. cerevisiae* (Ségault *et al.*, 1992). Les boîtes conservées de ce snoRNA sont indiquées, ainsi que les régions complémentaires aux oligonucléotides utilisés pour l'analyse des résidus Ψ contenus dans le snoRNA U3.

Panneau C: Analyse des résidus Ψ dans la région 5'-terminale du snoRNA U3 de *S. cerevisiae*. L'étude a été réalisée sur un mélange d'ARN totaux de *S. cerevisiae* et à partir d'un transcrit du gène U3 produit *in vitro* et servant de témoin, préparés comme indiqué dans la publication n°2. La localisation des résidus Ψ dans l'extrémité 5' du snoRNA U3 a été réalisée par la méthode de modification par le CMCT. Les ARN ont été traités au CMCT durant 2 min (piste 2), 10 min (piste 3) ou 20 min (piste 4). Une expérience avec omission du traitement au CMCT a été réalisée en tant que témoin (piste 1). Dans les pistes 1, 3 et 4, les ARN modifiés ou non au CMCT ont été soumis à un traitement alcalin à pH 10,4 pendant 3h.

Les résidus modifiés ont été identifiés par les arrêts qu'ils génèrent lors d'une extension d'amorce par la transcriptase inverse. Le site d'appariement sur le snoRNA U3 de *S. cerevisiae*, de l'oligonucléotide n° 112 utilisé, est indiqué dans le panneau B. Les ADNc obtenus ont été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 7%, en parallèle avec un séquençage réalisé avec la transcriptase inverse (pistes U, G, C et A).

L'absence d'arrêt d'élongation lors de l'analyse du snoRNA U3, après traitement au CMCT et hydrolyse alcaline, indique l'absence de résidus Ψ dans la région analysée.

2. RECHERCHE DE RESIDUS Ψ DANS LE snoRNA U3 DE *S. cerevisiae*

Comme nous l'avons vu dans l'introduction bibliographique, en plus des UsnRNA spliceosomaux, certains ARN nucléolaires (snoRNA), tels que le snoRNA U3, portent des modifications post-transcriptionnelles chez les vertébrés. Le snoRNA U3 est impliqué dans les étapes de clivage des pré-ARNr (pour revue, Maxwell et Fournier, 1995). Deux résidus Ψ ont été mis en évidence dans la région 5'-terminale du snoRNA U3 de rat (Reddy *et al.*, 1985). Cette région contient une boîte A' très conservée en séquence dans toutes les espèces, qui s'apparie au pré-ARNr (Fig. 46A). Cet appariement semble nécessaire aux étapes préliminaires de clivage assurant la maturation du pré-ARNr (Hughes, 1996; Méreau *et al.*, 1997). Ainsi, comme pour les UsnRNA spliceosomaux, des modifications post-transcriptionnelles sont localisées dans les segments fonctionnels des snoRNA.

Il était donc intéressant de savoir si les deux résidus Ψ , mis en évidence dans la séquence conservée A' du snoRNA U3 de rat, étaient également présents chez *S. cerevisiae*. Pour rechercher d'éventuels résidus Ψ dans le snoRNA U3 de *S. cerevisiae*, j'ai à nouveau utilisé la méthode d'identification de ces modifications par le CMCT. Cet ARN est un peu plus long que celui de vertébrés (333 nts au lieu de 214 nts), néanmoins j'ai pu analyser la totalité de la molécule, en utilisant différents oligonucléotides amorces de la transcriptase inverse (Fig. 46B). L'analyse de la région 5'-terminale de l'ARN est illustrée dans la Figure 46C. Les deux résidus Ψ détectés en position 8 et 12 dans l'ARN de rat ne sont pas présents dans le snoRNA U3 de *S. cerevisiae*. De plus, nous n'avons détecté aucun résidu Ψ dans l'ensemble de la molécule. La diminution du nombre de résidus Ψ dans les petits ARN nucléaires semble donc avoir un caractère général chez *S. cerevisiae*.

Comme nous l'avons mentionné dans la partie Introduction, les résidus Ψ peuvent stabiliser les hélices d'ARN qui les contiennent (pour revues, Davis, 1995; 1998; Auffinger et Westhof, 1998). Il est donc possible que le résidu Ψ_8 du snoRNA U3 de rat, localisé dans la partie de la boîte A' s'appariant au pré-ARNr (Fig. 46A), participe à la stabilisation de cette interaction. Chez *S. cerevisiae*, l'appariement avec le pré-ARNr, impliquant la boîte A' du snoRNA U3, contient trois paires de bases supplémentaires par rapport à celui formé chez les vertébrés (Fig. 46A). Cette extension, présente chez *S. cerevisiae*, a été retrouvée seulement chez des espèces eucaryotes très anciennes, telles que *T. thermophila* ou *T. brucei* (Fig. 4 de la revue) (Méréau *et al.*, 1997). Il serait intéressant d'étudier si la région 5'-terminale du snoRNA U3 de ces espèces est également dépourvue de résidus Ψ .

3. IDENTIFICATION DES RESIDUS Ψ DANS LES UsnRNA DES SPLICEOSOMES U12-DEPENDANTS DE CELLULES HeLa

Comme nous l'avons indiqué dans la partie Introduction de ce manuscrit, pendant que je réalisais mes travaux de thèse, une deuxième classe de spliceosomes a été mise en évidence (pour revue, Wu et Krainer, 1999). Ces spliceosomes contiennent quatre UsnRNA, qui leur sont spécifiques et qui sont essentiels à leur fonction (Tarn et Steitz, 1996a et 1996b). Deux de ces UsnRNA, les snRNA U11 et U12, avaient été mis en évidence, depuis longtemps, dans les noyaux de vertébrés, sans que leurs fonctions soient connues (Montzka et Steitz, 1988). Néanmoins, aucune analyse des résidus modifiés internes portés par ces ARN n'avait été réalisée. Seule la présence d'une coiffe $m^{2,2,7}G$ avait été mise en évidence (Montzka et Steitz, 1988). En ce qui concerne les deux autres UsnRNA spécifiques des spliceosomes U12-dépendants, les snRNA U4atac et U6atac, leur découverte est très récente (Tarn et Steitz, 1996b). Leur séquence a été établie à partir de l'ADN et, là encore, aucune analyse des résidus modifiés internes n'avait été réalisée. Dans la mesure où, aussi bien chez les vertébrés, les plantes, les insectes et les levures, les UsnRNA des spliceosomes U2-dépendants contiennent des résidus modifiés au sein de leurs séquences fonctionnelles, on pouvait penser qu'il en était de même pour les UsnRNA des spliceosomes U12-dépendants. Une analyse de ces modifications était susceptible d'apporter de nombreuses informations importantes:

⇒ Est-ce qu'en dépit de la forte différence de séquence entre UsnRNA des spliceosomes U12-dépendants et UsnRNA des spliceosomes U2-dépendants, il existe des modifications post-transcriptionnelles communes aux UsnRNA homologues des deux types de spliceosomes?

⇒ Nos études antérieures montraient une différence très importante du degré de modification des UsnRNA des spliceosomes U2-dépendants de vertébrés, par rapport aux UsnRNA de *S. cerevisiae*. Cette différence était-elle due à un requis plus élevé en modifications post-transcriptionnelles du système spliceosomal U2-dépendant de vertébrés et de plantes, par rapport au système levure? Ou, tout simplement, à la présence d'un plus grand nombre d'enzymes de modification chez les vertébrés et les plantes, qui aurait pour conséquence l'existence d'un plus grand nombre de modifications post-transcriptionnelles dans les ARN de ces organismes? Avec les UsnRNA des spliceosomes U12-dépendants, nous pouvions attendre une réponse intéressante sur ce point, puisque nous pouvions étudier deux ensembles de UsnRNA produits dans un même organisme et donc soumis à l'action possible du même ensemble d'enzymes de modification.

J'ai donc entrepris l'étude des modifications post-transcriptionnelles portées par les UsnRNA des spliceosomes U12-dépendants de cellules HeLa (U11, U12, U4atac et U6atac). Comme dans le cas des UsnRNA de *S. cerevisiae*, les UsnRNA des spliceosomes U12-

dépendants sont très peu abondants dans un extrait nucléaire, contrairement aux UsnRNA des spliceosomes U2-dépendants. Par contre, contrairement à ce qui se passe pour les UsnRNA de *S. cerevisiae*, leur taille est compatible avec une analyse globale, puisqu'elle est voisine de celle des UsnRNA des spliceosomes U2-dépendants.

Comme dans le cas des UsnRNA de *S. cerevisiae*, nous avons concentré nos efforts sur l'identification des résidus Ψ , ceci toujours en employant la méthode de modification par le CMCT. L'analyse par modification à l'hydrazine a été utilisée comme contrôle. Il faut souligner que certaines difficultés techniques ont été rencontrées, tout particulièrement pour les snRNA U4atac et U6atac, qui sont présents à de très faibles concentrations. En ce qui concerne le snRNA U4atac, j'ai dû modifier les conditions de modification par le CMCT, afin de pouvoir travailler sur de plus grandes quantités d'ARN totaux. Pour le snRNA U6atac, j'ai dû mettre au point une méthode de purification. J'ai employé la même approche que pour les snRNA U5 de *S. cerevisiae*: l'hybridation d'un oligonucléotide complémentaire biotinylé, ce qui a permis de réaliser une purification partielle du snRNA U6atac sur billes de streptavidine-agarose.

L'identification des résidus Ψ portés par les UsnRNA des spliceosomes U12-dépendants est présentée dans la publication n°3.

Publication n°3:

A limited number of pseudouridine residues in the human atac spliceosomal UsnRNAs as compared to human major spliceosomal UsnRNAs.

Massenet S. and C. Branlant

RNA, sous presse

Publication n°3

**A limited number of pseudouridine residues in the human
atac spliceosomal UsnRNAs as compared to human major
spliceosomal UsnRNAs.**

S. Massenet and C. Branlant.

RNA, sous presse.

A limited number of pseudouridine residues in the human atac spliceosomal UsnRNAs as compared to human major spliceosomal UsnRNAs.

Séverine Massenet and Christiane Branlant *

Laboratoire de Maturation des ARN et Enzymologie Moléculaire, UMR 7567 CNRS-UHP
Nancy I, Faculté des Sciences, BP 239, 54506 Vandoeuvre-les-Nancy Cedex, France

* Corresponding author : Christiane Branlant

Laboratoire de Maturation des ARN et Enzymologie Moléculaire,
UMR 7567 CNRS-UHP Nancy I
Faculté des Sciences, BP 239, 54506 Vandoeuvre-les-Nancy Cedex, France

Tel : 33.3.83.91.20.92

Fax : 33.3.83.91.20.93

e-mail : cbranlant@maem.uhp-nancy.fr

Running title : Pseudouridine residues in the HeLa atac UsnRNAs

Abbreviations : Ψ , pseudouridine; nt, nucleotides; RT, reverse transcription; CMCT, N-cyclohexyl-N'-(2-morpholinoethyl)-carbodiimid metho-p-toluolsulfonat; BS, branch site.

ABSTRACT

Two forms of spliceosomes were found in higher eukaryotes. The major form contains the U1, U2, U4, U5 and U6 snRNAs, the minor form contains the U11, U12, U4atac, U5 and U6atac snRNAs. Assembly and function of the major form are based on a complex dynamic of UsnRNA-UsnRNA and UsnRNA-premRNA interactions, and the involved UsnRNA segments are highly post-transcriptionally modified in plants and vertebrates. To further characterize the minor form of spliceosomes, we looked for the Ψ residues in HeLa cell's U11, U12, U4atac and U6atac snRNAs, using chemical approaches. Four Ψ residues were detected in total for these four atac UsnRNAs, against 20 in their counterparts of the major spliceosomes. The two Ψ residues detected in U12 are also found in U2 snRNAs. One of them belongs to the branch site-recognition sequence. It forms one of the base pairs that bulge out the A residue, responsible for the nucleophilic attack. Conservation of this strategic Ψ residue probably reflects a functional role. Another Ψ residue was detected in a U4atac snRNA segment involved in formation of helix II with U6atac. The fourth one was detected in the additional stem-loop structure present at the 3' end of U6atac snRNA. Differences in Ψ content of the atac and major UsnRNAs of human cells may participate to the differentiation of the two splicing systems. Based on secondary structure similarity, U2 and U12 snRNAs on the one hand, and U4 and U4atac snRNAs on the other hand, may share common Ψ synthases.

Key words: splicing/HeLa cells/post-transcriptional modifications/U12-dependent intron/atac spliceosome.

INTRODUCTION

For a very long time after their discovery, the introns, from the nuclear genes that encode proteins, were considered to be of a unique type, with 5'-GT and AG-3' dinucleotides at the borders (for review, Moore et al., 1993). They were thought to be all eliminated by the same processing machinery, the spliceosome, composed of 5 UsnRNAs (U1, U2, U4, U5 and U6) and a huge number of proteins (for review, Moore et al., 1993). The increasing number of available genomic and cDNA sequences over the last few years, has revealed the existence of introns that do not fulfill the GT-AG rule (Jackson, 1991 and for review, Wu & Krainer, 1999), and they were found to be spliced by a minor form of spliceosomes containing 4 specific UsnRNAs (U11, U12, U4atac and U6atac) and U5 snRNA (Tarn & Steitz, 1996a and b).

UsnRNAs were found to play a crucial role in the assembly and function of the major form of spliceosomes and, in the present stage of the knowledge, UsnRNAs from the minor form of spliceosomes are supposed to play very similar roles (for review, Wu & Krainer, 1999). Briefly, conventional GT-AG introns show the sequence GTRAGT at their 5' end, and YNYTRAC at their branch site (BS). These two sequences are essential for the early steps of assembly of the major form of spliceosomes: U1 snRNA interacts with the 5' sequence and U2 snRNA with the BS sequence (for review, Nilsen, 1998). Similarly, the AT-AC introns of class I (for review, Sharp & Burge, 1997) and some very recently discovered GT-AG introns (Dietrich et al., 1997) have the (G/A)TATCCTY sequence at their 5' end and the TCCTTRAY sequence at their BS. The former sequence is recognized by U11 snRNA (Kolossova & Padgett, 1997; Yu & Steitz, 1997), the later one by U12 snRNA (Tarn & Steitz, 1996b; Hall & Padgett, 1996). After binding of U1 and U2 snRNAs, assembly of the major spliceosomes is completed by the association of the (U4/U6, U5) tri-snRNP particle and protein factors (for review, Krämer, 1996). Similarly, a (U4atac/U6atac, U5) tri-snRNP is expected to join the minor form of pre-spliceosomes. At this stage of assembly, occurrence of conformational changes was demonstrated for the major spliceosomes (for review, Nilsen, 1998): the U4/U6 interaction is disrupted and replaced by a U2/U6 interaction, U1 snRNA is discarded and U6 snRNA interacts with the 5' extremity of the

intron. The resulting structure, that is formed by the pre-mRNA, U2 and U6 snRNAs, is supposed to play a direct role in catalysis. The 5'-terminal loop of U5 snRNA plays an important role for exon alignment.

Given the recent discovery of the minor form of spliceosomes, only a limited number of experimental evidences for rearrangements prior to catalysis, were obtained (Tarn & Steitz, 1996a; Yu & Steitz, 1997). However, based on potential base-pair interactions between U6atac and U12 snRNAs, on the one hand, and U6atac snRNA and the pre-mRNA, on the other hand, a model of interaction of these three RNAs, mimicking the interactions of U2 and U6 snRNAs with the pre-mRNA, was proposed. This model is in accord with recent cross-linking and genetic data (Tarn & Steitz, 1996a; Yu & Steitz, 1997; Incorvaia & Padgett, 1998; Shukla & Padgett, 1999).

According to the proposed UsnRNA-premRNA interactions, the major form of spliceosomes splices all the major conventional GT-AG introns and also the AT-AC introns of class II, that contain sequences recognized by U1, U2 and U6 snRNAs. The minor form of spliceosomes splices the AT-AC introns of class I, and also a few GT-AG introns, that contain sequences recognized by U11, U12 and U6atac snRNAs (for review, Wu & Krainer, 1999).

In plants and vertebrates, the UsnRNAs involved in the formation of the major spliceosomes carry numerous post-transcriptional modifications, primarily pseudouridines (Ψ) and 2'-O methylated residues. These modifications are concentrated in the segments involved in intermolecular interactions (for review, Massenet et al., 1998). Based on the high phylogenetic conservation of some of these modifications, they are supposed to play an important role in spliceosome assembly and/or function (Szkukalek et al., 1995; Massenet et al., 1999), i.e. stabilization of bimolecular interactions, recognition by protein factors, stabilization of RNA 3D structure, as already proposed for post-transcriptional modifications of other types of RNA (for review, Agris, 1996; Auffinger & Westhof, 1998; Davis, 1998). Experimental proofs of the importance of post-transcriptional modifications in the formation of active spliceosomes were obtained for the vertebrate U2 snRNA (Ségault et al., 1995; Yu et al., 1998).

Our recent mapping of Ψ residues in the *S. cerevisiae* spliceosomal UsnRNAs revealed a low number of Ψ residues in this organism, as compared to vertebrates (Massenet et al., 1999).

However, the detected Ψ residues were located in UsnRNA segments of extremely high functional importance. There are two possible explanations for the observed difference: either multicellular organisms produce a larger number of RNA-modification enzymes, leading to a higher level of post-transcriptional modification of the UsnRNAs, or, from a functional point of view, the *S. cerevisiae* splicing machinery is less dependent upon UsnRNA post-transcriptional modifications, as compared to the plant and the vertebrate major splicing machineries. We wanted to test whether the vertebrate UsnRNAs from the minor spliceosomes were as highly post-transcriptionally modified as vertebrate UsnRNAs from the major spliceosomes. As a first step, we mapped Ψ residues in the U11, U12, U4atac and U6atac snRNAs from HeLa cells. The data are presented in this report. They are discussed in terms of the possible role of Ψ residues in splicing and in connection with spliceosome evolution.

RESULTS AND DISCUSSION

Mapping of Ψ residues in the U11 and U12 snRNAs could be achieved on HeLa cell RNA extracts, using two complementary approaches, based on : i) the alkaline resistant modification of Ψ residues by CMCT (Bakin & Ofengand, 1993), and ii) the absence of modification of Ψ residues by hydrazine (Peattie, 1979). Due to the low abundance of the U4atac and U6atac snRNAs, only the CMCT analysis could be achieved successfully for these two RNAs, and this with the use of specific conditions for CMCT modification (see Materials and Methods). In this case, control experiments with U4atac and U6atac snRNAs produced by *in vitro* transcription were performed. In addition, for U6atac snRNA, a preliminary step of affinity selection, with a complementary biotinylated oligonucleotide retained on streptavidin-agarose beads, was required to get a clear CMCT analysis. Presence of alkaline resistant CMCT modifications and absence of hydrazine modification were detected by primer-extension analysis. Due to the difficulty of detecting CMCT modifications in U4atac snRNA, three different primers had to be used for reverse-transcriptase (RT) analysis. Representative examples of the analysis of atac UsnRNAs by the CMCT method, are illustrated in Figures 1 and 2. To save space, only the control experiment with the U4atac *in vitro* transcript is shown (Fig. 2A). For the same reason, hydrazine analysis is only illustrated for U12 snRNA (Fig. 1C).

No Ψ residue was detected in U11 snRNA (Fig. 1A). All the RT stops, detected after a one min CMCT incubation without further alkaline treatment, disappeared when the alkaline treatment was achieved. Indeed, no band was detected in high yield after a 10 or 20 min CMCT incubation followed by an alkaline treatment, except at natural RT stops (Lanes 3 and 4, Fig. 1A), and, this, even after a long exposure of the X-ray film (Lanes 2', 3' and 4'). Taking into account the RNA sequence, the natural pauses of RT, corresponding to positions 65 and 68 in the sequence ladder, cannot mask RT stops due to CMCT modification of Ψ residues. Only the natural pause, corresponding to position 18 in the sequence ladder, took place at a position where a RT stop would be expected if U₁₇ was converted into a Ψ residue. However, the intensity of the band

observed at this position clearly decreased after the alkaline treatment, instead of being increased as expected for a Ψ residue.

The absence of Ψ residue in U11 snRNA is in contrast to the high phylogenetic conservation of two Ψ residues in the 5'-terminal segment of U1 snRNA that interacts with the intron (Fig. 3) (for review, Massenet et al., 1998). In connection with this observation, it should be noticed that the analogy between U1 and U11 is not complete. Whereas, the U1 snRNA-intron interaction involves the whole sequence that is conserved at the 5' extremity of U2-dependent introns (for review, Nilsen, 1998), the interaction between the 5' extremity of U12-dependent introns and U11 snRNA only involves residues 4 to 8 of the intron (Kolossova & Padgett, 1997; Yu & Steitz, 1997). In addition, the U11 segment implicated in the interaction contains no U residue.

By inspection of the CMCT and hydrazine analyses of U12 snRNA and based on the same reasoning as for the analysis of U11 snRNA, we concluded that U12 snRNA contains 2 Ψ residues (Fig. 1B and 1C). Here again, this is in contrast with the 12 and 13 Ψ residues found in the rat hepatoma and the *Vicia faba* plant U2 snRNAs, respectively (Reddy et al., 1981, for review, Massenet et al., 1998). Interestingly, the Ψ residue found at position 19 is involved in one of the two base pairs that bulge out the A residue responsible for the nucleophilic attack in the first step of the splicing reaction (Fig. 3C). As far as post-transcriptional modifications of U2 snRNAs have been studied, the presence of a Ψ residue at the equivalent position of U2 snRNA seems to be conserved (position 34 in the human and plant U2 snRNAs and position 35 in the *S. cerevisiae* U2 snRNA, respectively). The presence of a Ψ residue at equivalent positions in the interactions formed between U12 snRNA and the BS sequence of U12-type introns on the one hand, and U2 snRNA and the BS sequence of U2-type introns on the other hand, is a strong hint for a functional importance of this post-transcriptional modification. Mutations at position 35 in the *S. cerevisiae* U2 snRNA resulted in a strong decrease of the *in vitro* splicing efficiency (McPheeters & Abelson, 1992) and is lethal *in vivo* (Pascolo & Séraphin, 1997). Effect of the generation of some compensatory mutations in the BS sequence and its U2 snRNA-recognition element was tested in *S. cerevisiae* (Pascolo & Séraphin, 1997). However further experiments are needed to get a clear answer on the functional importance of a Ψ residue at position 35. Presence

of a Ψ residue may stabilize the interaction between U2 or U12 snRNA with the BS sequence or may be involved in recognition of the heterologous helix by protein factors.

The second Ψ residue, that we detected in U12 snRNA (Ψ_{28} , Fig. 1) is located 9 nt upstream of the one found in the BS recognition sequence. A Ψ residue is also found at the equivalent position of all the studied U2 snRNAs (Ψ_{43} in the vertebrate RNA and Ψ_{44} in the *S. cerevisiae* RNA, for review, Massenet et al., 1998). Residue Ψ_{28} in U12 snRNA and the corresponding Ψ residues of the vertebrate and yeast U2 snRNAs belong to segments of the U12 and U2 snRNAs that can form an intermolecular helix III with the U6atac and U6 snRNAs, respectively (Fig. 3). A functional importance of helix III was demonstrated for the U2-type spliceosomes of HeLa cells (Sun & Manley, 1995), but not for those of *S. cerevisiae* (Yan & Ares, 1996). Up to now, helix III was not detected by crosslinking experiments in the U12-type spliceosomes (Tarn & Steitz, 1996a). In addition, recent data suggest that it cannot be formed in the plant atac UsnRNAs (Shukla & Padgett, 1999). Independent of helix III formation, the segments of the U2 and U12 snRNAs, containing the second highly conserved Ψ residue, have an important function for protein association. Indeed, this U2 snRNA segment binds proteins Prp9, Prp11, Prp21 and Cus1 in yeast and their counterparts, which form the SF3a and SF3b factors, in HeLa cells (for review, Krämer, 1996). Although, the absence of U to Ψ conversion at position 44 of the *S. cerevisiae* U2 snRNA was found to have no effect on growth (Massenet et al., 1999), the presence of a Ψ residue at this position in U2 and U12 snRNAs may facilitate protein recognition. The nucleotide sequence of the *Arabidopsis thaliana* U12 snRNA has been deduced from a cDNA sequence (Shukla & Padgett, 1999), no identification of post-transcriptional modification was achieved. Interestingly, a T residue is present at position 19, which opens the possibility of its conversion into a Ψ residue. However, a C residue is found at position 28, so that a Ψ residue is not expected at this position of the *A. thaliana* U12 snRNA.

Our Ψ mapping of the U4atac snRNA revealed the presence of a Ψ residue at position 12 (Fig. 2A). The RT stop at this position is followed by a stop of lower amplitude at the following U residue. Either this corresponds to a partial conversion of U into Ψ at position 11, or this is an

analysis artefact due to stuttering of RT, as already reported for Ψ residues located within a series of U residues (Bakin & Ofengand, 1993). The two stops were absent when the experiment was performed on an *in vitro* produced U4atac snRNA (Fig. 2A, Lanes 1 to 4 marked Transcript). As previously proposed for the U4 and U6 snRNAs, an heteroduplex can be formed between the U4atac and the U6atac snRNAs and this is supported by results of cross-linking and genetic experiments (Tarn & Steitz, 1996a; Shukla & Padgett, 1999). According to the proposed models (Fig. 4), stem II in the U4atac/U6atac duplex is extended compared to the U4/U6 stem II. Based on nucleotide sequence homology, the part of stem II formed by the U4atac sequence from positions 12 to 23 and the U6atac sequence from positions 28 to 39, is the counterpart of the U4/U6 stem II. Hence, residue Ψ_{12} , that we detected in the U4atac snRNA, is likely to be the counterpart of residue Ψ_4 in the vertebrate U4 snRNAs (Fig. 4). It should be noticed that formation of a Ψ residue was not observed at the equivalent position of the *S. pombe* and *S. cerevisiae* U4 snRNAs (Gu et al., 1996, Massenet et al., 1999). The two other Ψ residues, present in the vertebrate U4 snRNA segment extending from positions 70 to 80, have no counterpart in the U4atac snRNA.

In a previous review on UsnRNA post-transcriptional modifications, that we published before completion of this work, we mentioned that we had detected no Ψ residue in the U6atac snRNA (Massenet et al., 1998). However, by using the partial purification of U6atac snRNA described in this paper and an oligonucleotide primer that allowed analysis of the 3'-terminal region, one Ψ residue was detected at position 83 (Fig. 2B). As compared to U6 snRNA, U6atac snRNA contains an additional sequence in its 3'-terminal region. This additional sequence can be folded into a stem-loop structure (Tarn & Steitz, 1996a) and the Ψ residue, that we detected, is formed in one of the helices (Fig. 2B and 4). The Ψ residue, present in the 3'-terminal region of the vertebrate U6 snRNA, is located in a different context as referred to sequence and secondary structure (Fig. 4). Interestingly, the *Influenza* virus NS1 protein was recently shown to inhibit formation of the U12/U6atac complex (Wang & Krug, 1998). Protein NS1 binds to the additional stem-loop structure of the U6atac snRNA and Wang and Krug proposed that this additional stem-loop structure participates in additional interactions with U12 snRNA. The two other Ψ residues

found in vertebrate U6 snRNA, which have no counterpart in yeast U6 snRNA, are not found in the U6atac snRNA, either (Fig. 4).

Hence, our study reveals a low level of pseudouridylation of the atac UsnRNAs. Altogether, U11, U12, U4atac and U6atac snRNAs contain 4 Ψ residues, while their counterparts of the major spliceosomes contain 20 Ψ residues altogether. Three of the four detected Ψ residues correspond to post-transcriptional modifications common to the two splicing machineries, this in spite of a high level of sequence divergence of the two UsnRNA series.

Interestingly, the patterns of pseudouridylation of the *S. cerevisiae* UsnRNAs and vertebrate atac UsnRNAs show similarities. Tarn and Steitz (1996a) already pointed out, that the *S. cerevisiae* spliceosomes, which are memberships of the U2-type spliceosomes, however share several common features with the vertebrate U12-type spliceosomes. The list of common properties is increasing with the present data and other recent data: i) the U4atac snRNA and the *S. cerevisiae* U4 snRNA lack a 3' terminal stem-loop structure, ii) the secondary structure models, proposed for the isolated *S. cerevisiae* U6 snRNA and for the U6atac snRNA, show similarity and differ from that proposed for the vertebrate U6 snRNA (Tarn & Steitz, 1996a), iii) helix II formed between U2 and U6 snRNAs is essential for splicing in mammalian cells (Datta & Weiner, 1991; Wu & Manley, 1991), whereas it is only required when helix Ib cannot be formed in *S. cerevisiae* (Field & Friesen, 1996), it cannot be formed in the atac spliceosomes, as the 5' end of U12 is truncated compared to U2 snRNA (Fig. 3), iv) helix III, formed between U2 and U6 snRNAs (Fig. 3), is needed for active spliceosome formation in HeLa cells (Sun & Manley, 1995), but not in *S. cerevisiae* (Yan & Ares, 1996), this helix was not detected by crosslinking experiments in the vertebrate atac spliceosomes (Tarn & Steitz, 1996a) and it cannot be formed with the plant atac UsnRNAs (Shukla & Padgett, 1999), v) both *S. cerevisiae* and atac UsnRNAs contain a limited number of Ψ residues, as compared to vertebrate UsnRNAs from the main form of spliceosomes, vi) the 5' and BS sequences of the *S. cerevisiae* introns and of the atac introns show a very low level of deviation, when compared to the consensus sequences, whereas a high degree of divergence is observed for the vertebrate U2-type introns. Points v) and vi) may be related one to the other: a high degree of post-transcriptional modification of the UsnRNAs may

be required in vertebrate and plant UsnRNAs of the major spliceosomes, in order to stabilize the interactions between the pre-mRNA and the UsnRNAs, which are of low energy in the U2-type spliceosomes. Such common properties may reflect characteristic features of a common ancestor for the U2-dependent and the U12-dependent spliceosomes. According to this hypothesis, the U2-dependent spliceosomes of vertebrates would have diverge more rapidly from the common ancestor, than the vertebrate U12-dependent spliceosomes and the *S. cerevisiae* U2-dependent spliceosomes. Based on the observation of U12-dependent introns in the plant *A. thaliana*, it appears that the two spliceosomal machineries have existed side by side for more than billion years (for review, Sharp & Burge, 1997). This means, that over all this time, the two systems should have preserved enough distinctive features to ensure fidelity of intron excision. Differences in the pattern of UsnRNA post-transcriptional modifications may participate to ensure specificity to each splicing machinery.

How can there be numerous conversions of U residues into Ψ residues in the UsnRNAs from the major form of spliceosomes and a low number of such conversions in the atac UsnRNAs, whereas both RNA series are synthesized in the same cell? One main explanation is the replacement of the U residues, that are converted into Ψ residues in the major UsnRNAs, with other residues in the atac UsnRNAs. Another important question is whether a given Ψ synthase can act both on one of the main UsnRNAs and on its counterpart of the atac spliceosomes, for instance at positions 19 of U12 snRNA and 34 of U2 snRNA. Up to now, little is known about UsnRNA- Ψ synthases. Only one of them was identified in yeast, the Pus1p tRNA- Ψ synthase, that catalyzes Ψ_{44} formation in the *S. cerevisiae* U2 snRNA (Massenet et al., 1999). Based on *in vitro* assays, performed in the presence of 5-fluorouridine with HeLa cell nuclear extract or S100 extract, Patton concluded that distinct Ψ synthases are needed for formation of the various Ψ residues of the main UsnRNAs (for review, Patton, 1994). Furthermore, formation of Ψ residues in U4 and U6 snRNA transcripts was found to be increased upon formation of the U4/U6 snRNA duplex (Zerby & Patton, 1996; Zerby & Patton, 1997). The involvement of distinct Ψ synthases for Ψ formation in UsnRNAs is also supported by the absence of consensus sequence at the site of pseudouridylation. The UsnRNA- Ψ synthase specificity may depend upon RNA secondary-

structure recognition. Hence, as in spite of nucleotide sequence dissimilarity, U2 and U12 snRNAs can be folded into very similar structures (Wassarman & Steitz, 1992), the enzyme that converts U₃₄ into Ψ in U2 snRNA may be the same that converts U₁₉ into Ψ in U12 snRNA. The same is true for Ψ₄₃ formation in U2 snRNA and Ψ₂₈ formation in U12 snRNA. As mentioned above, Ψ₁₂ in the U4atac snRNA has a position in stem II of the U4atac/U6atac duplex very similar to that of Ψ₄ of U4 snRNA in the U4/U6 duplex. Here also, a common Ψ synthase recognizing the heterologous stem II, may modify both U4 and U4atac snRNAs. It is difficult to extend this reasoning to the formation of Ψ₈₃ in the U6atac snRNA, as the Ψ residue found in the counterpart region of U6 snRNA is located in a completely different context of sequence and secondary structure. Interestingly, we noticed that Ψ₈₃ in U6atac snRNA and Ψ₁₉ in U12 snRNA are formed in sequences showing a strong degree of similarity (UApyGPuNΨAAGGNA). This opens the question of the existence of a Ψ synthase common to U12 and U6atac snRNAs. Can it be a RNA guided Ψ synthase? Answers to these various questions on U snRNA-Ψ synthases require further progress in the identification of these enzymes in vertebrates.

MATERIALS AND METHODS

RNA Preparation

HeLa nuclear extracts from the Computer Cell Culture Center S.A. (Belgium) were used as the source of atac UsnRNAs. To this end, the nuclear extract was successively treated by equal volumes of phenol, phenol/chloroform (1 :1) and chloroform/isoamyl alcohol (24 :1). The extracted total nuclear RNA was ethanol precipitated and used for CMCT or hydrazine modifications or for purification of the U6atac snRNA.

U6atac snRNA purification

U6atac snRNA was purified from a total nuclear RNA extract, by affinity selection on streptavidin-agarose beads, using the following oligonucleotide: 5'-(dU*)₅GAAGTAGGTGGCAATGCCTTAACCGT-3' (* denotes a biotinylated 2'-deoxyuridine), that is complementary to nucleotide positions 78 to 103 of U6atac snRNA. The affinity selection was performed as described in Ségault et al. (1995), using 50 µg of total nuclear RNA and 1 µg of oligonucleotide. Hybridization of the oligonucleotide was performed at 65°C for 10 min. After affinity selection, the streptavidin-agarose beads were treated with 160 µg/ml of proteinase K (Boehringer) in 50 µl of buffer containing 5 mM EDTA, 300 mM NaCl, 1.5% SDS, 50 mM Tris-HCl, pH 7.5. Then, RNA was phenol extracted, ethanol precipitated and quantified with a Gene Quant apparatus

Identification of Ψ residues in the atac UsnRNAs

The CMCT-modification protocol was adapted from that of Bakin & Ofengand (1993), as previously described (Massenet et al. 1999). CMCT modifications were performed on 10 µg of total nuclear RNA for 1, 10 and 20 min for U11 and U12 snRNA analyses, on 50 µg of total nuclear RNA for 2, 10, 20 and 30 min for U4atac snRNA analysis, on about 100 ng of affinity purified U6atac snRNA for 1, 10 and 20 min for U6atac snRNA analysis and on 1 µg of *in vitro* produced U4atac and U6atac UsnRNAs for 1 or 2, 10 and 20 min. The alkaline treatment was for

3 hours, as described previously (Massenet et al., 1999). Hydrazine modifications were performed as previously described in Massenet et al. (1999), on 10 µg of total nuclear RNA. Positions of CMCT and hydrazine modifications were identified by primer-extension analyses, using the AMV RT (Life Science, USA), in the conditions described in Mougin et al. (1996). The oligonucleotides, complementary to the following regions of the atac UsnRNAs, were used as primers: U4atac : nts 41 to 61, nts 70 to 82 and nts 108 to 131, U6atac : nts 106 to 125, U11 : nts 95 to 110 and U12 : nts 60 to 79 and nts 121 to 146. Primers were 5'-end labeled with [γ -³²P] ATP (3000Ci/mmol) and T4 polynucleotide kinase. RNA sequencing ladders were done on 40 µg of total nuclear RNA for U11 and U12 snRNAs and on *in vitro* produced U4atac and U6atac snRNAs. U4atac and U6atac coding regions, under the control of a T7 promoter, were provided by A. Mougin (unpublished results).

ACKNOWLEDGMENTS

Pr J. Steitz is acknowledged for information on atac UsnRNA sequences given before publication. A. Mougin is thanked for providing us with atac UsnRNA coding sequences under the control of T7 promoters. This work was supported by laboratory funds from the « Ministère de la Recherche et de l'Enseignement supérieur » and the French « Centre National de la Recherche Scientifique ». S. Massenet is a pre-doctoral fellow supported by a fellowship from the « Ministère de la Recherche et de l'Enseignement supérieur ».

LEGENDS TO FIGURES

Figure 1

A/ Absence of Ψ residue in the HeLa cell's U11 snRNA. Ten μg of total nuclear RNA were modified by CMCT for 1, 10 and 20 min (Lanes 2, 3 and 4, respectively), with (Lanes 3 and 4) or without (Lane 2) alkaline treatment at pH 10.4, conditions are described in Materials and Methods. A control primer extension analysis was made without CMCT treatment (Lane 1). Primer-extension analysis was made with an oligonucleotide complementary to positions 95 to 110 of U11 snRNA. No RT stop was detected after the alkaline treatment (Lanes 3 and 4), even after a longer exposure of the X-ray film (Lanes 2', 3' and 4' in the insert on the right of Panel A). Lanes U, G, C, A correspond to the RNA sequencing ladder. Nucleotide positions, starting from the 5'-terminal nucleotide bound to the cap structure, are indicated on the left of the Panel.

B/ Localization of Ψ residues in HeLa cell's U12 snRNA. As in Panel A, CMCT modifications were for 1, 10 and 20 min (Lanes 2, 3 and 4, respectively), with (Lanes 3 and 4) or without (Lane 2) alkaline treatment. The oligonucleotide primer used for RT analyses was complementary to U12 snRNA from positions 60 to 79. The two RT stops, corresponding to residues Ψ_{19} and Ψ_{28} , are indicated.

C/ Verification of the presence of Ψ residues at positions 19 and 28, using hydrazine treatment. Ten μg of total RNA extracted from a HeLa cell nuclear extract were treated with hydrazine, in the conditions described in Materials and Methods (Lane 5). A control-extension experiment was made in the absence of hydrazine (Lane 1). The absence of hydrazine reactivity at positions 19 and 28 confirmed the presence of Ψ residues at these two positions.

Figure 2

A/ Localization of Ψ residues in the HeLa cell's U4atac snRNA. Fifty μg of total nuclear RNA (Lanes marked U4atac snRNA) or 2 μg of *in vitro* produced U4atac snRNA (Lanes marked Transcript) were modified by CMCT for 2, 10, 20 and 30 min (Lanes 2, 3, 4 and 5, respectively), with (Lanes 3, 4 and 5) or without (Lane 2) alkaline treatment. Primer-extension analyses were

made with an oligonucleotide complementary to positions 70 to 82 of U4atac snRNA. The identified Ψ_{12} residue is indicated. Note the increased alkaline resistance of the CMCT modification at position 11, that can reflect either a partial conversion of U into Ψ or a stuttering phenomenon (Bakin and Ofengand, 1993).

B/ Localization of Ψ residues in the HeLa cell's U6atac snRNA. One hundred ng of purified U6atac UsnRNA (see Materials and Methods) were modified by CMCT for 1, 10 and 20 min (Lanes 2, 3 and 4, respectively), with (Lanes 3 and 4) or without (Lane 2) alkaline treatment. Primer-extension analyses were made with an oligonucleotide complementary to positions 106 to 125 of the U6atac snRNA. The identified Ψ_{83} residue is indicated.

Figure 3

Positions of Ψ residues within the network of RNA-RNA interactions of the *S. cerevisiae* spliceosomes (Panel A), the vertebrate major form of spliceosomes (Panel B) and the vertebrate minor form of spliceosomes (Panel C). The interaction of the 5'- splice site with U1 (Panels A and B) or U11 snRNAs (Panel C) is shown (Scheme I of each Panel). UsnRNA-UsnRNA and UsnRNA-premRNA interactions at the catalytic center of the spliceosomes are shown in schemes II of the three Panels. Helices I, II and III formed between the U2 and U6 snRNAs or between the U12 and U6atac snRNAs are represented, as well as, the base-pair interactions between U2 or U12 snRNAs and the BS sequences. The interaction between U6atac snRNA and a sequence close to the 5'-splice site is indicated by overlined residues joined by an arrow. The interaction between the terminal loop I of U5 snRNA and the exon extremities is also shown for the minor form of spliceosomes (Panel C) (for a review on these interactions in the major form of spliceosomes, see Nilsen, 1998, and in the minor form of spliceosomes, see references in the text). The Ψ residues are boxed, base and ribose methylations are indicated in bold (this study, Massenet et al., 1999 and for review, Reddy, 1988 and Massenet et al., 1998). Nucleotide positions, starting from the 5' extremity of each snRNA, are indicated.

Figure 4

Positions of Ψ residues in the U4-U6 (Panel A) and U4atac-U6atac (Panel B) snRNA duplces. For the U4/U6 RNA duplex, the nucleotide sequences and post-transcriptional modifications shown are from rat (Harada et al., 1980; Krol et al., 1981; for review, Reddy, 1988), the heterologous stems I and II are as proposed by Brown and Guthrie (1988). Pseudouridine residues are squared in black. For the U4atac/U6atac RNA duplex, nucleotide sequences are from human as determined by Tarn & Steitz (1996a), Ψ residues (this study) are squared in black. The U4atac/U6atac heterologous stems I and II were proposed by Tarn & Steitz (1996a).

REFERENCES

- Agris PF. 1996. The importance of being modified: roles of modified nucleosides and Mg^{2+} in RNA structure and function. *Prog. Nucleic Acid. Res. Mol. Biol.* 53:79-129.
- Auffinger P, Westhof E. 1998. Effects on pseudouridylation on tRNA hydration and dynamics: a theoretical approach. In: Grosjean H, Benne R, eds. *The modification and editing of RNA*. Washington, DC: ASM Press. pp 103-112.
- Bakin A, Ofengand J. 1993. Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique. *Biochemistry* 32:9754-9762.
- Brow DA, Guthrie C. 1988. Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature* 334:213-218.
- Datta B, Weiner AM. 1991. Genetic evidence for base pairing between U2 and U6 snRNAs in mammalian mRNA splicing. *Nature* 352:821-824.
- Davis DR. 1998. Biophysical and conformational properties of modified nucleosides in RNA (nuclear magnetic resonance studies). In: Grosjean H, Benne R, eds. *The modification and editing of RNA*. Washington, DC: ASM Press. pp 85-102.
- Dietrich RC, Incorvaia R, Padgett RA. 1997. Terminal intron dinucleotide sequences do not distinguish between U2- and U12-dependent introns. *Mol. Cell.* 1:151-160.
- Field DJ, Friesen JD. 1996. Functionally redundant interactions between U2 and U6 spliceosomal snRNAs. *Genes Dev.* 10:489-501.
- Gu J, Patton JR, Shimba S, Reddy R. 1996. Localization of modified nucleotides in *Schizosaccharomyces pombe* spliceosomal small nuclear RNAs: modified nucleotides are clustered in functionally important regions. *RNA* 2:909-918.
- Hall SL, Padgett RA. 1996. Requirement of U12 snRNA for *in vivo* splicing of a minor class of eukaryotic nuclear pre-mRNA introns. *Science* 271:1716-1718.
- Harada F, Kato N, Nishimura S. 1980. The nucleotide sequence of nuclear 4.8S RNA of mouse cells. *Biochem. Biophys. Res. Commun.* 95:1332-1340.

- Incorvaia R, Padgett RA. 1998. Base pairing with U6atac snRNA is required for 5' splice site activation of U12-dependent introns *in vivo*. *RNA* 4:709-718.
- Jackson IJ. 1991. A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Res.* 19:3795-3798.
- Kolossova I, Padgett RA. 1997. U11 snRNA interacts *in vivo* with the 5' splice site of U12-dependent (AU-AC) pre-mRNA introns. *RNA* 3:227-233.
- Krämer A. 1996. The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu. Rev. Biochem.* 65:367-409.
- Krol A, Branlant C, Lazar E, Gallinaro H, Jacob M. 1981. Primary and secondary structures of chicken, rat and man nuclear U4 RNAs. Homologies with U1 and U5 RNAs. *Nucleic Acids Res.* 9:2699-2716.
- 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.* 19:2142-2154.
- Massenet S, Mougín A, Branlant C. 1998. Posttranscriptional modifications in the U Small Nuclear RNAs. In: Grosjean H, Benne R, eds. *The modification and editing of RNA*. Washington, DC: ASM Press. pp 201-228.
- McPheeters DS, Abelson J. 1992. Mutational analysis of the yeast U2 snRNA suggests a structural similarity to the catalytic core of group I introns. *Cell* 71:819-831.
- Moore JM, Query CC, Sharp PA. 1993. Splicing of precursors to messenger RNAs by the spliceosome. In: Gesteland R, Athuis J, eds. *The RNA world*. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press. pp 303-358.
- Mougín A, Grégoire A, Banroques J, Ségault V, Fournier R, Brulé F, Chevrier-Miller M, Branlant C. 1996. Secondary structure of the yeast *Saccharomyces cerevisiae* pre-U3A snoRNA and its implication for splicing efficiency. *RNA* 2:1079-1093.

- Nilsen TW. 1998. RNA-RNA interactions in nuclear pre-mRNA splicing. In: Simos RW, Grunberg-Manago M, eds. *RNA structure and fonction*. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press. pp 279-308.
- Pascolo E, Séraphin B. 1997. The branchpoint residue is recognized during commitment complex formation before being bulged out of the U2 snRNA-pre-mRNA duplex. *Mol. Cell. Biol.* 17:3469-3476.
- Patton JR. 1994. Pseudouridine formation in small nuclear RNAs. *Biochimie* 76:1129-1132.
- Peattie DA. 1979. Direct chemical method for sequencing RNA. *Proc. Natl. Acad. Sci. USA* 76:1760-1764.
- Reddy R. 1988. Compilation of small RNA sequences. *Nucleic Acids Res.* 16:71-85.
- Reddy R, Henning D, Epstein P, Busch H. 1981. Primary and secondary structure of U2 snRNA. *Nucleic Acids Res.* 9:5645-5658.
- Ségault V, Will CL, Sproat BS, Lührmann R. 1995. *In vitro* reconstitution of mammalian U2 and U5 snRNPs active in splicing: Sm proteins are functionally interchangeable and are essential for the formation of functional U2 and U5 snRNPs. *EMBO J.* 14:4010-4021.
- Sharp PA, Burge CB. 1997. Classification of introns: U2-type or U12-type. *Cell* 91:875-879.
- Shukla GC, Padgett RA. 1999. Conservation of functional features of U6atac and U12 snRNAs between vertebrates and higher plants. *RNA* 5:525-538.
- Sun JS, Manley JL. 1995. A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes Dev.* 9:843-854.
- Szkukalek A, Myslinski E, Mougin A, Lührmann R, Branlant C. 1995. Phylogenetic conservation of modified nucleotides in the terminal loop 1 of the spliceosomal U5 snRNA. *Biochimie* 77:16-21.
- Tarn WY, Steitz JA. 1996a. Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science* 273:1824-1832.
- Tarn WY, Steitz JA. 1996b. A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron *in vitro*. *Cell* 84:801-811.

- Wang W, Krug RM. 1998. U6atac snRNA, the highly divergent counterpart of U6 snRNA, is the specific target that mediates inhibition of AT-AC splicing by the *influenza* virus NS1 protein. *RNA* 4:55-64.
- Wassarman KM, Steitz JA. 1992. The low-abundance U11 and U12 small nuclear ribonucleoproteins (snRNPs) interact to form a two-snRNP complex. *Mol. Cell. Biol.* 12:1276-1285.
- Wu JA, Manley JL. 1991. Base pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. *Nature* 352:818-821.
- Wu Q, Krainer AR. 1999. AT-AC pre-mRNA splicing mechanisms and conservation of minor introns in voltage-gated ion channel genes. *Mol. Cell. Biol.* 19:3225-3236.
- Yan D, Ares M, Jr. 1996. Invariant U2 RNA sequences bordering the branchpoint recognition region are essential for interaction with yeast SF3a and SF3b subunits. *Mol. Cell. Biol.* 16:818-828.
- Yu YT, Shu MD, Steitz JA. 1998. Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J.* 17:5783-5795.
- Yu YT, Steitz JA. 1997. Site-specific crosslinking of mammalian U11 and U6atac to the 5' splice site of an AT-AC intron. *Proc. Natl. Acad. Sci. USA* 94:6030-6035.
- Zerby DB, Patton JR. 1996. Metabolism of pre-messenger RNA splicing cofactors: modification of U6 RNA is dependent on its interaction with U4 RNA. *Nucleic Acids Res.* 24:3583-3589.
- Zerby DB, Patton JR. 1997. Modification of human U4 RNA requires U6 RNA and multiple pseudouridine synthases. *Nucleic Acids Res.* 25:4808-4815.

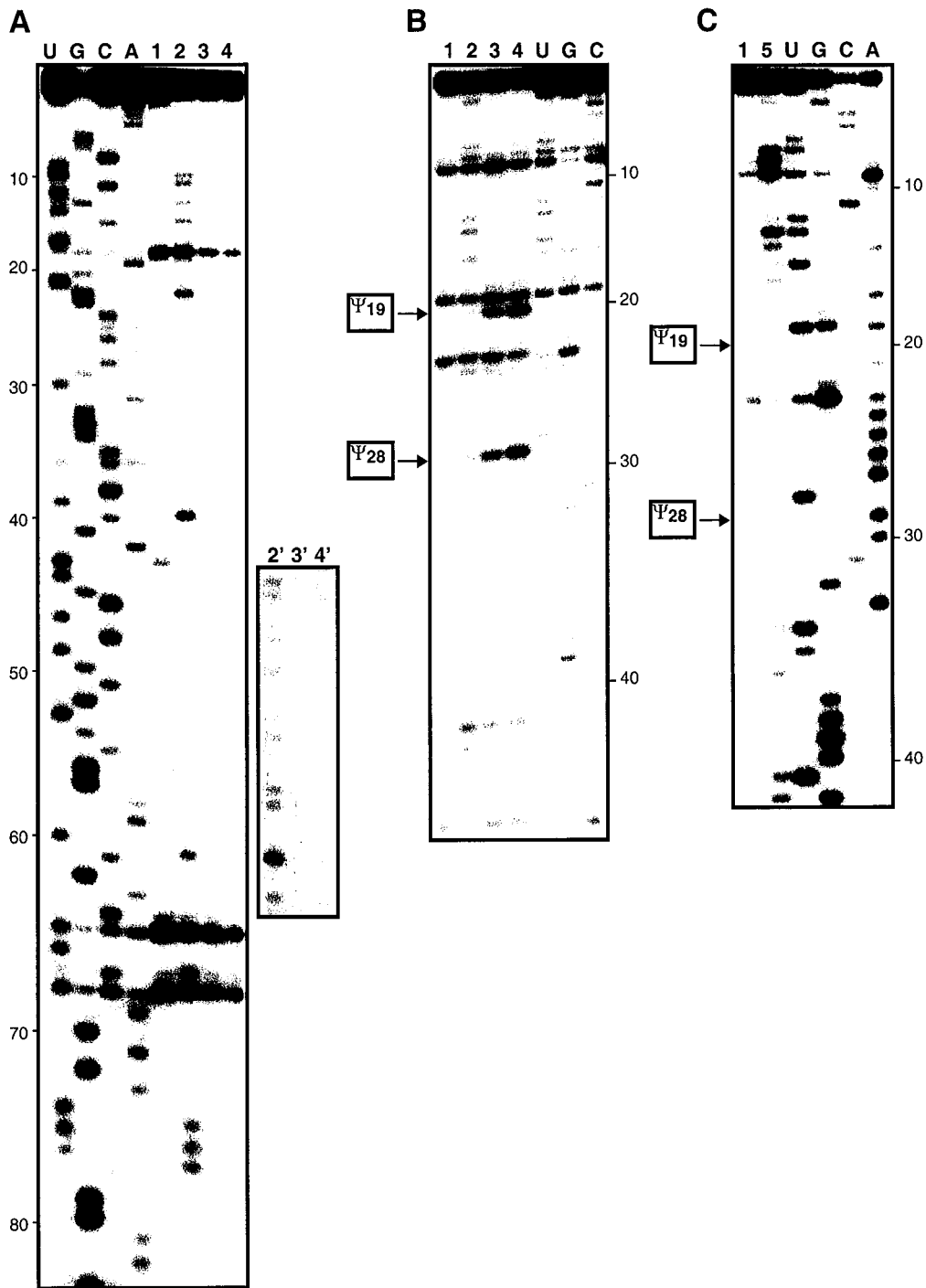


Fig. 1
S. Massenet and C. Branlant

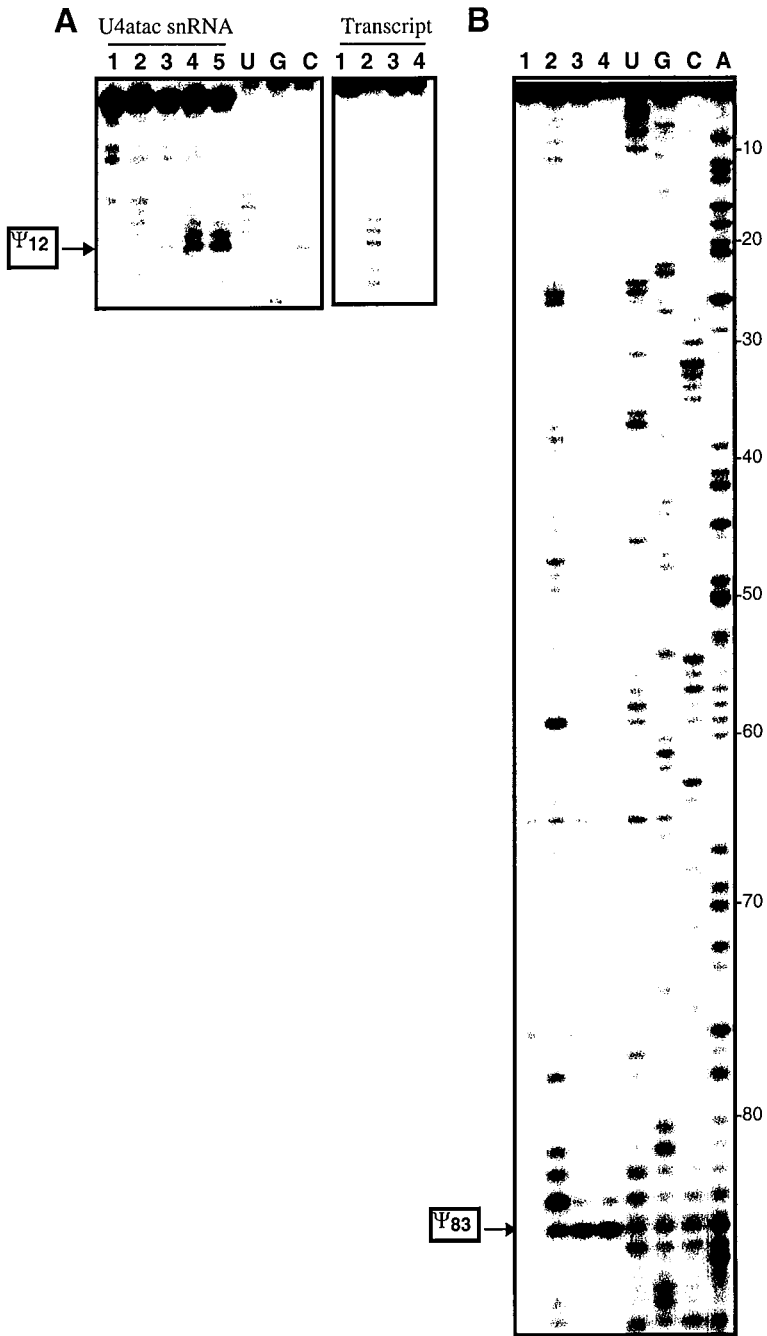


Fig. 2
S. Massenet and C. Branlant

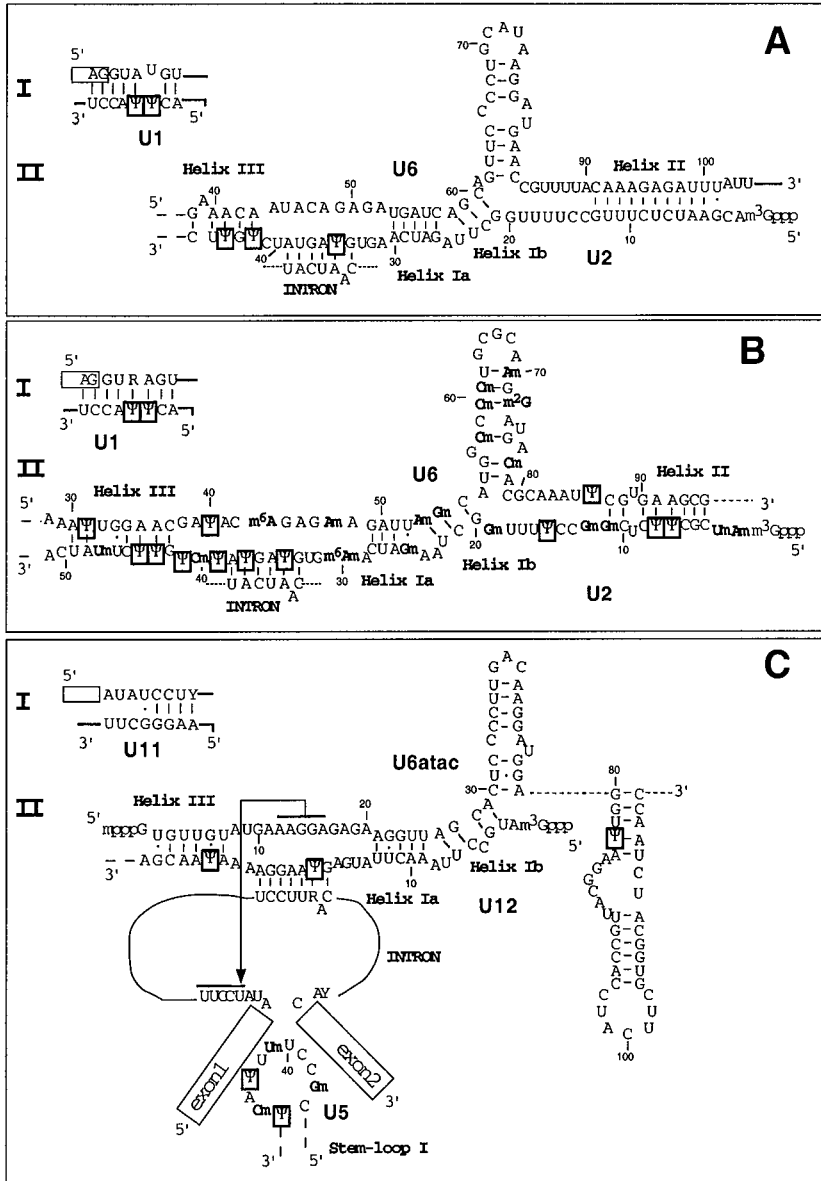


Fig. 3
 S. Massenet and C. Branlant

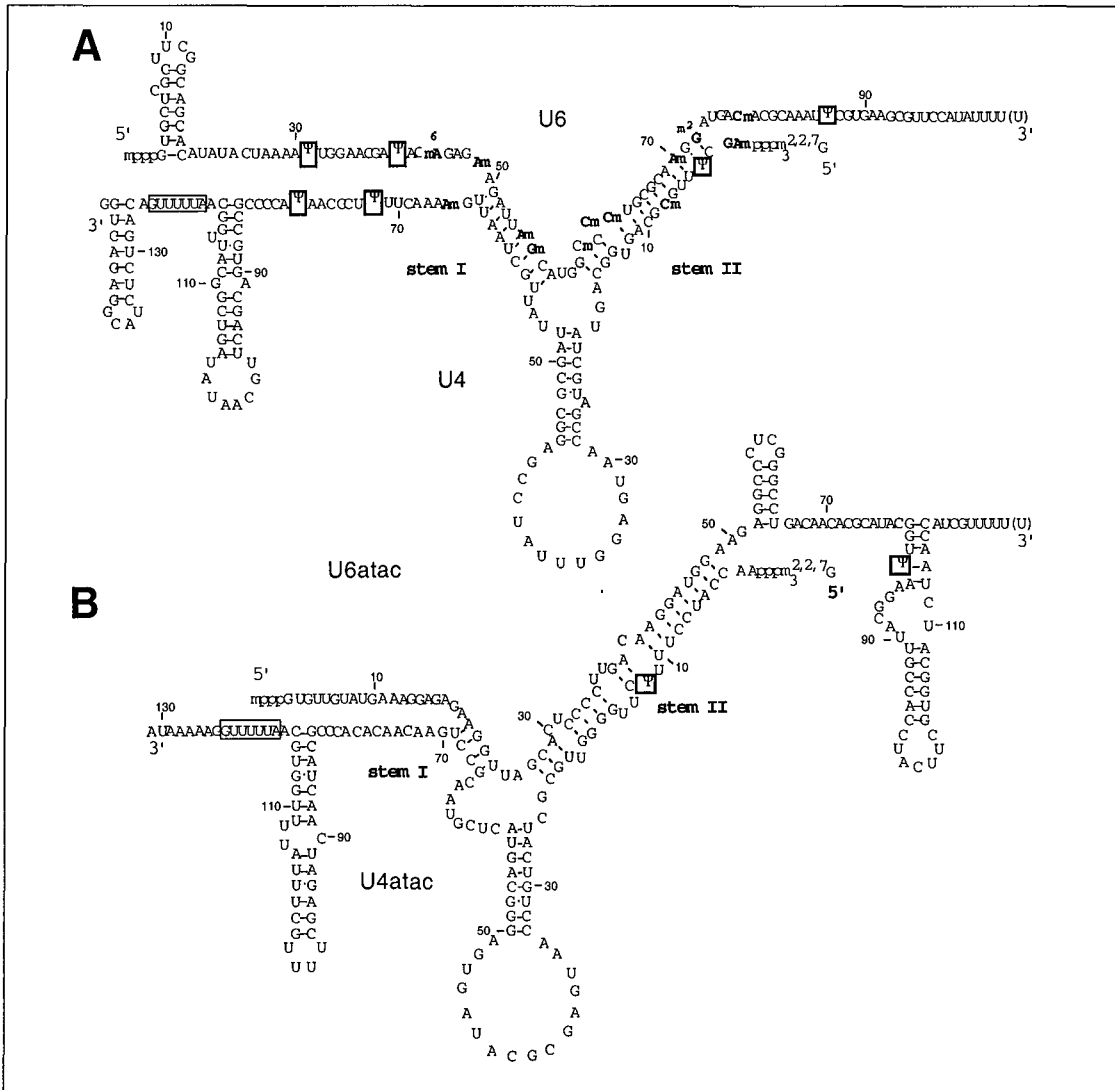


Fig. 4
S. Massenet and C. Branlant

B/ DISCUSSION ET PERSPECTIVES

1. CERTAINS RESIDUS Ψ SONT TRES CONSERVES DANS LES UsnRNA DES SPLICEOSOMES U2-DEPENDANTS ET SONT PRESENTS DANS LES UsnRNA DES SPLICEOSOMES U12-DEPENDANTS.

Un des objectifs de notre étude des résidus Ψ portés par les UsnRNA des spliceosomes U12-dépendants de cellules HeLa était de savoir si, du fait d'une plus grande multiplicité d'enzymes catalysant la modification des ARN chez les vertébrés, tout ARN métaboliquement stable portait, dans ces organismes, un nombre élevé de modifications post-transcriptionnelles. Clairement, nos résultats montrent que la réponse est négative. En effet, alors que les UsnRNA des spliceosomes U12-dépendants et U2-dépendants sont produits dans la même cellule, le degré de pseudouridylation des UsnRNA des spliceosomes U12-dépendants est beaucoup plus faible que celui des UsnRNA des spliceosomes U2-dépendants. Cette observation est un argument en faveur d'une importance fonctionnelle des nombreuses modifications post-transcriptionnelles portées par les UsnRNA des spliceosomes U2-dépendants.

Comme nous l'avons déjà discuté à la fois dans les publications 2 et 3 et dans la revue, la nécessité de nombreux résidus Ψ dans ces UsnRNA pourrait s'expliquer par le fort degré de dégénérescence des séquences fonctionnelles des introns U2-dépendants. En effet, aussi bien les spliceosomes de *S. cerevisiae* que les spliceosomes U12-dépendants, dont les UsnRNA portent peu de résidus Ψ , reconnaissent un nombre limité d'introns dont les régions fonctionnelles ont une séquence extrêmement conservée.

Ce fait pourrait s'expliquer sur une base évolutive. Si l'on admet que les eucaryotes primitifs contenaient des introns dans leurs gènes nucléaires, on peut imaginer que, du fait de la nécessité d'une multiplication très rapide des cellules, lorsque les conditions environnementales sont favorables, les levures aient réduit la taille de leur génome en perdant un grand nombre d'introns de gènes nucléaires. Une autre alternative est que les eucaryotes primitifs ne possédaient, au contraire, que peu d'introns, et que ces derniers se seraient ensuite multipliés au cours de l'évolution. On pourrait alors imaginer que, toujours par nécessité de minimisation de l'ADN génomique, la sélection naturelle ait limité l'expansion des introns dans les gènes nucléaires de levure. Quelle que soit l'hypothèse, ce faible nombre d'introns dans les gènes nucléaires de levure n'a manifestement pas conduit à un fort brassage entre ces introns, par recombinaison homologue entre des séquences introniques répétées, ce qui a dû être le cas pour les introns U2-dépendants de métazoaires. En d'autres termes, il est possible que les introns des gènes nucléaires de levure, ainsi que la machinerie d'épissage qui les

élimine, aient conservé certains caractères ancestraux, perdus dans le système spliceosomal U2-dépendant de métazoaires. En effet, les introns de *S. cerevisiae* présentent plus de points communs avec les introns autoépissables de groupe II, dont ils sont supposés dériver, que les introns U2-dépendants de métazoaires: forte conservation des séquences fonctionnelles et dépendance vis-à-vis de la structure secondaire de l'intron et des bordures exoniques (Goguel et Rosbash, 1993; Libri *et al.*, 1995; Charpentier et Rosbash, 1996; Mouglin *et al.*, 1996; Fournier, 1997). Sur cette base, on pourrait donc penser que le passage des introns de groupe II aux introns épissés par des spliceosomes se serait accompagné, dans un premier temps, de l'intervention d'un nombre limité de protéines et de la mise en place de quelques modifications post-transcriptionnelles dans les UsnRNA issus de la fragmentation des introns de groupe II. Ensuite, au cours de l'évolution des introns U2-dépendants, qui s'est accompagnée d'une grande dégénérescence des séquences fonctionnelles et de l'absence de contraintes au niveau structure secondaire, le nombre de protéines impliquées et de modifications post-transcriptionnelles des UsnRNA aurait augmenté.

L'hypothèse d'un plus faible nombre de résidus modifiés dans les UsnRNA ancestraux pourrait également être une explication valable pour les UsnRNA des spliceosomes U12-dépendants de métazoaires, dans la mesure où il est actuellement proposé que ce système spliceosomal soit également un système vestige, dont les introns seraient peu à peu convertis en introns U2-dépendants, comme nous l'avons mentionné dans l'introduction bibliographique (Burge *et al.*, 1998).

Il était intéressant de constater que, parmi toutes les positions de conversion de résidus U en résidus Ψ , seules deux sont conservées dans tous les systèmes (UsnRNA des spliceosomes U2-dépendants et U12-dépendants de vertébrés, UsnRNA des spliceosomes U2-dépendants de plantes (*V. faba*), UsnRNA de *S. cerevisiae* et *S. pombe*). Ces deux résidus modifiés sont portés par le snRNA U2 et son homologue dans les spliceosomes U12-dépendants, le snRNA U12. Cette conservation phylogénétique est à corrélérer au fait que seul un snRNA U2 naturel, portant les modifications post-transcriptionnelles, est capable de participer à la reconstitution d'un spliceosome actif, à partir d'un extrait nucléaire de cellules HeLa où le snRNA U2 endogène a été éliminé (Ségault *et al.*, 1995). Un transcrit U2, produit *in vitro*, dépourvu de modifications post-transcriptionnelles, ne conduit pas, dans les mêmes conditions, à la formation d'un spliceosome actif, alors qu'un transcrit du snRNA U5 produit *in vitro* est fonctionnel dans le même type d'expériences (Ségault *et al.*, 1995).

L'un des deux résidus Ψ conservés dans les snRNA U2 et U12 est localisé dans la séquence en amont de la séquence de reconnaissance de la boîte de branchement (Fig. 47A). L'autre est situé dans cette séquence de reconnaissance, au niveau de la paire de bases qui jouxte le

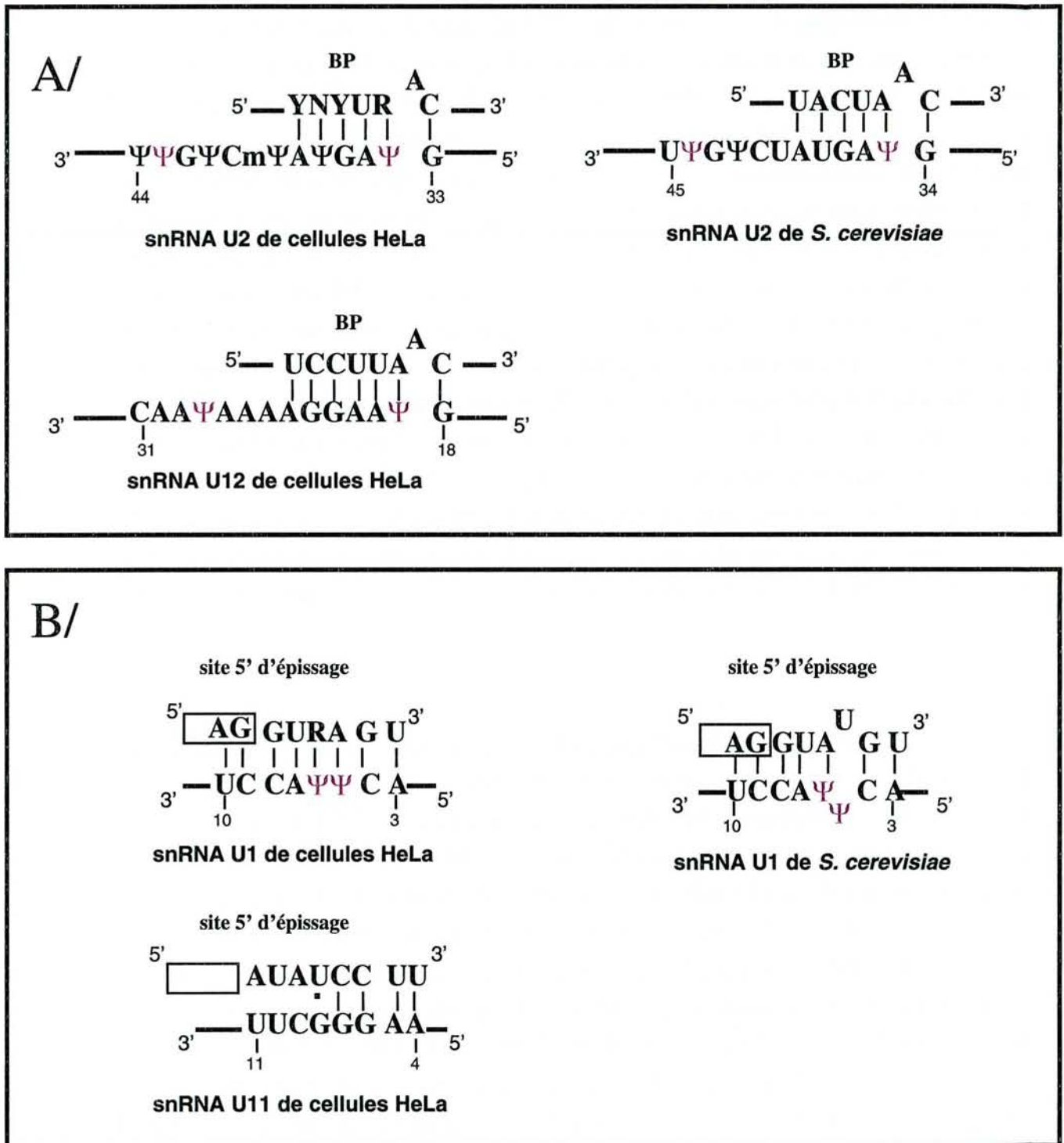


Figure 47: Interactions entre UsnRNA et pré-ARNm impliquant des résidus Ψ conservés.

A/ Appariement entre les boîtes de branchement et les snRNA U2 ou U12 chez les vertébrés et chez *S. cerevisiae*.

B/ Appariement entre les extrémités 5' des introns et les snRNA U1 ou U11 chez les vertébrés et chez *S. cerevisiae*.

Les résidus Ψ phylogénétiquement conservés sont indiqués en violet.

BP: boîte de branchement.

Y: pyrimidine

R: purine

résidu A responsable de la première attaque nucléophile (Fig. 47A). Une telle conservation de ces deux modifications dans les snRNA U2 et U12 suggère fortement qu'ils apportent un avantage sélectif important aux deux systèmes spliceosomaux. L'absence de phénotype que nous avons observée, lorsque le résidu Ψ_{44} du snRNA U2 de *S. cerevisiae* n'est pas formé, indique que cette modification n'est cependant pas essentielle. Comme nous l'avons discuté dans les publications n°1 et 2, aucune information précise n'est pour l'instant disponible sur l'importance fonctionnelle du résidu Ψ conservé, contenu dans la séquence de reconnaissance de la boîte de branchement (Ψ_{35} dans le cas du snRNA U2 de *S. cerevisiae*, Fig. 47A). Une approche pour obtenir des informations sur ce point est d'identifier la Ψ -synthase responsable de cette modification chez *S. cerevisiae* et d'interrompre sa séquence codante au niveau chromosomique. Cette étude fait partie de l'objectif plus large que nous nous sommes fixé, à la suite des travaux que nous venons de décrire. En effet, nous avons identifié, parmi les Ψ -synthases déjà connues, une seule enzyme impliquée dans la formation des résidus Ψ dans les UsnRNA de *S. cerevisiae*. Cette enzyme, Pus1p, forme le résidu Ψ_{44} dans le snRNA U2. Il restait donc à essayer d'identifier la ou les enzymes catalysant la formation des 5 autres résidus Ψ présents dans les UsnRNA de *S. cerevisiae*. Le travail réalisé, dans cette optique, sera décrit à la fin de la partie Résultats/Discussion.

Deux autres résidus Ψ , qui pourraient être fonctionnellement importants, sont les résidus Ψ_5 et Ψ_6 , fortement conservés dans tous les snRNA U1, mais pas dans le snRNA U11 (Fig. 47B). L'importance de ces deux résidus Ψ pourra également être étudiée par délétion du (ou des) gène(s) codant la (ou les) Ψ -synthase(s) responsable(s) de leur formation. En fait, suivant les espèces, le résidu Ψ_5 pourrait intervenir de manière différente dans la mise en place de l'appariement entre le snRNA U1 et le site 5' d'épissage. En effet, dans toutes les espèces, sauf chez *S. cerevisiae*, le résidu Ψ_5 est apparié au nucléotide en position 4 de l'intron (Fig. 47B). Ce résidu Ψ_5 pourrait donc, chez la plupart des espèces, stabiliser l'appariement du snRNA U1 avec l'extrémité 5' de l'intron. Il peut en être de même pour le résidu Ψ_6 . Il est à noter que l'appariement $\Psi_5:A_4$ est remplacé par un appariement $G_8:U_4$ dans les spliceosomes U12-dépendants (Fig. 47B). Dans le cas des spliceosomes de *S. cerevisiae*, la situation est différente. En effet, le résidu Ψ_5 n'est pas apparié à un résidu de l'intron et un mésappariement de type Ψ_5/U_4 est toujours observé (Fig. 47B). Comme nous l'avons mentionné dans la partie Introduction de ce manuscrit, la mutation du résidu Ψ_5 en résidu A dans le snRNA U1 de *S. cerevisiae*, permettant la formation d'un appariement $A_5:U_4$ avec le site 5', conduit à une forte diminution de la croissance (Siliciano et Guthrie, 1988; Séraphin *et al.*, 1989b). Ces résultats avaient conduit les auteurs à conclure que l'existence de cet appariement supplémentaire chez *S. cerevisiae* devait avoir un effet négatif sur la réaction

d'épissage. On pourrait imaginer qu'une trop forte stabilisation de l'appariement U1/site 5' gêne le relargage du snRNA U1, lors de la mise en place de l'interaction entre le snRNA U6 et le site 5'. Cependant, une autre explication est possible. En effet, selon des travaux non publiés de D. Libri, le mésappariement Ψ_5/U_4 augmenterait la stabilité de l'interaction formée entre le snRNA U1 et le site 5' dans un extrait nucléaire, et ceci probablement par reconnaissance de ce mésappariement par une ou plusieurs protéines, qui auraient un rôle stabilisateur. Ces observations nous amènent à poser la question suivante: est-ce qu'un mésappariement U_5/U aurait les mêmes capacités de stabilisation de l'interaction que le mésappariement Ψ_5/U ? La seule manière de répondre à cette question sera, là encore, d'empêcher la formation de cette modification par délétion du gène codant la Ψ -synthase responsable de la catalyse.

L'identification de la (ou des) Ψ -synthase(s), responsable(s) de la formation des résidus Ψ , autres que le résidu Ψ_{44} , dans les UsnRNA de *S. cerevisiae*, est donc essentielle afin de tester l'importance fonctionnelle de ces modifications. De plus, cette identification permettra de savoir si chacune de ces modifications est catalysée par une enzyme distincte, ou si certaines enzymes sont responsables de la formation de plusieurs résidus Ψ portés par les UsnRNA, ou bien encore si une seule enzyme est responsable de la formation de tous les résidus Ψ dans le snRNA U2, à l'exception du résidu Ψ_{44} . Il serait également intéressant de savoir si la (ou les) enzyme(s) agissant sur les UsnRNA pourrait présenter une double spécificité de substrats, comme c'est le cas pour Pus1p. Cette caractéristique de Pus1p méritait, par elle-même, des études ultérieures que nous avons entreprises à l'aide de différentes approches, qui sont explicitées et discutées dans le chapitre suivant.

2. A QUELLES ETAPES DE LA BIOGENESE DES UsnRNP SONT FORMEES LES MODIFICATIONS POST-TRANSCRIPTIONNELLES DES UsnRNA?

Presque aucune donnée n'est à ce jour disponible sur l'ordre d'apparition des différentes modifications portées par les UsnRNA et sur les stades de la biogenèse des UsnRNP où elles ont lieu. Les modifications post-transcriptionnelles des UsnRNA sont-elles réalisées dans le noyau après transcription de ces ARN? Dans le cytoplasme avant, pendant ou après la formation des pré-particules U1, U2, U4 et U5? Ou bien, encore, après retour dans le noyau des pré-particules, lors de l'association des protéines spécifiques à chaque UsnRNP? Est-ce que certaines modifications portées par les UsnRNAs sont synthétisées grâce à un système d'ARN guides?

Nous n'avons détecté aucun résidu Ψ dans le snRNA U6 de *S. cerevisiae*, alors que le snRNA U6 humain porte 3 résidus Ψ , 8 résidus 2'-O-méthylés, ainsi que deux résidus méthylés sur la base. Des résultats inattendus viennent d'être obtenus sur le mode de synthèse des 2'-O-méthylations dans le snRNA U6 de vertébrés et les observations faites pourraient également être valables pour les pseudouridylations de cet ARN. En effet, il a été montré très récemment, chez la souris, *C. elegans* et *X. laevis*, que deux snoRNA C/D sont nécessaires à la méthylation en 2' du ribose des résidus A₄₇ et C₇₇ des snRNA U6 de vertébrés (ces deux ARN guides sont appelés snoRNA mgU6-47 et mgU6-77, respectivement, mg pour "methylation guide") (Tycowski *et al.*, 1998). Un troisième snoRNA C/D pourrait aussi guider la méthylation du résidu A₅₃ dans le snRNA U6 humain (snoRNA mgU6-53) (Ganot *et al.*, 1999). Ces snoRNA présentent les mêmes caractéristiques que les snoRNA C/D guides de méthylation des pré-ARNr eucaryotes: présence des boîtes C, C', D et D', association à la fibrillarine et localisation nucléolaire (Tycowski *et al.*, 1998). Le snoRNA mgU6-77 est d'ailleurs nécessaire à la 2'-O-méthylation de la position 2970 dans l'ARNr 28S, en plus de son action sur le snRNA U6.

Comme nous l'avons décrit dans l'introduction de ce manuscrit, la biogenèse de la snRNP U6 se fait de manière différente de celle des autres UsnRNP. En effet, les snRNA U1, U2, U4 et U5 sont transcrits par l'ARN polymérase II, puis transitent dans le cytoplasme où ils s'associent aux protéines Sm. Ils sont ensuite réimportés dans le noyau, où a lieu l'association des protéines spécifiques. Le snRNA U6, quant à lui, est transcrit par l'ARN polymérase III et ne transite pas dans le cytoplasme. Il est associé dans le noyau à un autre ensemble de protéines, qui sont les protéines Lsm. La découverte de snoRNA guides nécessaires à la 2'-O-méthylation des snRNA U6 de vertébrés était inattendue car aucun fait expérimental ne suggérait jusque là que des UsnRNA spliceosomaux puissent transiter dans le nucléole. Cependant, très récemment, l'équipe de T. Kiss a montré que des fragments du snRNA U6 de souris, transcrits dans le nucléole par l'intermédiaire de l'ARN polymérase I, sont efficacement pseudouridylés et 2'-O-méthylés (Ganot *et al.*, 1999). De plus, après injection dans le nucléoplasme d'oocytes de *X. laevis*, un transcrit du snRNA U6 produit *in vitro* s'accumule transitoirement dans le nucléole (Terns et coll., résultats non publiés). Il semblerait donc que les résidus Ψ et 2'-O-méthylés portés par les snRNA U6 de vertébrés soient formés dans le nucléole de manière snoRNA dépendante, comme c'est le cas pour la modification des pré-ARNr, bien qu'à l'heure actuelle aucun snoRNA H/ACA spécifique du snRNA U6 n'ait été décrit.

Une telle modification du snRNA U6 dans le nucléole n'est pas en accord avec des résultats antérieurs de J. Patton. En effet, ces résultats, obtenus *in vitro*, suggéraient que la pseudouridylation du snRNA U6 de cellules HeLa dépendait de son interaction avec le snRNA U4 et de la fixation des protéines Sm sur le site Sm du snRNA U4 (Zerby et Patton,

1996). Cette pseudouridylation pouvait avoir lieu dans un extrait S100 ou un extrait nucléaire de cellules HeLa (Zerby et Patton, 1996). Il faut noter que le taux de modification d'un snRNA U6, produit *in vitro*, après incubation dans ces différents extraits, était très faible (de l'ordre de 10% par rapport au pourcentage théorique attendu de résidus Ψ par mole de snRNA U6). Les positions modifiées n'avaient pas été identifiées avec précision. Comme nous le verrons par la suite, dans le cas de l'étude de la Ψ -synthase Pus1p, certaines modifications des ARN peuvent être obtenus *in vitro*, sans que ces mêmes modifications soient présentes *in vivo*.

Etant donné que nous n'avons pas détecté de résidus Ψ dans le snRNA U6 de *S. cerevisiae*, il est possible qu'aucun homologue des snoRNA guides de pseudouridylation du snRNA U6 de vertébrés n'existe chez *S. cerevisiae*. Du fait que nous n'avons pas analysé les 2'-O-méthylations portées par cet UsnRNA, il est difficile de savoir s'il peut, par contre, exister des snoRNA guides de méthylation du snRNA U6 chez *S. cerevisiae*. D'après les données de la littérature, chez *S. cerevisiae*, le résidu de l'ARNr 26S correspondant au résidu en position 2970 de l'ARNr 28S de *X. laevis* n'est pas 2'-O-méthylé. Donc, chez *S. cerevisiae*, le snoRNA mgU6-77 qui, chez les vertébrés, est responsable de cette 2'-O-méthylation ainsi que de la méthylation en position 77 du snRNA U6 est différent ou n'existe pas.

Une explication possible pour l'absence de résidus Ψ dans le snRNA U6 de *S. cerevisiae* pourrait être que, dans cette espèce, cet UsnRNA ne transite pas dans le nucléole. Il n'est d'ailleurs pas encore sûr que, dans cet organisme, les snRNA U1, U2, U4 et U5 transitent dans le cytoplasme.

Quoiqu'il en soit, le mode de synthèse des résidus Ψ et 2'-O-méthylés dans les snRNA U1, U2, U4 et U5 semble différent de celui du snRNA U6. Cette hypothèse est basée sur plusieurs observations:

- ⇒ des fragments du snRNA U2 de souris transcrits dans le nucléole ne sont ni 2'-O-méthylés, ni pseudouridylés dans ce compartiment, contrairement au snRNA U6 (Ganot *et al.*, 1999),
- ⇒ nous avons montré que l'absence d'expression du gène codant la Ψ -synthase nucléolaire Cbf5p de *S. cerevisiae*, responsable de la formation snoRNA dépendante des résidus Ψ dans les pré-ARNr, ne modifiait pas le taux de pseudouridylation des snRNA U1, U2 et U5,
- ⇒ nous avons également montré que l'enzyme Pus1p modifiait *in vitro* un transcrit correspondant au snRNA U2 de *S. cerevisiae* et ceci sans l'aide d'ARN guide.

Ces résultats sont en faveur de l'idée que la formation des résidus Ψ et 2'-O-méthylés dans les snRNA U1, U2 et U5 (et par extrapolation dans le snRNA U4) n'a pas lieu dans le nucléole et ne repose pas sur le système des snoRNA guides. L'hypothèse la plus probable est qu'un ensemble de 2'-O-méthylases et de Ψ -synthases soit responsable de la formation de ces

modifications, comme c'est le cas pour la modification des ARNt. En partant de cette hypothèse, nous pouvons nous demander dans quel compartiment cellulaire ces modifications sont réalisées, et à quelle étape de la maturation des UsnRNP elles ont lieu.

L'équipe de J. Patton avait observé que certaines modifications des UsnRNA pouvaient être formées dans des transcrits, produits *in vitro*, après incubation soit en présence d'un extrait nucléaire, soit en présence d'un extrait S100 de cellules HeLa (pour revue, Patton, 1994a). Ces observations suggèrent que certaines des activités Ψ -synthases, agissant sur les UsnRNA de cellules HeLa, sont localisées dans le cytoplasme et d'autres dans le noyau.

Nous avons montré que la Ψ -synthase Pus1p pouvait modifier un transcrit, produit *in vitro*, du gène codant le snRNA U2 de *S. cerevisiae*, et ceci sans qu'il soit nécessaire d'ajouter des protéines Sm ou des protéines spécifiques de la snRNP U2 de levure. L'enzyme Pus1p est localisée dans le noyau (Simos *et al.*, 1996). Elle pourrait donc modifier le snRNA U2, soit immédiatement après sa transcription, soit après formation de la snRNP U2, sans pour autant être dépendante de la présence des protéines. Une situation différente avait été observée, par l'équipe de J. Patton, pour le snRNA U5 de cellules HeLa. Dans des expériences de modification *in vitro* avec des extraits cellulaires, la présence des protéines Sm semblait nécessaire. Les Ψ -synthases concernées semblent donc avoir pour substrat la pré-particule U5 ou la snRNP U5 (Patton, 1991; Patton, 1994b). De même, J. Patton a observé que la présence d'un site Sm intact augmente l'efficacité de formation, *in vitro*, de résidus Ψ dans le snRNA U4 de cellules HeLa (Zerby et Patton, 1997).

Il a été montré par d'autres auteurs (Ségault *et al.*, 1995; Yu *et al.*, 1998) que la présence de certaines des modifications du snRNA U2 est requise pour la formation de la snRNP U2 17S de vertébrés. La formation de ces modifications doit donc précéder la fixation des facteurs SF3a et SF3b sur le snRNA U2.

Les données actuelles sur les étapes où ont lieu les modifications post-transcriptionnelles sont donc encore très imprécises. Afin de mieux connaître l'ordre chronologique de formation des résidus Ψ et 2'-O-méthylés dans les UsnRNA, notre équipe va donc débiter une étude basée sur la micro-injection d'ARN dans les oocytes de *X. laevis*, ceci en collaboration avec l'équipe de R. Lührmann. Des UsnRNA, produits par transcription *in vitro* qui pourront, lorsque ce sera nécessaire, être marqués en 3', seront injectés dans le noyau et leur transport vers le cytoplasme sera inhibé par mutation de la coiffe m⁷G. Ceci permettra d'analyser les résidus Ψ spécifiquement formés dans le noyau. De la même manière, des transcrits, produits *in vitro* seront injectés dans le cytoplasme et l'import de ces transcrits dans le noyau sera inhibé, afin de cartographier les résidus Ψ spécifiquement formés dans ce compartiment.

Un autre volet de recherche, ouvert par nos résultats et qui sera abordé par l'équipe, est l'étude de la formation des résidus Ψ dans les précurseurs des UsnRNA. En effet, comme nous l'avons mentionné dans l'introduction, chez *S. cerevisiae*, les snRNA U1, U2 et U5 sont transcrits sous forme de précurseurs et sont ensuite maturés par un processus faisant intervenir l'enzyme Rnt1, homologue à la RNase III bactérienne (Chanfreau *et al.*, 1997; Abou Elela et Ares, 1998; Seipelt *et al.*, 1999). G. Chanfreau nous a déjà fourni une souche de *S. cerevisiae* dans laquelle le gène *RNT1* a été délété. Le précurseur du snRNA U2 s'accumule dans cette souche (Abou Elela et Ares, 1998), ce qui va nous permettre d'étudier la formation des résidus Ψ dans cet ARN précurseur. Le précurseur du snRNA U1 ne pourra pas être analysé par cette approche, car une deuxième voie de maturation, pour cet ARN, coexiste avec le système Rnt1 (Seipelt *et al.*, 1999). L'analyse du précurseur du snRNA U5 sera également difficile. En effet, dans une souche *rnt1*-, la voie de maturation conduisant à la production de la forme longue du snRNA U5 (U5L) est bloquée, mais pas la voie de maturation conduisant à la production de la forme courte du snRNA U5 (U5S) (Chanfreau *et al.*, 1997). Ainsi, seul le précurseur du snRNA U2 de *S. cerevisiae* pourra être étudié par cette approche. Etant donné que la formation de la snRNP U2 se fait au même taux que dans une souche sauvage, l'analyse ne pourra porter que sur le précurseur contenu dans la UsnRNP.

*PARTIE 3: ETUDE DE LA RECONNAISSANCE DE L'ARN PAR LA
Ψ-SYNTHASE Pus1p DE S. cerevisiae. COMPARAISON DES ENZYMES
Pus1p DE LEVURE ET DE SOURIS.*

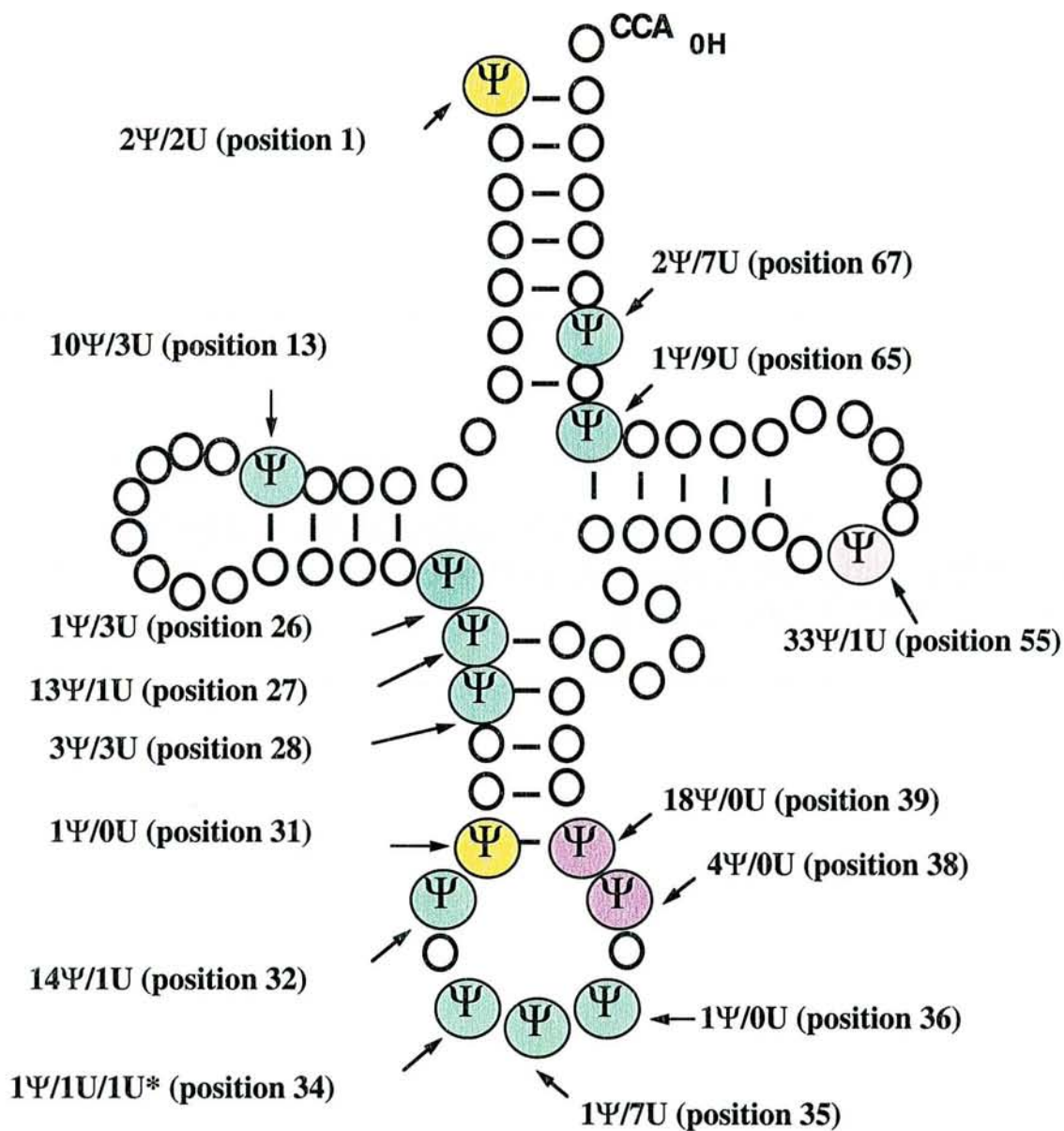


Figure 48: Fréquence de la présence de résidus Ψ et U, aux différentes positions des ARNt où des résidus Ψ peuvent être présents.

L'analyse concerne les 34 ARNt cytoplasmiques de *S. cerevisiae* séquencés au niveau ARN. Le nombre de résidus U, de résidus Ψ et de résidus U modifiés (autres que des résidus Ψ) (U*) est indiqué. Le résidu Ψ₅₅ formé par Pus4p est indiqué en rose (Becker *et al.*, 1997a) et les résidus Ψ₃₈ et Ψ₃₉ formés par Pus3p sont montrés en violet (Lecoite *et al.*, 1998).

L'action de Pus1p a été testée sur toutes les positions où des conversions de résidus U en résidus Ψ ont lieu naturellement dans les ARNt cytoplasmiques de *S. cerevisiae* (Motorin *et al.*, 1998), à l'exception des positions 1 et 31 indiquées en jaune.

Ce schéma est adapté de celui de Motorin *et al.* (1998).

A/ RECHERCHE D'UN MOTIF ARN RECONNU PAR L'ENZYME Pus1p DE *S. cerevisiae*.

1. PRESENTATION DU PROBLEME

1.1. L'enzyme Pus1p de *S. cerevisiae* est responsable de la formation de plusieurs résidus Ψ dans les ARNt

Des résidus Ψ peuvent être trouvés à 15 positions distinctes dans les ARNt cytoplasmiques de *S. cerevisiae*. Certaines conversions de résidus U en résidus Ψ sont observées dans presque tous les ARNt (comme la Ψ_{55}). D'autres ne sont présentes que dans un seul des ARNt (Ψ en positions 26, 31, 34, 35, 36 et 65) (Fig. 48). Afin d'identifier les positions des ARNt de *S. cerevisiae* où la Ψ -synthase Pus1p catalyse la conversion de résidus U en résidus Ψ , Motorin *et al.* (1998) ont utilisé une large gamme de transcrits, produits *in vitro*. Des analyses ont été réalisées sur des transcrits entiers, après modification par Pus1p, ou sur des fragments de ces transcrits modifiés, produits par hydrolyse par la RNase T1, les résidus Ψ étant identifiés, soit par la méthode de modification par le CMCT, soit par celle du plus proche voisin (voir ci-dessous).

Les résultats obtenus montrent qu'*in vitro*, la Ψ -synthase Pus1p est capable de catalyser la formation de résidus Ψ en positions 27, 28, 34, 35 et 36 des ARNt cytoplasmiques de *S. cerevisiae*. L'analyse des résidus Ψ formés dans les ARNt d'une souche de *S. cerevisiae pus1⁻*, a permis de vérifier que Pus1p était bien responsable de la formation des résidus Ψ à 4 de ces positions (27, 28, 34 et 36). Cependant, même en absence de Pus1p, le résidu Ψ_{35} est toujours formé *in vivo*. Cette analyse montre également que dans une souche de *S. cerevisiae pus1⁻*, les résidus Ψ aux positions 26, 65 et 67 ne sont pas formés (Motorin *et al.*, 1998). L'enzyme Pus1p ne catalyse pas *in vitro* la formation de ces résidus Ψ . Différentes explications sont possibles: i) la structure des différents transcrits utilisés *in vitro* peut différer de la structure adoptée par ces ARNt *in vivo* et ne serait pas appropriée pour la reconnaissance par Pus1p *in vitro* ou ii) la formation de ces 3 résidus Ψ , par une ou plusieurs Ψ -synthases non caractérisées, nécessite la formation préalable de certains résidus Ψ par Pus1p.

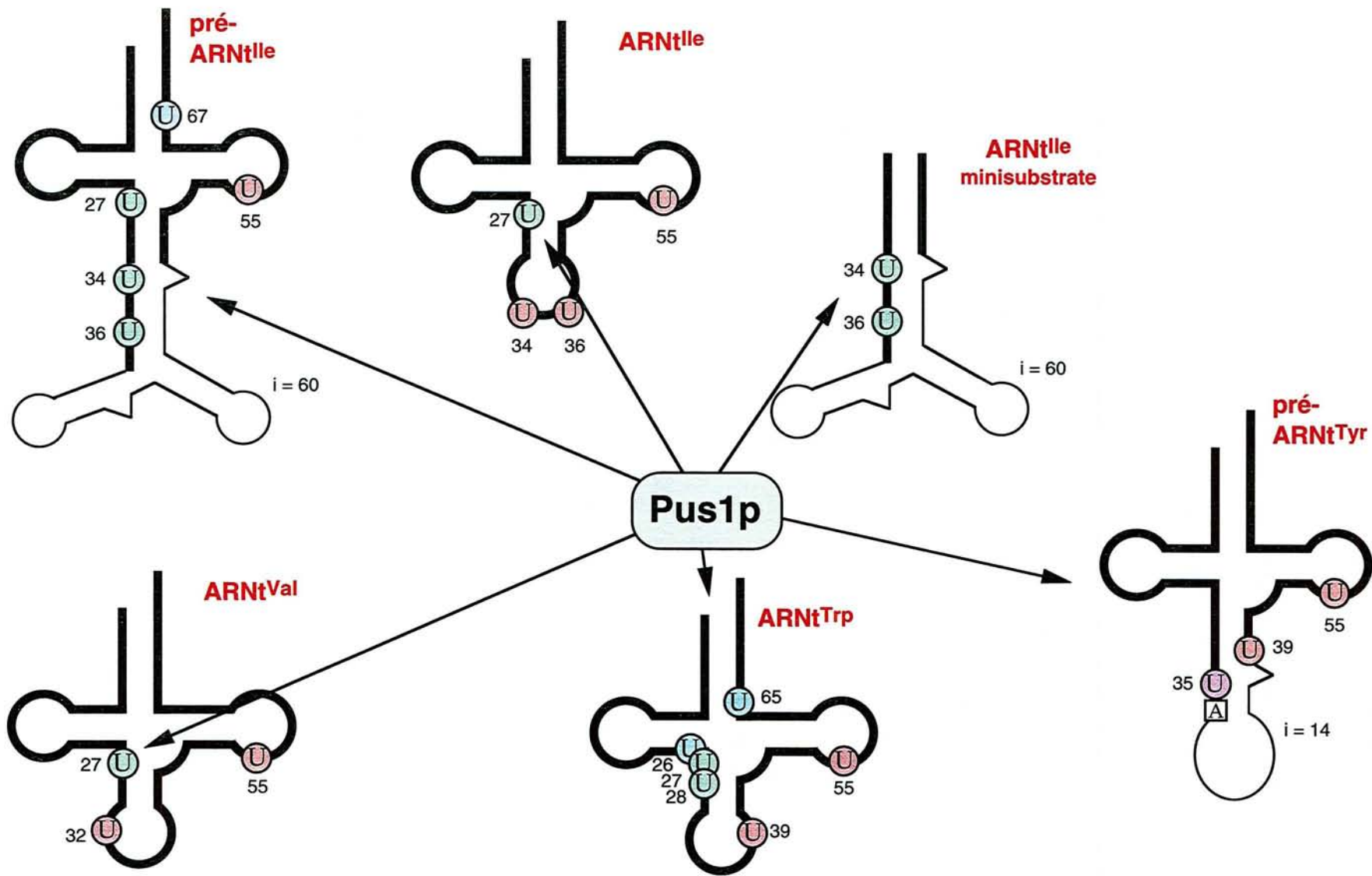


Figure 49: Exemples d'ARNt étudiés dans le cadre de la recherche des sites de pseudouridylation de l'enzyme Pus1p (Motorin *et al.*, 1998). Pour chaque ARNt ou pré-ARNt, les résidus U convertis en résidus Ψ , *in vivo* et *in vitro*, par Pus1p sont en vert. Les résidus U, qui ne sont pas convertis en résidus Ψ *in vivo* après abolition de l'expression du gène *PUS1*, et qui ne sont pas formés *in vitro* par Pus1p, sont indiqués en bleu. Le résidu Ψ_{35} converti *in vitro* par Pus1p est indiqué en violet. Les résidus U non convertis en résidus Ψ par Pus1p et convertis *in vivo* et *in vitro* par d'autres enzymes sont indiqués en rose. Les tailles des introns sont données (i).

1.2. La formation des résidus Ψ par Pus1p dans les ARNt ne dépend pas de la structure 3D des ARNt. Pus1p peut agir sur des petits fragments d'ARNt

La Ψ -synthase Pus1p est capable de convertir les résidus U₂₇ et U₂₈ en résidus Ψ dans l'ARNt mature. Cependant, la formation de résidus Ψ en positions 34, 35 et 36 est intron dépendante. Chez *S. cerevisiae*, la formation de résidus Ψ aux positions 34 et 36 n'est observée que dans l'ARNt^{lle} mineur. Des expériences de modification *in vitro*, avec un extrait cellulaire de levure, ont montré la nécessité de l'intégrité des 60 nts de l'intron pour qu'il y ait modification. *In vitro*, un ARN ne contenant que la structure anticodon avec l'intron est modifié aux positions 34 et 36 (Szweykowska-Kulinska *et al.*, 1994; Simos *et al.*, 1996; Motorin *et al.*, 1998) (Fig. 49). Il a aussi été montré que la formation de chacune de ces deux modifications se fait de manière indépendante (Szweykowska-Kulinska *et al.*, 1994). Ainsi, l'enzyme Pus1p pourrait former certains résidus avant l'élimination de l'intron et, dans ce cas, la structure 3D de l'ARNt n'est pas requise. Elle pourrait catalyser la synthèse des autres résidus Ψ soit dans le pré-ARNt, soit dans l'ARNt mature.

La formation du résidu Ψ_{35} dans l'ARNt^{Tyr} cytoplasmique de *S. cerevisiae* est également intron dépendante (Johnson et Abelson, 1983). D'une manière générale, dans toutes les espèces où un résidu Ψ est présent en position 35, la formation de ce résidu a lieu sur l'ARNt précurseur contenant l'intron (pour revue, Grosjean *et al.*, 1997). Pus1p est capable de catalyser *in vitro* la formation du résidu Ψ_{35} dans le pré-ARNt^{Tyr} de *S. cerevisiae* et la présence de l'intron est nécessaire. Cependant, comme nous l'avons déjà mentionné, lorsque le gène *PUS1* est délété, ce résidu Ψ est toujours formé. Ces observations indiquent donc que, chez *S. cerevisiae*, une Ψ -synthase (ARNt₃₅: Ψ -synthase), non caractérisée à ce jour, modifie la position 35 du pré-ARNt^{Tyr} *in vivo*. La question qui se pose maintenant est de savoir si Pus1p peut également former la Ψ_{35} dans le pré-ARNt^{Tyr} *in vivo* et, donc, si deux Ψ -synthases ont une activité redondante *in vivo*, comme ceci a été proposé par l'équipe d'H. Grosjean (Motorin *et al.*, 1998).

1.3. Pus1p peut agir *in vitro* sur des ARN présentant un repliement de type ARNt

Pus1p est également capable d'agir sur des ARN pouvant adopter une structure de type ARNt, comme c'est le cas pour la région 3'-terminale de l'ARN génomique du virus de plantes TYMV (Turnip Yellow Mosaic Virus), qui présente une structure 3D très similaire à celle des ARNt et qui peut être aminoacylé par la valyl-ARNt synthétase *in vitro* et *in vivo* et l'histidiny-ARNt synthétase *in vitro* (Rudinger *et al.*, 1992; pour revue, Florentz et Giegé, 1995). Pus1p catalyse, en effet, la conversion de l'uridine en position 65, de ce segment

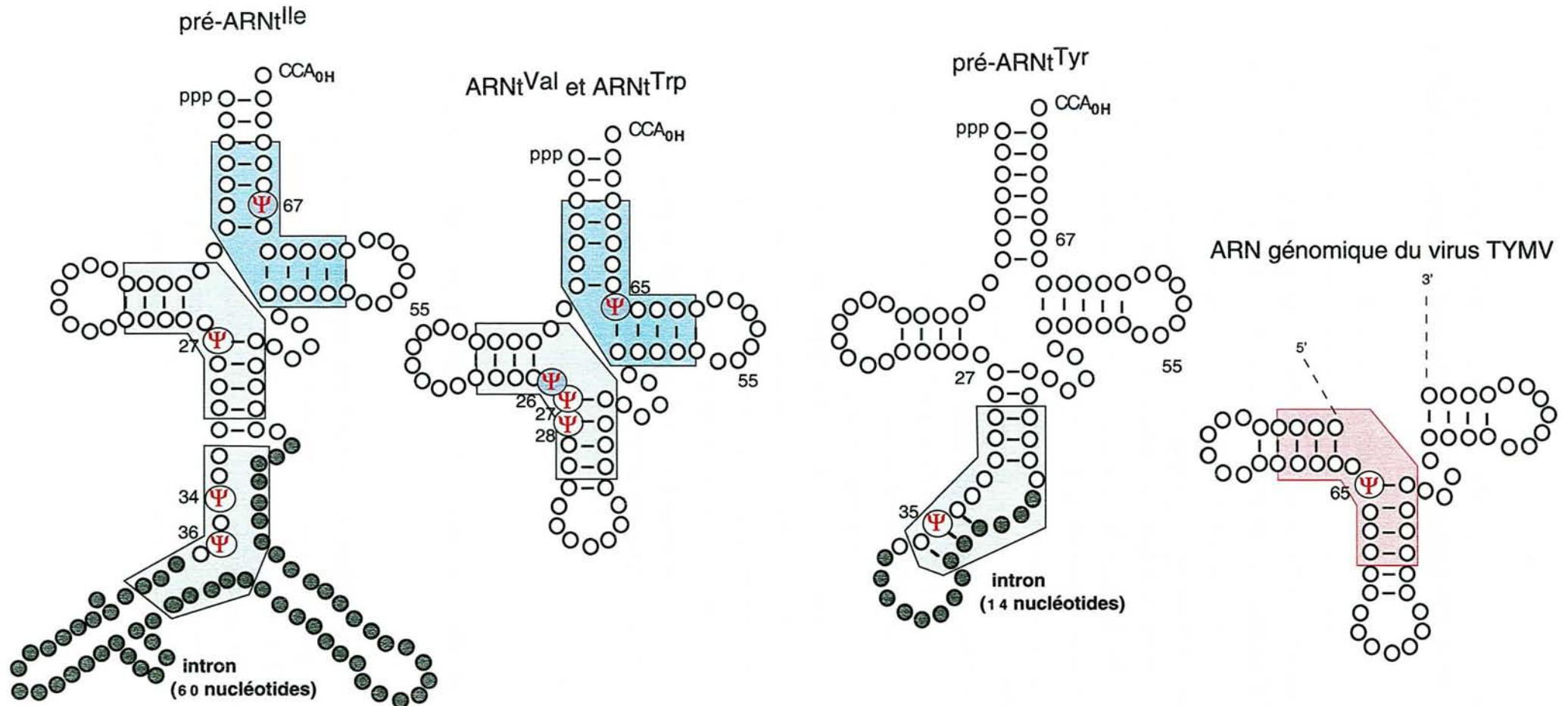


Figure 50: Représentation des domaines structuraux, formés de deux hélices, reconnus par la Ψ -synthase Pus1p de *S. cerevisiae* dans différents substrats ARN (ARNt, pré-ARNt et ARN génomique du virus TYMV).

Le domaine structural communs aux différents substrats est encadré. Pour les substrats ARNt ou pré-ARNt, ce domaine est représenté en bleu lorsque seules des informations obtenues *in vivo* suggèrent que Pus1p est impliquée dans la modification (en bleu). Le domaine est représenté en vert lorsque la position (en vert) est modifiée par Pus1p à la fois *in vivo* et *in vitro* (Motorin *et al.*, 1998).

Le domaine modifié *in vitro* par Pus1p, dans l'ARN génomique du virus TYMV, est en rouge (Becker *et al.*, 1998).

d'ARN, en résidu Ψ (Fig. 50) (Becker *et al.*, 1998). Dans la structure de type ARNt, formée par l'extrémité 3' de cet ARN viral, l'uridine 65 correspond à la position 27 des ARNt, également cible d'une conversion en résidu Ψ dans certains ARNt (Fig. 49 et 50) (Motorin *et al.*, 1998).

1.4. L'enzyme Pus1p reconnaît les ARNt de manière large

L'interaction entre Pus1p et les ARNt semble impliquer des interactions électrostatiques et des interactions hydrophobes et la constante de dissociation du complexe est de l'ordre de 15 à 150 nM (Arluison *et al.*, 1999a, 1999b). Comme la Ψ -synthase TruA d'*E. coli*, la Ψ -synthase Pus1p se fixe sur tous les ARNt, indépendamment de la présence d'un résidu U cible (Kammen *et al.*, 1988; Arluison *et al.*, 1999a). Cependant, l'examen des positions modifiées par cette Ψ -synthase, au sein de la structure 2D des ARNt, montre que le ou les résidus Ψ formés se trouvent toujours à proximité de la jonction de deux structures tige-boucles (Fig. 50).

1.5. Quel est le requis en structure et en séquence pour, qu'après s'être lié à Pus1p, un ARN soit modifié par cette Ψ -synthase?

L'ensemble des travaux sur les ARNt et les mini-substrats dérivés d'ARNt, réalisé par l'équipe d'H. Grosjean, suggérait que Pus1p modifiait des résidus U situés dans une région charnière entre deux hélices, ou à l'extrémité d'une hélice bordant une telle région charnière. Comment préciser cette spécificité supposée de Pus1p? Tous les substrats ARNt étaient très similaires en structure. Une idée a été de tester l'action de l'enzyme Pus1p recombinante sur différents ARN n'étant pas des substrats naturels de cette enzyme et dont la structure 3D était connue ou pouvait être déterminée expérimentalement. Cette approche semblait raisonnable étant donné que H. Grosjean et Y. Motorin avait montré que l'enzyme Pus1p recombinante formait *in vitro* des résidus Ψ dans une grande variété de substrats. La formation de ces résidus Ψ , par la Ψ -synthase Pus1p, avait été suivie par chromatographie bidimensionnelle après digestion des ARN modifiés par les RNases P1 ou T2 (Fig. 51). L'utilisation de la méthode au CMCT nous a ensuite permis de localiser précisément les positions modifiées. Les domaines structuraux de l'ARN ribosomique 26S de *S. cerevisiae*, produits par transcription *in vitro*, constituaient un bon modèle d'étude. En effet, il y a longtemps, un modèle de structure 2D avait été proposé pour ces domaines par notre équipe (Veldman *et al.*, 1981) et des vérifications expérimentales de cette structure 2D pouvaient être réalisées sur les transcrits. Nous avons choisi d'étudier l'action de l'enzyme Pus1p recombinante sur les domaines II, IV et V, qui portent naturellement des résidus Ψ (Veldman *et al.*, 1981; pour

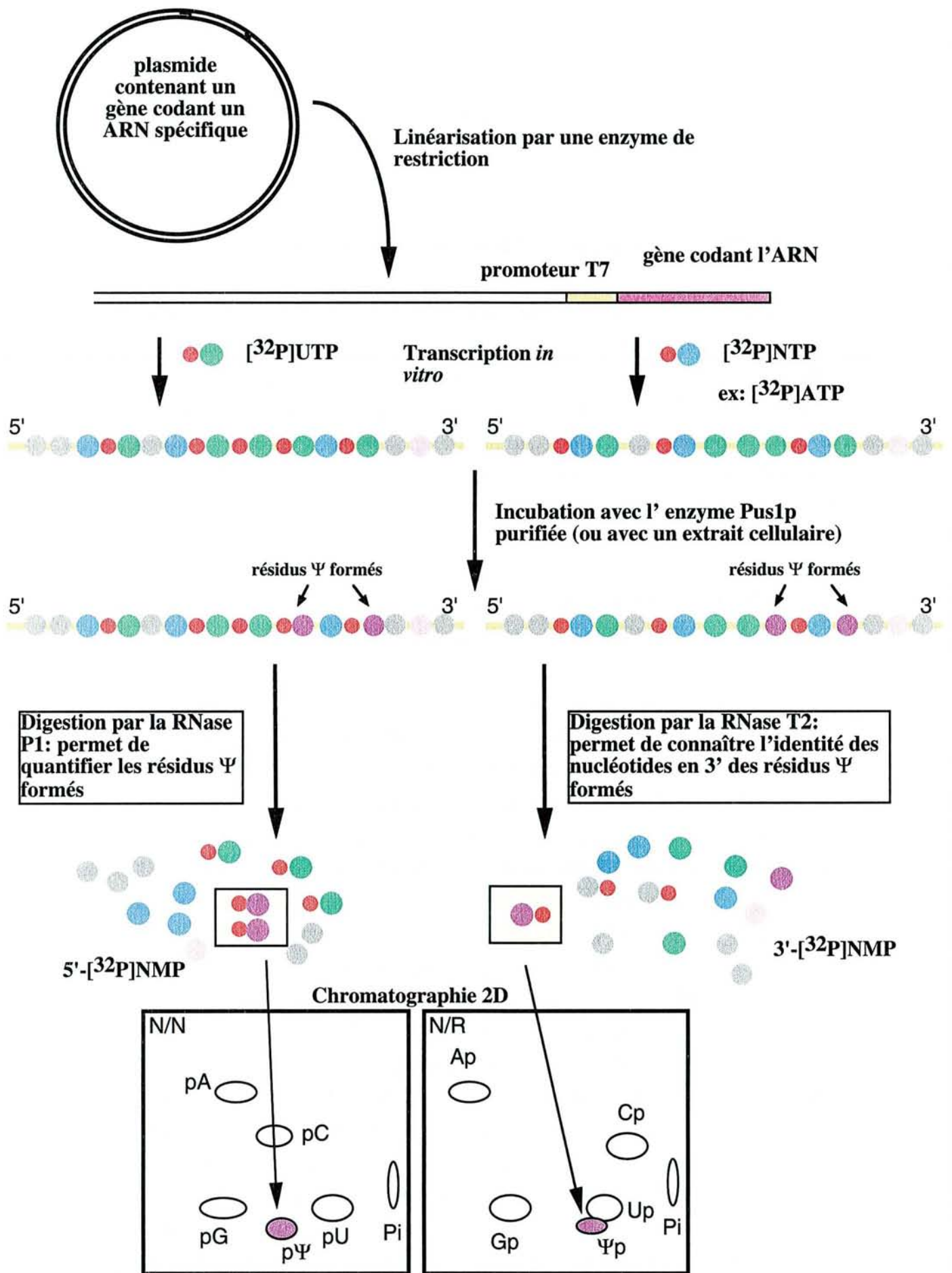


Figure 51: Principe de la quantification des résidus Ψ, basée sur l'hydrolyse totale par la RNase P1, et principe de la localisation des résidus Ψ par la méthode dite "du plus proche voisin", basée sur l'hydrolyse totale de l'ARN par la RNase T2.
bleu: résidu A; violet: résidu Y; vert: résidu U; gris: résidu G; rose: résidu C.

revue, Ofengand et Fournier, 1998). Les régions de l'ARNr 26S étudiées ont été sélectionnées de manière à ce que, produites individuellement par transcription *in vitro*, elles conservent leur structure 2D. En utilisant la méthode basée sur la modification par le CMCT, j'ai identifié les résidus Ψ formés dans les domaines IV et V, après incubation avec l'enzyme Pus1p recombinante. En parallèle, j'ai analysé la structure 2D de ces domaines, produits par transcription *in vitro*, ceci dans les conditions de tampon utilisées pour l'incubation avec Pus1p. Pour cela, j'ai utilisé des sondes enzymatiques et chimiques:

- La RNase V1, qui coupe spécifiquement les régions structurées des ARN,
- La RNase T1, qui clive spécifiquement les guanosines situées dans des régions en simple brin, en conditions limitantes,
- Le diméthylsulfate (DMS), qui modifie les positions N₇-G, N₁-A et, plus lentement, les positions N₃-C dans les ARN,
- Le CMCT, qui modifie les positions N₃-U et plus lentement les positions N₁-G dans les ARN.

Ces deux sondes chimiques permettent donc d'identifier les nucléotides non impliqués dans une interaction de type Watson et Crick (Fig. 52).

L'analyse de la structure 2D de la région 3' du transcrit correspondant au domaine IV est présentée dans la publication n°4. Quelques différences sont observées par rapport au modèle de structure 2D préalablement établi. L'analyse du fragment étudié du transcrit correspondant au domaine V est présentée dans la Figure 53. Dans ce cas, la structure observée est en accord avec le modèle précédemment proposé (Veldman *et al.*, 1981).

L'examen des positions modifiées mettait en évidence un requis en séquence. Tous les résidus U convertis en résidus Ψ étaient précédés d'une purine. L'identité du résidu en aval de la position modifiée était par contre variable. Enfin, d'après la structure 2D des deux domaines, les résidus modifiés étaient, dans la plupart des cas, situés à l'extrémité d'une hélice ou en bordure d'une hélice. L'idée que Pus1p modifie des régions charnières entre des structures tige-boucles, ressortie des études réalisées sur les ARNt par l'équipe d'H. Grosjean, était en accord avec les données obtenues sur ces domaines de l'ARNr 26S.

Pour essayer de mieux définir le motif structural reconnu, H. Grosjean et Y. Motorin ont ensuite étudié l'action de Pus1p sur un transcrit correspondant à l'ARN génomique du virus de plantes BMV (Brome Mosaic Virus). Cet ARN, comme l'ARN génomique du virus TYMV, contient un domaine 3'-terminal dont le repliement présente des homologies avec les ARNt (Felden *et al.*, 1994), cette structure est aminoacylable par la tyrosyl-ARNt synthétase

Sondes chimiques

- DMS (Diméthylsulfate)
- CMCT

N_3 -C, N_1 -A et N_7 -G
 N_3 -U et N_1 -G

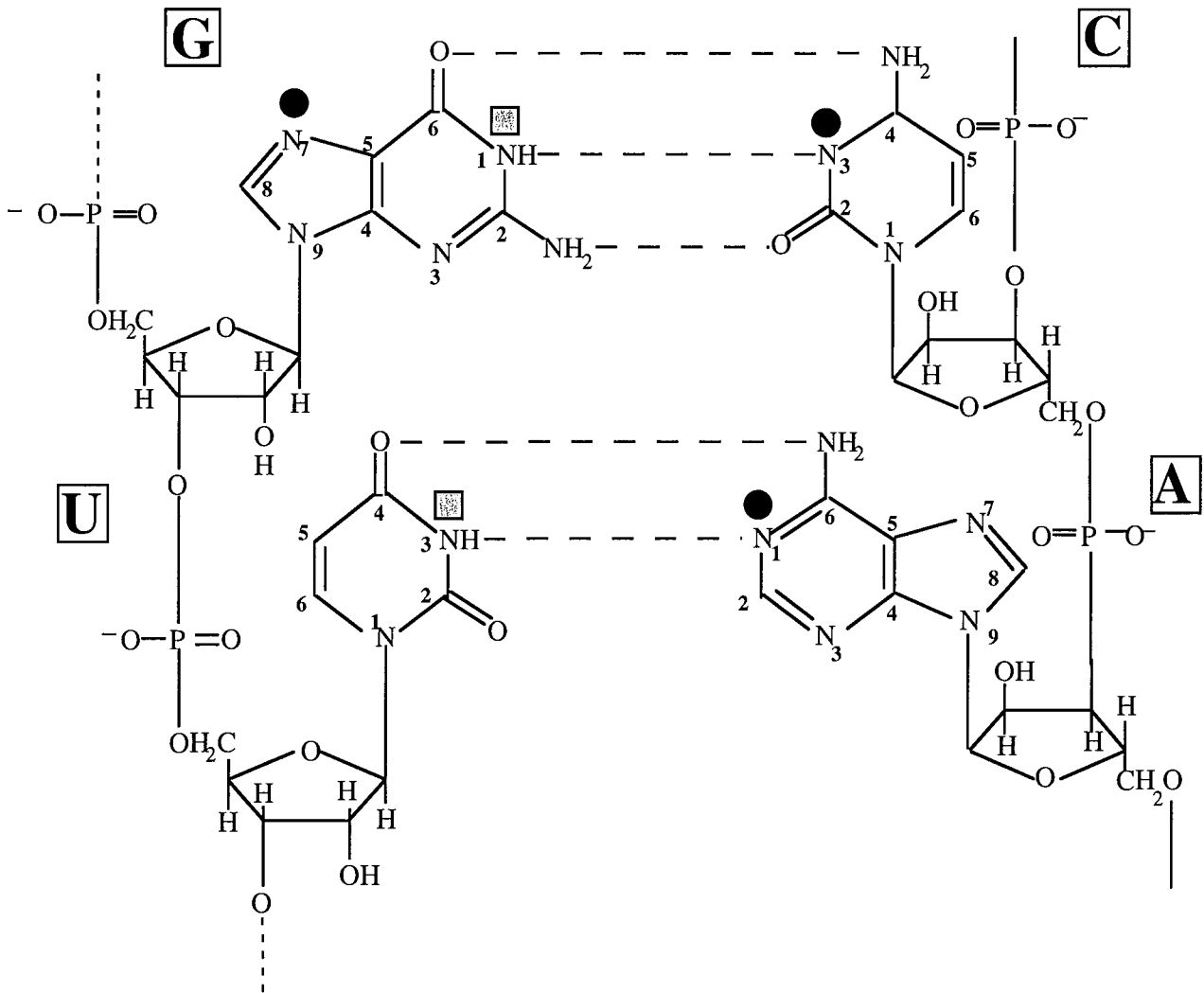


Figure 52 : Localisation des sites de modification des bases par les sondes chimiques.

◻ : site de modification par le CMCT,

● : site de modification par le DMS.

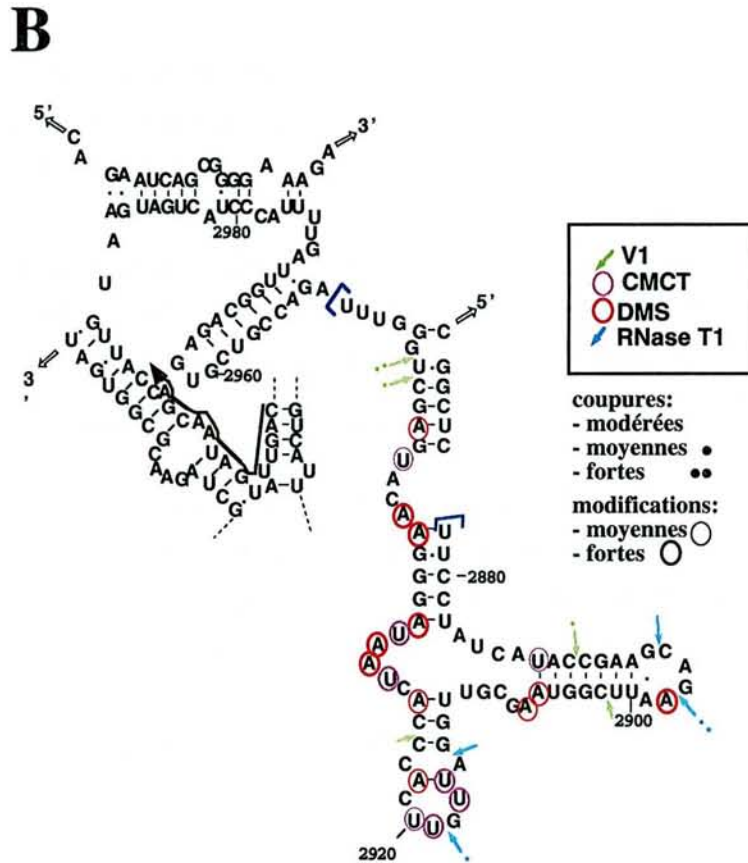
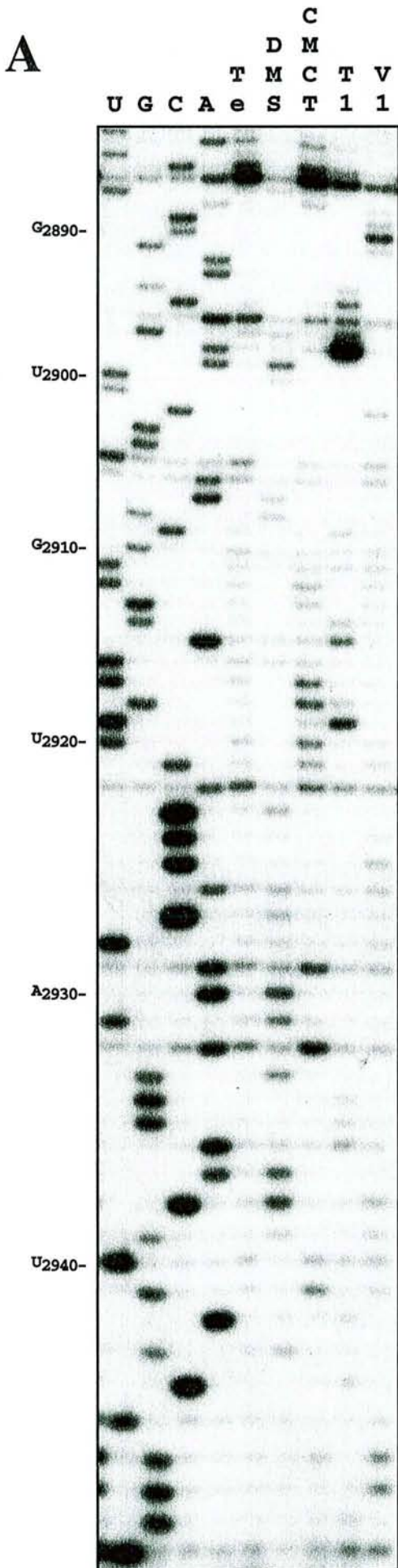


Figure 53: Etude de la structure secondaire de la région 3'-terminale, du transcrit produit *in vitro*, correspondant au domaine V de l'ARNr 26S de *S. cerevisiae*.

Panneau A: Analyse par la transcriptase inverse des modifications chimiques (colonnes DMS et CMCT) et des coupures par les RNases T1 et V1 (colonnes T1 et V1, respectivement) portées par la région 3'-terminale du transcrit correspondant au domaine V de l'ARNr 26S de *S. cerevisiae*. Un transcrit non modifié a servi de témoin pour détecter les pauses naturelles de la transcriptase inverse (colonne Te). Les ADNc obtenus ont été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 7%, en parallèle avec un séquençage par la transcriptase inverse (colonnes U, G, C et A). L'oligonucléotide, dont le site d'appariement sur le transcrit est indiqué dans le Panneau B, a servi d'amorce. La position des nucléotides dans la séquence de l'ARNr 26S de *S. cerevisiae* est indiquée à gauche de la fluorographie.

Panneau B: Représentation schématique des modifications chimiques par le DMS et le CMCT et des coupures par les RNases T1 et V1 portées par le transcrit analysé. La structure du fragment du domaine V de l'ARNr 26S de *S. cerevisiae* représentée est celle décrite par Veldman *et al.* (1981). Les différentes modifications et coupures sont montrées, respectivement, par les cercles et les flèches, dont le code couleur est indiqué. La région analysée est située entre les crochets.

in vivo et *in vitro* (Kohl et Hall, 1974; pour revue, Florentz et Giegé, 1995), ainsi que par l'histidinyl-ARNt synthétase et la valyl-ARNt synthétase *in vitro* (Felden *et al.*, 1998). De manière inattendue, la conversion du résidu U en résidu Ψ, observée dans le domaine 3'-terminal de l'ARN du virus BMV, n'a pas lieu dans la région "ARNt-like", mais dans une extension. De manière intéressante, la structure formée par cette extension comporte deux hélices reliées par un court lien flexible et la modification a lieu dans la dernière paire de bases d'une des hélices. Nous retrouvons l'idée que les modifications ont lieu dans un motif comprenant deux hélices et, plus précisément, au niveau de la charnière entre ces deux hélices.

Pour aller plus loin dans la définition du motif reconnu, des mini-substrats ARNt et BMV ont alors été produits par l'équipe d'H. Grosjean. Les résultats montrent la nécessité d'une flexibilité de la région charnière. L'ensemble de cette étude nous a conduit à proposer un motif structural préférentiellement reconnu par l'enzyme Pus1p de *S. cerevisiae*.

L'ensemble de ces données est présenté dans un article soumis pour publication à RNA.

Publication n°4:

Recognition of RNA substrates by yeast pseudouridine-synthase I (Pus1p) *in vitro*.

Motorin Y., Massenet S., Becker H.F., Raué A., Branlant C. and Grosjean H.

Soumis à RNA

Publication n°4

**Recognition of RNA substrates by yeast pseudouridine
synthase I (Pus1p) *in vitro*.**

Y. Motorin, S. Massenet, H. F. Becker, A. Raué, C. Branlant and H. Grosjean.

Soumis à RNA

Recognition of RNA substrates by yeast pseudouridine synthase 1 (Pus1p)
in vitro.

Y. Motorin^{1*+}, S. Massenet²⁺, H. Becker¹, H.A. Raué³, C. Branlant², H. Grosjean^{1*}

¹ CNRS, Laboratoire d'Enzymologie et Biochimie Structurales,
1 avenue de la Terrasse, F-91198 Gif-sur-Yvette, France

² Laboratoire de Maturation des ARN et Enzymologie moléculaire, UMR7567 CNRS-UHP
Nancy I, Faculté des Sciences, BP 239, 54506 Vandoeuvre-les-Nancy Cedex, France

³ Department of Biochemistry and Molecular Biology, IBMW, Biocentrum Amsterdam, Vrije
Universiteit, de Boelelaan 1083, 1081 HV Amsterdam, The Netherlands

+ The two first authors contributed equally

*Authors for correspondence:

Y.Motorin and H.Grosjean

CNRS, Laboratoire d'Enzymologie et Biochimie Structurales,
Avenue de la Terrasse, Bat. 34,
F-91198 Gif-sur-Yvette, FRANCE

Tel. (33)-1-69.82.34.98

FAX (33)-1-69.82.31.29

e-mail Yuri.Motorin@maem.uhp-nancy.fr

Abbreviations: Ψ, pseudouridine; tRNA, transfer RNA; 3D, three dimensional; CMCT, 1-cyclohexyl-3-(2-morpholinoethyl)-carbodiimide metho-*p*-toluene-sulfonat; DMS, dimethylsulfate; BMV, Brome Mosaic Virus; TYMV, Turnip Yellow Mosaic Virus.

ABSTRACT

Yeast RNA:pseudouridine synthase Pus1p is a multisite-specific modification enzyme responsible for the formation of pseudouridine residues at a number of specific positions in tRNA and U2 snRNA. We found that recombinant Pus1p is also able to modify non-natural substrates *in vitro* and used this possibility to refine the knowledge on yeast Pus1p specificity. We first studied the *in vitro* action of Pus1p on synthetic domains II, IV and V of *S. cerevisiae* 26S rRNA. The extent of U to Ψ conversion was established by thin-layer chromatography and modified positions were identified by reverse transcription-based chemical mapping. Surprisingly, three out of eight target sites identified in domains IV and V corresponded to natural sites of pseudouridylation in 26S rRNA. However, the disruption of *PUS1* gene did not affect the modification of these sites *in vivo*. It is noteworthy that all detected modifications took place at U residues 3'-adjacent to a purine and that Purine-U dinucleotide target is located at the extremity or the border of a helix. For further definition of the recognition motif, the action of yeast Pus1p was tested on the aminoacylable 3'-terminal region (206 nt) of Brome Mosaic Virus (BMV) RNA. A unique target site was detected outside the tRNA-like structure in a two-helix-structure motif. Then, using minisubstrates, we confirmed that RNA consisting of two stem-loop structures joined by a flexible short linker are efficiently modified by yeast Pus1p *in vitro*, provided that a Purine-U dinucleotide is located at the junction of one of the helix.

Keywords: tRNA, rRNA, pseudouridine synthase, maturation, yeast, *Saccharomyces cerevisiae*

INTRODUCTION

The modified nucleotide pseudouridine (ribo-furanosyl- β -uracyl, Ψ) is found in most of the metabolically stable RNAs (namely tRNAs, rRNAs, UsnRNAs and snoRNAs). The formation of Ψ residues is a posttranscriptional process and is catalyzed by specific RNA: Ψ -synthases. These proteins constitute a large family of enzymes sharing some common sequence elements but differing in their substrate specificities. All presently identified RNA: Ψ -synthases catalyze the intramolecular isomerization of U to Ψ residues without requirement for energy source or additional cofactor. One important question is to know how many enzymes (RNA:pseudouridine synthases) are required for formation of Ψ residues in RNAs.

Recently, several tRNA: Ψ -synthases from the yeast *Saccharomyces cerevisiae* have been cloned, expressed in *E. coli* and their recognition properties towards tRNA substrates have been identified (Simos et al., 1996; Becker et al., 1997a; 1997b; Lecointe et al., 1998; Motorin et al., 1998). While all these enzymes use tRNA as substrate, their requirements at the level of RNA 2D and 3D structures appeared to be quite variable. Some enzymes catalyze U to Ψ conversion only at a unique position in tRNAs, while others are capable of acting at several distinct sites. For example, the tRNA: Ψ -synthase Pus3p (previously called Deg1, encoded by ORF YFL001w) catalyzes Ψ formation in the anticodon loop of tRNAs and proximal anticodon stem (positions 38/39), and tRNA recognition by Pus3p depends on the correct tRNA 3D structure since the disruption of tertiary base-pair interactions considerably decreases the modification efficiency (Lecointe et al., 1998). In contrast, the enzyme that is responsible for the site-specific formation of the universally conserved Ψ residue at position 55 (Pus4p, ORF YNL292w) (Becker et al., 1997a) does not require the tRNA 3D structure for recognition and is capable of acting on simple stem-loop minisubstrates (Becker et al., 1997b). The third known enzyme (Pus1p, ORF YPL212c) has the peculiarity to display an unexpected multisite- and dual-substrate specificity. Indeed, yeast Pus1p is implicated in the formation of Ψ residues located at 8 different positions in

tRNAs (Motorin et al., 1998) and in the formation of Ψ_{44} in snRNA U2 (Massenet et al., 1999). It was also shown that Pus1p can recognize and modify *in vitro* the "tRNA-like" 3'-end of Turnip Yellow Mosaic Virus (TYMV) RNA (Becker et al., 1998). Taken together these results suggest that yeast Pus1p behaves as a more general RNA-modification enzyme than initially anticipated (Simos et al., 1996) and this point deserves further investigation.

In addition to tRNA-specific Ψ -synthases, several rRNA-specific Ψ -synthases have been recently identified. It is noteworthy that *Bacteria* and *Eukarya* use quite different strategies for defining the sites of Ψ formation in rRNAs. In contrast to eukaryotes where formation of the numerous Ψ residues in rRNAs is snoRNA-guided (see below), only a few Ψ residues are present in bacterial rRNAs (Branlant et al., 1981; Ofengand & Bakin, 1997; Ofengand & Fournier, 1998) and they are formed by site- (or multisite-) specific enzymes. Several such enzymes have been recently purified from *E. coli* and *B. subtilis* and characterized (Wrzesinski et al., 1995a; 1995b; Conrad et al., 1998; Huang et al., 1998; Raychaudhuri et al., 1998; 1999; Conrad et al., 1999; Niu & Ofengand, 1999). Among them, RluA catalyzes the site-specific modification of U₇₄₆ in 23S rRNA and is also capable of modifying U₃₂ in tRNA (Wrzesinski et al., 1995b; Raychaudhuri et al., 1999). Two multisite-specific *E. coli* enzymes (RluC and RluD) account for the formation of 6 Ψ residues in 23S rRNA (Conrad et al., 1998; Huang et al., 1998; Raychaudhuri et al., 1998). In all cases the disruption of the corresponding genes leads to the disappearance of Ψ residues at the expected positions in rRNA. Unexpectedly, when tested *in vitro*, the recombinant enzymes RluC and RluD were shown to modify rRNA transcripts at several positions where Ψ residues are not found in authentic rRNAs (Huang et al., 1998). Thus, *in vitro* these two enzymes possess a much more relaxed substrate specificity than *in vivo*, indicating that factor(s) other than the intrinsic properties of enzymes and RNA substrates contribute to the definition of RNA modification sites *in vivo*. Some other Ψ -synthases may also have such relaxed specificity *in vitro*.

Despite the larger number of Ψ residues in the rRNA of cytoplasmic ribosomes from *Eukarya* as compared to *Bacteria* (Veldman et al., 1981; Ofengand & Fournier, 1998), the number of enzymes, required for synthesis of these Ψ residues seems to be rather limited. A unique enzyme (Cbf5p in yeast, or its homologs in higher eukaryotes) catalyzes U to Ψ conversion at numerous sites in the nucleolar rRNA precursors, through its interaction with the H/ACA class of snoRNAs (Lafontaine et al., 1998; Watkins et al., 1998). So far, specific snoRNA guides have been found for at least 26 out of a total of 44 pseudouridylation sites in yeast rRNAs (Ofengand & Bakin, 1997; Samarsky & Fournier, 1999). In a large number of cases the direct involvement of the snoRNA in modification at a given site was proven experimentally either by the abrogation of Ψ formation upon deletion of the snoRNA gene (Ganot et al., 1997) or by the selection of a different site of modification upon changing the guide sequence complementary to the rRNA (Ni et al., 1997). However, at the present state of knowledge, one cannot formally exclude the possibility that a few U to Ψ conversions in the cytoplasmic rRNAs of eukaryotes are catalyzed by another enzymatic system, for instance, an enzyme with a broad substrate specificity like Pus1p.

Altogether this shows the high complexity of the mechanisms that determine which U residues in RNAs are converted into Ψ residues *in vivo*. Progress in the field requires a better understanding of both the intrinsic properties of the RNA: Ψ -synthases and of the additional factors involved in the selection of the modification site in the living cell. Based on the dual-substrate specificity of the *S. cerevisiae* RNA: Ψ -synthase Pus1p (Motorin et al., 1998; Massenet et al., 1999) and on its capability of modifying minisubstrates derived from minor tRNA^{Ile} precursor (Szweykowska-Kulinska et al., 1994), a study of this enzyme was expected to bring interesting information in the field. In preliminary assays, we found that in addition to tRNA and snRNA U2, the recombinant *S. cerevisiae* Pus1p enzyme also acts on other RNA substrates *in vitro*. Hence, we decided to use this peculiarity to get a more precise definition of its substrate specificity. Various RNA substrates with defined or experimentally identified secondary structures were produced *in vitro*, namely *S. cerevisiae*

26S rRNA domains, Brome Mosaic Virus (BMV) RNA 3'-terminal domain and truncated tRNAs. They were subjected to the action of the recombinant Pus1p *in vitro*, and the modified positions were mapped by the CMCT/RT method. Comparison of the structural environment of the *in vitro* modified nucleotides allowed us to define the major structural requirements for efficient recognition of an RNA substrate by Pus1p. The results demonstrate a broad substrate specificity of Pus1p *in vitro* and imply the existence of additional factors limiting its action *in vivo*.

Materials and Methods

Chemicals, enzymes and materials

[α -³²P]-radiolabeled nucleotide triphosphates (400 Ci/mmol) were from Amersham (UK), *Penicillium citrinum* nuclease P1 and *Aspergillus oryzae* RNase T2 from Sigma (USA), RNase V1 from Pharmacia (Sweden), RNase T1 from Boehringer (Germany), 1-cyclohexyl-3-(2-morpholinoethyl)-carbodiimide metho-*p*-toluene-sulfonat (CMCT) and dimethylsulfate (DMS) from Aldrich (USA), bacteriophage T7 RNA polymerase and restriction enzymes from MBI Fermentas (Vilnius, Lithuania). RNasin and AMV reverse transcriptase were from Promega (USA). Synthetic oligodeoxynucleotides used as primers for reverse transcription were purchased from MWG-Biotech (Germany) and used without further purification. Thin layer cellulose plates were from Schleicher & Schuell (Germany) and all other chemicals from Merck Biochemicals (Germany).

N-terminally His₆-tagged recombinant *S. cerevisiae* RNA:pseudouridine synthase (Pus1p) was purified as described previously (Motorin et al., 1998).

DNA template preparation and in vitro transcription

All the RNA coding sequences used in this study were cloned in plasmid pGEM3 or pGEM4 downstream from the T7 RNA polymerase promoter. Plasmids containing different structural domains of yeast 26S rRNA were constructed by subcloning the appropriate restriction fragments of the *S. cerevisiae* rDNA unit present in plasmid pJV12 (Venema et al., 1995) (domain II (nt 650-1453): *Ava*II-*Ssp*I cloned blunt into the *Sma*I site of pGEM3; domain IV (nt 1873-2391): *Xba*I-*Pvu*II cloned between the *Xba*I and *Sma*I sites of pGEM4; domain V (nt 2361-3243): *Sty*I-*Sty*I cloned blunt into the *Sma*I site of pGEM4). For domains II and IV, transcription with T7 RNA polymerase was achieved after linearization of the recombinant plasmids with the *Hind*III or *Eco*RI, respectively. For domain V, a subregion (nt 2362-3023) was produced by PCR amplification (30 cycles 1 min at 94°C, 1 min at 55°C and 2 min at 74°C) using a 3'-primer (CGTACTAAGTTCAATTACTATTGC)

complementary to nt 2999-3023 and a primer corresponding to T7 promoter sequence of the vector plasmid. After verification of its size by agarose electrophoresis, the PCR product was used as template for T7 RNA transcription without further purification.

The 3'-terminal region (206 nt) of the Brome Mosaic Virus (BMV) RNA was transcribed from plasmid BMV RNA3/4 linearized by *Tth1111* (kindly provided by Dr. T.Dreher, Oregon University, USA) (Dreher & Hall, 1988; Felden et al., 1994).

DNA templates for tRNA minisubstrates (mini-#1, #2 and #3) were prepared by template-independent PCR amplification of two partially overlapping oligonucleotides (8 bases overlap at 3'-ends). PCR amplification was performed in 30 cycles (1 min at 94°C, 1 min at 50°C and 1 min at 74°C) using 200 pmol of each of the two oligonucleotide primers. DNA for the minisubstrate #4 was prepared by PCR amplification using primers specific for the sequence of BMV RNA. In all cases, the 5'-primer contained the sequence of the T7 RNA polymerase promoter. The size of the PCR products was verified by electrophoresis in 2% agarose. They were used as templates for T7 RNA polymerase transcription after extraction with the PCR-Clean extraction kit (Qiagen, Germany).

In vitro T7 RNA-polymerase transcription using [α -³²P]-radiolabeled nucleotide triphosphates and purification of the resulting T7-runoff RNA transcripts by denaturing gels electrophoresis were performed as described (Jiang et al., 1997). Unlabeled RNA transcripts were prepared using the standard approach (for details, see Massenot et al., 1999).

Formation and quantification of pseudouridine in radiolabeled T7 RNA transcripts

The activity of purified yeast recombinant Pus1p (Motorin et al., 1998) was tested at 30°C. The standard incubation mixture (50 μ l) contained 100 mM Tris-HCl, pH 8.0, 100 mM ammonium acetate, 5 mM MgCl₂, 2 mM DTT, 0.1 mM EDTA, 1-2 fmol of appropriately [³²P]-radiolabeled T7-runoff transcripts as substrate and 20-300 ng (0.32 - 4.76 pmol) of purified Pus1p. After incubation, the pseudouridine content in the radiolabeled transcripts was analyzed using nuclease P1 or RNase T2 digestion as

described previously (Motorin et al., 1998). The standard deviation of the method was found to be about of +/- 0.1 mol Ψ /mol for tRNA and minisubstrates and of 0.3-0.5 mol/mol in the case of long rRNA transcripts.

Modification of cold RNA transcripts in vitro

Preparative modification of rRNA transcripts *in vitro* by recombinant Pus1p was performed under conditions identical to those used above for radiolabeled transcripts, except that 2 μ g of RNA transcripts and 1 or 2 μ g of recombinant Pus1p were used per 10 μ l of final incubation mixture. Before modification rRNA transcripts were incubated for 3 min at 80°C and cooling down to 37°C for renaturation. Modification of rRNA transcripts by Pus1p was also performed at three different MgCl₂ concentrations (5, 10 and 30 mM). The incubation was done for 2 h at 37°C and the reaction was stopped by phenol-extraction and ethanol precipitation. Modification of the *in vitro* produced BMV transcript was carried out for 2 hours at 37°C in a standard incubation buffer.

Localization of Ψ residues by CMCT mapping

Mapping of the pseudouridine residues introduced into RNA transcripts by Pus1p was performed by the CMCT/RT technique as described previously (Bakin & Ofengand, 1993), with the modifications described in Massenet et al. (1999). Oligonucleotides, complementary to the following regions of 26S rRNA, were used as primers for reverse transcription: nt 1143 to 1163 (domain II), nt 2368 to 2392 (domain IV) and nt 2999-3023 (domain V). Reverse transcription on the BMV RNA fragment was done with an oligonucleotide primer complementary to the 25 3'-terminal nucleotides of the molecule (TGGTCTCTTTTAGAGATTTACAGTG).

Structural probing of RNA transcripts

Chemical modification and enzymatic digestion were carried out in the Pus1p standard reaction buffer (see above), on ~2 μ g of *in vitro* synthesized rRNA fragments

(domains IV and subregion of domain V) in a final volume of 10 μ l. Prior to the modification, the transcripts were heated to 80°C and slowly cooled down to 37°C. Chemical modifications were done using 1 or 2 μ l of a dimethylsulfate (DMS) solution in ethanol (1:8 vol:vol) or with 0.84 or 1.68 mg of CMCT per assay. The incubations were done for 10 min at 37°C and the reactions were stopped by ethanol precipitation, with 2.5 M sodium acetate and 2 M β -mercaptoethanol for DMS modification and with 3 M sodium acetate for the CMCT reaction. Modified RNAs were dissolved in 5 μ l of H₂O, 1 μ l of the solution was used for reverse transcriptase elongation assay (Massenet et al., 1999).

Enzymatic probing was performed at 37°C for 5 min with 0.01 or 0.02 units of RNase V1 or with 0.15 or 0.25 units of RNase T1. The reactions were stopped by addition of 1 μ l of 100 mM EDTA, followed by phenol extraction. After ethanol precipitation, the same procedure as for chemical modifications was used.

Preparation of RNA from a PUS1-disrupted S. cerevisiae strain

Total ribosomal RNAs from a wild-type and a *PUS1*-disrupted yeast strain (Simos et al., 1996) were prepared using the method described in Massenet et al. (1999) with the following modifications. The ribosomal pellet obtained after high-speed centrifugation was resuspended in lysis buffer and further extracted by equal volumes of phenol, phenol/chloroform (1:1) and chloroform/isoamyl alcohol (24:1). The ethanol-precipitated rRNAs were dissolved in water and used for CMCT-treatment and reverse transcription.

RESULTS

General strategy

The preliminary results suggested that Pus1p is capable of acting on a wide variety of RNA substrates *in vitro*. We hypothesized that the identification of *in vitro* targets for Pus1p modification in such molecules with known secondary structure, should help in understanding of Pus1p substrate specificity. We selected substrates corresponding to well-defined structural domains of the *S. cerevisiae* 26S rRNA, the Brome Mosaic Virus (BMV) RNA and tRNAs. As no experimental analysis of the *S. cerevisiae* 26S rRNA secondary structure was previously performed, the folding of the *in vitro* produced rRNA domains was investigated by chemical and enzymatic probes.

Recombinant yeast Pus1p catalyzes the pseudouridylation of rRNA transcripts *in vitro*

We tested the activity of recombinant Pus1p on structural domains of the *S. cerevisiae* 26S rRNA where Ψ residues are naturally encountered, namely domains II, IV and V (Veldman et al., 1981; Ofengand & Fournier, 1998). Whole domains (II and IV) or portion of domain (V) were produced *in vitro* by T7 transcription, and the design of the transcripts was made in order to preserve the authentic secondary structure of these rRNA regions. The maintenance of the correct secondary structure was verified experimentally for the portions of domain IV and V, where modified positions were identified (see below).

As the selected rRNA domains had very large sizes (domains II, 803 nt, IV, 446 nt and subregion of domain V, 707 nt), [α - ^{32}P]UTP-labeled RNA transcripts were first tested as substrates for pseudouridylation by Pus1p *in vitro*. After incubation with the enzyme, the transcripts were digested with either nuclease P1 or RNase T2 and the resulting nucleotides were separated by 2D-thin-layer chromatography. The results of 2D-chromatographic separation of modified 5'-mononucleotides synthesized in rRNA

transcripts after 60 min are shown in Figure 1, and the corresponding time courses are illustrated in Figure 2 (A, B).

Quantification of the Ψ and U spots on the chromatograms of the nuclease P1 digests allowed calculation of the total molar fraction of U residues converted into Ψ residues upon incubation. For domains II and IV this number was of about 8 Ψ mol/mol of the transcript, while for rRNA domain V subregion, it was higher, about 12 mol of Ψ /mol of transcript (Fig. 2A). This number is even higher than the 8 Ψ residues that occur naturally in *S. cerevisiae* 26S rRNA domain V. The analysis of RNase T2 digests (Fig. 2B) shows that about 30 % of the formed Ψ residues are 5'-adjacent to uridines. This number suggests that the identity of the residue downstream from the uridine is not important for recognition by Pus1p, since on a statistical basis, one can expect that about 25% of the Ψ residues would be 5'-adjacent to a U residue. Similar results have been obtained using ATP-labelling followed by T2 digestion (data not shown).

Chemical mapping of Ψ residues in rRNA transcripts modified *in vitro*

As the studied rRNA domains were very large, we restrained our mapping of Ψ residues formed by the recombinant Pus1p to the regions of the rRNA domains IV and V which contain Ψ residues in naturally occurring 26S rRNA (nt 2220-23260 in domain IV and nt 2849-2990 in domain V). To this end, the CMCT/RT approach was used. The modification and mapping experiments were performed several times and reproducible results were obtained. Figure 3 shows representative examples of CMCT/RT analyses of domain IV and subregion of domain V after their incubation with the Pus1p enzyme. CMCT/RT analysis of the same rRNA domains incubated in the absence of Pus1p was always performed as a control. Two examples of such controls are given in Figure 3.

Even though the CMCT/RT approach is not suitable for the precise quantification of modification extent at a given site, the results of primer extension analysis illustrated in Figure 3 show that the efficiency of Ψ residue formation at different locations varied considerably. Strong signals indicating efficient modification were reproducibly observed

at positions 2335, 2911 and 2940. In contrast, several other sites were modified to a much lesser extent as attested by the rather weak primer extension arrests.

This observed variable efficiency of modification at different sites is in good agreement with the rather high number of Ψ residues detected by P1 and T2 digestions and with the "tailing" behaviour of the kinetic curves of modification (see Fig. 2). Apparently, only a few uridines in each RNA domain act as 'strong' modification sites that are rapidly converted to Ψ residues with a high yield, while many others constitute 'weak' sites that are modified only partially. Furthermore, it may be that several uridines modified to a low extent were not detected by the CMCT/RT mapping. One possible explanation for the presence of weak modifications sites was a structural heterogeneity of the RNA substrates under experimental conditions. To address this question, the 2D structure of the portions of domains IV and V containing Ψ residues was studied experimentally. Figure 4 shows an example of chemical and enzymatic probing of domain IV. Although one cannot exclude some degree of structural heterogeneity which cannot be detected by chemical and enzymatic probing, the results obtained suggest that the two rRNA domains adopted a well defined 2D-structure, in the conditions used for the incubation with Pus1p (100 mM Tris-HCl, pH 8.0, 100 mM ammonium acetate, 5 mM MgCl₂, 2 mM DTT, 0.1 mM EDTA). Interestingly, whereas the increasing of MgCl₂ concentration to 30 mM had no marked effect on the RNA 2D-structure (not shown), such high MgCl₂ concentration considerably decreased the extend of modification by recombinant Pus1p. No formation of Ψ residues was detected in these conditions by CMCT/RT mapping (Fig. 3C).

U to Ψ conversion in rRNA domains occurs 3'-adjacent to purine residues in the vicinity of helix extremities

As mentioned above, the secondary structure of *S. cerevisiae* 26S rRNA had not yet been studied experimentally. Long ago, we proposed a secondary structure model based on a comparative analysis of potential secondary structures in large rRNAs (Veldman et

al., 1981). This model has been largely validated by the increasing knowledge on rRNA nucleotide sequences. Only a few improvements were proposed later (Gutell et al., 1994; Schnare et al., 1996). The region of domain IV analyzed for formation of Ψ residues by Pus1p was expected to contain the helices numbered 61, 64, 67, 69, 70 and 71 according to the nomenclature of Leffers et al. (1987) (Fig. 4). Our experimental analysis of the transcript confirmed the presence of most of these helices (61, 64, 67, 69 and 71). However, instead of helix 70, an extended helix 71 was found. Indeed, the strong T1 RNase cleavages 3' to G residues 2271 and 2272 is a strong argument against the formation of helix 70. The extension of helix 71 involves residues 2280 to 2283 and 2306 to 2309 (Fig. 4). Such structure with an extended helix 71 and no helix 70 may be stabilized by formation of pseudoknot structures. One possibility is a base-pair interaction between nucleotides 2256-2258 in terminal loop of helix structure 69 and residues 2273-2275. No modification by DMS and no cleavage by RNase T1 were detected in these two segments. Only CMCT modification of U residues was observed. Other possibility of independent formation of pseudoknot structures by the terminal loop of stem-loop structure 69 and the segment 2274-2278 with other segments of domain IV cannot be excluded. A complete analysis of the entire domain IV produced *in vitro* would be required to answer this question. Concerning the studied portion of domain V, the presence of all the expected helices (90, 91 and 92) (Fig. 5) was confirmed experimentally (not shown).

Inspection of the RNA sequence around the detected pseudouridine in domains IV and V, shows that all the modified U residues have a 5'-adjacent purine residue. In agreement with results of P1 digestion, the identity of the downstream residue is variable. Our experimental study of the secondary structure of the *in vitro* produced domains IV and V, allowed a clear proposal for the secondary structure of 7 of the 8 identified target sites. Only the nature of the secondary structure of the segment containing the detected Ψ_{2259} was not determined. The inspection of the location in the 2D structure of the 7 other detected Ψ residues (Fig. 5), shows that Pus1p was capable of modifying residues located

in single-stranded or double-stranded segments. However, the modification sites are never located right in the center of a helical.

Analysis of naturally modified 26S rRNA in *PUS1*-disrupted strains.

Three out of the 8 Ψ residues detected after incubation of recombinant Pus1p with 26S rRNA domains are also found in naturally modified 26S ribosomal RNA (positions 2259 and 2348 in domain IV and position 2940 in domain V). If Pus1p modifies its RNA substrates on a statistical basis, one could not expect such a high proportion of sites, where Ψ is naturally found in rRNA. The isomerization of U₂₉₄₀ in *S. cerevisiae* is known to be guided by the snR37 snoRNA and disruption of the *SNR37* gene abolishes modification at this site *in vivo* (Ni et al., 1997). However, no snoRNA guides have as yet been reported for positions 2259 and 2348 (Ofengand & Fournier, 1998; Samarsky & Fournier, 1999), see also http://www.bio.umass.edu/biochem/rna-sequence/Yeast_snoRNA_Database/snoRNA_DataBase.html). Hence, in spite of the idea that most Ψ residues in rRNAs are likely to be formed by the snoRNA-guided mechanism, it was important to control whether Pus1p is involved in the formation of these two naturally occurring Ψ residues.

To that end, 26S rRNAs, extracted from a yeast strain that carries a disruption of the *PUS1* gene (Simos et al., 1996) and the parent wild-type strain, were subjected to CMCT modification. The regions encompassing nt 2180-2320 in domain IV and nt 2850-2990 in domain V were analyzed by primer extension. No differences were found in the pattern of pseudouridylation between the two samples (data not shown). These results demonstrate that, despite its ability to modify positions 2259, 2348 and 2940 *in vitro*, Pus1p is not likely to be involved in modification of the 26S rRNA at these positions *in vivo*.

Formation of Ψ residue in a T7 transcript corresponding to the 3'-terminal region of Brome Mosaic Virus RNA

Taken together, the results obtained earlier with various yeast tRNAs (Motorin et al., 1998), the tRNA-like fragment of Turnip Yellow Mosaic Virus (TYMV) RNA (Becker et al., 1998) and the present results on Ψ formation in yeast rRNA transcripts suggested that

substrate recognition by Pus1p is based upon relatively simple conformational features, rather than sequence specificity. To obtain further evidence for this idea we explored the modification of another natural RNA with known three-dimensional structure, namely the 3'-terminal fragment of BMV RNA. Similarly to TYMV RNA, the 3'-end of BMV RNA folds into a tRNA-like structure (Felden et al., 1994). However, due to the presence of a large insertion in the region corresponding to the tRNA T Ψ -loop, the overall folding of the BMV fragment is more complex (*cf.* Fig. 6A).

When an [α -³²P]UTP-labeled T7 transcript corresponding to the 3'-terminal 206 nt of BMV RNA was incubated with Pus1p *in vitro* we found 1 mol of Ψ /mol of transcript to be present after 20 min (Fig. 6B). CMCT/RT analysis (Fig. 6C) identified a single site of modification at position 40, which, to our surprise, was located not in the 'tRNA-like' part of BMV, but between the two coaxially stacked helices (B1 and B2) that form part of the insertion into the T Ψ -like loop (Felden et al., 1994). Interestingly, this structural environment closely resembles that of position 27, the most frequently modified U residue in tRNAs (Motorin et al., 1998). It also resembles the structural environment of the more efficiently modified position in rRNA domain IV (position 2335) and of one of the less efficiently modified position in domain V (2348) (Fig. 5).

Modification of RNA minisubstrates by Pus1p *in vitro*

The results presented above suggested that the 'minimal' set of structural features that is efficiently recognized by Pus1p might encompass two short helical stems connected by a flexible linker, the converted U residue being preceded by a purine residue and located either at the extremity of one stem or at its border in the flexible link. To verify if Pus1p *in vitro* can modify such minimal molecules, we prepared and tested several minisubstrates, whose structures are shown in Fig. 7. Three minisubstrate molecules (#1 - #3) were derived from the sequences of naturally modified tRNAs, while the minisubstrate #4 was based on BMV RNA.

The first minisubstrate (#1) was derived from yeast tRNA^{Ile}. The D-loop was deleted and the region encompassing the TΨ-stem/loop (C₄₄) and acceptor stem (U₈) was replaced by a CG dinucleotide. Design of minisubstrate #2 was based upon the observation that introduction of a tertiary G₂₆...A₄₄ base-pair into tRNA^{Ile} improves its binding to Pus1p (Arluisson et al., 1999) and that the sites, modified *in vitro* are 3'-adjacent to purine. Additional mutations were introduced into the bulge connecting the two helices (*cf.* Fig. 7) to improve minisubstrate stability and to avoid the possibility of alternatively folded structures. Design of minisubstrate #3 was based on the sequence of yeast tRNA^{Val} but had an 'open' structure that should allow more freedom of movement of the two RNA stems. Finally, minisubstrate #4 corresponds to the portion of BMV RNA (encircled in Fig. 6) that encompasses stems B1 and B2. Helix B1 was closed by a GUAA tetraloop. Note that this molecule also has an 'open' structure similar to that of minisubstrate #3.

Each [α -³²P]UMP labeled minisubstrate was incubated with increasing amounts of recombinant Pus1p for 90 min at 30°C and the extent of modification was assessed by analysis of the nuclease P1 digest. As shown in Figure 8, no modification of minisubstrate #1 could be detected, but minisubstrates #2, #3 and #4 were modified, albeit at different levels. Minisubstrate #3 was efficiently modified even at the lowest Pus1p concentration used. In contrast, the minisubstrate #2 was rather inefficiently modified even at high concentration of Pus1p. The last minisubstrate (#4) was modified to intermediate level.

DISCUSSION

Substrate specificity of yeast Pus1p *in vivo* and *in vitro*

Of the three RNA:pseudouridine synthases so far identified in yeast, Pus1p shows by far the broadest substrate specificity *in vivo*. Previous experiments have shown this enzyme to be involved in Ψ formation at least eight different sites in various tRNAs and tRNA precursors (Motorin et al., 1998) and in the modification of U₄₄ in U2 snRNA (Massenet et al., 1999). The data presented in this paper demonstrate an even larger capability of modification of Pus1p *in vitro*. The recombinant enzyme was able to modify numerous positions in 26S rRNA domains as well as a single U residue in the 3'-terminal region of BMV RNA. This modified position is located outside the tRNA-like structure.

Interestingly, 3 out of 8 Ψ residues introduced into the T7 rRNA transcripts *in vitro* by Pus1p occur at positions that are naturally modified (Fig. 5). However, we showed that the extent of pseudouridylation at these sites *in vivo* was not affected by deletion of *PUS1* gene. This reinforced the idea that Ψ formation in rRNAs synthesized in the nucleolus is essentially based on the snoRNA-guided mechanism. Hence, our data shows that despite being not involved in rRNA biogenesis, Pus1p recognizes numerous sites in rRNA *in vitro*.

As mentioned in the introduction, several examples of rRNA modification at 'cryptic' sites have also been reported for prokaryotes. While *in vivo* the *E. coli* rRNA:pseudouridine synthases RluC and RluD each convert only 3 uridines in *E. coli* 23S rRNA, and none in 16S rRNA, *in vitro* the recombinant enzymes hypermodify both types of rRNA, to the extent where ~5% of the U residues are converted (Huang et al., 1998). The rRNA:pseudouridine synthase YpuI (RluB) from *B. subtilis* also modifies 'cryptic' sites in both 16S and 23S rRNA *in vitro* (Niu & Ofengand, 1999). Finally, the full length 16S rRNA as well as its domains serve as substrates for the site-specific *E. coli* tRNA:m⁵U(T)₅₄₇-methyltransferase *in vitro* although the naturally modified 16S rRNA in *E. coli* does not contain m⁵U (Gu et al., 1994).

All these examples show that an isolated modification enzyme *in vitro* may act upon a substantially wider range of RNA substrates than in its natural cellular environment. Caution should, therefore, be taken in extrapolating the results obtained *in vitro* to the *in vivo* function of such enzymes. Restriction of the modifying activity to only the natural substrate(s) *in vivo* may be due to different cellular compartmentalization of the enzyme and its non-natural RNA substrate (e.g. Pus1p probably resides in the nucleoplasm, while pre-rRNA intermediates and mature rRNA are present in the nucleolus). Another likely explanation is the masking of 'cryptic' sites by interactions with (ribosomal) proteins or steric hindrance (e.g. for prokaryotic rRNA). Based on the present results it clearly appears that the *in vivo* specificity of Pus1p results both from its intrinsic properties and from several other cellular determinants.

Recognition of RNA by Pus1p

Previous comparative analysis of the known natural RNA substrates of Pus1p failed to reveal any common recognition sequence (Motorin et al., 1998; Massenet et al., 1999), suggesting that substrate recognition by Pus1p relies on features of spatial, rather than primary structure. Efficient modification of minisubstrate RNA molecules composed of an anticodon stem-loop prolonged with its natural intron showed that the overall 3D structure of tRNA is dispensable for recognition by Pus1p (Szweykowska-Kulinska et al., 1994; Motorin et al., 1998). Nevertheless, binding of Pus1p is improved in the presence of the correct tertiary interactions in tRNA (Arluison et al., 1999). Furthermore, almost any tRNA was found to bind to Pus1p with a K_D ranging from 15 to 150 nM independently of the presence of a target uridine residue. Finally, the presence of a non-Watson-Crick G...A base pair 5'-adjacent to U₂₇ decreased the binding constant (Arluison et al., 1999).

The present extensive analysis of Pus1p action on 26S rRNA domains shows that, in this non-natural substrate, Pus1p modifies only those U residues which are 3' adjacent to purine and are located at the border of helix extremity. Modification occurred either in the helical part or in the single-stranded part. Such specificity is in a good agreement with the

location of naturally modified positions in tRNAs. None of the modified positions are found in terminal loops, they are all located at a junction between two helical regions and preferentially preceded by purine (Motorin et al., 1998). However, this preference for uridine residues 3'-adjacent to purines seems to be less strict in the case of tRNAs. Indeed, based on the inspection of yeast *S. cerevisiae* tRNA sequences, only 60% of the sites modified by Pus1p are 3'-adjacent to a purine. Interestingly, the residue Ψ_{44} in *S. cerevisiae* snRNA U2, which is formed by Pus1p *in vivo*, is 3'-adjacent to a guanine. The unique residue formed by Pus1p *in vitro* in the 3'-terminal region of BMV RNA is also 3'-adjacent to a G residue. Altogether, this suggests a strong preference of Pus1p for uridine residues in the Purine-U dinucleotides. This preference was further confirmed by the absence of reactivity of the minisubstrate #1, derived from tRNA^{Ile}, where the putative modification site (position 27 as referred to tRNA numbering) is 3'-adjacent to U residue.

Minisubstrate #1 was designed on the basis of the tRNA^{Ile} structure. In spite that this tRNA lacks the consensus purine residue 5'-adjacent to Ψ_{27} position, it is modified by Pus1 both *in vitro* and *in vivo*. However, the binding constant of tRNA^{Ile} to Pus1p is almost 10 times higher than for other tRNA bearing G residue at position 26 (Arluison et al., 1999). The present data shows a greater dependence of the action of Pus1p for a 5' purine when the substrate is not the entire tRNA. Indeed, after introduction of a 5'-adjacent G residue and substitution of 3 nucleotides in the internal loop of minisubstrate #1 to avoid undesirable base-pairing (minisubstrate #2), a modification of position 27 by Pus1p was detected. However, the Ψ formation was still rather inefficient. In contrast, the minisubstrate #3, derived from tRNA^{Val} naturally modified by Pus1p at position 27, is highly modified *in vitro*. In this minisubstrate, the D stem-loop and anticodon stem-loop of tRNA^{Val} are connected by a G residue 5'-adjacent to the modified uridine. In minisubstrate #3 the two helical stems should be able to rotate freely around G₂₆ and this may facilitate the adaptation of the substrate to the catalytic site of the enzyme. Minisubstrate #4, which is based on the 3'-terminal region of BMV RNA and has a similar degree of structural flexibility, was also modified with high efficiency. The study of minisubstrates revealed the

complex specificity of Pus1p: U residues 3' adjacent to purine and located at the junction of two helical regions are targets of the enzyme, however, the efficiency of their conversion into Ψ residues strongly depends upon the conformation of the segments linking the helices.

Altogether the results confirm our hypothesis that a simple RNA structure composed of two stems connected by a linker forms an efficient substrate for Pus1p. They also show that freedom of movement of the two helices relative to each other is required and that the presence of a 5'-adjacent purine residue (G or A) is important for efficient recognition of RNA substrate by Pus1p. It remains to be determined whether the nature of 5'-adjacent base and the global flexibility of the molecule improve the binding of the substrate to Pus1 or are required for correct positioning the target residue in active center of the enzyme.

Acknowledgments

This work was supported by laboratory funds from the "Ministère de la Recherche et de l'Enseignement Supérieur", the CNRS, specific research grants from CNRS ("Programme Physique et Chimie du vivant 1997") and ARC ("Action de Recherche sur le Cancer") and the Human Frontier Science Program (grant #RG-360/93 to HAR). We thank Drs G. Simos and E. Hurt (Heidelberg University, Germany) for providing *PUS1*-disrupted yeast strain, Dr R.E. Jeeninga (Vrije Universiteit, Amsterdam, The Netherlands) for rRNA-containing plasmids and Dr. T. Dreher (Oregon University, USA) for the plasmid containing 3'-fragment of BMV RNA. S. Massenet is a pre-doctoral fellow supported by a fellowship from the "Ministère de la Recherche et de l'Enseignement Supérieur". Y. Motorin is the recipient of an Associated Researcher position at the CNRS ("Poste Rouge"). We thank Dr J-P. Waller (CNRS, Gif-sur-Yvette) for critical reading of the manuscript and useful comments.

Legends to figures

Fig. 1. Pseudouridylation of yeast 26S rRNA transcripts by recombinant Pus1p *in vitro*. [α - 32 P]UMP-labeled T7 transcripts of domain II (panels A and B), domain IV (panels C and D) and subregion of domain V (panels E and F) were incubated at 30°C for 60 min with recombinant Pus1p, completely digested with nuclease P1 (panels A, C and E) or RNase T2 (panels B, D and F) and the digests were analyzed by 2D-thin-layer chromatography. Representative autoradiograms of each of the digests are shown.

Fig. 2. Time courses for the formation of pseudouridine in the rRNA domain II, domain IV and subregion of domain V incubated at 30°C with purified recombinant Pus1p (20 ng/10 μ l). Panel A presents the total number of formed pseudouridine residues (as measured by nuclease P1 digestion); Panel B presents the number of pseudouridine residues 5'-adjacent to U (measured by RNase T2 digestion). 26S rRNA domain II (open circles), domain IV (filled circles) and subregion of domain V (open squares).

Fig. 3. Primer extension mapping of Ψ residues in 26S rRNA domain IV (Panels A to C) and subregion of domain V (Panel D) transcripts modified *in vitro* by recombinant Pus1p. Unmodified (control) and *in vitro* Pus1p-modified transcripts were subjected to treatment with CMCT for 2 min (lanes 2 and 3) or for 10 and 20 minutes (lanes 4 and 5, respectively). Lane 1 shows a control extension on RNA not treated with CMCT. Sequencing ladders (lanes UGCA) are shown for reference. Experimental conditions are described in Materials & Methods. In lanes 3, 4 and 5, the CMCT-treated RNA was subjected to an alkaline treatment for 2 hours in 50 mM sodium-bicarbonate buffer at pH 10.4. Arrows indicate the reverse-transcription stops corresponding to pseudouridine residues.

Fig. 4. Secondary structure analysis of an *in vitro* produced domain V of *S. cerevisiae* 26S rRNA. Panel A: Transcripts corresponding to the entire domain V of 26S rRNA were subjected to chemical (DMS and CMCT) and enzymatic (V1 and T1 nucleases) probing as described in Materials and Methods. An unmodified transcript was used as a control (Ct). The 26S rRNA region complementary to the primer is shown in Panel B as an arrow covering helix 75. Lanes U, G, C and A correspond to the RNA sequencing ladder. Nucleotide positions, starting from the 5'-terminal nucleotide, are indicated on the left of the Panel. Panel B: The results of probing experiments are schematically represented on the structure previously proposed for the 26S rRNA domain V (Veldman et al., 1981; Gutell et al., 1994; Schnare et al., 1996) (insert in Panel B) and on the alternative structure that we found to be formed by the region encompassing nts 2267 to 2313 in the T7 transcript. Brackets delimit the portion of domains IV that was investigated in this study. Helices are numbered according to Leffers et al. (1987).

Fig. 5. Positions of Ψ residues formed *in vitro* in yeast 26S rRNA domains IV (A) and V (B). The portions of domains IV and V that were investigated for the presence of Ψ residues (in brackets) are folded according to the secondary structure established experimentally (see Fig. 4 for domain IV and not shown for domain V). Nucleotides are numbered starting from 5'-end. Naturally occurring pseudouridine residues (Ofengand & Fournier, 1998) are shown as Ψ . Arrows indicate the positions of pseudouridine residues formed in the transcripts upon incubation with recombinant Pus1p. Major modification sites are shown in bold.

Fig. 6. Pseudouridine formation in the tRNA-like structure at the 3'-end of BMV RNA. Panel A shows a schematic representation of the secondary structure of the 3'-end fragment of BMV RNA (adapted from Felden et al., 1994). The tRNA-like structure is shown in grey, the helices B1 and B2 used in the minisubstrate #4 (see Fig. 7) are encircled. Nucleotides are numbered starting from 3'-end. The position of Ψ_{40}

synthesized upon incubation of the T7 transcript with recombinant Pus1p is indicated by an arrow. Panel B: time course of pseudouridylation of the BMV transcript by recombinant Pus1p (20 ng/10 μ l). Panel C: Mapping of Ψ residues in the 3'-end fragment of the BMV RNA transcript modified *in vitro* by recombinant Pus1p. Pus1p-modified transcript was treated with CMCT (no treatment, lane 1; 2 min, lane 3; 20 min, lane 5) and analyzed by reverse transcription as described in Materials & Methods. Sequencing ladders (lanes UGCA) are shown for reference. The arrow points to the strong stop corresponding to CMC-modified pseudouridine at position 40.

Fig. 7. The structures of minisubstrate RNA molecules used in this study. At left is a schematic representation of the general 2D/3D structure of tRNA. The portion used for construction of the minisubstrates is encircled. The position of pseudouridine-27 is shown. Minisubstrates #1 and #2 were designed based on the structure of yeast tRNA^{Ile} and minisubstrate #3 was derived from yeast tRNA^{Val}. Minisubstrate #4 was derived from the structure of the 3'-terminal fragment of BMV RNA (see Fig. 6). The potential modification site in each minisubstrate is boxed, the presence of pseudouridine at this position after incubation with Pus1p is indicated by an arrow. Residues in minisubstrate #2 different from those in #1 are indicated in bold. The G...A base pair 5'-adjacent to the modification site is indicated in the structure of minisubstrate #2. Possible U-A Watson-Crick base-pair 5'-adjacent to U₂₇ in mini#1 is indicated by dots.

Fig. 8. Modification of RNA minisubstrates by Pus1p. Minisubstrate RNA molecules (about 2 nM final concentration) labeled with [α -³²P]UTP were incubated in 10 μ l of standard buffer (see Materials and Methods) with various amounts of Pus1p for 90 min at 30°C. The amount of pseudouridine formed was measured after nuclease P1 digestion as described in Materials & Methods. The extent of

modification for minisubstrates #1 (open squares), #2 (filled squares), #3 (open circles) and #4 (filled circles) are shown.

References

- Arluison V, Buckle M, Grosjean H. 1999. Pseudouridine synthase Pus1 of *S.cerevisiae*: kinetic characterization, tRNA structural requirement and real time analysis of its complex with tRNA. *J. Mol. Biol.* 289:491-502.
- Bakin A, Ofengand J. 1993. Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center - analysis by the application of a new sequencing technique. *Biochemistry* 32:9754-9762.
- Becker HF, Motorin Y, Florentz C, Giegé R, Grosjean H. 1998. Pseudouridine and ribothymidine formation in the tRNA-like domain of Turnip Yellow Mosaic virus RNA. *Nucleic Acids Res* 26:3991-3998.
- Becker HF, Motorin Y, Planta RJ, Grosjean H. 1997a. The yeast gene YNL292w encodes a pseudouridine synthase (Pus4) catalyzing the formation of Ψ_{55} in both mitochondrial and cytoplasmic tRNAs. *Nucleic Acids Res.* 25:4493-4499.
- Becker HF, Motorin Y, Sissler M, Florentz C, Grosjean H. 1997b. Major identity determinants for enzymatic formation of ribothymidine and pseudouridine in the T Ψ -loop of yeast tRNAs. *J. Mol. Biol.* 274:505-518.
- 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.* 9:4303-24.
- Conrad J, Niu L, Rudd K, Lane BG, Ofengand J. 1999. 16S ribosomal RNA pseudouridine synthase RsuA of *Escherichia coli*: Deletion, mutation of the conserved Asp102 residue, and sequence comparison among all other pseudouridine synthases. *RNA* 5:761-763.

- Conrad J, Sun D, Englund N, Ofengand J. 1998. The *rluC* gene of *Escherichia coli* codes for a pseudouridine synthase that is solely responsible for synthesis of pseudouridine at positions 955, 2504, and 2580 in 23 S ribosomal RNA. *J. Biol. Chem.* 273:18562-18566.
- Dreher TW, Hall TC. 1988. Mutational analysis of the sequence and structural requirements in brome mosaic virus RNA for minus strand promoter activity. *J. Mol. Biol.* 201:31-40.
- Felden B, Florentz C, Giegé R, Westhof E. 1994. Solution structure of the 3'-end of Brome Mosaic Virus genomic RNAs. Conformational mimicry with canonical transfer RNAs. *J. Mol. Biol.* 235:508-531.
- Ganot P, Bortolin ML, Kiss T. 1997. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* 89:799-809.
- Gu X, Ofengand J, Santi DV. 1994. *In vitro* methylation of *Escherichia coli* 16S rRNA by tRNA:m⁵U₅₄-methyltransferase. *Biochemistry* 33:2255-2261.
- Gutell RR, Larsen N, Woese CR. 1994. Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. *Microbiol Rev* 58:10-26.
- Huang L, Ku J, Pookanjanatavip M, Gu X, Wang D, Greene PJ, Santi DV. 1998. Identification of two *Escherichia coli* pseudouridine synthases that show multisite specificity for 23S RNA. *Biochemistry* 37:15951-15957.
- Jiang H-Q, Motorin YA, Jin Y-X, Grosjean H. 1997. Pleiotropic effects of intron removal on base modifications pattern of yeast tRNA^{Phe}: an *in vitro* study. *Nucleic Acids Res* 25:2694-2701.
- Lafontaine DLJ, 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.* 12:527-537.

- Lecoite F, Simos G, Sauer A, Hurt EC, Motorin Y, Grosjean H. 1998. Characterization of yeast protein Deg 1 as Pseudouridine synthase (Pus 3) catalysing the formation of Ψ_{38} and Ψ_{39} in tRNA anticodon loop. *J. Biol. Chem.* 273:1316-1323.
- Leffers H, Kjems J, Ostergaard L, Larsen N, Garrett RA. 1987. Evolutionary relationships amongst *archaeobacteria*. A comparative study of 23 S ribosomal RNAs of a sulphur-dependent extreme thermophile, an extreme halophile and a thermophilic methanogen. *J. Mol. Biol.* 195:43-61.
- Massenet S, Motorin Y, Lafontaine DLJ, Hurt EC, Grosjean H, Branlant C. 1999. Pseudouridine mapping in the *Saccharomyces cerevisiae* spliceosomal UsnRNAs reveals that pseudouridine synthase Pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA. *Mol. Cell. Biol.* 19:2142-2154.
- Motorin Y, Keith G, Simon C, Foiret D, Simos G, Hurt E, Grosjean H. 1998. The yeast tRNA:pseudouridine synthase Pus1p displays a multisite substrate specificity. *RNA* 4:856-869.
- Ni J, Tien AL, Fournier MJ. 1997. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* 89:565-573.
- Niu L, Ofengand J. 1999. Cloning and characterization of the 23S RNA pseudouridine 2633 synthase from *Bacillus subtilis*. *Biochemistry* 38:629-635.
- Ofengand J, Bakin A. 1997. Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, *archaeobacteria*, mitochondria and chloroplasts. *J Mol Biol* 266:246-268.
- Ofengand J, Fournier MJ. 1998. The pseudouridine residues of rRNA: number, location, biosynthesis, and function. In: Grosjean, H & Benne, R, eds. *The Modification and Editing of RNA*. New York: ASM Press. pp 229-254.

- Raychaudhuri S, Conrad J, Hall BG, Ofengand J. 1998. A pseudouridine synthase required for the formation of two universally conserved pseudouridines in ribosomal RNA is essential for normal growth of *Escherichia coli*. *RNA* 4:1407-1417.
- Raychaudhuri S, Niu L, Conrad J, Lane BG, Ofengand J. 1999. Functional effect of deletion and mutation of the *Escherichia coli* ribosomal RNA and tRNA pseudouridine synthase RluA. *J. Biol. Chem.* 274:18880-18886.
- Samarsky DA, Fournier MJ. 1999. A comprehensive database for the small nucleolar RNAs from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 27:161-164.
- Schnare MN, Damberger SH, Gray MW, Gutell RR. 1996. Comprehensive comparison of structural characteristics in eukaryotic cytoplasmic large subunit (23 S-like) ribosomal RNA. *J. Mol. Biol.* 256:701-19.
- Simos G, Tekotte H, Grosjean H, Segref A, Sharma K, Tollervy D, Hurt EC. 1996. Nuclear pore proteins are required for the biogenesis of functional tRNA. *EMBO J.* 15:2270-2284.
- Szweykowska-Kulinska Z, Senger B, Keith G, Fasiolo F, Grosjean H. 1994. Intron-dependent formation of pseudouridines in the anticodon of *Saccharomyces cerevisiae* minor tRNA^{Leu}. *EMBO J.* 13:4636-4644.
- 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.* 9:6935-52.
- Venema J, Dirks-Mulder A, Faber AW, Raué HA. 1995. Development and application of an *in vivo* system to study yeast ribosomal RNA biogenesis and function. *Yeast* 11:145-156.
- Watkins NJ, Gottschalk A, Neubauer G, Kastner B, Fabrizio P, Mann M, Lührmann 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* 4:1549-1568.

Wrzesinski J, Bakin A, Nurse K, Lane BG, Ofengand J. 1995a. Purification, cloning, and properties of the 16S RNA pseudouridine 516 synthase from *Escherichia coli*. *Biochemistry* 34:8904-8913.

Wrzesinski J, Nurse K, Bakin A, Lane BG, Ofengand J. 1995b. A dual-specificity pseudouridine synthase: An *Escherichia coli* synthase purified and cloned on the basis of its specificity for Ψ_{746} in 23S RNA is also specific for Ψ_{32} in tRNA^{Phe}. *RNA* 1:437-448.

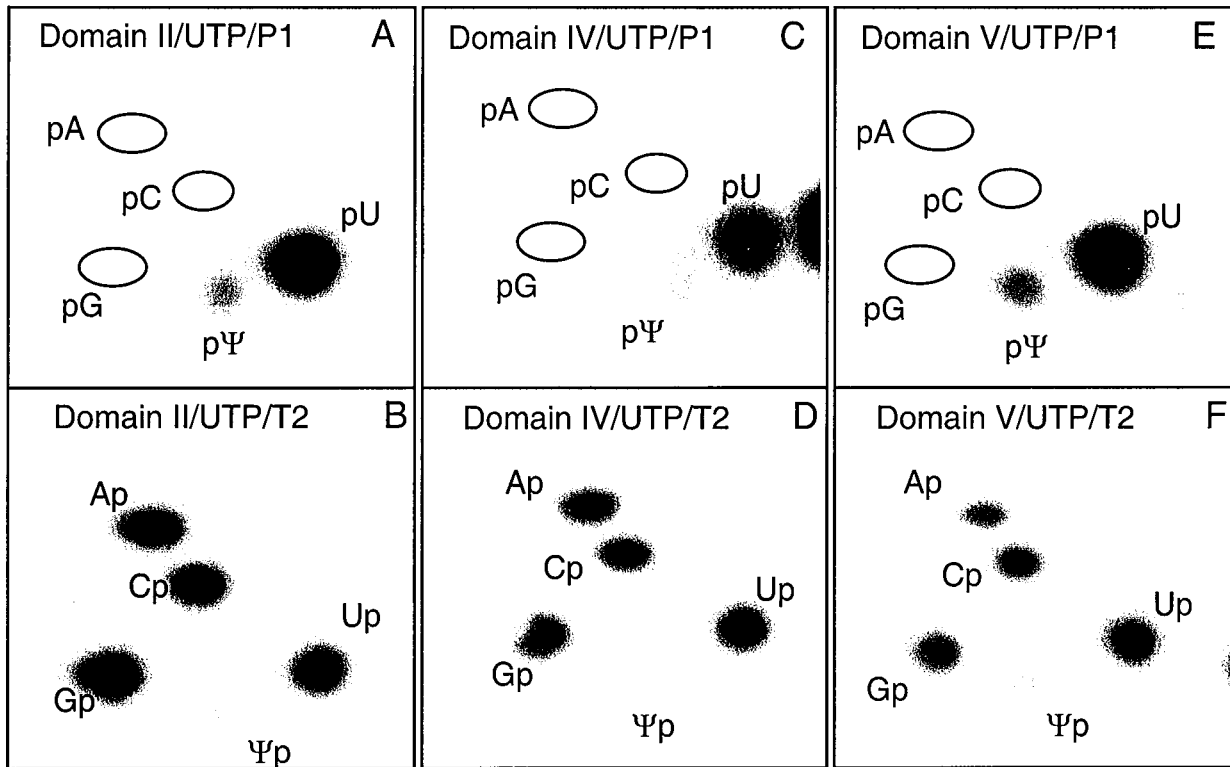
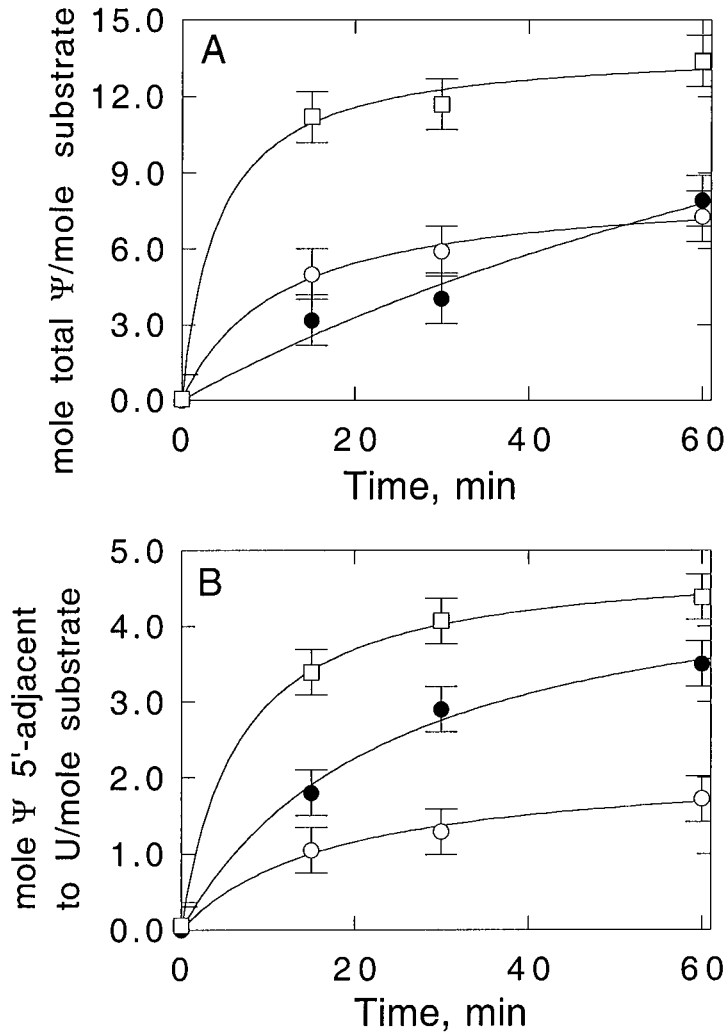


Fig. 1 Motorin *et al*



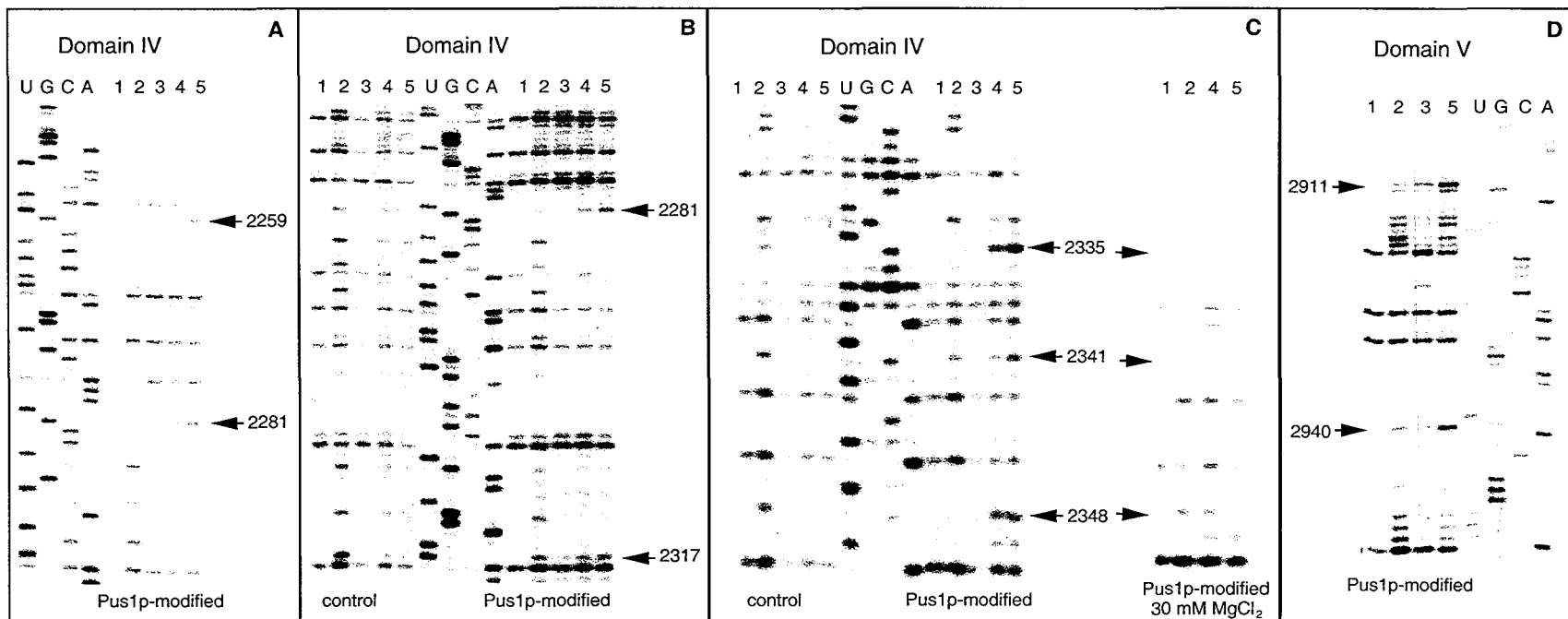


Fig.3 Motorin *et al*

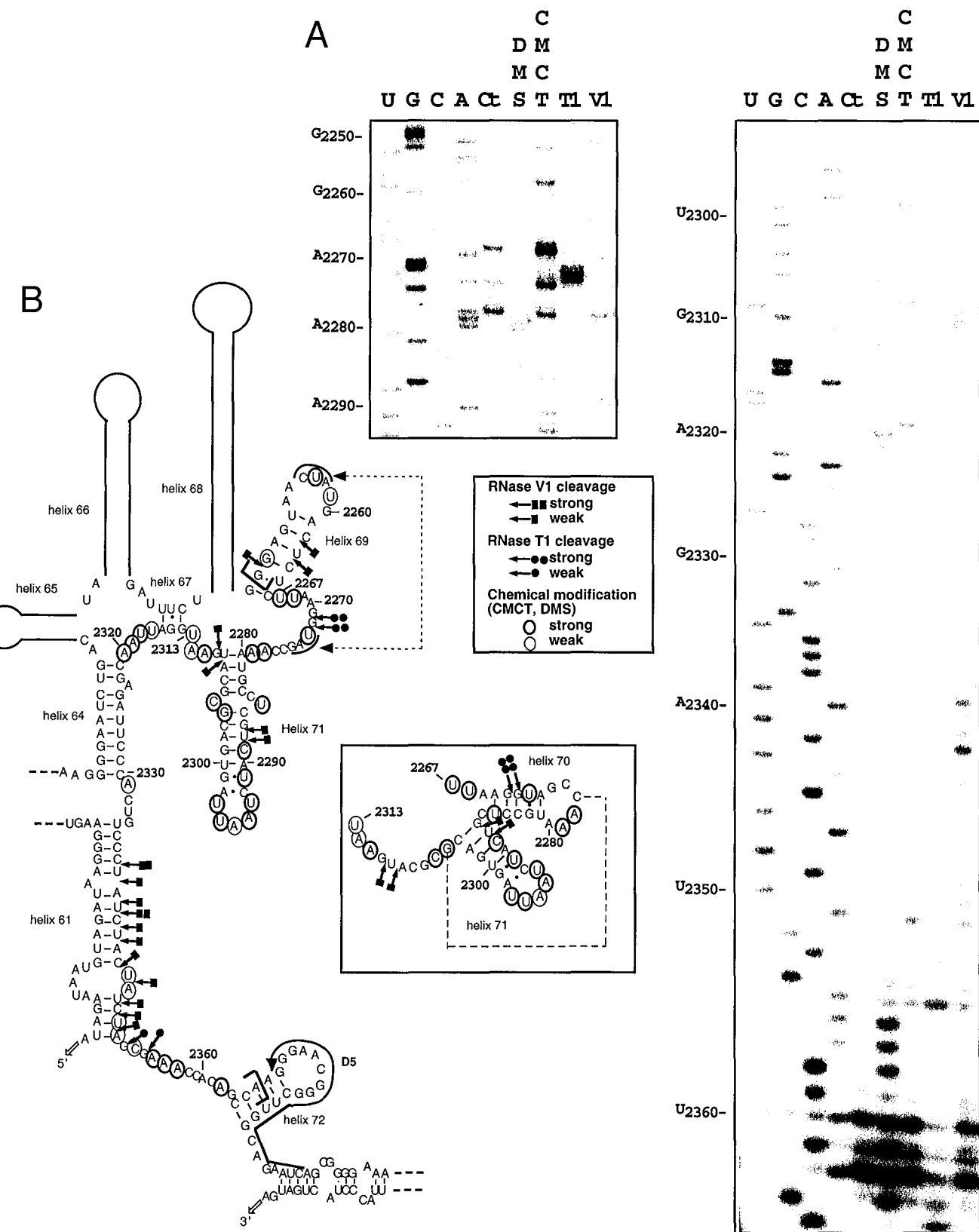


Fig. 4 Motorin *et al*

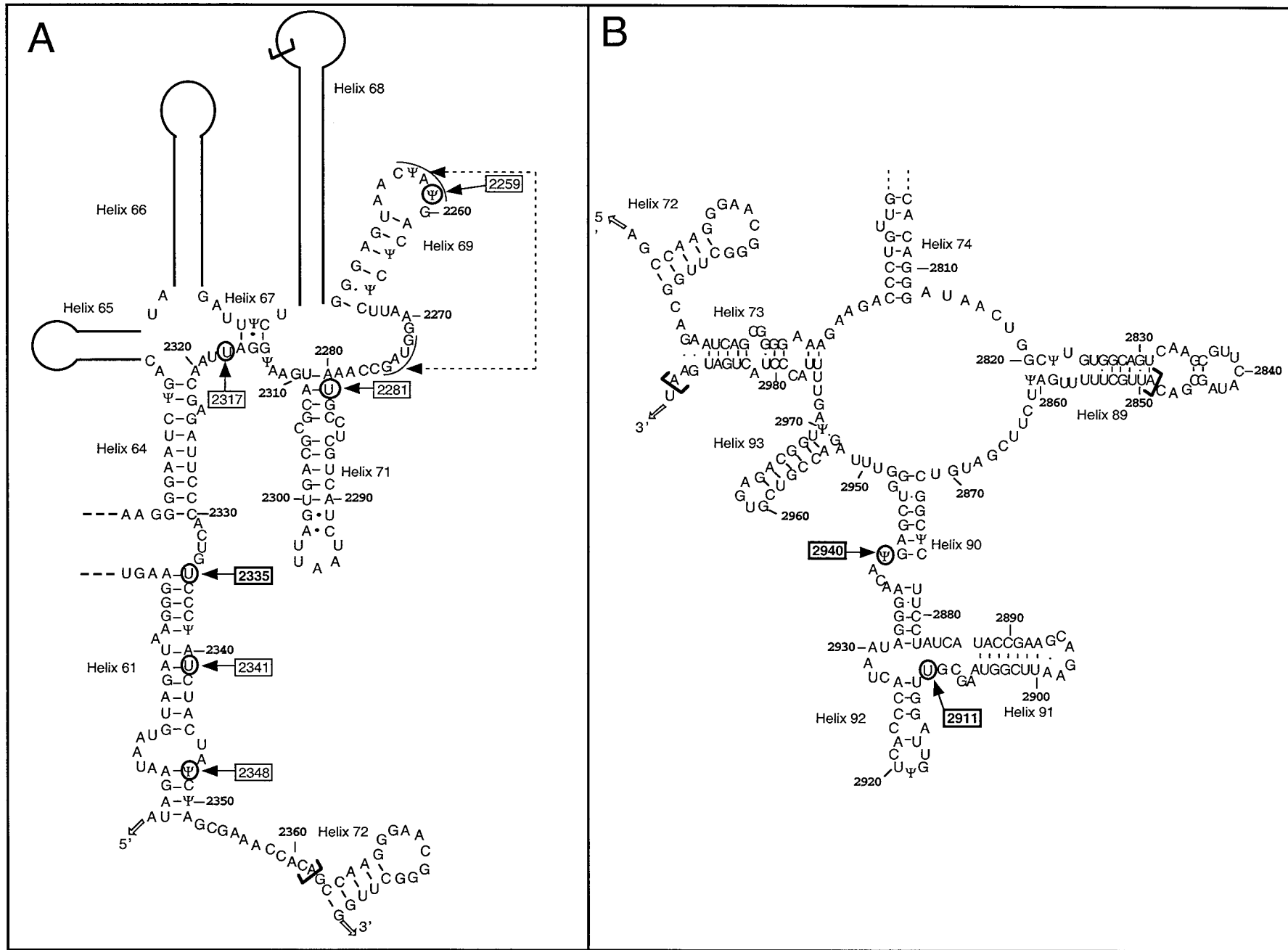


Fig. 5 Motorin *et al*

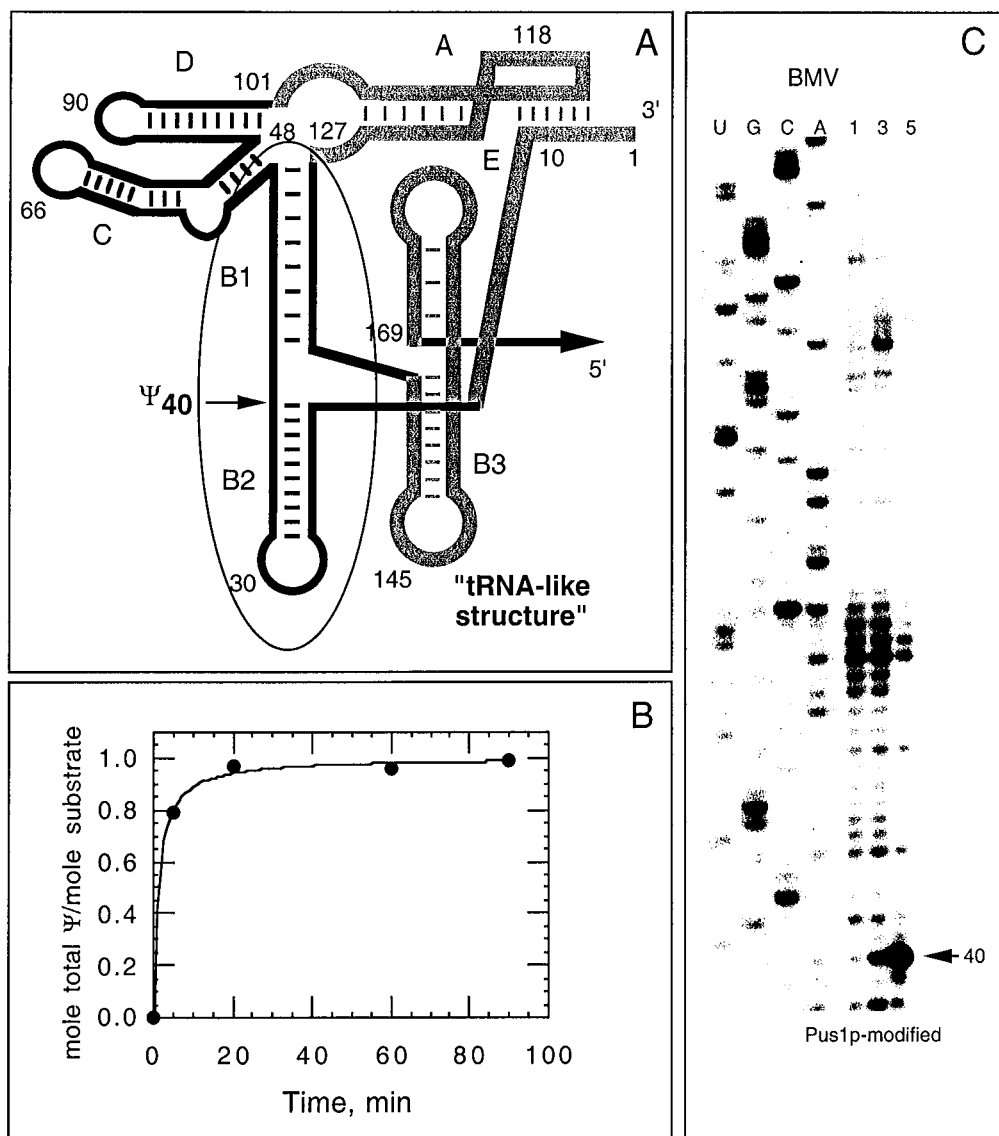


Fig. 6 Motorin *et al*

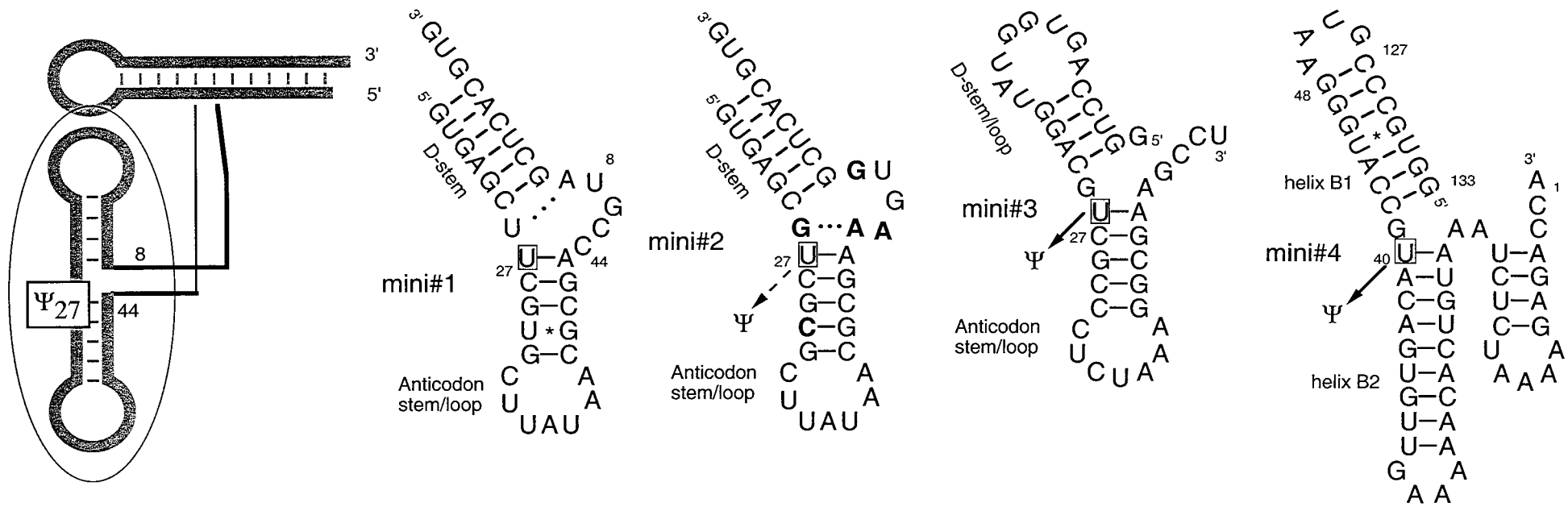
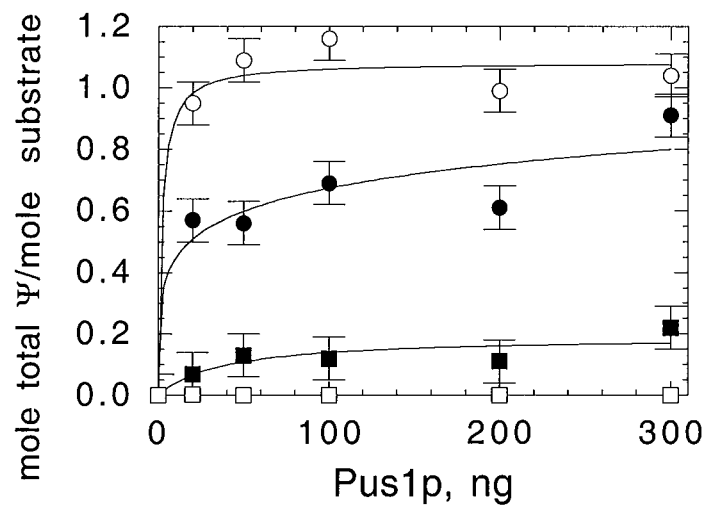


Fig. 7 Motorin *et al*



2. DISCUSSION

2.1. Pus1p modifie *in vitro* des ARN qu'elle ne modifie pas *in vivo*

En utilisant des transcrits, produits *in vitro*, correspondant à deux des domaines de l'ARNr 26S de levure, nous avons démontré que l'enzyme Pus1p modifie *in vitro* certains résidus U qui ne sont pas des cibles naturelles pour cette Ψ -synthase *in vivo*. Il avait déjà été démontré que l'ARN du virus de plantes TYMV était un bon substrat pour Pus1p, *in vitro* (Becker *et al.*, 1998). Nous montrons qu'il en est de même pour la région 3'-terminale du virus BMV, mais dans ce cas la modification a lieu hors de la structure "ARNt-like". La capacité de former des résidus Ψ dans des ARN qui ne sont pas des substrats naturels n'est pas spécifique à la Ψ -synthase Pus1p. En effet, il a été démontré que d'autres enzymes de modification possèdent également cette propriété. C'est le cas des Ψ -synthases RluC et RluD d'*E. coli* qui sont responsables *in vivo* de la formation de 6 résidus Ψ dans l'ARN 23S (en positions 955, 2504 et 2580 dans le cas de RluC et en positions 1911, 1915 et 1917 dans le cas de RluD) (Huang *et al.*, 1998a). *In vitro*, ces deux enzymes catalysent chacune la conversion en résidus Ψ d'environ 5% des résidus U dans des transcrits des gènes des ARN 16S et 23S d'*E. coli*, produits *in vitro* (Huang *et al.*, 1998a). La localisation des sites modifiés n'a pas été analysée. De même, la Ψ -synthase RluB de *B. subtilis* forme *in vitro* un et trois résidus Ψ dans des transcrits, produit *in vitro*, correspondant, respectivement, à l'ARN 16S et à l'ARN23S d'*E. coli* (Niu et Ofengand, 1999). Les 4 régions modifiées, *in vitro*, par RluB ont été comparées et une séquence consensus a pu être établie pour les sites de modification identifiés (5' RRCGU Ψ 3'). De plus, les résidus Ψ modifiés sont toujours localisés dans une région double brin. Comme nous l'avons montré, les sites modifiés par Pus1p *in vitro* présentent également des caractéristiques communes. *In vitro*, les Ψ -synthases RluB, C et D ne modifient pas des transcrits correspondant à des ARNt. Pus1p est donc la Ψ -synthase qui présente la spécificité la plus large *in vitro*, car elle peut modifier pratiquement tous les ARN, produits *in vitro*, quelle que soit leur nature.

Pourquoi ces sites, modifiés *in vitro*, ne sont-ils pas modifiés *in vivo*, dans le cas où l'enzyme et l'ARN cible coexistent dans la même cellule?

⇒ une première explication peut être que certains sites sont reconnus avec une moins bonne efficacité car ils ne présentent pas tous les signaux requis pour une reconnaissance optimale. Ils seraient utilisés *in vitro* seulement lorsque l'ARN est placé en présence d'une concentration d'enzyme beaucoup plus élevée que la concentration présente *in vivo*. Même *in vitro*, en présence d'une concentration élevée en enzyme, nous avons constaté que le taux de modification variait beaucoup suivant les sites de modification. Nous avons clairement observé la présence de deux catégories de sites, les sites fortement modifiés par Pus1p et les

sites faiblement modifiés. L'enzyme Pus1p ayant une affinité pour l'ARN en général, cette différence de taux de modification, entre les différents résidus U cibles, pourrait provenir d'une différence de positionnement de ces résidus dans le site actif de l'enzyme.

De la même manière, un des 4 résidus U modifiés *in vitro* dans les ARNr d'*E. coli* par la Ψ -synthase RluB de *B. subtilis*, est modifié beaucoup plus efficacement et rapidement que les 3 autres sites (Niu et Ofengand, 1999). Ce site est l'homologue du site où l'enzyme de *B. subtilis* formerait un résidu Ψ *in vivo* dans l'ARN 23S (Niu et Ofengand, 1999). Un résidu A, situé en boucle latérale dans l'hélice où est localisé le résidu fortement modifié et qui n'est pas présent dans les 3 autres hélices plus faiblement modifiées *in vitro* par RluB, pourrait augmenter l'affinité de l'enzyme pour son substrat et expliquer le fort degré de modification à cette position (Niu et Ofengand, 1999).

⇒ une deuxième explication pour l'observation de sites de modification *in vitro* non observés *in vivo* est que ces sites seraient masqués *in vivo*, soit par des protéines, soit par le repliement tertiaire de l'ARN.

⇒ enfin, une dernière explication est l'existence d'une différence de localisation cellulaire entre l'enzyme et l'ARN.

2.2. Motif de reconnaissance de Pus1p sur l'ARN

2.2.1. La structure 3D des ARNt n'est pas nécessaire pour la reconnaissance par Pus1p

Il avait déjà été montré qu'un mini-substrat correspondant à une partie du pré-ARNt^{Ile} mineur de *S. cerevisiae*, la tige-boucle de l'anticodon et son intron, est un excellent substrat pour la formation par Pus1p des résidus Ψ_{34} et Ψ_{36} (Fig. 49) (Motorin et al., 1998). Nous avons montré qu'un mini-substrat renfermant le bras anticodon et l'hélice D de l'ARNt^{Val} de *S. cerevisiae* est également efficacement modifié *in vitro* à la position 27 par Pus1p. Clairement, l'intégrité de la structure 3D des ARNt n'est pas nécessaire à la reconnaissance et à la modification par Pus1p. Cependant, des études enzymatiques ont montré que la fixation de Pus1p, aux ARNt^{Asp} et ARNt^{Phe} de *S. cerevisiae*, est améliorée quand ceux-ci présentent une structure tertiaire correcte (Arluison *et al.*, 1999). Par ailleurs, comme nous avons pu le constater, l'ARNt^{Ile} est modifié à un taux faible *in vitro* par Pus1p, alors que le mini-substrat 1 que nous avons utilisé, renfermant le bras anticodon et l'hélice D de cet ARNt n'est pas modifié du tout par Pus1p *in vitro*.

Parmi les autres ARNt: Ψ -synthases caractérisées chez *S. cerevisiae*, Pus4p ne requiert pas non plus la structure 3D de l'ARNt pour la reconnaissance et la modification de la position 55. En effet, un mini-substrat correspondant au bras T Ψ C est efficacement modifié par cette enzyme *in vitro* (Becker *et al.*, 1997a). C'est également le cas de son homologue chez *E. coli*,

	position 26	position 27	position 28	position 34	position 35	position 36	position 65	position 67
ARNt ^{Ile} mineur (ΨAΨ)		CUΨCG		CUΨAΨ (intron)		ΨAΨAG (intron)		CCΨCG
ARNt ^{Lys} (CUU)		Cm ² ₂ GΨAU						
ARNt ^{Lys} (cmnm ⁵ UmUU)		Cm ² ₂ GΨΨC	m ² ₂ GΨΨCG					UAΨGA
ARNt ^{Leu} (UAG)		Cm ² ₂ GΨCA						
ARNt ^{Met} (CAU)		Cm ² ₂ GΨCA						
ARNt ^{Arg} (ICG)		Cm ² ₂ GΨCU						
ARNt ^{Arg} (mcm ⁵ UCU) ₁		Cm ² ₂ GΨCU						
ARNt ^{Arg} (mcm ⁵ UCU) ₂		Cm ² ₂ GΨCU						
ARNt ^{Val} (IAC)		CAΨCU						
ARNt ^{Val} (CAC)		Cm ² ₂ GΨΨG	m ² ₂ GΨΨGC					
ARNt ^{Val} (ncm ⁵ UAC)		Cm ² ₂ GΨCG						
ARNt ^{Trp} (CmCA)	GCΨΨΨ	CΨΨΨC	ΨΨΨCG				UGΨCC	
ARNt ^{Tyr} (GΨA)					UGΨAi ^b A (intron)			
ARNt ^{Glu} (mcm ⁵ s ² UUC)		CAΨCA						

Tableau 11: Séquences bordant les résidus Ψ formés par Pus1p dans les ARNt cytoplasmiques de *S. cerevisiae*.

Les résidus Ψ aux positions indiquées en violet ne sont pas formés *in vitro* par l'enzyme Pus1p recombinante, mais leur formation n'a pas lieu *in vivo* lorsque l'expression du gène *PUS1* est inhibée. Les résidus Ψ aux positions indiquées en noir sont formés *in vivo* et *in vitro* par Pus1p. Ces résidus Ψ sont indiqués en rouge dans les séquences. Le résidu Ψ en position 35 dans l'ARNt^{Tyr} est formé *in vitro* par Pus1p, mais sa formation a toujours lieu *in vivo* lorsque l'expression du gène *PUS1* est inhibée. Les résidus Ψ, pour lesquels le terme intron est indiqué, nécessitent la présence de l'intron pour être formés par Pus1p.

la Ψ -synthase TruB (Gu *et al.*, 1998). Par contre, l'action de Pus3p, qui catalyse la formation des résidus Ψ_{38} et Ψ_{39} , ou de son homologue TruA chez *E. coli*, dépend de l'intégrité de la structure tertiaire des ARNt (Edqvist *et al.*, 1993; Grosjean *et al.*, 1996; Lecointe *et al.*, 1998).

2.2.2. Un motif structural spécifique est reconnu par Pus1p

Bien que toutes deux indépendantes de la structure 3D des ARNt, les Ψ -synthases Pus1p et Pus4p diffèrent cependant par le nombre de sites qu'elles modifient dans les ARN. En effet, alors que la Ψ -synthase Pus4p reconnaît spécifiquement la structure de la boucle T Ψ C et modifie exclusivement l'uridine présente en position 55 dans cette boucle (Becker *et al.*, 1997a et 1997b), Pus1p semble modifier 8 sites différents dans les ARNt. De plus, sa capacité à catalyser, *in vitro*, la formation de résidus Ψ dans des ARN qui ne sont pas ses substrats naturels montre que le spectre de substrats potentiels de cette enzyme est beaucoup plus large que celui de Pus4p. En effet, notre analyse des différentes positions modifiées par Pus1p *in vitro* dans des transcrits correspondant aux domaines IV et V de l'ARN 26S de *S. cerevisiae* ou aux ARN génomiques des virus TYMV et BMV montre que toute uridine, précédée d'une purine et localisée dans une région flexible à la jonction de deux hélices ou proche de cette jonction, peut être un substrat potentiel pour Pus1p. Ces caractéristiques sont retrouvées pour la plupart des positions modifiées, *in vivo* et *in vitro*, par Pus1p dans les ARNt. Le requis d'une purine en 5' du résidu modifié semble cependant moins strict pour la modification des ARNt. En effet, chez *S. cerevisiae*, 35% des résidus U supposés convertis en résidus Ψ par Pus1p dans les ARNt sont localisés en 3' de pyrimidines (Tableau 11). Il est également intéressant de constater que, parmi les 65% de résidus U supposés convertis en Ψ par Pus1p dans les ARNt et précédés d'une purine, il s'agit dans 75% des cas d'un résidu G (et notamment m²G) (Tableau 11). Pus1p pourrait donc avoir une légère préférence pour les dinucléotides GU. Par ailleurs, en ce qui concerne les positions modifiées par Pus1p dans les ARNt, un résidu C en 3' est observé dans 56% des cas, alors que d'après les études que nous avons réalisées sur les substrats non naturels, aucun requis n'a été observé pour le résidu en 3' du résidu U modifié. Il faut bien sûr tenir compte du fait qu'une statistique établie à partir de la comparaison de la situation dans différents ARNt est biaisée par la conservation de l'identité des résidus à certaines positions des ARNt et notamment en position 27 (Tableau 11). Néanmoins, il semble ressortir que Pus1p pourrait avoir une spécificité d'action particulière sur les ARNt, par rapport à celle qu'elle a sur des ARN autres que les ARNt.

Pour approfondir les connaissances sur la spécificité de Pus1p, nous avons abordé une étude des structures et des séquences reconnues dans le snRNA U2 par cette Ψ -synthase et nous allons la décrire dans la suite de cette partie des Résultats.

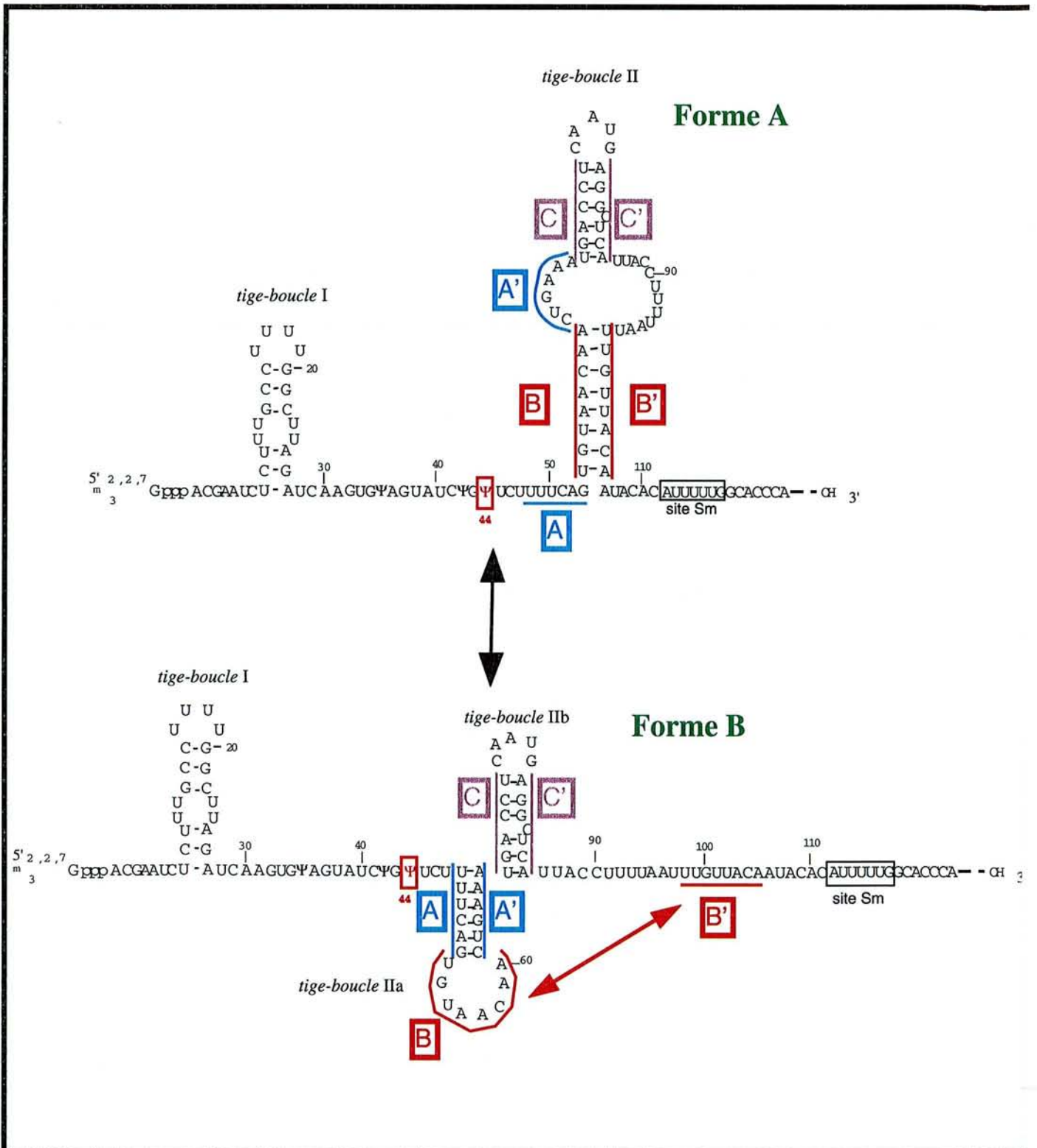


Figure 54: Position du résidu Ψ_{44} dans les deux structures alternatives possibles du snRNA U2 de *S. cerevisiae*. Les structures sont celles proposées par Ares et Igel, 1990; Zavanelli et Ares, 1991; Zavanelli *et al.*, 1994. Le résidu Ψ_{44} formé par Pus1p est indiqué en rouge.

B/ RECHERCHE DES MOTIFS DU snRNA U2 RECONNUS PAR LA Ψ-SYNTHASE Pus1p

1. INTRODUCTION

Comme nous l'avons vu, Pus1p modifie *in vivo* les ARNt à différentes positions et également la position 44 du snRNA U2 de *S. cerevisiae*. Nous nous sommes donc demandé si le mode de reconnaissance du snRNA U2 par Pus1p est le même que celui des ARNt et des différents substrats non naturels que nous avons étudiés. Pour répondre à cette question, nous avons débuté la recherche des éléments du snRNA U2 impliqués dans la reconnaissance par Pus1p. Dans les deux structures secondaires possibles du snRNA U2 (Fig. 54), le segment contenant le résidu U₄₄ est une longue séquence simple brin, encadrée par deux structures tige-boucles. Vu la distance élevée entre le résidu cible de Pus1p et ces structures tige-boucles, le site modifié dans le snRNA U2 ne semblait pas conforme aux sites sélectionnés pour Pus1p, selon l'étude que nous avons décrite précédemment. Le seul point en accord avec les résultats précédents était la présence d'une purine, et plus précisément d'un résidu G, en amont du résidu cible.

Il était donc important de réaliser une étude de la structure 2D du transcrit du snRNA U2, produit *in vitro*, et utilisé pour les expériences de modification *in vitro* par l'enzyme Pus1p recombinante. Dans un premier temps, j'ai donc réalisé cette étude de structure. I. Ansmant, que j'ai encadrée pour son stage de DEA, a poursuivi l'étude en produisant des transcrits U2 où plusieurs structures tige-boucles, localisées à proximité de la position 44, étaient délétées de manière individuelle ou en combinaison. Elle a ensuite testé l'action de Pus1p sur ces ARN tronqués. Par ailleurs, afin d'étudier l'importance de l'identité des nucléotides bordant le résidu modifié, I. Ansmant a réalisé une mutagenèse statistique du segment contenant le nucléotide cible, puis elle a étudié l'action de Pus1p, *in vitro*, sur certains des ARN mutants.

Comme ces résultats ne sont pas présentés sous forme de publication, nous décrirons dans un premier temps les méthodes employées pour cette étude puis, dans un deuxième temps, les résultats obtenus.

1. Réaction PCR.

Matrice ADN pT7-U2	5 ng
Tampon Pfu 10x	5 µl
(KCl 10 mM, (NH ₄) ₂ SO ₄ 100 mM, MgSO ₄ 20 mM, Triton X-100 1%, BSA 1 mg/ml, Tris-HCl 200 mM pH 8,8)	
Oligonucléotides	100 ng de chaque
Pfu (Stratagène)	2,5 unités
εdNTP (1,25 mM de chacun des nucléotides)	8 µl
H ₂ O stérile	qsp 50 µl

Cycles utilisés: ↪ Dénaturation : 5' à 94°C

↪ 25 cycles d'amplification : Dénaturation : 1 min à 94°C

Hybridation : 1 min 30 sec à 48°C

Elongation : 7 min à 72°C

↪ Extension finale : 10 min à 72°C

2. Déprotéinisation.

Les produits PCR subissent 1 extraction avec un mélange phénol-chloroforme (vol/vol), suivie d'une extraction au chloroforme seul. Ils sont ensuite précipités à l'éthanol en présence d'AcNH₄ 2M. Les culots sont repris dans 43,5 µl d'H₂O stérile.

3. Phosphorylation des produits PCR.

Sambrook *et al.*, 1989.

4. Purification sur papier NA45 (DEAE membrane S&S).

La totalité du produit PCR phosphorylé est déposé sur gel d'agarose 0,8%. Après migration, la bande d'ADN est localisée sous UV à 256nm. Le gel est alors coupé juste en-dessous de la bande, et un fragment de membrane est inséré. L'électrophorèse est poursuivie 5 à 10 minutes, le temps que l'ADN migre dans la membrane, ce qui est vérifié aux UV. La membrane est ensuite placée pendant 30' à 65°C dans 400µl de tampon d'éluat (NaCl 1 M, EDTA 1 mM, Tris 20 mM pH8,0). L'éluat est centrifugé 1' à 12000 rpm, afin d'éliminer les déchets de membrane, et transféré dans un autre tube eppendorf. Deux extractions au butanol (vol/vol) sont ensuite réalisées et l'ADN est précipité à l'éthanol. Le culot est repris dans 31 µl d'H₂O stérile. Cinq µl sont déposés sur gel d'agarose, afin de vérifier l'efficacité de la purification.

5. Ligation.

Sambrook *et al.*, 1989.

6. Transformation de cellules.

La transformation se fait avec une demi-ligation (15 µl) et 100 µl de cellules compétentes (TG1 ou XL1 blue). Les cellules sont étalées sur boîte LBA (ampicilline 100 mg/l).

Tableau 12: Protocole de mutagenèse par PCR.

2. METHODES

2.1. Etude de la structure secondaire de l'extrémité 3'-terminale d'un transcrit, produit *in vitro*, correspondant au snRNA U2 de *S. cerevisiae*

Deux µg de snRNA U2 de *S. cerevisiae*, produits par transcription *in vitro* comme décrit dans la publication n°2, ont été incubés dans 10 µl de tampon RB (AcNa 100 mM; MgCl₂ 5 mM, DTT 2 mM, EDTA 0,1 mM, Tris HCL 100 mM pH 8) et chauffés pendant 3 min à 80°C. Le mélange a été amené doucement à 37°C. Les analyses de la structure secondaire de ce transcrit ont été faites dans le tampon RB et à 37°C, afin de réaliser l'analyse de la structure dans les conditions utilisées pour l'incubation avec Pus1p.

Les conditions d'utilisation des sondes chimiques et enzymatiques étaient les mêmes que celles utilisées dans la publication n°4.

2.2. Production des variants du snRNA U2 de *S. cerevisiae*

2.2.1. Délétion des différentes tige-boucles du snRNA U2 par la méthode PCR

Le protocole est donné dans le Tableau 12. La mutagenèse a été réalisée à partir du plasmide pT7-U2 (voir publication n°2). Des oligonucléotides, s'hybridant de part et d'autre de la (ou des) structure(s) tige-boucle(s) à déléter du snRNA U2 de *S. cerevisiae*, ont été utilisés afin de produire par PCR un plasmide linéarisé (Tableau 13). Ce produit linéaire, après avoir subi une déprotéinisation, une phosphorylation et avoir été purifié sur papier NA45, a été circularisé puis préparé en grande quantité après transformation de cellules *E. coli* TG1 ou XL1 blue.

2.2.2. Génération de mutations ponctuelles dans le snRNA U2 de *S. cerevisiae*

La région codante du snRNA U2 de *S. cerevisiae* a été clonée dans le phagemide pBS KS-. La méthode de Kunkel a été utilisée pour réaliser la mutagenèse dirigée (Kunkel, 1985; Kunkel *et al.*, 1991). L'oligonucléotide utilisé était dégénéré à 6 positions (Tableau 13). Les différents mutants du snRNA U2 obtenus ont été séquencés (Tabor et Richardson, 1987) et clonés dans le plasmide PUC18, sous contrôle du promoteur de l'ARN T7 polymérase.

Oligonucléotide N°	Séquence de l'oligonucléotide	Segments du snRNA U2 complémentaires à l'oligonucléotide	Méthode de mutagenèse utilisée
1224	5'GTTACTGAAANNANNNATACTA CACTTG 3'	(3' →5') nucléotides 29 à 59	Kunkel
1372	5'TTACCTTTTAATTTGTTACAA 3'	(3' →5') nucléotides 566 à 585 (délétion de la tige-boucle 2b)	PCR
1373	5'TTTCAGTTGTTACTAG AAA 3'	(5' →3') nucléotides 537 à 546 (délétion de la tige-boucle 2b)	"
1374	5'TGACCTCAATGAGGCTCATT 3'	(3' →5') nucléotides 547 à 567 (délétion de la tige-boucle 2a)	"
1375	5'AGAACAGATACTACTTGA 3'	(5' →3') nucléotides 507 à 526 (délétion de la tige-boucle 2a)	"
1376	5'TCAAGTGTAGTATCTGTTCTT 3'	(3' →5') nucléotides 507 à 527 (délétion de la tige-boucle 1)	"
1377	5'GATTCGTCCTATAGTGAGTCGT 3'	(5' →3') nucléotides 465 à 486 (délétion de la tige-boucle 1)	"

Tableau 13: Oligodéoxynucléotides utilisés pour les mutagenèses dirigées du gène du snRNA U2 de *S. cerevisiae*, qui ont été réalisées par la méthode de Kunkel ou par PCR.

L'oligonucléotide 1224 est dégénéré sur 6 positions (N).

2.2.3. Modification *in vitro* des variants obtenus par Pus1p et analyse des résidus Ψ formés

Les différents mutants du gène du snRNA U2 de *S. cerevisiae* ont été transcrits *in vitro* selon la méthode décrite dans la publication n°2. Deux µg de transcrits ont été chauffés à 80°C et incubés en présence d'environ 0,6 µg d'enzyme Pus1p purifiée pendant 2h à 37°C et les résidus Ψ formés ont été localisés par modification par le CMCT. Les conditions expérimentales utilisées étaient les mêmes que dans la publication n°2.

3. RESULTATS

3.1. Analyse de la structure du snRNA U2 de *S. cerevisiae*, produits *in vitro*, dans les conditions où les modifications par Pus1p sont réalisées

Comme nous l'avons mentionné dans l'introduction bibliographique (§B2.2.2b), sur la base de la séquence établie pour les snRNA U2 de différentes espèces, la possibilité de replier le snRNA U2 selon deux structures secondaires alternatives est conservée dans toutes les espèces. D'après des études chez *S. cerevisiae*, une seule de ces deux structures aurait un caractère essentiel sur le plan fonctionnel (forme B). Il n'est pas sûr que la forme A existe réellement *in vivo*. La conservation des appariements mis en jeu dans la forme A pourrait être liée à la nécessité de pouvoir former, dans la forme B, une structure en pseudonoeud (Fig. 54). Néanmoins, quelle que soit la structure du snRNA U2 (A, B ou B+pseudonoeud), la séquence renfermant le résidu U₄₄ converti en résidu Ψ par Pus1p est toujours en simple brin. Cette observation était étonnante sur la base des résultats que nous avons obtenus avec différents transcrits correspondant à des substrats naturels ou non naturels de Pus1p.

Aucune étude expérimentale de la structure secondaire d'un transcrit produit *in vitro* et correspondant au snRNA U2 de *S. cerevisiae* n'avait encore été réalisée. Les données de structure sur cet ARN provenaient d'une étude par le DMS, réalisée *in vivo*, qui suggérait l'existence d'une hétérogénéité *in vivo* (Ares et Igel, 1990; Zavanelli *et al.*, 1994). Des informations complémentaires avaient été obtenues par approche génétique et par comparaison phylogénétique. La structure 2D d'un ARN pouvant varier en fonction de la concentration en cations mono ou divalents, il était important d'étudier la structure 2D du transcrit que nous avons utilisé pour les études de modification *in vitro* par Pus1p, dans les conditions expérimentales utilisées pour ces études.

Comme pour l'étude de la structure des domaines de l'ARNr 26S de *S. cerevisiae*, produits *in vitro*, j'ai utilisé une sonde chimique (DMS) et deux sondes enzymatiques (nucléases T1 et V1). Un exemple d'analyse des positions de modification et de clivage, par extension d'amorce par la transcriptase inverse, est donné dans la Figure 55A. L'ensemble des résultats obtenus est représenté de manière schématique dans la Figure 55B.

Une forte sensibilité au DMS et à la RNase T1 est observée pour le segment compris entre les positions 30 et 47, qui renferme le résidu modifié par Pus1p. Nous confirmions bien son état simple brin dans le transcrit produit *in vitro*. Par ailleurs, la présence simultanée de modifications fortes par le DMS et de coupures fortes par la RNase V1, dans le segment compris entre les positions 54 à 60, reflétait la coexistence de transcrits présentant des structures alternatives. Les deux formes, A et B, devaient être présentes en équilibre avec, probablement, la forme comportant la structure en pseudonoeud, assurant la transition entre ces deux formes.

Sur la base de cette forte hétérogénéité des transcrits, nous ne pouvons pas exclure l'existence d'une forme où le dinucléotide UC en aval du résidu U₄₄ serait apparié avec le dinucléotide GA (positions 69 et 70), ceci dans le prolongement de l'hélice qui peut être formée entre les segments A et A' dans la forme B. Dans cette configuration, le résidu U₄₄ borderait une région hélicoïdale. Les résultats obtenus avec les sondes de structure ne permettent pas d'apporter de preuves claires de l'existence d'un tel conformère. En effet, le résidu C₄₆ est modifié par le DMS. Il est donc en simple brin dans au moins une partie des transcrits. Elles ne permettent pas non plus de rejeter cette hypothèse.

3.2. Les 139 nts en 5' du snRNA U2 de *S. cerevisiae* sont substrats de Pus1p

Le snRNA U2 de *S. cerevisiae* a la particularité d'être très long (1175 nt) par rapport aux snRNA U2 de vertébrés (187 nt) (pour revue, Guthrie et Patterson, 1988). La région 5'-terminale portant les nucléotides 1 à 124 est essentielle à l'activité du snRNA U2 dans l'épissage (Igel et Ares, 1988; McPheeters *et al.*, 1989). Nous avons testé si cette région était suffisante pour que Pus1p forme le résidu Ψ en position 44. Pour cela, nous avons produit par transcription *in vitro* un ARN correspondant à cette région 5'-terminale du snRNA U2 de *S. cerevisiae*. Cette transcription a été réalisée, par l'ARN T7 polymérase, à partir du plasmide pT7-U2 (voir publication n°2), linéarisé par l'enzyme EarI, dont le site de coupure se trouve dans la région codante du snRNA U2. Un ARN correspondant aux 139 nts en 5' du snRNA U2 a ainsi pu être produit.

Ces transcrits U2 tronqués ont été incubés avec Pus1p, une expérience contrôle étant réalisée en parallèle avec un transcrit U2 entier. Les résidus Ψ formés ont été localisés par extension d'amorce, après modification par le CMCT. Les résultats obtenus sont illustrés dans la Figure

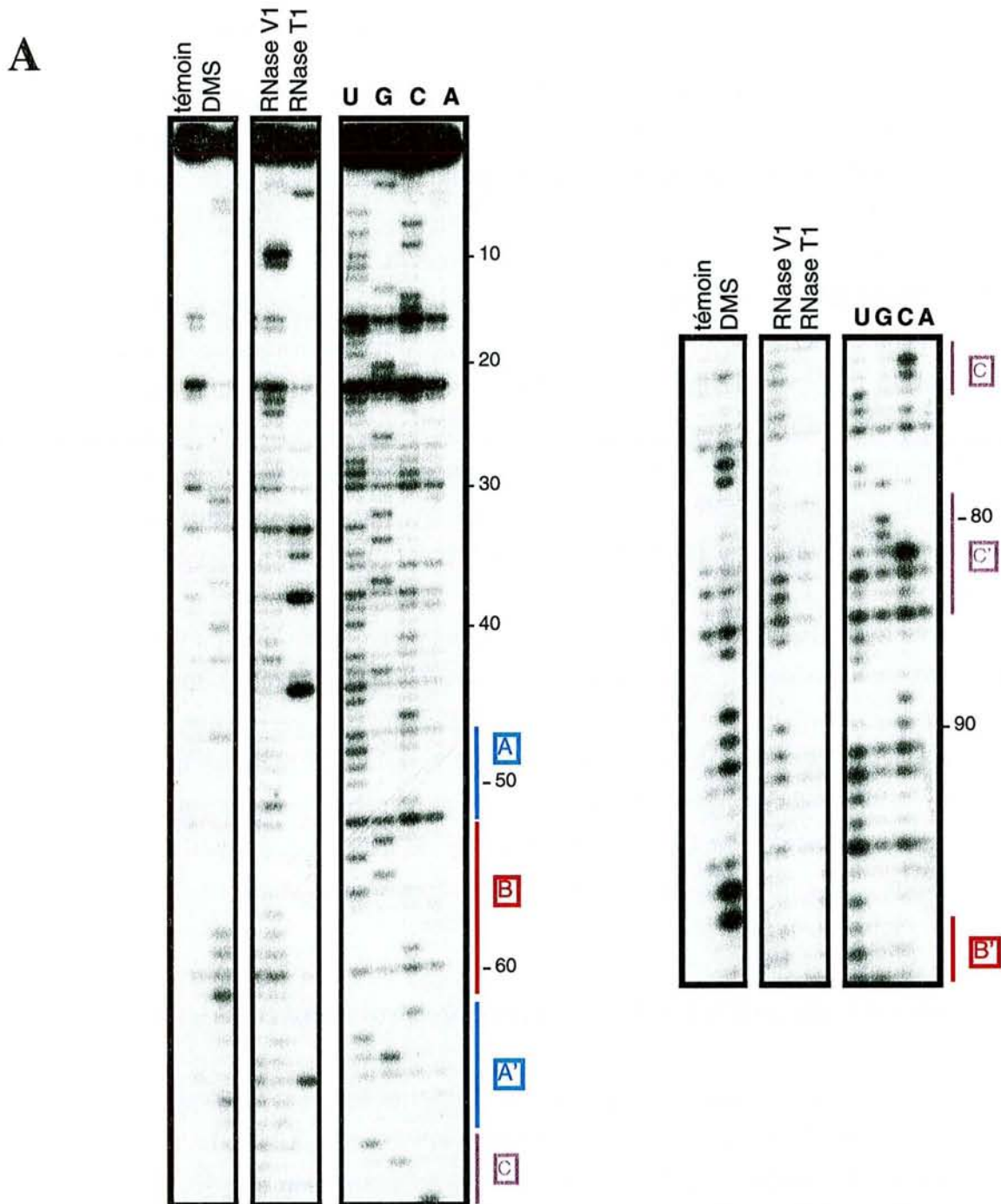
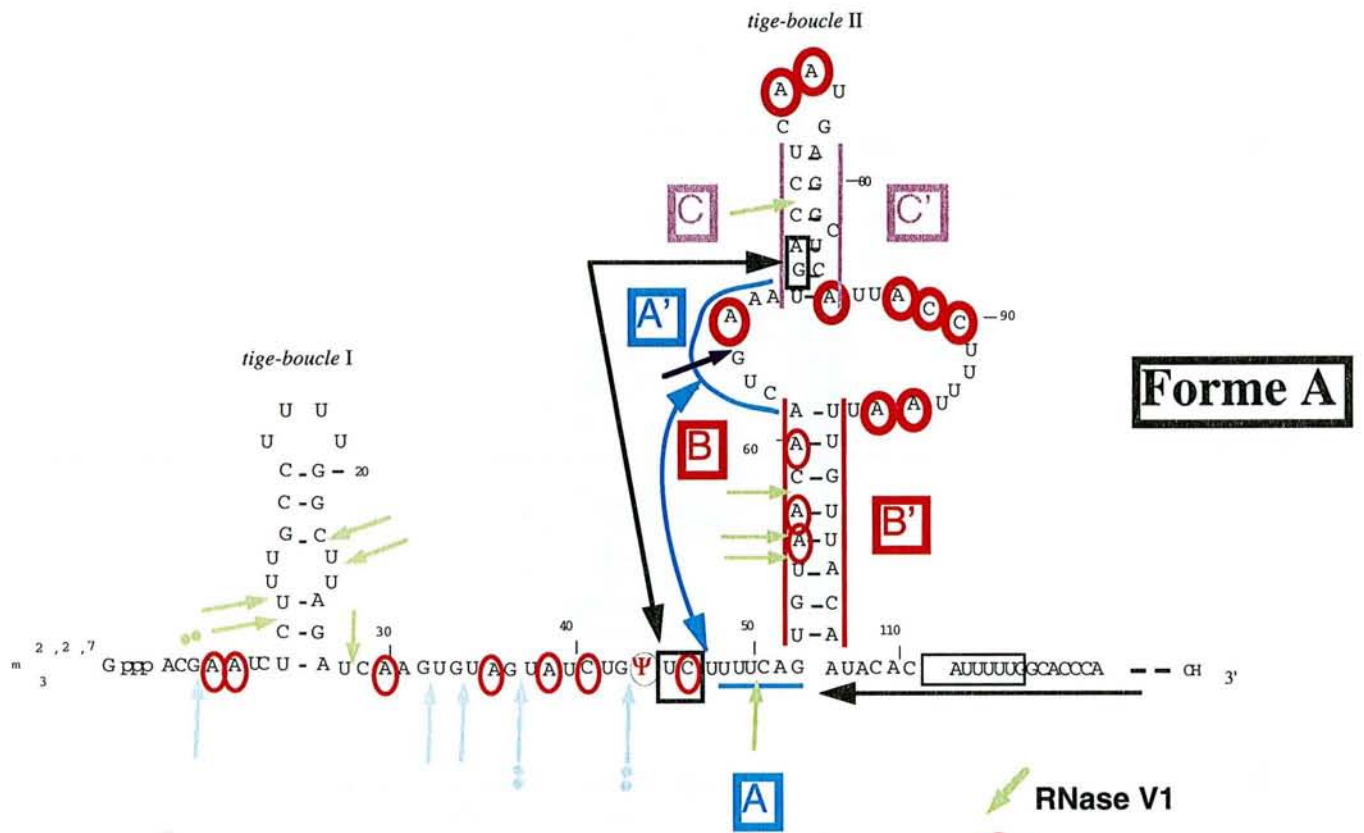


Figure 55: Etude de la structure secondaire de la région 3'-terminale du transcrit, correspondant au snRNA U2 de *S. cerevisiae*, dans les conditions utilisées pour les incubations avec Pus1p.

Panneau A: Analyse par la transcriptase inverse des modifications chimiques (colonnes DMS) et des coupures par le RNases T1 et V1 (colonnes T1 et V1, respectivement), portées par la région 3'-terminale du transcrit correspondant au snRNA U2 de *S. cerevisiae*. Un transcrit non modifié a servi de témoin pour détecter les pauses naturelles de la transcriptase inverse (colonne témoin). Les ADNc obtenus ont été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 7%, en parallèle avec un séquençage réalisé avec la transcriptase inverse (colonnes U, G, C et A) en utilisant le même oligonucléotide amorce, dont le site d'appariement sur le transcrit est indiqué dans le Panneau B. La position de nucléotides dans la séquence du snRNA U2 de *S. cerevisiae* est indiquée à droite de la fluorographie. Les positions de segments complémentaires (A-A', B-B', C-C') sont indiquées.

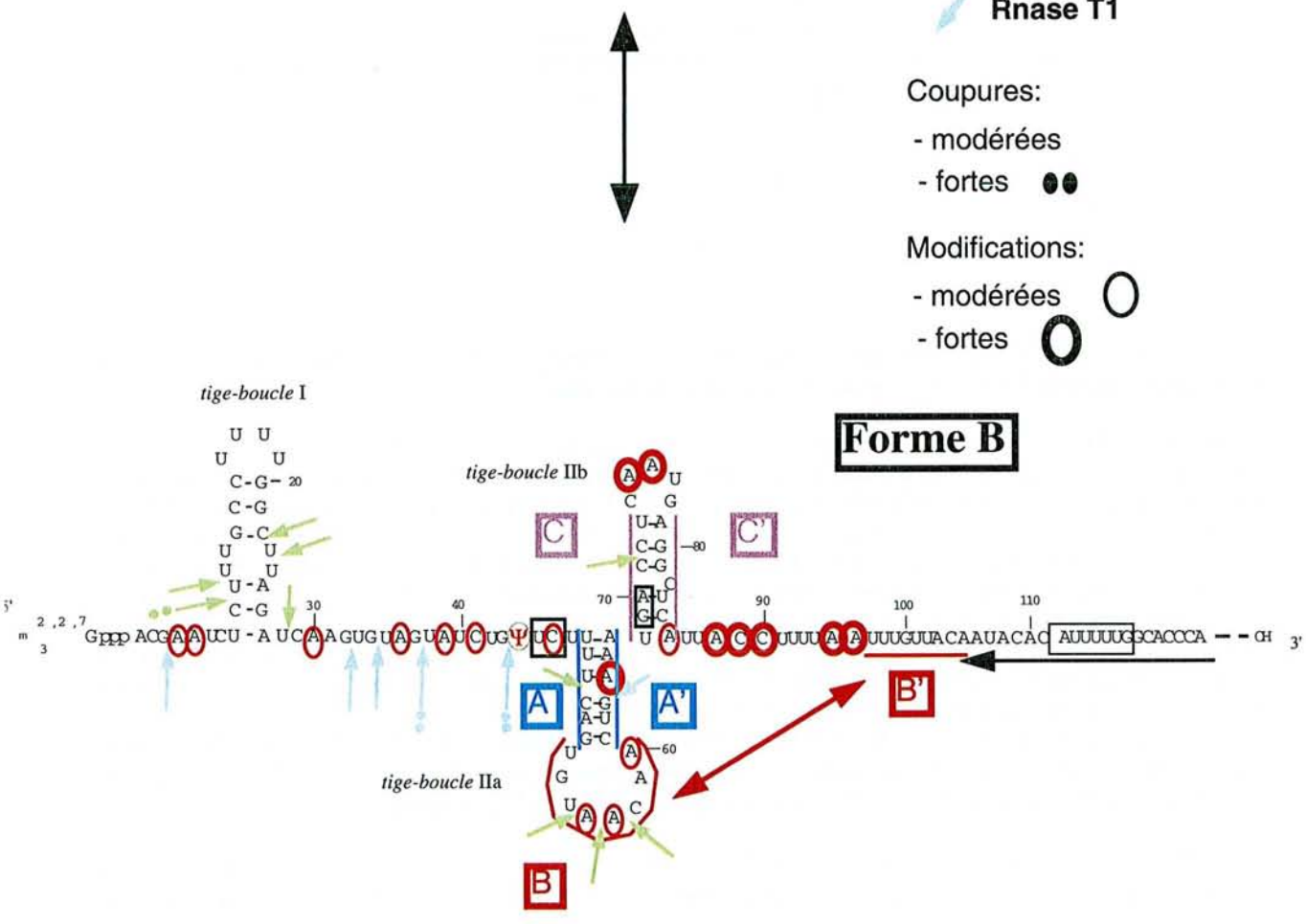
Panneau B: Représentation schématique des résultats obtenus après modification chimique par le DMS et coupure par le RNases T1 et V1 du transcrit correspondant au snRNA U2 de *S. cerevisiae*. Les deux structures possibles du snRNA U2 de *S. cerevisiae* sont représentées, ainsi que les possibilités d'interaction de type pseudonoeud (Ares et Igel, 1990; Zavanelli et Ares, 1991; Zavanelli *et al.*, 1994). Les différentes modifications et coupures sont indiquées. Le site d'appariement de l'oligonucléotide utilisé comme amorce est représenté.



RNase V1
 DMS
 Rnase T1

Coupures:
 - modérées
 - fortes ●●

Modifications:
 - modérées ○
 - fortes ○○



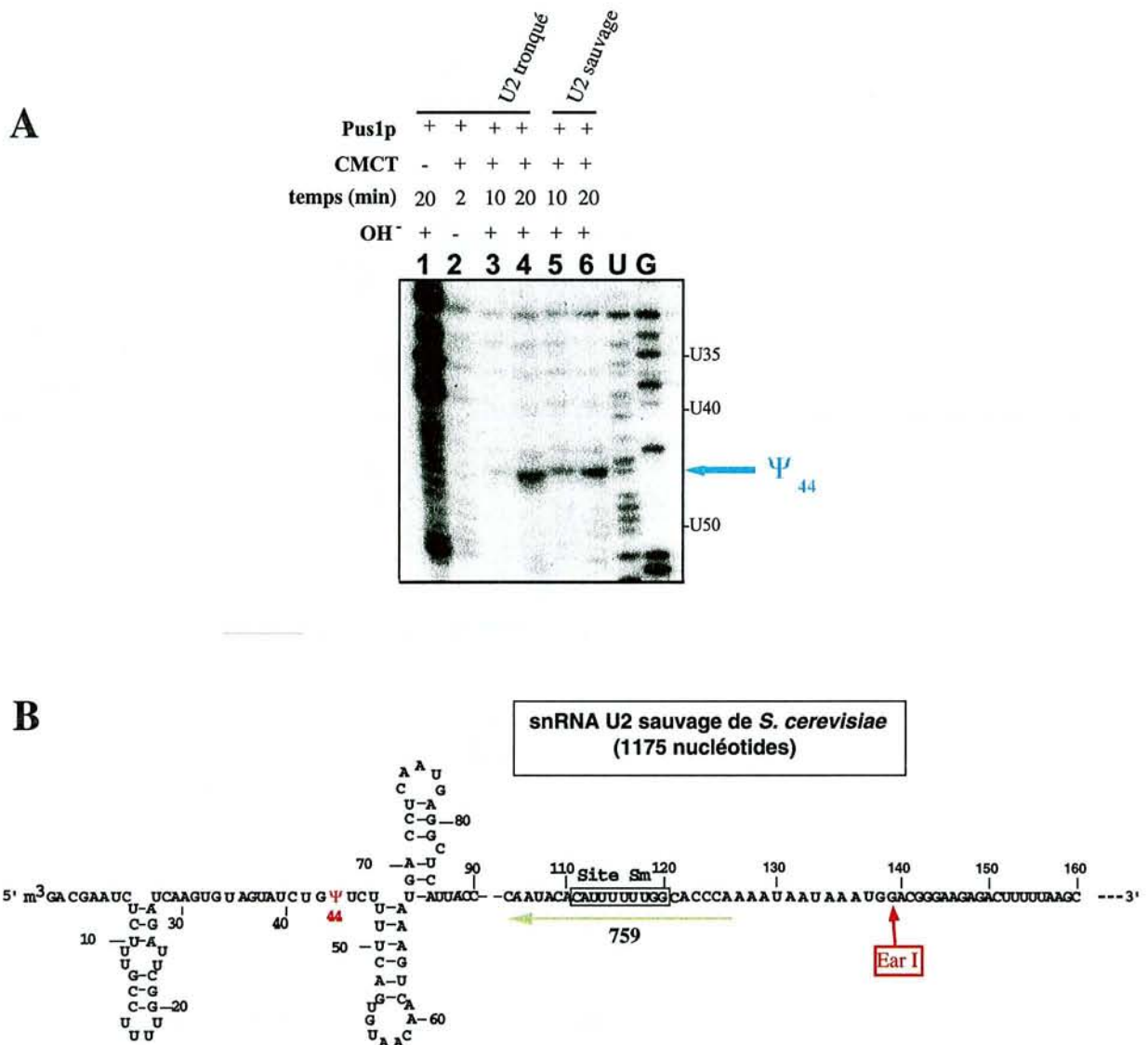


Figure 56: Formation des résidus Ψ par l'enzyme Pus1p dans des transcrits produits *in vitro* correspondant au snRNA U2 de *S. cerevisiae* entier ou tronqué.

Panneau A: L'étude a été réalisée sur des transcrits, produits *in vitro*, correspondants au snRNA U2 de *S. cerevisiae* entier (pistes 5 et 6) ou tronqué (139 nts 5'-terminaux, pistes 1 à 4). Les transcrits ont été incubés en présence de Pus1p dans le tampon RB5x. Un traitement par le CMCT a ensuite été réalisé pendant 2 min (piste 2), 10 min (pistes 3 et 5) ou 20 min (pistes 4 et 6). Dans les pistes 1 et 3 à 6, l'ARN modifié ou non par le CMCT a été soumis à un traitement alcalin à pH 10,4 pendant 3h. Une expérience avec omission du traitement par le CMCT a été réalisée en tant que témoin (piste 1). Les bases modifiées par le CMCT ont été identifiées par les arrêts qu'elles génèrent lors d'une extension d'amorce par la transcriptase inverse. L'oligonucléotide 759 était utilisé comme amorce. Les ADNc obtenus ont été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 7%, en parallèle avec les produits des réactions de séquençage, du transcrit U2 de *S. cerevisiae*, par la transcriptase inverse permettant l'identification des résidus U et G.

Après traitement de 3h au carbonate de sodium pH 10,4, seule la modification portée par le résidu 44 (indiqué par la flèche) persiste à un taux élevé et conduit à un arrêt de la transcriptase inverse. Un résidu Ψ est donc présent à cette position, à la fois dans le transcrit U2 entier et dans le transcrit U2 tronqué.

Panneau B: Structure de la région 5'-terminale du transcrit U2 entier (1175 nts). La région où s'apparie l'oligonucléotide, utilisé comme amorce, est soulignée par une flèche. Le résidu Ψ_{44} et la séquence Sm sont indiqués. La position 139, correspondant à l'extrémité 3' des transcrits produits après coupure du plasmide PT7-U2 par l'enzyme EarI, est indiquée par une flèche rouge.

56. Ils montrent que le résidu Ψ_{44} est formé à un taux similaire dans le snRNA U2 tronqué et dans le snRNA U2 entier. Les 139 nts en 5' du snRNA U2 sont donc suffisants à la reconnaissance par Pus1p et pour la catalyse de la formation du résidu Ψ en position 44. Comme pour la modification des ARNt, la structure globale du snRNA U2 n'est donc pas nécessaire à sa modification par Pus1p.

3.3. Influence des structures tige-boucles localisées à proximité de la position 44 sur la reconnaissance et la catalyse par Pus1p

Etant donné que Pus1p semble reconnaître des structures tige-boucles dans les ARNt, nous avons testé l'effet de la délétion des structures tige-boucles localisées à proximité de la position 44, sur la reconnaissance par Pus1p du snRNA U2. Les régions codant les 3 structures tige-boucles, présentes dans la partie 5'-terminale du snRNA U2 de *S. cerevisiae*, ont été délétées une à une ou en combinaison, au sein du gène U2 porté par le plasmide PT7-U2. La stratégie PCR utilisée pour réaliser les délétions successives des régions codant les structures tige-boucles I, IIa et IIb du snRNA U2 est représentée dans la Figure 57. Après linéarisation des plasmides, 4 transcrits différents correspondant au snRNA U2 entier portant seulement certaines structures tige-boucles ont été produits. L'ARN U2 Δ I ne comporte plus la structure tige-boucle I, l'ARN U2 Δ IIa, la structure tige-boucle IIa, l'ARN U2 Δ IIb, la structure tige-boucle IIb et l'ARN U2 Δ IIaIIb ne comporte plus que la structure tige-boucle I. Les positions modifiées *in vitro* par Pus1p, dans les ARN variants produits par transcription *in vitro*, ont été analysées par extension d'amorce, après traitement par le CMCT. Les résultats obtenus pour les transcrits U2 Δ IIa, U2 Δ IIb et U2 Δ IIaIIb sont illustrés dans les Figures 58 à 60. La Figure 61 résume l'ensemble des données obtenues par cette approche.

Le résidu Ψ_{44} est formé en absence de la structure tige-boucle I, en absence de la structure tige-boucle IIa et en absence de la structure tige-boucle IIb. Cependant, dans ce dernier cas, nous constatons la formation d'un second résidu Ψ en position 56. Enfin, aucun résidu n'est formé lorsque les deux structures tige-boucles IIa et IIb sont délétées. Il semblerait que la formation *in vitro* du résidu Ψ_{44} dans le snRNA U2 par Pus1p nécessite la présence d'au moins une structure tige-boucle à proximité du site cible.

3.4. Influence de l'identité des nucléotides autour de la position 44 sur la reconnaissance et la catalyse par Pus1p

Afin de tester l'importance, sur la reconnaissance par Pus1p, de l'identité des résidus situés autour de la position 44, nous avons produit une série de variants du snRNA U2 de

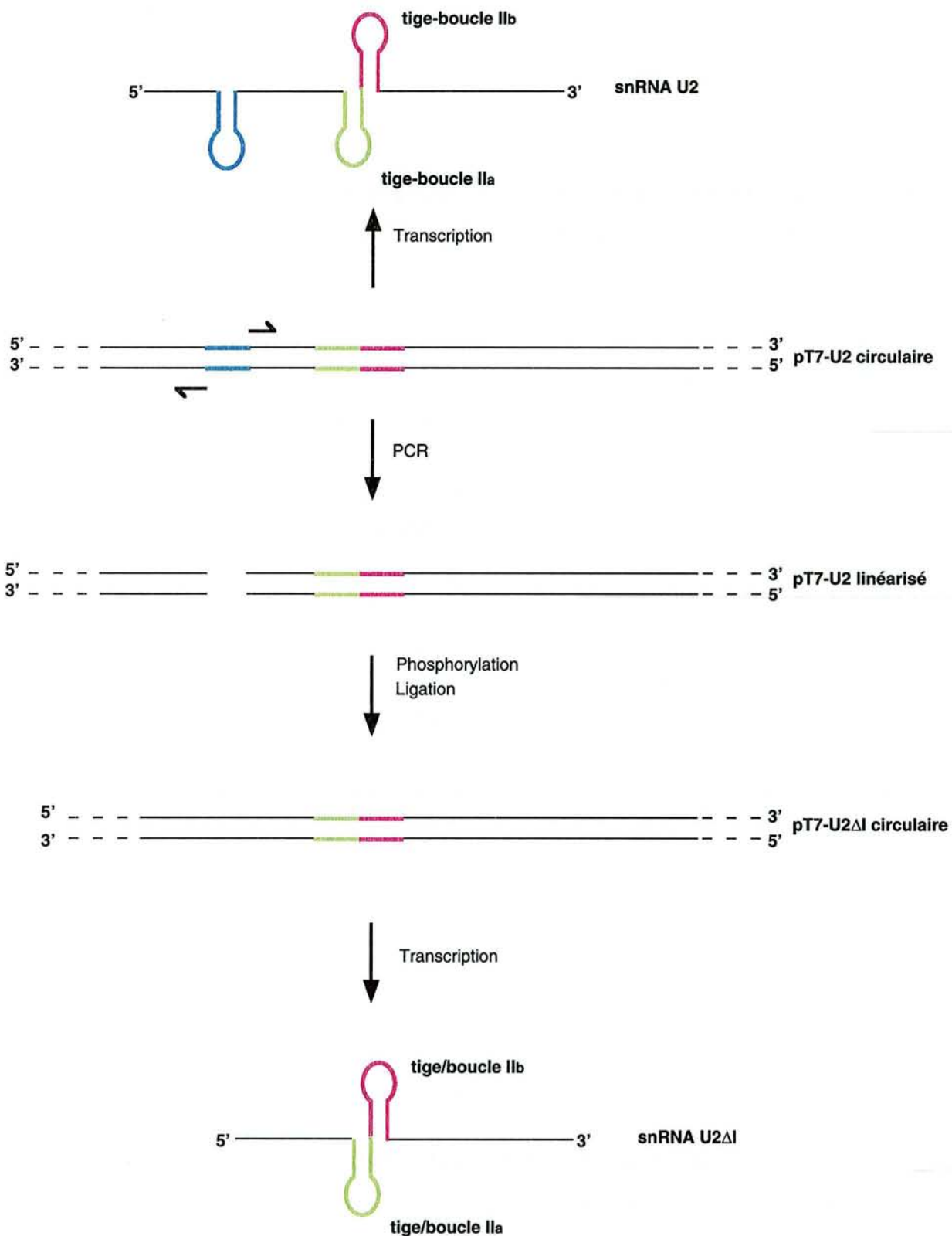


Figure 57: Principe de la délétion des régions codant les structures tige-boucles du snRNA U2 de *S. cerevisiae* par la méthode PCR.

Les séquences plasmidiques colorées correspondent aux séquences codant les structures tige-boucles dans le snRNA U2. Le code couleur est conservé dans la représentation des ARN: bleu, structure tige-boucle I; vert, structure tige-boucle IIa; rose, structure tige-boucle IIb.

A.

Pus1p	-	-	-	-	+	+				
CMCT	-	+	+	+	+	+				
temps (min)	20	2	10	20	10	20				
OH ⁻	+	-	+	+	+	+				
	1	2	3	4	5	6	U	G	C	A

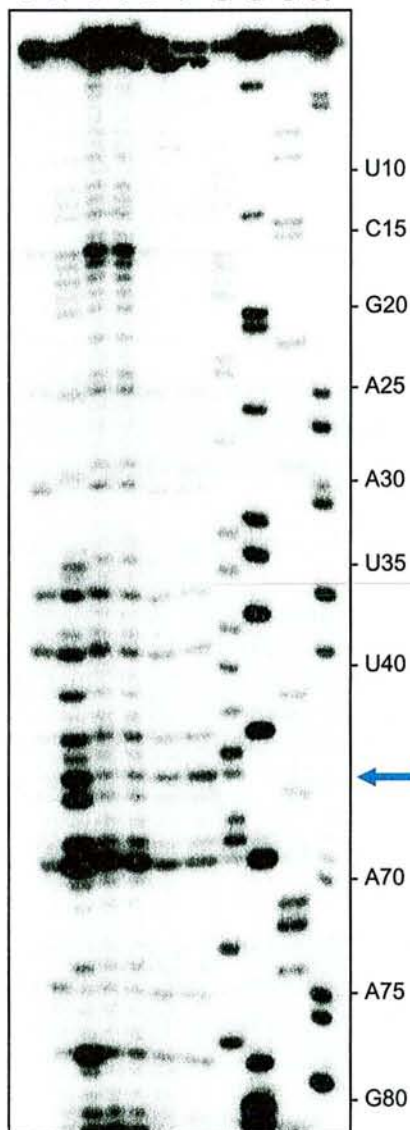
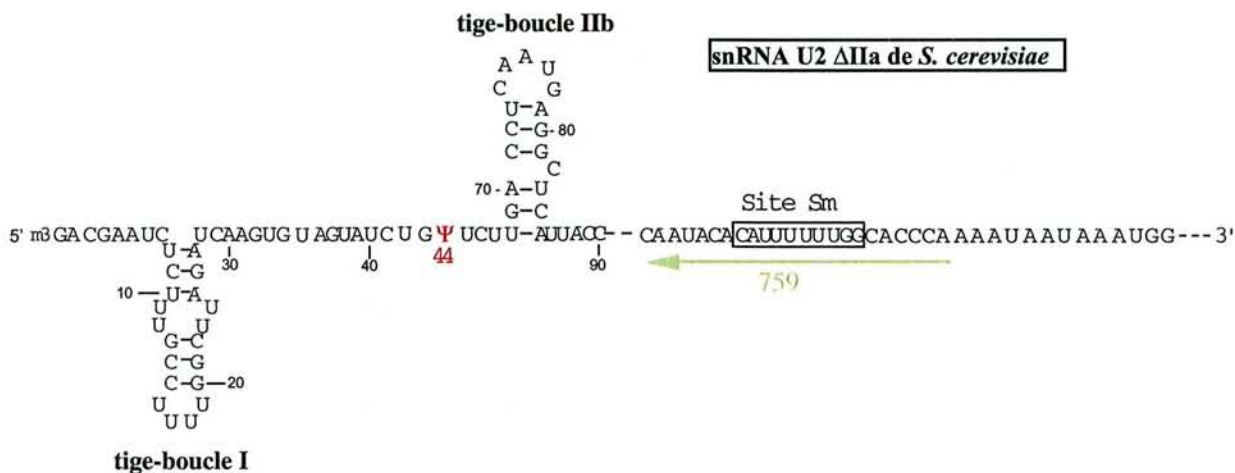


Figure 58: Formation des résidus Ψ par l'enzyme Pus1p dans le transcrit U2 Δ IIa de *S. cerevisiae*, dans lequel la tige-boucle IIa a été déléetée.

Panneau A: Les analyses ont été réalisées sur le transcrit U2 Δ IIa (représenté dans le panneau B), produit par transcription *in vitro* par l'ARN T7 polymérase. Ce transcrit a été incubé en présence (pistes 5 et 6) ou en absence (pistes 1 à 4) de Pus1p recombinante. Le traitement par le CMCT a été réalisé pendant 2 min (piste 2), 10 min (pistes 3 et 5) ou 20 min (pistes 4 et 6), accompagné (pistes 1 et 3 à 6) ou non (piste 2) d'un traitement alcalin. La présence de résidus Ψ a été détectée par extension d'amorce par la transcriptase inverse. La numérotation adoptée est celle du snRNA U2 sauvage.

Panneau B: Structure proposée pour le transcrit U2 Δ IIa. La région d'appariement de l'oligonucléotide utilisé est représentée, ainsi que le résidu Ψ_{44} .

B.



A. Pus1p + + +
 CMCT - + +
 temps (min) 20 10 20
 OH⁻ + + +

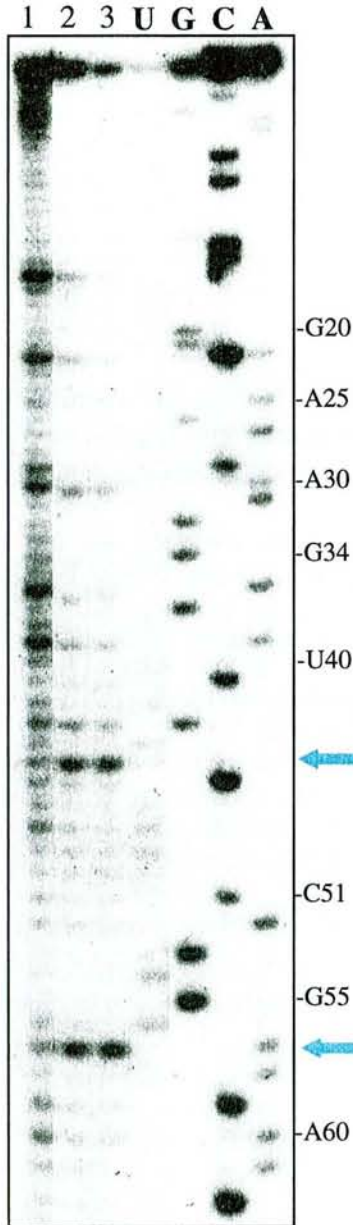
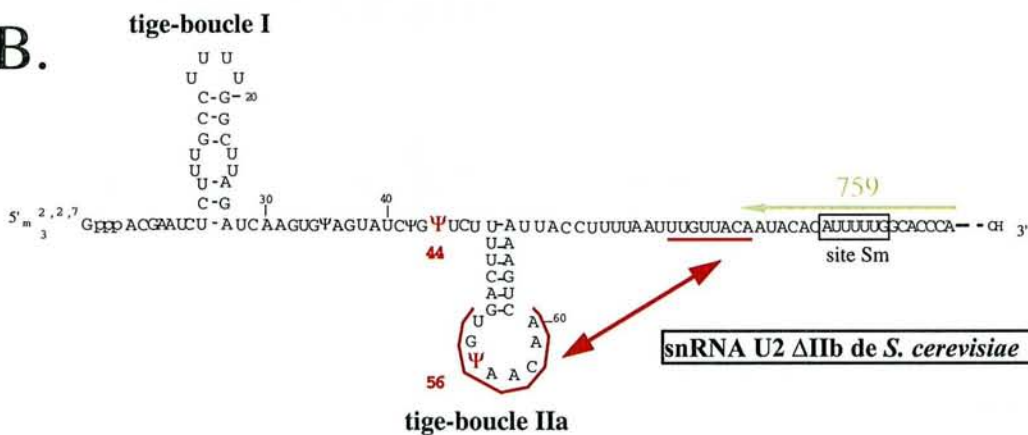


Figure 59: Formation des résidus Ψ par l'enzyme Pus1p dans le transcrit U2 Δ IIb de *S. cerevisiae*, dans lequel la tige-boucle IIb a été déléetée.

Panneau A: Les analyses ont été réalisées sur le transcrit U2 Δ IIb (représenté dans le panneau B), produit par transcription *in vitro* par l'ARN T7 polymérase. Ce transcrit a été incubé en présence de Pus1p recombinante. Le traitement par le CMCT a été réalisé pendant 10 min (piste 2) ou 20 min (piste 3), accompagné d'un traitement alcalin. La présence de résidus Ψ a été détectée par extension d'amorce par la transcriptase inverse. La numérotation adoptée est celle du snRNA U2 sauvage.

Panneau B: Structure proposée pour le transcrit U2 Δ IIb. La région d'appariement de l'oligonucléotide utilisé est représentée, ainsi que les résidus Ψ_{44} et Ψ_{56} .

B.



A.

Pus1p	-	-	-	-					+	+	+	+
CMCT	-	+	+	+					-	+	+	+
	20	2	10	20					20	2	10	20
OH	+	-	+	+					+	-	+	+

1 2 3 4 U G C A 5 6 7 8

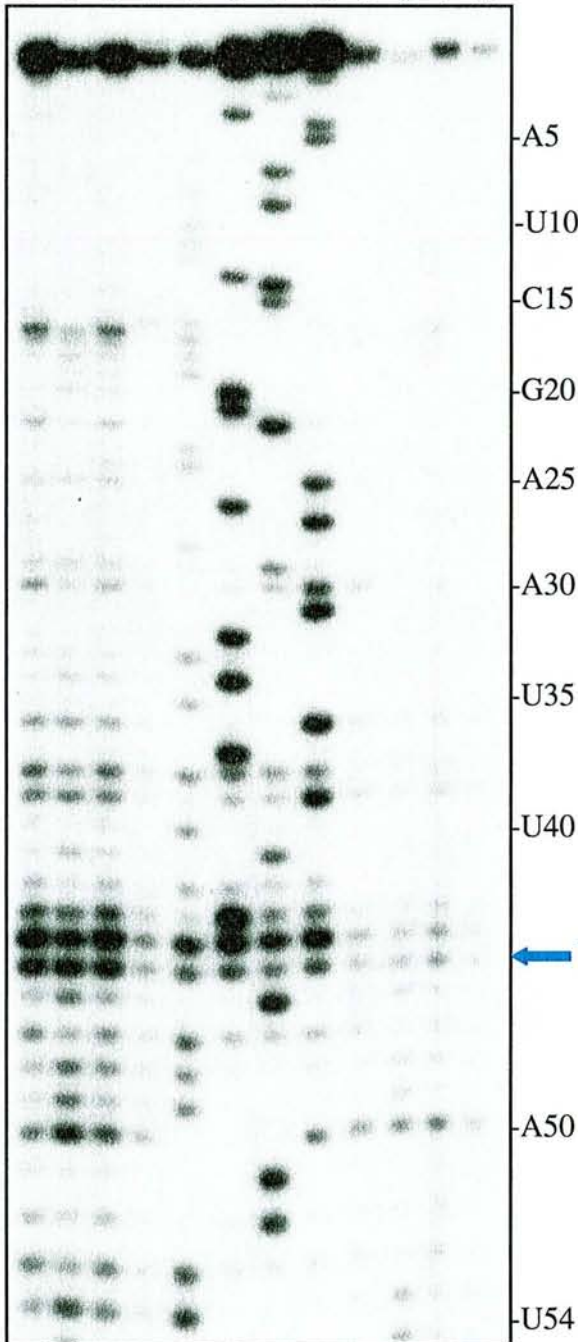
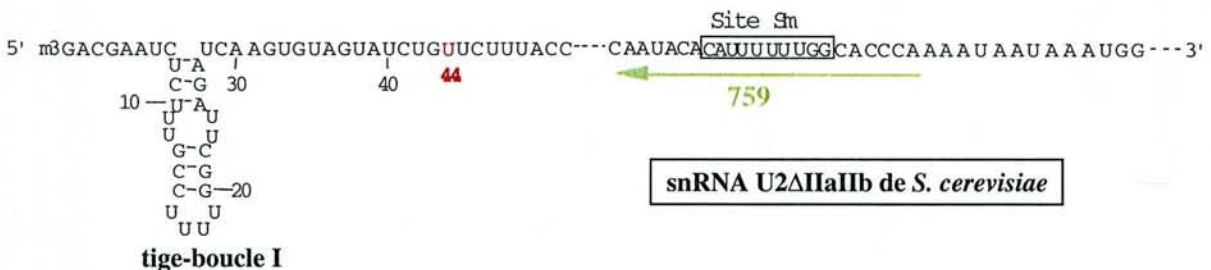


Figure 60: Formation des résidus Ψ par l'enzyme Pus1p dans le transcrit $U2\Delta IIa IIb$ de *S. cerevisiae*, dans lequel les tige-boucles IIa et IIb ont été délétées.

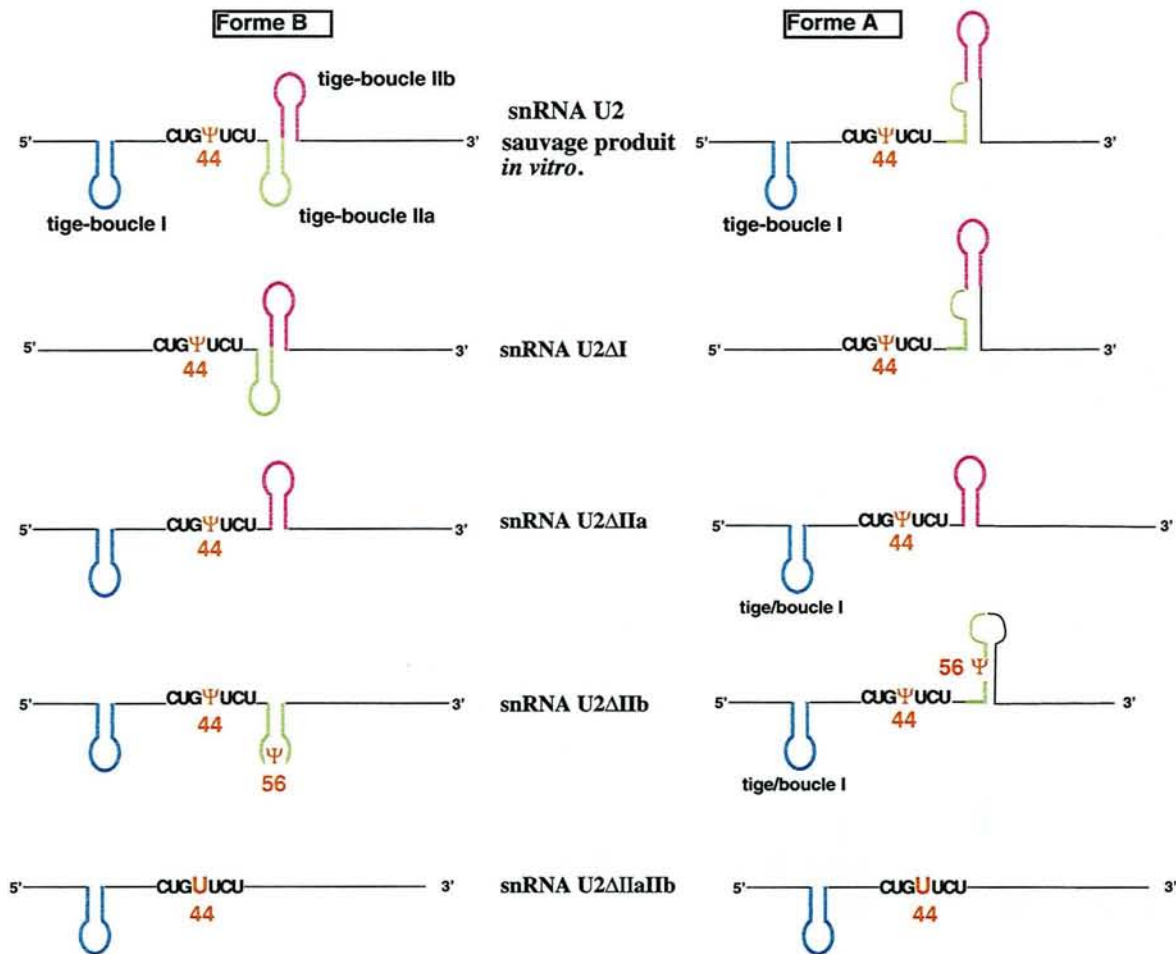
Panneau A: Les analyses ont été réalisées sur le transcrit $U2\Delta IIa IIb$ (représenté dans le panneau B), produit par transcription *in vitro* par l'ARN T7 polymérase. Ce transcrit a été incubé en présence (pistes 5 à 8) ou en absence (pistes 1 à 4) de Pus1p recombinante. Le traitement par le CMCT a été réalisé pendant 2 min (pistes 2 et 6), 10 min (pistes 3 et 7) ou 20 min (pistes 4 et 8), accompagné (pistes 3, 4, 7 et 8) ou non (pistes 2 et 6) d'un traitement alcalin. La présence de résidus Ψ a été détectée par extension d'amorce par la transcriptase inverse. La numérotation adoptée est celle du snRNA U2 sauvage.

Panneau B: Structure proposée pour le transcrit $U2\Delta IIa IIb$. La région d'appariement de l'oligonucléotide utilisé est représentée.

B.



A/ Délétions des structures tige-boucles localisées à proximité de la position 44.



B/ Mutations des nucléotides encadrant la position 44.

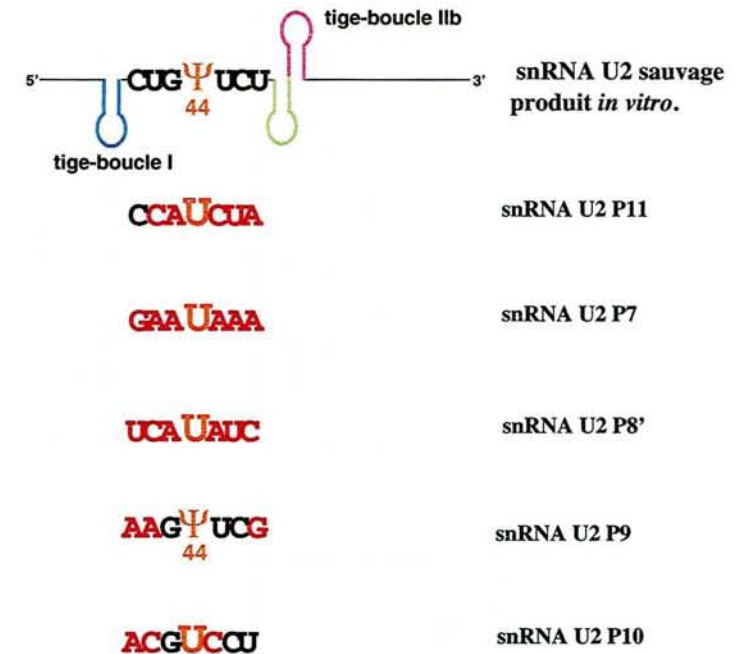


Figure 61: Récapitulatif des résultats obtenus lors de l'étude de la spécificité de la pseudouridine synthase recombinante Pus1p de *S. cerevisiae*. Les résidus Ψ formés, *in vitro*, par la protéine Pus1p recombinante sont indiqués en orange. La position des modifications détectées sont indiquées en se référant à la numérotation du snRNA U2 sauvage. Dans le panneau A, les différents mutants de délétion sont représentés en utilisant les structures A et B du snRNA U2 sauvage et en supposant que les structures tige-boucles ne sont pas modifiées par les délétions. Dans le panneau B, les positions mutées sont indiquées en rouge et les positions identiques à la séquence sauvage en noir.

S. cerevisiae par mutagenèse dirigée en utilisant la méthode de Kunkel et un oligonucléotide dont les séquences faisant face aux positions 41 à 43 et 45 à 47 étaient dégénérées. Nous avons ainsi obtenu un grand nombre de variants, mais le degré de dégénérescence de l'oligonucléotide étant élevé, tous les variants obtenus renfermaient plusieurs substitutions de bases (Fig. 62). De ce fait, seuls quelques-uns de ces variants nous ont semblés intéressants pour étudier la spécificité de Pus1p. Nous avons choisi d'étudier, dans un premier temps, des variants présentant une purine en 5' du résidu U₄₄ cible de Pus1p. Le but de cette étude était de savoir si la présence d'un résidu purine en position 43 était suffisante pour la modification de la position 44 par Pus1p, ou si l'identité d'autres nucléotides autour du résidu cible était également importante.

Pour cela, nous avons analysé deux variants (U2 P7 et U2 P8') pour lesquels des substitutions étaient présentes aux 6 positions cibles de la mutagenèse, un variant (U2 P11) où 5 positions présentaient une substitution de base et 2 variants (U2 P9 et U2 P10) ne présentant que 3 substitutions de bases (Fig. 61 et 62). L'enzyme Pus1p n'est pas active sur les transcrits P7, P8', P10 et P11 (Fig. 61). Par contre, Pus1p modifie le transcrit P9, le taux de modification est néanmoins plus faible que dans le cas de l'ARN ayant la séquence sauvage. Ces résultats suggèrent que, dans le cas du snRNA U2, la présence d'un résidu G est nécessaire mais pas suffisante. En effet, la modification n'est observée que pour l'un des deux mutants (P9) où le résidu G en 5' est préservé (Fig. 61). En plus, dans ce mutant P9, le résidu en 3' est conservé, ce qui n'est pas le cas dans le mutant P10 non modifié. Il semble donc que la présence d'un résidu U en 3' soit également importante.

Séquence U2 sauvage: 5'... .G U A U C U G U U C U U U U...3'

44

Séquence des variants
du snRNA U2:

P3'	...C U G U G C U...	
P4'	...C U A U U C G...	
P5'	...C U G U A C A..	
P6'	...G C C U U C U...	
P7'	...C C A U U U U...	
P8'	...U C A U A U C...	*
P9'	...C A G U G C G...	
P10'	...G G U U G C A..	
P6	...C U G U C G A..	
P7	...G A A U A A A..	*
P8	...C U G U A U U...	
P9	...A A G U U C G...	*
P10	A C G U C C U...	*
P11	...C C A U C U A..	*
P12	...C C C U A C A..	

Figure 62: Les snRNA U2 variants obtenus par mutagenèse avec un oligonucléotide dégénéré.
Le résidu U₄₄ est indiqué en vert. Les résidus modifiés par rapport à la séquence sauvage sont indiqués en rose.

* variants étudiés lors de ce travail.

4. DISCUSSION ET PERSPECTIVES

L'étude de la spécificité de la Ψ -synthase Pus1p, vis-à-vis des ARNt et de leurs précurseurs, ainsi que de différents ARN qui ne sont pas des cibles naturelles, mettait en évidence, comme nous l'avons vu, la nécessité d'une structuration particulière de l'ARN, à savoir deux structures tige-boucles contiguës, pour qu'il y ait reconnaissance et modification de l'ARN. La présence d'un résidu purine en 5' du résidu U cible semblait également importante pour la reconnaissance *in vitro* par Pus1p. Dans l'état actuel de notre étude, il ne semble pas que ces critères de reconnaissance s'appliquent dans le cas du snRNA U2.

Il semble donc que la spécificité de Pus1p vis-à-vis du snRNA U2 repose sur des bases différentes de celle observée vis-à-vis des ARNt. En premier lieu, la dépendance vis-à-vis de l'identité des résidus en bordure de l'uridine en position 44 serait plus grande. La Ψ -synthase semblerait en effet sensible à l'identité du résidu en 5' de la position U₄₄, mais également à celle du résidu en 3'. De plus, la présence en 5' d'un résidu G semble plus favorable que la présence d'un résidu A. Jusqu'à présent, dans les mutants produits, plusieurs substitutions de bases simultanées étaient présentes. Il sera important de réaliser des mutations ponctuelles des résidus G₄₃, U₄₅ et C₄₆ et d'analyser la modification de ces variants par Pus1p *in vitro*. Il restera à savoir si la dépendance de la modification par Pus1p vis-à-vis des nucléotides en bordure du résidu U₄₄ est une dépendance vis-à-vis de la séquence primaire ou vis-à-vis de la structure secondaire du snRNA U2. En effet, selon les modèles de structure secondaire admis pour le snRNA U2 de *S. cerevisiae* (Ares et Igel, 1990; Zavanelli et Ares, 1991; Zavanelli *et al.*, 1994), les trois résidus en aval du résidu U₄₄ sont en simple brin. Ceci a été confirmé par notre étude de la structure secondaire du transcrit du gène du snRNA U2, produit *in vitro*. Cependant, comme nous l'avons mentionné, on ne peut pas exclure la possibilité d'un appariement des résidus U₄₅C₄₆ et G₆₉A₇₀ dans certaines des molécules. Dans cette configuration, le résidu 44 est situé immédiatement en bordure d'une l'hélice, comme c'est le cas pour certains résidus cibles de Pus1p dans les ARNt. Cependant, un argument contre l'importance que pourrait avoir une telle structuration est apporté par le variant U2 Δ Iib. En effet, dans ce mutant, l'appariement entre les résidus U₄₅C₄₆ et G₆₉A₇₀ ne peut se former puisque l'ensemble de la structure tige-boucle Iib a été éliminé, pourtant le résidu Ψ ₄₄ est formé, *in vitro*, par Pus1p.

Le requis au niveau structure secondaire pour la reconnaissance et la modification des ARNt et du snRNA U2 semble donc différent. Nos résultats suggèrent bien la nécessité d'au moins une structure tige-boucle en aval du site cible dans le snRNA U2, puisqu'aucune modification n'est observée après délétion des structures tige-boucles Iia et Iib. Mais, aucune des structures alternatives que peut prendre le snRNA U2 de *S. cerevisiae* ne comporte le motif en deux structures tige-boucles encadrant le site cible, observé pour les ARNt.

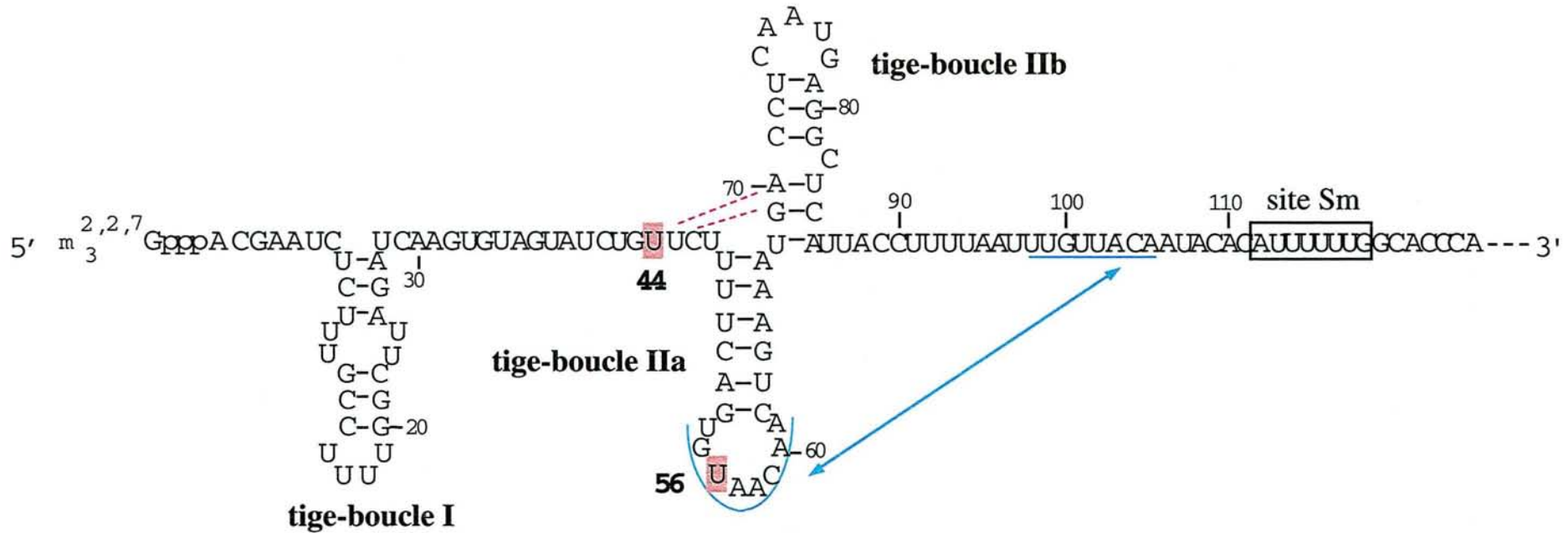


Figure 63: Appariements possibles au sein du snRNA U2 de *S. cerevisiae*, pouvant être impliqués dans la reconnaissance du résidu 44 comme cible de Pus1p dans l'ARN sauvage et du résidu 56 comme cible dans le variant U2ΔIIb.

(D'après Igel et Ares, 1988; Ares et Igel, 1990; Zavanelli *et al.*, 1994)

La possibilité de former un appariement entre les résidus 45/46 et 69/70 est indiquée en rose, la structure de type pseudonoed qui peut être formée entre les résidus 54 à 61 et 98 à 105 est indiquée en bleu.

Avec le variant U2 Δ I**b**, nous avons observé la formation *in vitro* par Pus1p du résidu Ψ_{44} , mais également d'un autre résidu Ψ en position 56. Ce résidu Ψ_{56} est bordé d'un résidu G en 5' et d'un résidu A en 3'. L'analyse expérimentale de la structure secondaire du mutant U2 Δ I**b**, permettra de déterminer quel est l'environnement structural de la position 56. En effet, la boucle comprenant les nucléotides 54-61 peut interagir avec la séquence 98-105, formant ainsi une structure en pseudonœud dans la forme B ou une hélice dans la forme A (Fig. 63). Il nous faudra donc déterminer si, en éliminant la structure tige-boucle I**b**, l'appariement entre les segments 54-61 et 98-105 est renforcé ou, au contraire, limité. Le renforcement de la formation d'une structure de type pseudonœud est vraisemblable, ce qui expliquerait peut-être la génération d'un site similaire à ceux que nous avons détectés dans les domaines des ARNr, produits *in vitro*.

Les travaux que nous avons réalisés sur la spécificité de Pus1p vis-à-vis du snRNA U2 sont préliminaires. Il est important de les conforter par une étude de la structure secondaire de chaque variant et par la production et l'analyse de variants, pour lesquels les nucléotides entourant la position 44 seront mutés ponctuellement. Par ailleurs, nous avons pour l'instant uniquement testé l'effet de délétions simultanées des structures tige-boucles I**a** et I**b**. Nous avons déjà produit les variants où les structures tige-boucles I et I**a**, d'une part, et I et I**b**, d'autre part, ont été délétées. Nous devons donc tester l'action de Pus1p sur ces variants. Enfin, dans ces travaux, nous avons testé si l'enzyme est active ou inactive vis-à-vis des différents variants. Il serait important, lorsque l'enzyme est active sur un variant, d'apporter une information sur la variation d'efficacité. En effet, il semblerait que la formation du résidu Ψ_{44} dans les variants U2 Δ I et U2 Δ I**a** par Pus1p ait lieu avec une efficacité moindre par rapport au snRNA U2 sauvage. Nous pouvons envisager de réaliser des expériences de compétition entre variants en insérant des étiquettes différentes à l'extrémités 3' de chaque variant, permettant ainsi leur analyse spécifique au sein d'un mélange de variants modifiés par Pus1p, ceci en utilisant des oligonucléotides complémentaires de chaque étiquette. Il serait également intéressant de tester la formation *in vivo* du résidu Ψ_{44} dans ces différents variants, en transformant une souche sauvage de *S. cerevisiae* avec des plasmides portant un des variants du gène du snRNA U2. La pseudouridylation *in vivo* de ce variant par Pus1p pourra être également analysée avec un oligonucléotide spécifique d'une étiquette portée par ce variant.

En conclusion, la Ψ -synthase Pus1p semble donc se fixer sur tous les ARNt et ceci, même en l'absence de résidus U cibles (Arluison *et al.*, 1999). Ensuite, elle modifie les résidus U localisés à la jonction de deux hélices. Cette Ψ -synthase modifie également le snRNA U2, mais les règles qui définissent la sélection du seul résidu U modifié *in vivo* et *in vitro* ne sont pas encore claires. On peut espérer qu'une étude cristallographique des complexes formés par Pus1p et ces différents substrats permettra un jour de mieux comprendre les bases de cette différence de reconnaissance de substrat entre Pus1p et les ARNt, d'une part, et Pus1p et le snRNA U2, d'autre part. Jusqu'à présent, seule la Ψ -synthase d'*E. coli* RluC a été cristallisée, et ceci en absence de son substrat, mais l'analyse des cristaux obtenus n'a pas encore été réalisée (Corollo *et al.*, 1999).

C/ ETUDE COMPARATIVE DES Ψ -SYNTHASES Pus1p DE *S. cerevisiae* ET DE SOURIS

1. INTRODUCTION

Très récemment, l'équipe de J. Patton a cloné et caractérisé une Ψ -synthase chez la souris, présentant 37% d'identité avec l'enzyme Pus1p de *S. cerevisiae* (Chen et Patton, 1999). Nous nommerons, par la suite, cette enzyme mPus1p, et l'enzyme correspondante de *S. cerevisiae* ScPus1p. La Ψ -synthase mPus1p a été ensuite exprimée chez *E. coli*. L'enzyme recombinante obtenue porte une étiquette polyhistidine à son extrémité N-terminale, ce qui a permis à Chen et Patton de la purifier sur colonne d'agarose Ni-NTA. L'activité de cette Ψ -synthase a ensuite été testée *in vitro* sur différents ARNt de *S. cerevisiae* et sur le pré-ARNt^{Ser} humain, produits *in vitro*, ceci afin de comparer sa spécificité avec celle de ScPus1p (Fig. 64). Cette analyse montre que mPus1p catalyse *in vitro* la formation du résidu Ψ_{27} dans les ARNt^{Val} et pré-ARNt^{Ile} de levure, des résidus Ψ_{34} et Ψ_{36} dans le pré-ARNt^{Ile} de levure, ainsi que du résidu Ψ_{28} dans le pré-ARNt^{Ser} humain (Chen et Patton, 1999). Ces 4 positions sont également modifiées *in vitro* par ScPus1p (Simos *et al.*, 1996; Motorin *et al.*, 1998). De plus, mPus1p ne catalyse pas *in vitro* la formation des résidus Ψ en positions 26 et 67 dans les différents ARNt de *S. cerevisiae* analysés (Chen et Patton, 1999). Ces positions ne sont également pas modifiées *in vitro* par l'enzyme ScPus1p, mais cette enzyme pourrait être responsable de leur formation *in vivo* (Motorin *et al.*, 1998). Chez *S. cerevisiae*, les résidus Ψ en positions 39 et 55 sont formés, respectivement, par les Ψ -synthases Pus3p et Pus4p (Becker *et al.*, 1997a; Lecoite *et al.*, 1998). Comme son homologue chez *S. cerevisiae*, l'enzyme mPus1p ne catalyse pas *in vitro* la conversion de résidus U en résidus Ψ à ces positions (Motorin *et al.*, 1998; Chen et Patton, 1999).

Les spécificités des enzymes mPus1p et ScPus1p *in vitro*, vis-à-vis de substrats ARNt de *S. cerevisiae*, sont donc quasiment identiques. La seule différence est la formation *in vitro* par mPus1p du résidu Ψ en position 32 dans les ARNt, modification que ScPus1p ne catalyse ni *in vitro*, ni *in vivo* (Motorin *et al.*, 1998). Il faut cependant noter que le taux de résidus Ψ formés par mPus1p à la position 32 est tellement faible qu'il n'est pas quantifiable (Chen et Patton, 1999). Chen et Patton (1999) en ont donc conclu que les enzymes mPus1p et ScPus1p sont homologues et seraient responsables, *in vivo*, des conversions de résidus U en résidus Ψ aux mêmes positions des ARNt cytoplasmiques de souris et de *S. cerevisiae*.

Le gène codant l'enzyme Pus1p humaine a également été cloné (Chen et Patton, 1999) et l'enzyme correspondante présente 92% d'identité avec mPus1p, ce qui montre l'existence d'un

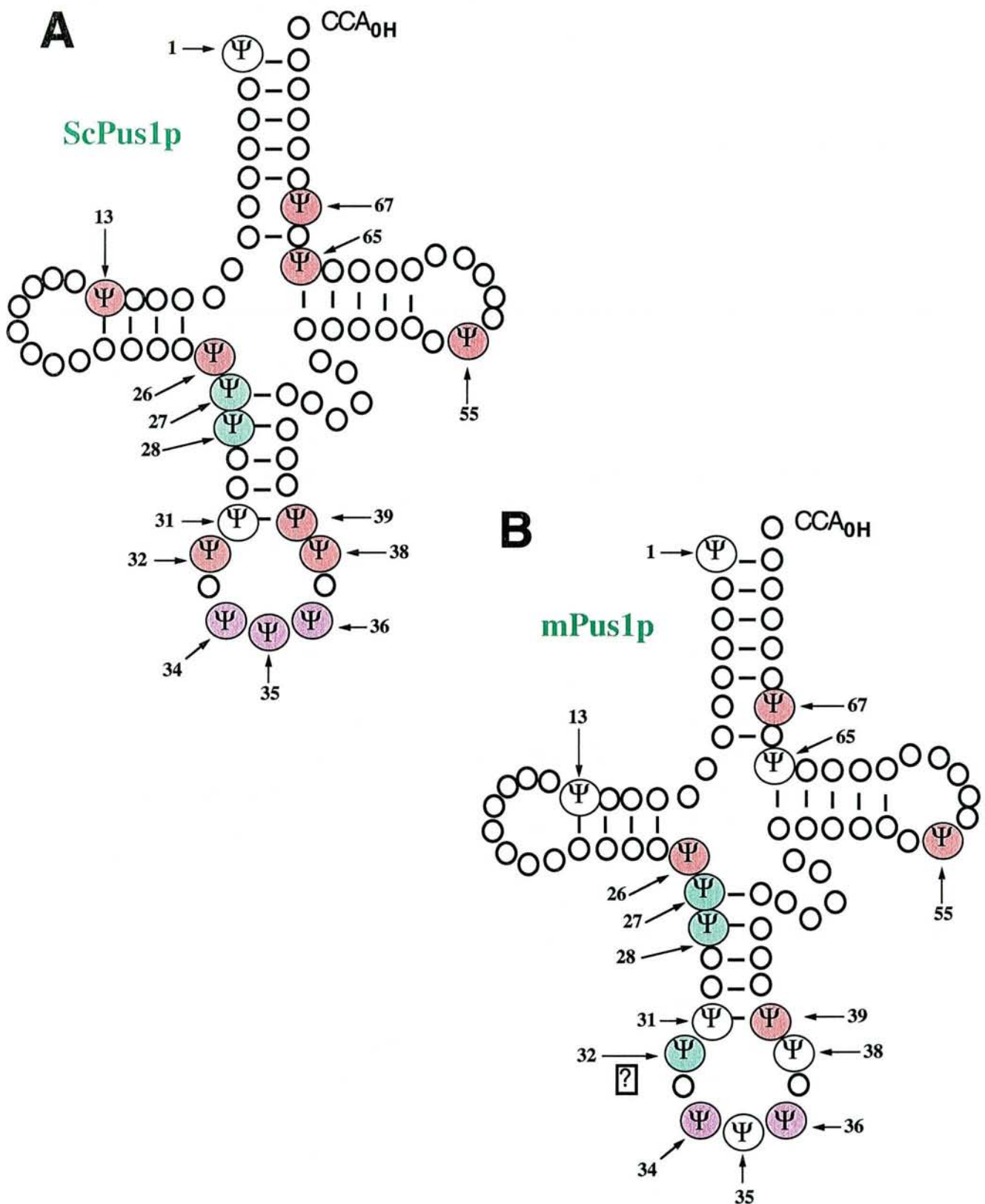


Figure 64: Les résidus Ψ formés *in vitro* par les enzymes ScPus1p (A) et mPus1p (B) dans des ARNt de *S. cerevisiae*, produits par transcription *in vitro*.

Les résidus Ψ, formés *in vitro* par l'une ou l'autre des Ψ-synthases, sont indiqués dans des cercles verts, dans le cas d'une modification intron indépendante, et violets dans le cas d'une modification intron dépendante. Il n'a pas été recherché si les enzymes Pus1p forment *in vitro* les résidus Ψ indiqués dans des cercles blancs. Les résidus Ψ, indiqués dans des cercles rouges, sont présents *in vivo* dans les ARNt cytoplasmiques, mais ne sont pas modifiés *in vitro* par les Ψ-synthases Pus1p.

D'après Simos *et al.*, 1996; Motorin *et al.*, 1998; Chen et Patton, 1999.

fort degré de conservation de la Ψ -synthase Pus1p chez les vertébrés. L'enzyme humaine n'a pas encore été purifiée.

Afin de savoir si mPus1p présentait, comme ScPus1p, une double spécificité de substrat et modifiait également le snRNA U2, nous avons débuté une collaboration avec l'équipe de J. Patton. Cette équipe nous a fourni l'enzyme mPus1p purifiée. Etant donné que mPus1p modifiait *in vitro* les mêmes sites que ScPus1p dans les ARNt, nous avons testé si, *in vitro*, mPus1p pouvait également catalyser la formation du résidu Ψ_{44} dans un snRNA U2 de *S. cerevisiae*, produit *in vitro*. Nous avons ensuite comparé l'action des enzymes de levure et de vertébrés sur un snRNA U2 humain, produit *in vitro*. Dans ce cas, nous avons testé l'action de l'enzyme sur l'ARN libre et sur la pré-particule U2 reconstituée *in vitro* avec un mélange de protéines Sm.

2. METHODES

2.1. Identification des résidus Ψ portés par le snRNA U2 de cellules HeLa

Les ARN totaux de cellules HeLa ont été préparés par extraction phénolique d'un extrait nucléaire de cellules HeLa (Computer Cell Culture Center SA, Belgique). La concentration en ARN a été déterminée par mesure de l'absorption à 260 nm et 10 μ g d'ARN totaux ont ensuite été modifiés par le CMCT, dans les conditions décrites dans la publication n°2. Les modifications ont ensuite été analysées par transcription inverse, à l'aide d'un oligonucléotide complémentaire aux nucléotides du snRNA U2 humain compris entre les positions 153 et 170.

2.2. Production des snRNA U2 humain et de *S. cerevisiae in vitro*

Le protocole de transcription, ainsi que le plasmide contenant la séquence codante du snRNA U2 de *S. cerevisiae* sous contrôle du promoteur T7 (PT7-U2), sont décrits dans la publication n°2. Le plasmide PMRG₃U2-27 contient la région codante du snRNA U2 humain, sous contrôle du promoteur T7; il nous a été fourni par T. Pederson (Worcester, USA) (Jacobson *et al.*, 1993). Avant transcription, ce plasmide a été linéarisé par l'enzyme de restriction SmaI.

La production de transcrits uniformément marqués au ³²P, correspondant au snRNA U2 humain, a été obtenue à partir de 500 ng de plasmide linéarisé, dans le mélange réactionnel suivant:

Tampon de transcription 5x (Tris HCl 0,2 M pH 7,9; MgCl ₂ 30 mM; NaCl 50 mM)	2 µl
RNasin (Pharmacia)	10 U
DTT 200 mM	0,5 µl
Mélange d'ATP, GTP et CTP (5 mM)	1 µl
UTP 0,5 mM	1 µl
UTP [α - ³² P] (800 mCi/mmol; 10 Ci/l)	0,5 µl
ARN T7 polymérase (Pharmacia)	3,5 U
BSA (2 mg/ml)	0,5 µl
H ₂ O	qsp 10 µl

Le mélange a été incubé 1h à 37°C, la matrice a ensuite été détruite par traitement à la DNase (10 U, incubation 30 min à 37 °C). Les ARN ont été extraits par traitement au phénol et précipités deux fois en présence d'acétate d'ammonium 2M, afin d'éliminer les nucléotides non incorporés. La radioactivité du précipité a été mesurée au compteur à scintillation (Beckman).

2.3. Modification des snRNA U2 humain et de *S. cerevisiae* produits par transcription *in vitro*, par les Ψ -synthases mPus1p et ScPus1p purifiées

L'équipe de J. Patton nous a fourni 2 lots de protéines mPus1p. Le premier lot d'enzyme avait été uniquement purifié sur colonne de nickel (m₁Pus1p). Le deuxième lot d'enzyme avait été purifié sur colonne de nickel, puis fractionné sur colonne échangeuse d'anions (m₂Pus1p). Ces deux lots ont toujours été utilisés en parallèle et des résultats similaires ont été obtenus dans les deux cas.

Un à 2 µg de snRNA U2 humain ou de *S. cerevisiae*, produits *in vitro*, ont été incubés à 37°C, pendant 30 à 60 min, en présence de; i) environ 1,2 µg de m₁Pus1p ou 0,6 µg de m₂Pus1p dans un tampon contenant 50 mM Tris HCl pH 7,5; 2 mM MgCl₂; 10mM NH₄Cl, 5mM DTT et 100 µg/ml BSA, ou de ii) environ 1,2 µg de ScPus1p dans le tampon RB, précédemment décrit. L'ensemble a ensuite été soumis à une extraction au phénol et les résidus Ψ formés ont été analysés par transcription inverse après modification par le CMCT, comme décrit dans la publication n°2.

2.4. Formation *in vitro* de particules ribonucléoprotéiques contenant le snRNA U2 humain et modification du snRNA U2 au sein de ces particules, par les Ψ -synthases mPus1p et ScPus1p purifiées

La pré-particule U2 de cellules HeLa a été reconstituée, *in vitro*, à partir de snRNA U2 humain, produit par transcription *in vitro*, et d'un mélange de protéines des UsnRNP spliceosomales, fourni par R. Lührmann (Marburg, Allemagne). Ce mélange contient toutes les protéines Sm (E, F, G, D1, D2, D3, B et B'), les protéines U1A et U1C de la snRNP U1 et, à un faible taux, les protéines U2A' et U2B'' de la snRNP U2. Ce mélange de protéines (TP pour "Total Protein") a été conservé à une concentration de 0,7 $\mu\text{g}/\mu\text{l}$.

L'assemblage de la pré-particule U2 humaine a été réalisé avec 0,5 μg d'ARN, produits par transcription *in vitro*, et 0,6 μl de TP dans le tampon de dilution approprié (20 mM Hepes pH 7,9; 5 mM MgCl_2 ; 50 mM KCl; 0,2 mM EDTA; 0,5 mM DTT; 5% glycérol), en présence de 10 U de RNasin et dans un volume final de 8 μl . Le mélange a été incubé 10 min dans la glace, 2 μl de tampon de réaction 5x concentré ont été ajoutés (100 mM Hepes pH 7,9; 25 mM MgCl_2 ; 250 mM KCl; 2,5 mM DTT; 5% glycérol), puis l'ensemble a été incubé d'abord 30 min à 30°C, puis 15 min à 37°C.

Une réaction contenant, en plus du mélange réactionnel décrit ci-dessus, environ 10000 cpm de transcrit U2 humain uniformément radioactif, nous a permis de vérifier l'assemblage des pré-particules. Pour cela, après la réaction, 2 μl de tampon de dépôt CSB (20 mM Hepes pH 7,9; 40% glycérol; 0,03% bleu bromophénol; 0,03% bleu xylène cyanol) ont été ajoutés au mélange réactionnel et l'ensemble a été déposé sur gel de polyacrylamide natif à 6%.

La modification, par les deux Ψ -synthases, de la pré-particule U2 reconstituée et l'analyse des résidus Ψ formés ont ensuite été effectuées dans les mêmes conditions que pour le snRNA U2 seul, produit par transcription *in vitro*.

2.5. Etude de la structure du snRNA U2 humain produit *in vitro*

L'analyse de la structure secondaire du snRNA U2 humain, produit *in vitro*, a été réalisée dans les mêmes conditions que celle du transcrit correspondant au snRNA U2 de *S. cerevisiae* (page 127). La RNase T2, qui coupe en conditions limitantes, de manière préférentielle, les régions simple brin de l'ARN, a été utilisée comme sonde des régions simple brin, en plus de la RNase T1 et du DMS. L'incubation a été réalisée pendant 5 min dans le tampon RB à 37°C, 0,25 à 1U d'enzyme ont été utilisées par réaction.

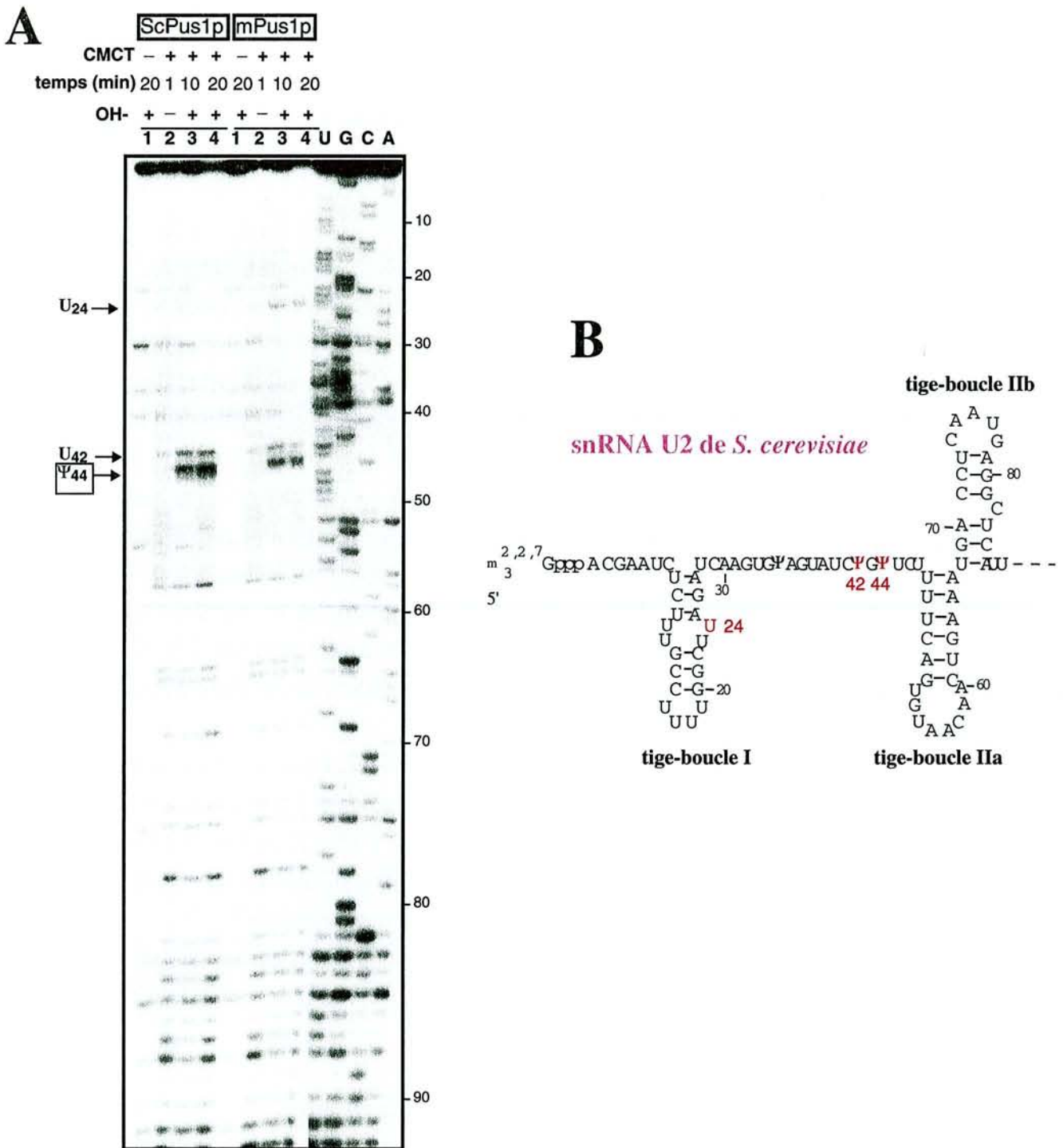


Figure 65 : Formation de résidus Ψ par incubation du snRNA U2 de *S. cerevisiae* produit *in vitro*, avec les enzymes ScPus1p ou mPus1p purifiées.

Panneau A : L'étude a été réalisée sur un transcrite correspondant au snRNA U2 de *S. cerevisiae*, incubé en présence des Ψ-synthases ScPus1p ou mPus1p purifiées. Après incubation avec l'une ou l'autre des Ψ-synthases, le transcrite a été traité par le CMCT durant 1 min (piste 2), 10 min (piste 3) ou 20 min (piste 4). Une expérience avec omission du traitement par le CMCT a été réalisée en tant que témoin (piste 1). Dans toutes les pistes, les ARN modifiés ou non par le CMCT ont été soumis à un traitement alcalin à pH 10,4 pendant 3 h.

Les résidus modifiés ont été identifiés par les arrêts qu'ils génèrent lors d'une réaction d'extension d'amorce par la transcriptase inverse. L'oligonucléotide utilisé est le même que dans la Figure 55. Les ADNc obtenus ont été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 7%, en parallèle avec une réaction de séquençage par la transcriptase inverse (pistes U, G, C et A).

Panneau B : Région 3'-terminale du snRNA U2 de *S. cerevisiae*. Les 3 résidus Ψ formés *in vivo* dans le snRNA U2 sont indiqués dans la séquence. Les nucléotides, localisés aux 3 positions où des résidus Ψ sont formés *in vitro* par ScPus1p ou mPus1p, sont indiqués en rouge.

3. RESULTATS

3.1. *In vitro*, la Ψ -synthase mPus1p catalyse la formation du résidu Ψ_{44} dans le snRNA U2 de *S. cerevisiae* produit *in vitro*

Le snRNA U2 de *S. cerevisiae*, produit par transcription *in vitro*, a été incubé en présence des Ψ -synthases ScPus1p ou mPus1p. Les résidus Ψ formés ont été localisés par modification par le CMCT. Les résultats obtenus sont illustrés dans la Figure 65.

Le fort arrêt d'élongation à la position 44 (Pistes 3 et 4, mPus1p) indique que la Ψ -synthase mPus1p, comme son homologue ScPus1p (Pistes 3 et 4, ScPus1p), forme très efficacement le résidu Ψ_{44} *in vitro* dans le snRNA U2 de *S. cerevisiae*. On peut noter un faible arrêt d'élongation en position 42, que ce soit pour l'analyse des transcrits modifiés par mPus1p ou par ScPus1p, ce qui suggère un faible taux de pseudouridylation à ce site. Bien que nous ne l'ayons pas mentionné précédemment, car le degré de modification était très faible, un faible taux de conversion du résidu U_{42} était aussi observé dans les analyses précédentes réalisées sur le snRNA U2 de *S. cerevisiae*, produit par transcription *in vitro* et modifié *in vitro* par ScPus1p (voir également la Fig. 5 de la publication n°2). Le taux de modification à ce site est donc variable mais toujours très faible.

Comme la Ψ -synthase ScPus1p, la Ψ -synthase mPus1p modifie donc *in vitro* très efficacement la position 44 et beaucoup plus faiblement la position 42. On peut noter que mPus1p est également capable de catalyser faiblement la formation d'un résidu Ψ en position 24 dans le snRNA U2 de *S. cerevisiae*, ce qui n'est pas le cas pour l'enzyme ScPus1p. La Ψ -synthase mPus1p semble donc avoir une spécificité de substrat légèrement plus large que ScPus1p *in vitro*.

3.2. Contrairement à la Ψ -synthase ScPus1p, la Ψ -synthase mPus1p ne catalyse pas *in vitro* la formation de résidus Ψ dans un snRNA U2 humain produit *in vitro*

La Ψ -synthase mPus1p catalysant *in vitro* la formation du résidu Ψ_{44} dans le snRNA U2 de *S. cerevisiae*, la question se posait alors de savoir si cette enzyme pouvait également former un résidu Ψ , à la même position dans un snRNA U2 de vertébrés, voir même plusieurs résidus. Afin de répondre à cette question, la démarche consistait à analyser les résidus Ψ formés dans un snRNA U2 humain, produit *in vitro*, après incubation en présence de mPus1p. Bien que les snRNA U2 humain et de rat soient identiques à 97% (seuls les résidus aux positions 108 à 111 et 116 diffèrent entre les deux organismes) (pour revue, Reddy, 1988), il

Positions	Résidus détectés par Reddy <i>et al.</i> , 1981 (1)	Résidus détectés par Branlant <i>et al.</i> , 1982 (2)	Résidus détectés au cours de ce travail (3)
6	Ψ	-	Ψ
7	Ψ	-	Ψ
15	Ψ	-	Ψ(4)
34	Ψ	-	Ψ
37	Ψ	-	Ψ
39	Ψ	Ψ	Ψ
41	Ψ	Ψ	Ψ
43	Ψ	Ψ	Ψ
44	Ψ	Ψ	Ψ
54	Ψ	Ψ	Ψ
58	Ψ	Ψ	Ψ
60	U	Ψ	Ψ
69	U	Ψ	U(5)
72	U	Ψ	U(5)
89	Ψ	Ψ	Ψ
91	Ψ	Ψ	Ψ
130	U	U	U(6)

Tableau 14: Comparaison des résidus Ψ détectés dans le snRNA U2 de cellules HeLa, avec les résultats antérieurs de Reddy *et al.* (1981) sur le snRNA U2 d'hépatome de rat et ceux de Branlant *et al.* (1982) sur le snRNA U2 de rat et de cellules HeLa.

- 1) Résidus Ψ identifiés dans le snRNA U2 d'hépatome de rat.
- 2) Résidus Ψ identifiés dans la région 3'-terminale du snRNA U2 de cellules HeLa et le snRNA U2 de foie et de cerveau de rat.
- 3) Résidus Ψ identifiés, au cours de ce travail, dans le snRNA U2 de cellules HeLa.
- 4) Ce résidu Ψ est détecté à très faible taux.
- 5) Ces résidus U ne sont pas modifiés par le CMCT.
- 6) Ce résidu U a un comportement particulier après modification par le CMCT suivi d'un traitement alcalin, que ce soit dans le snRNA U2 de cellules HeLa ou dans le transcrit produit *in vitro*.

était important dans un premier temps d'identifier les résidus Ψ portés naturellement par le snRNA U2 humain, car cette étude n'avait été réalisée que de manière partielle (Branlant *et al.*, résultats non publiés).

3.2.1. Localisation des résidus Ψ dans le snRNA U2 de cellules HeLa

Il y a une vingtaine d'années, l'analyse complète des résidus Ψ du snRNA U2 de rat avait été faite par Reddy *et al.* (1981), lors du séquençage d'un snRNA U2 isolé d'hépatome de rat. Une analyse de la région 3'-terminale de snRNA U2 de rat extrait de foie et de cerveau avait aussi été réalisée par Branlant *et al.* (1982). En ce qui concerne le snRNA U2 humain, seul le séquençage de l'extrémité 3'-terminale de l'ARN extrait de cellules HeLa avait été réalisé (Branlant *et al.*, résultats non publiés). Les résultats obtenus par les équipes de C. Branlant et R. Reddy différaient à 3 positions (Tableau 14). Afin d'avoir une identification complète des résidus Ψ portés par le snRNA U2 de cellules HeLa, nous avons modifié cet ARN par le CMCT, au sein d'un mélange d'ARN issus d'un extrait nucléaire. Les résidus Ψ ont ensuite été identifiés par extension d'amorce. Les résultats obtenus sont illustrés dans les Figures 66 et 67. Ils nous ont permis de mettre en évidence 14 résidus Ψ dans le snRNA U2 humain. La plupart de ces résidus avait déjà été détectée par Reddy *et al.* (1981) et Branlant *et al.* (1982). Il faut noter que le résidu Ψ en position 15 ne conduit qu'à un arrêt très faible de la transcriptase inverse dans nos analyses. La modification pourrait être partielle.

Nos résultats différaient seulement des résultats antérieurs aux positions 60, 69 et 72 (Tableau 14). Nous confirmons la présence du résidu Ψ en position 60, qui avait été détecté par Branlant *et al.* (1982), mais pas par Reddy *et al.* (1981). Par contre, les uridines en positions 69 et 72 ne sont pas des résidus Ψ , comme proposé précédemment par Branlant *et al.* (1982) sur la base d'une absence de réactivité à l'hydrazine. Néanmoins, il doit s'agir de résidus modifiés, étant donné que nous n'avons observé aucune modification de ces résidus par le CMCT (piste 2, ARN, positions U₆₉ et U₇₂, Fig. 66) et que Branlant *et al.* n'avaient observé aucune réaction vis-à-vis de l'hydrazine. A part les résidus Ψ , Um et Ψ_m , aucun autre dérivé de l'uridine n'a encore été mis en évidence dans les UsnRNA. L'utilisation de plus grandes quantités de snRNA U2, avec fractionnement d'un segment contenant ces deux résidus et analyse par spectrométrie de masse, seront nécessaires pour vérifier la présence de résidus modifiés et pour les identifier.

Nous avons également observé un comportement inhabituel vis-à-vis du CMCT du résidu en position 130. En effet, cette uridine est modifiée par le CMCT, mais l'adduit ne semble pas hydrolysé par le traitement alcalin, que ce soit dans le snRNA U2 naturel modifié ou dans le transcrit produit *in vitro* (position U₁₃₀, Fig. 66). En effet, un arrêt de la transcriptase inverse

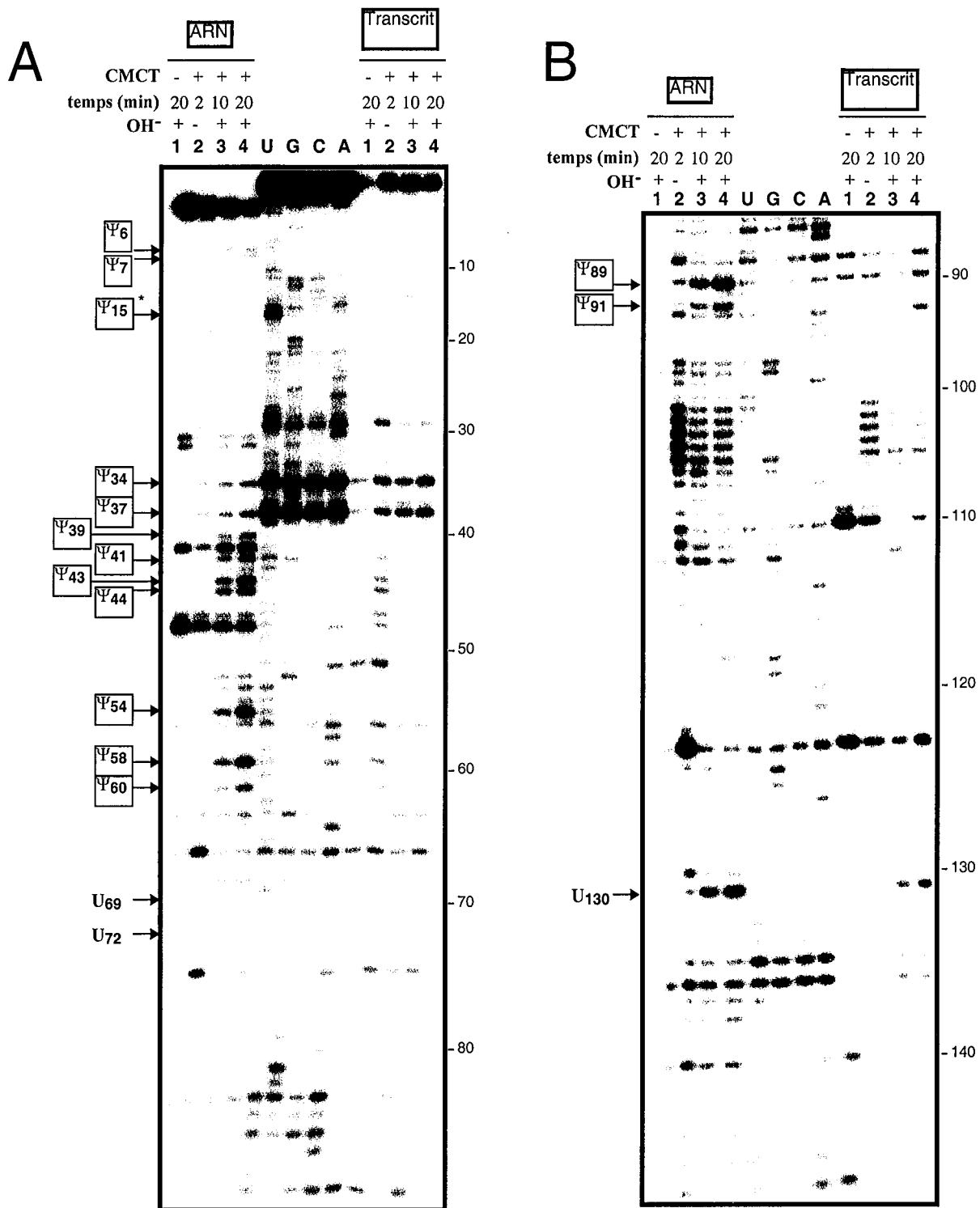


Figure 66: Localisation des résidus Ψ dans le snRNA U2 extrait de cellules HeLa.

L'étude a été réalisée sur un mélange d'ARN totaux de cellules HeLa (ARN), préparé comme indiqué dans la publication n°3 et, pour contrôle, sur un ARN produit par transcription *in vitro*, correspondant au snRNA U2 humain (Transcrit). Les résidus Ψ portés par le snRNA U2 humain ont été identifiés par la méthode de modification par le CMCT. Les ARN ont été traités par le CMCT durant 2 min (piste 2), 10 min (piste 3) ou 20 min (piste 4). Une expérience avec omission du traitement par le CMCT a été réalisée en tant que témoin (piste 1). Dans toutes les pistes, les ARN modifiés ou non par le CMCT ont été soumis à un traitement alcalin à pH 10,4 pendant 3 h. Les panneaux A et B correspondent à deux temps de migration différents permettant d'étudier, respectivement, les séquences allant des positions 1 à 88 et des positions 96 à 145.

Les résidus modifiés ont été identifiés par les arrêts qu'ils génèrent lors d'une extension d'amorce par la transcriptase inverse. Ils sont indiqués par une flèche à gauche de l'autoradiographie. Le site d'appariement sur le snRNA U2 humain de l'oligonucléotide amorce est indiqué dans la Figure 67. Les ADNc obtenus ont été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 7%, en parallèle avec une réaction de séquençage par la transcriptase inverse (pistes U, G, C et A). Les résidus uridines en positions 69 et 72, non modifiés par le CMCT, sont indiqués, ainsi que le résidu U en position 130 dont l'adduit CMCT n'est pas éliminé lors du traitement alcalin. Le résidu Ψ en position 15 (marqué d'une étoile) n'est que faiblement détecté.

est observé après modification par le CMCT et traitement alcalin. Une autre explication possible pourrait être l'apparition d'un clivage en milieu alcalin. Ce résidu U est localisé dans une boucle à 4 nucléotides (Fig. 67) qui pourrait avoir un fort degré de structuration. La question est de savoir si le comportement inhabituel de ce résidu est lié à un type particulier de structure de la boucle qui le contient. Il nous est difficile, dans ces conditions, de déterminer si le résidu U₁₃₀ est converti *in vivo* en résidu Ψ. Néanmoins, cette possibilité est peu vraisemblable, puisque les études antérieures de Branlant *et al.*, montraient que ce résidu est modifié par l'hydrazine.

Notre étude met donc en évidence la présence de 14 résidus Ψ dans le snRNA U2 humain, soit un résidu supplémentaire par rapport à ce qui avait été décrit pour le snRNA U2 d'hépatome de rat (Reddy *et al.*, 1981). Parmi ces résidus, le résidu Ψ₄₃ correspond au résidu Ψ₄₄ formé par l'enzyme ScPus1p dans le snRNA U2 de *S. cerevisiae* (Fig. 67).

3.2.2. Analyse des résidus Ψ formés *in vitro* par les enzymes ScPus1p et mPus1p dans le snRNA U2 humain, produit *in vitro*, libre ou associé aux protéines Sm

Aucun résidu Ψ n'a pu être détecté dans le snRNA U2 humain, produit *in vitro*, après son incubation avec mPus1p (Fig. 68). Par contre, après incubation avec l'enzyme ScPus1p, nous avons détecté la formation de 3 résidus Ψ aux positions 2, 43 et 49 (Fig. 69). La position 43 du snRNA U2 humain correspond à la position 44 du snRNA U2 de *S. cerevisiae*. Nous formons donc avec l'enzyme ScPus1p le même résidu Ψ dans les snRNA U2 humain et de *S. cerevisiae* produits *in vitro*. La différence, dans le cas du snRNA U2 humain, est la formation de résidus Ψ à deux autres positions, où des résidus Ψ ne sont pas trouvés dans l'ARN naturel. Il faut néanmoins remarquer que le taux de formation du résidu Ψ en position 2 est très faible (Fig. 69).

La question était de savoir ce qui pouvait expliquer la différence de réactivité de l'enzyme de *S. cerevisiae*, vis-à-vis du snRNA U2 humain et du snRNA U2 de *S. cerevisiae*. Des études antérieures menées au laboratoire par A. Szkukàlek sur le snRNA U2 humain, produit *in vitro*, et par K. Hartmuth sur le snRNA U2 de rat, après extraction au phénol, suggéraient qu'en absence des protéines des UsnRNP, ces UsnRNA formaient des appariements supplémentaires par rapport aux deux structures secondaires admises pour le snRNA U2. Une explication possible de la différence de modification des snRNA U2 humain et de *S. cerevisiae* par ScPus1p *in vitro* pouvait donc être l'existence de différences structurales. Afin de tester cette hypothèse, j'ai réalisé une étude expérimentale de la structure secondaire du snRNA U2 humain, produit par transcription *in vitro*. Pour cela, j'ai utilisé,

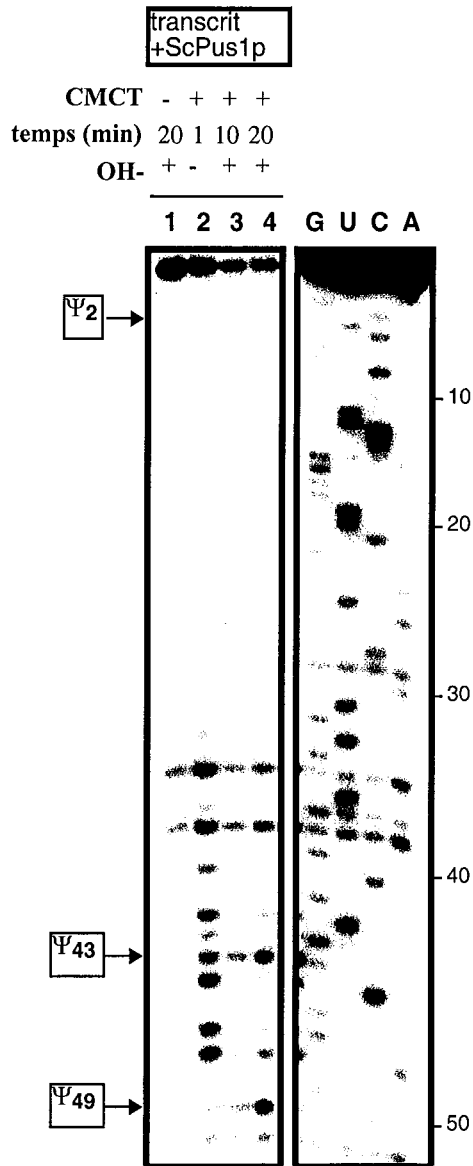


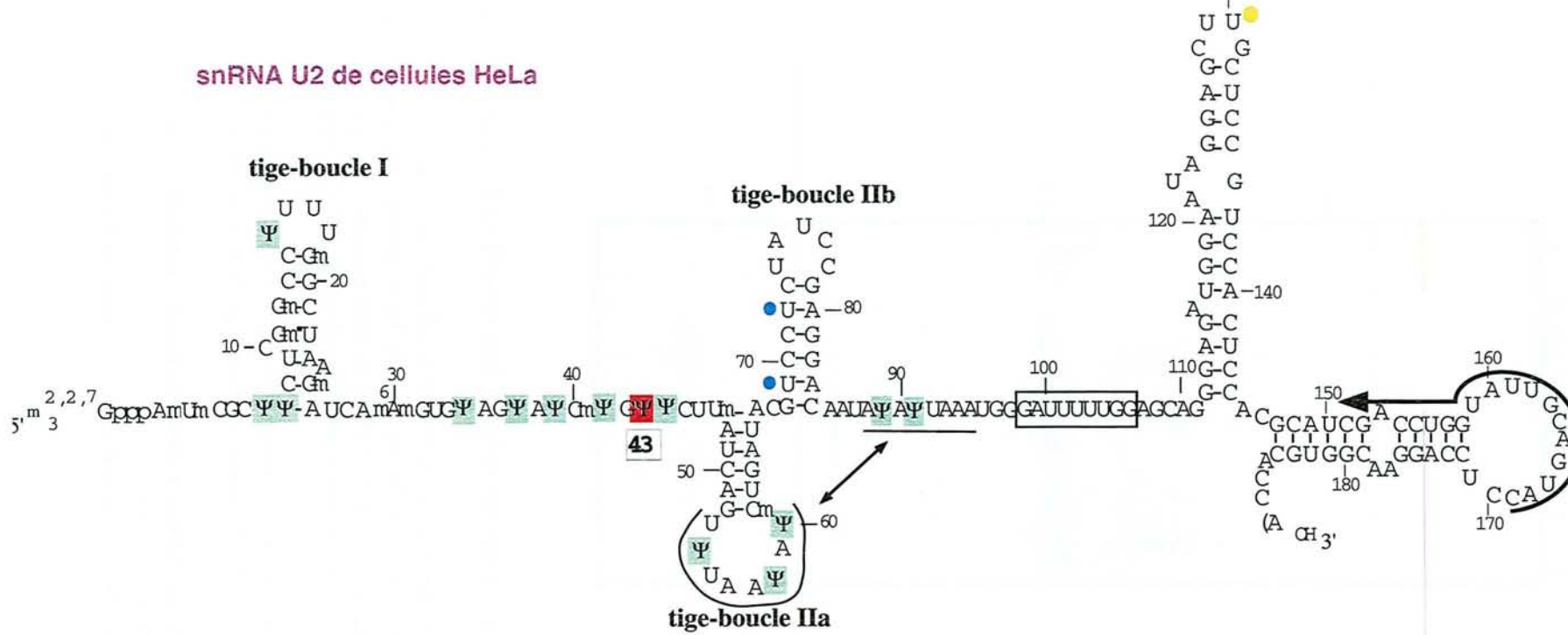
Figure 69 : Identification des résidus Ψ formés *in vitro* par l'enzyme ScPus1p dans un snRNA U2 humain produit par transcription *in vitro*.

Le snRNA U2 humain, produit par transcription *in vitro* a été incubé en présence de la Ψ -synthase ScPus1p purifiée (transcrit+ScPus1p), puis traité par le CMCT durant 1 min (piste 2), 10 min (piste 3) ou 20 min (piste 4). Une expérience avec omission du traitement par le CMCT a été réalisée en tant que témoin (piste 1). Dans toutes les pistes, les ARN modifiés ou non par le CMCT ont été soumis à un traitement alcalin à pH 10,4 pendant 3 h.

L'oligonucléotide utilisé est le même que dans la Figure 67. Les ADNc obtenus ont été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 7%, en parallèle avec les produits des 4 réactions de séquençage par la transcriptase inverse (pistes U, G, C et A).

Les arrêts d'élongation dans les colonnes 3 et 4 indiquent la formation, *in vitro*, par ScPus1p de trois résidus Ψ dans le snRNA U2 humain.

snRNA U2 de cellules HeLa



snRNA U2 de *S. cerevisiae*

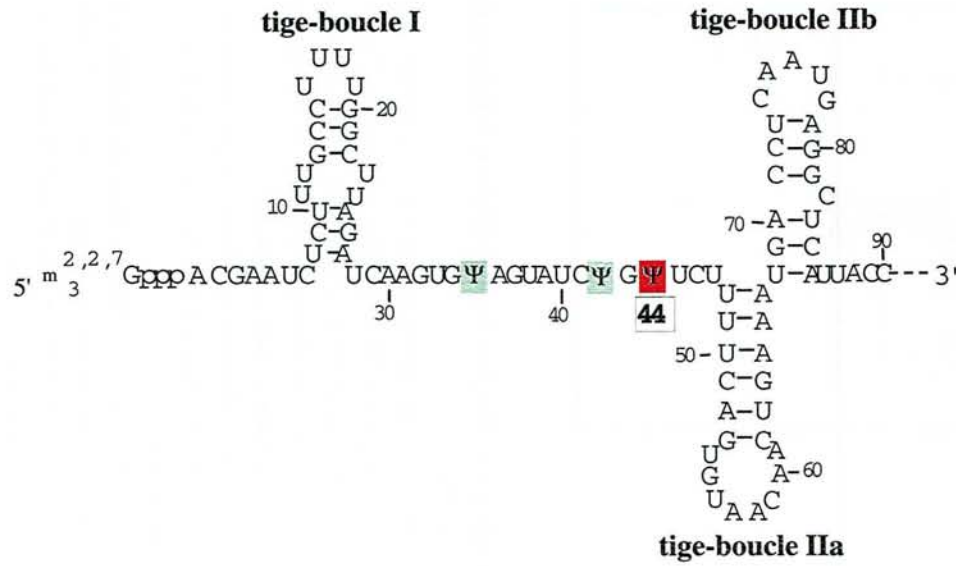


Figure 67: Localisation des résidus Ψ dans le snRNA U2 de cellules HeLa, comparaison avec la localisation des résidus Ψ dans le snRNA U2 de *S. cerevisiae*.

Les résidus U non modifiés par le CMCT sont indiqués par un point bleu. Le résidu U, indiqué par un point jaune, est modifié par le CMCT mais l'adduit CMCT ne semble pas hydrolysé par traitement alcalin. Le résidu Ψ en position 44 dans le snRNA U2 de *S. cerevisiae*, qui est modifié *in vivo* et *in vitro* par ScPus1p, est indiqué en rouge, ainsi que son homologue dans le snRNA U2 de cellules HeLa. La région de complémentarité entre le snRNA U2 humain et l'oligonucléotide utilisé pour l'analyse par extension d'amorce (Figure 66) est indiquée.

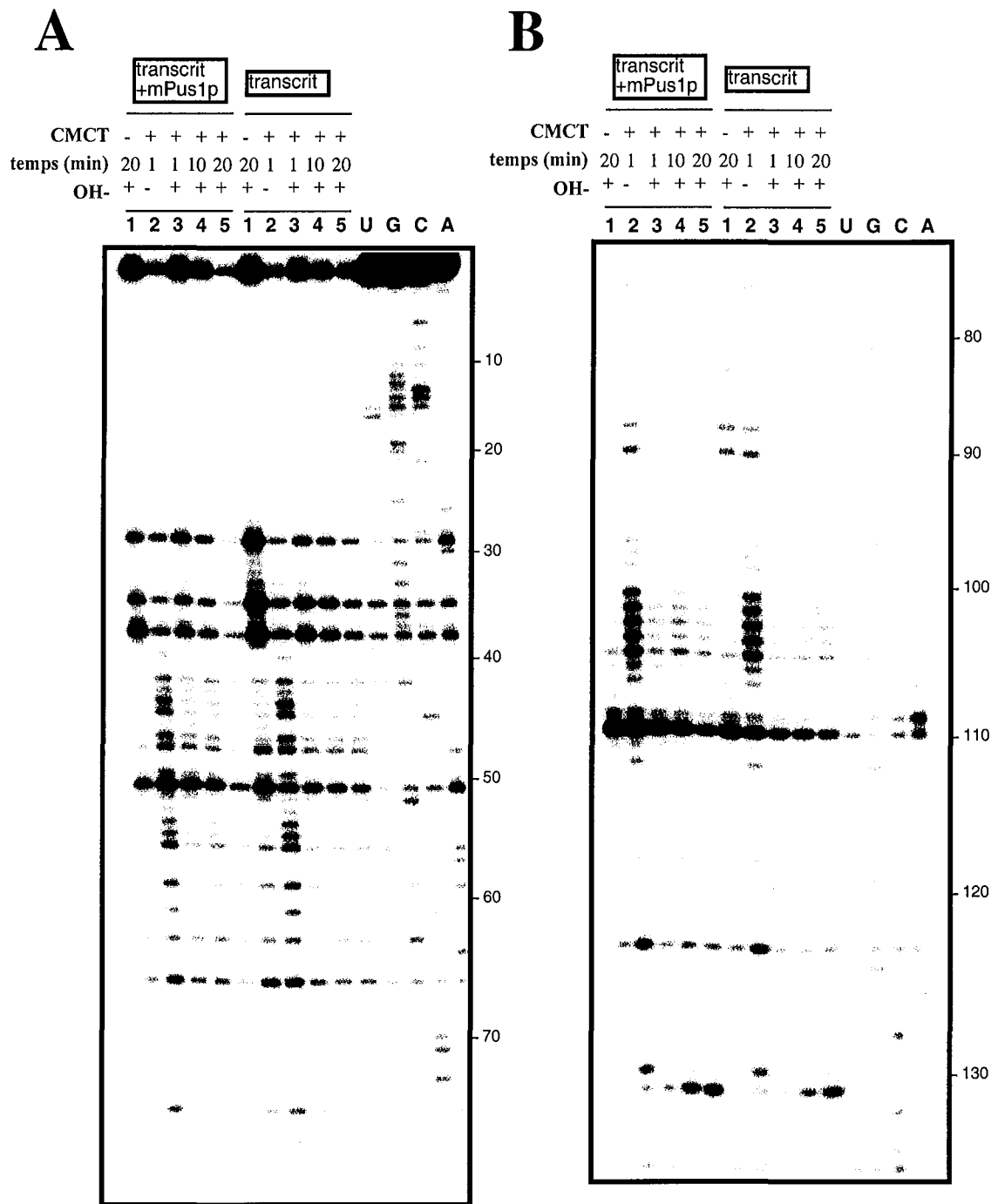


Figure 68 : Recherche des résidus Ψ formés *in vitro* par l'enzyme mPus1p dans un snRNA U2 humain, produit *in vitro*.

Le transcrit correspondant au snRNA U2 humain a été incubé en présence de la Ψ -synthase mPus1p purifiée (transcrit+mPus1p), puis traité par le CMCT pendant 1 min (piste 2 et 3), 10 min (piste 4) ou 20 min (piste 5). Une expérience avec omission du traitement par le CMCT a été réalisée en tant que témoin (piste 1). Dans toutes les pistes, les ARN modifiés ou non par le CMCT ont été soumis à un traitement alcalin à pH 10,4 pendant 3 h. Une analyse d'un transcrit non incubé avec la Ψ -synthase mPus1p a été utilisée comme témoin (transcrit).

L'oligonucléotide utilisé était le même que dans la Figure 67. Les ADNc obtenus ont été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 7%, en parallèle avec les produits des 4 réactions de séquençage par la transcriptase inverse (pistes U, G, C et A).

L'absence d'arrêts d'élongation dans les colonnes 3, 4 et 5 (transcrit+mPus1p) montre que dans les conditions utilisées, l'enzyme mPus1p ne catalyse pas la formation de résidus Ψ dans le snRNA U2 humain.

Les panneaux A et B correspondent à deux temps de migration différents pour étudier, respectivement, les séquences allant des nucléotides 1 à 78 et 75 à 135.

comme décrit précédemment, des sondes chimiques et enzymatiques. Les résultats obtenus sont représentés dans la Figure 70. La présence simultanée de coupures par la RNase V1 et de modifications par le DMS, dans la région 53 à 60, montre l'existence d'un équilibre entre les formes A et B. De plus, ces résultats sont en faveur de l'existence d'une structure de type A avec une extension de la structure tige-boucle II, impliquant les segments 40 à 51 et 96 à 111 (en violet sur la Fig. 70B). La présence d'une telle extension est en accord avec la détection de coupures par la RNase V1, qui est spécifique des régions structurées, en position 106 et dans les segments compris entre les positions 95 et 100 et entre les positions 40 et 45. Néanmoins, la présence simultanée de coupures par les RNases V1, T1 et T2, dans les segments 40 à 43 et 95 à 111 suggère que cet appariement supplémentaire n'est pas présent dans toutes les molécules.

Nous détectons donc une différence structurale entre les snRNA U2 humain et de *S. cerevisiae*, qui pourrait expliquer une différence de réactivité de l'enzyme ScPus1p vis-à-vis de ces deux ARN. Par ailleurs, la structure particulière prise par le snRNA U2 humain, produit *in vitro*, pouvait peut-être expliquer pourquoi l'enzyme mPus1p, qui agit sur le snRNA U2 de *S. cerevisiae* produit *in vitro*, n'agit pas sur le snRNA U2 humain produit *in vitro*. Dans la conformation du snRNA U2 humain où la structure tige-boucle II présente une extension, le résidu U₄₃ est situé dans l'hélice supplémentaire (Fig. 71). Une possibilité était que ce type de structuration limite l'action de l'enzyme mPus1p.

Il est à noter que dans la structure II étendue, le résidu U₄₉ modifié par ScPus1p remplit les critères que nous avons mis en évidence pour les sites cibles de cette enzyme. Ce résidu est localisé en 3' d'un résidu A et est localisé en bordure d'une hélice.

Il n'est pas exclu qu'une extension de la structure tige-boucle I puisse être formée dans certaines molécules de snRNA U2 humain produites *in vitro*. Cet appariement supplémentaire impliquerait les séquences 3 à 6 et 30 à 33. Ce type d'interaction pourrait expliquer la formation par l'enzyme ScPus1p d'un résidu Ψ en position 2. Dans cette extension de l'hélice I, ce résidu est situé en bordure d'hélice (Fig. 71). Par ailleurs, il est précédé d'un résidu A.

Seule la formation du résidu U₄₃ par ScPus1p est difficile à expliquer dans ce contexte structural, puisque ce résidu est localisé dans une hélice. Néanmoins, comme le montre notre analyse structurale, l'hélice contenant le résidu U₄₃ n'est pas formée dans toutes les molécules.

Etant donné que les appariements supplémentaires qui semblaient se former dans le snRNA U2 humain, produit *in vitro*, encadraient le site Sm (Fig. 71), nous avons pensé que la fixation des protéines Sm sur cet ARN devrait empêcher la formation de ces appariements supplémentaires, vraisemblablement en entraînant une transition de la structure du snRNA U2 vers la forme B active. Si l'hypothèse était exacte, nous nous attendions à observer, en présence de protéines Sm, un changement de la réactivité de l'enzyme ScPus1p vis-à-vis du

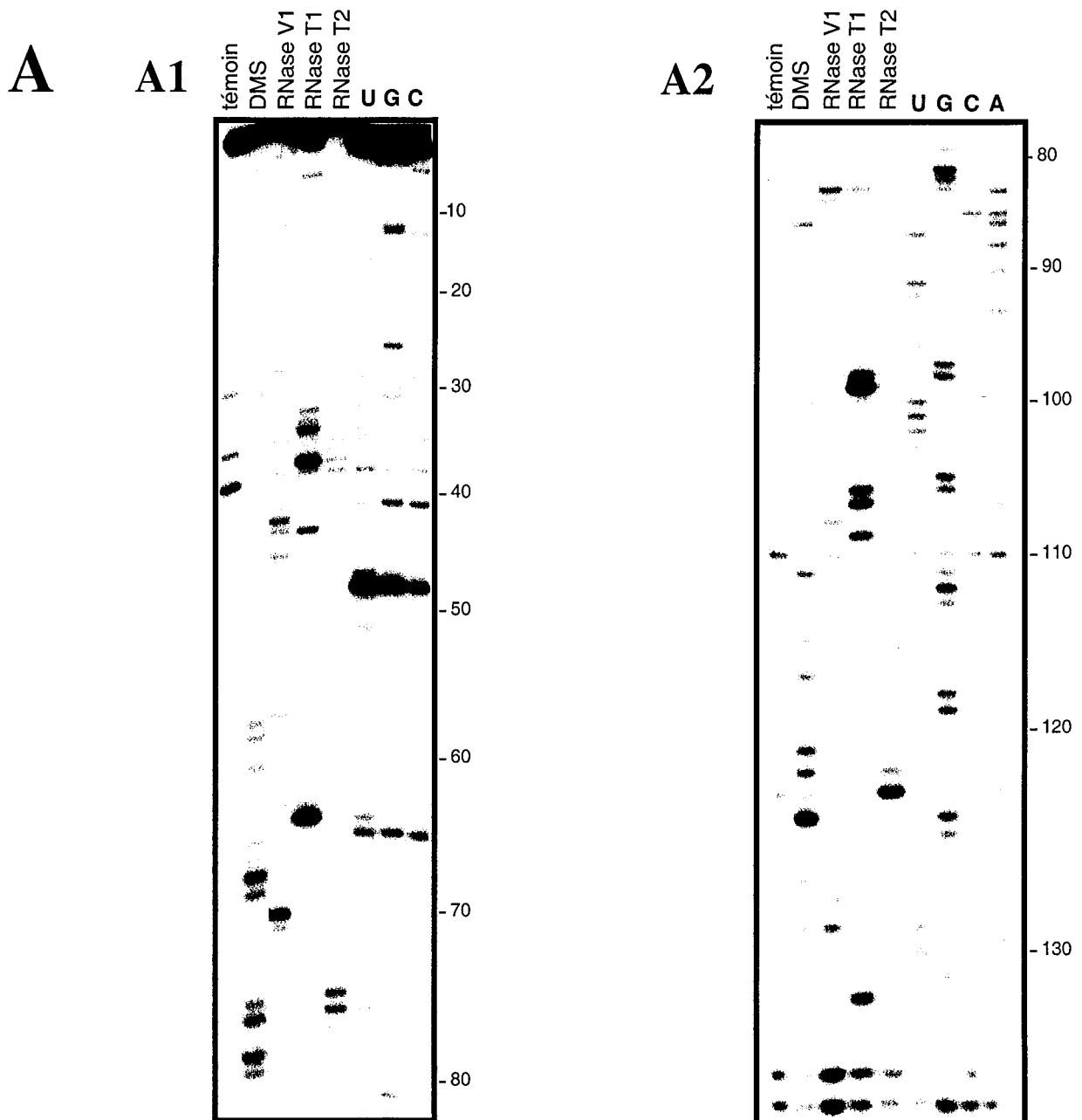
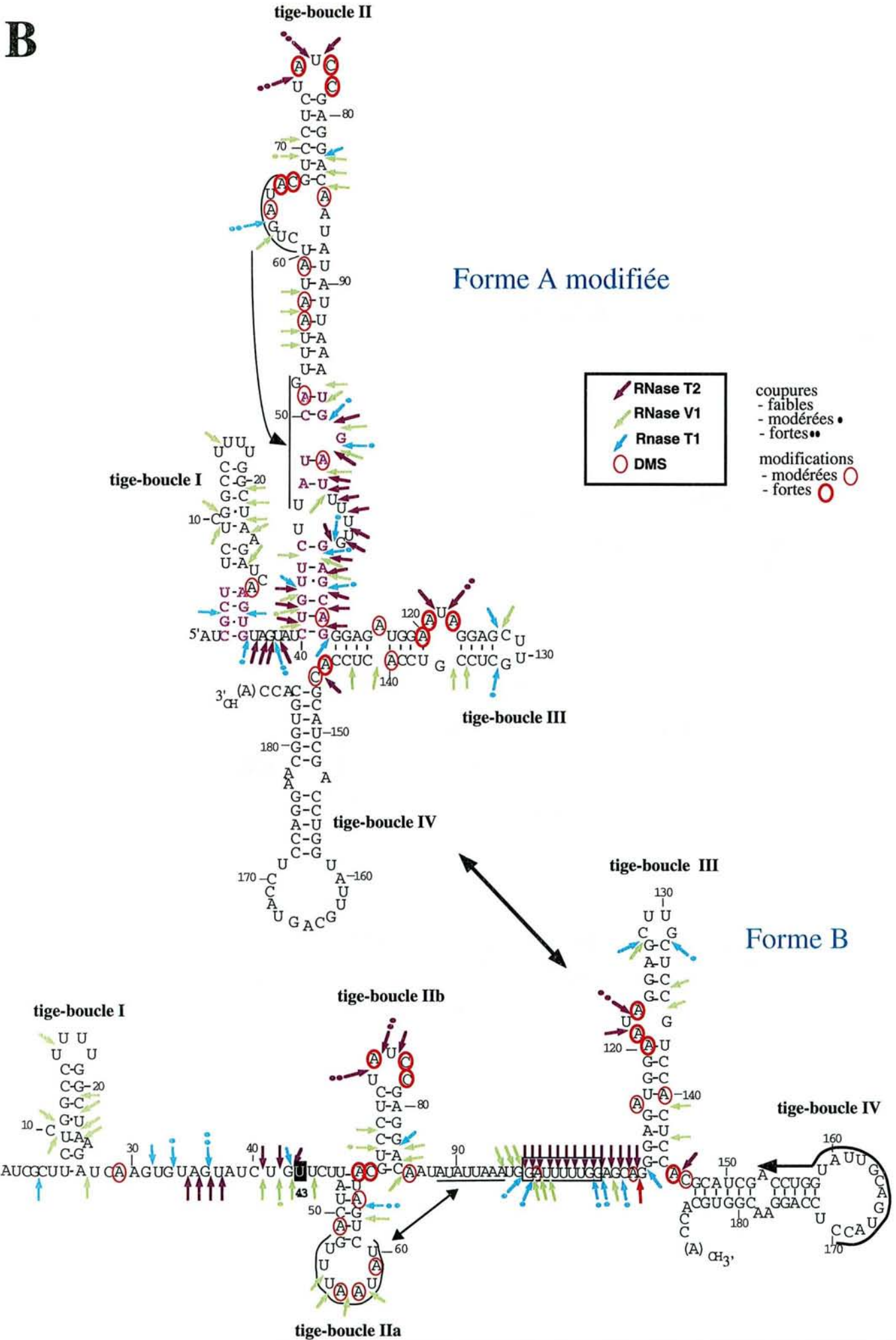


Figure 70: Étude de la structure secondaire du snRNA U2 humain produit par transcription *in vitro*.

Panneau A: Analyse par la transcriptase inverse des modifications chimiques (pistes DMS) et des coupures par les RNases T1, V1 et T2 (pistes T1, V1 et T2, respectivement), générées dans le snRNA U2 humain, produit par transcription *in vitro*. Un transcrit non modifié a servi de témoin afin de détecter les pauses naturelles de la transcriptase inverse (piste témoin). Les ADNc obtenus ont été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 7%, en parallèle avec un séquençage par la transcriptase inverse (colonnes U, G, C et A). L'oligonucléotide, dont le site d'appariement sur le transcrit est indiqué dans le Panneau B, a servi d'amorce. La position des nucléotides dans la séquence du snRNA U2 humain est indiquée à droite de la fluorographie. Les panneaux A1 et A2 correspondent à deux temps de migration différents permettant d'étudier, respectivement, les séquences allant des positions 1 à 82 et des positions 78 à 147.

Panneau B: Représentation schématique des résultats obtenus par modification chimique du snRNA U2 humain par le DMS et par coupure par les RNases T1, T2 et V1. Les résultats obtenus sont reportés sur les deux structures de cet ARN. Les extensions proposées des hélices I et II sont indiquées en violet. Les différentes modifications et coupures sont montrées respectivement par les cercles et les flèches, dont le code de couleur est indiqué. Le site d'appariement de l'oligonucléotide utilisé est représenté.

B

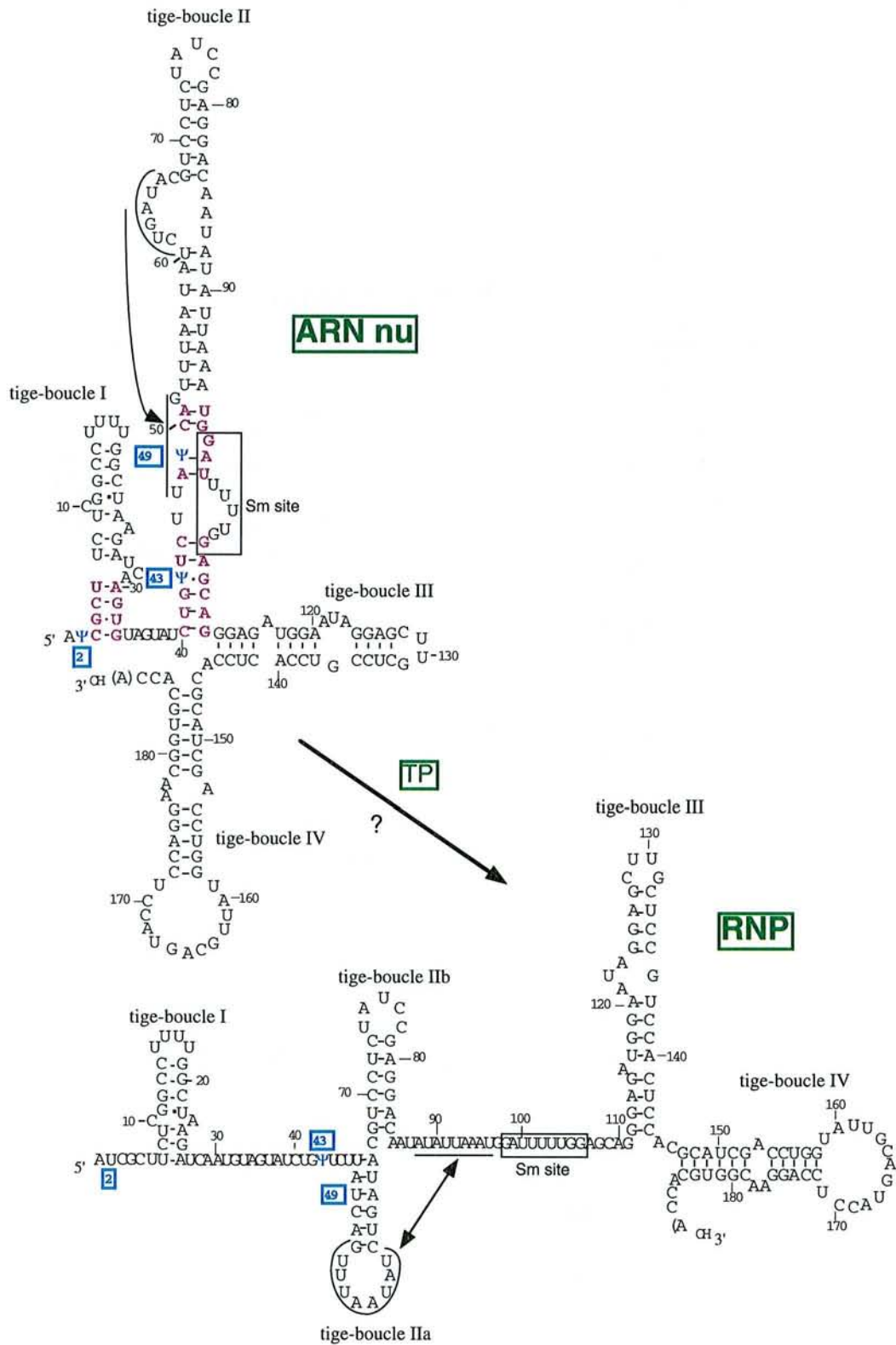


Figure 71: La structure A du snRNA U2 humain avec une extension de l'hélice II, que nous avons mise en évidence dans le transcrit, et sa conversion possible suite à la fixation des protéines Sm.

Les appariements supplémentaires par rapport à la structure A sont indiqués en violet. Les résidus Ψ formés, *in vitro*, par ScPus1p sont indiqués en bleu.

snRNA U2 humain, produit *in vitro*. Nous pouvions également attendre une activité de l'enzyme mPus1p vis-à-vis de cet ARN. L'équipe de R. Lührmann nous a fourni un mélange de protéines extraites d'une fraction qui contenait l'ensemble des UsnRNP spliceosomales. L'équipe de R. Lührmann avait déjà montré que, par incubation de cette fraction avec le snRNA U2, les protéines Sm s'associent à ce UsnRNA. Nous avons incubé le snRNA U2 humain, produit par transcription *in vitro*, avec ces protéines et vérifié par électrophorèse en gel non dénaturant que nous avons bien formation d'un complexe susceptible de correspondre aux pré-particules U2 (Fig. 72). Ce complexe a été soumis à l'action de l'enzyme ScPus1p et à celle de l'enzyme mPus1p. La formation de résidu Ψ a été analysée par modification par le CMCT. Les résultats obtenus sont illustrés dans la Figure 73. Comme on peut le constater, après association du snRNA U2 humain avec les protéines Sm, l'enzyme ScPus1p ne modifie plus cet ARN aux positions 2 et 49. Seule la position 43 est modifiée. Ceci conforte les hypothèses que nous avons formulées sur l'influence de la structure secondaire du snRNA U2 humain sur son degré de modification par ScPus1p. Dans ces conditions, comme pour le snRNA U2 de *S. cerevisiae*, un seul résidu du snRNA U2 humain est modifié par ScPus1p et c'est le résidu homologue à celui modifié dans le snRNA U2 de *S. cerevisiae*. Par contre, même en présence des protéines Sm, aucune action de l'enzyme mPus1p n'a été détectée sur le snRNA U2 humain produit par transcription *in vitro*.

4. DISCUSSION ET PERSPECTIVES

4.1. La Ψ -synthase ScPus1p purifiée modifie le snRNA U2 humain produit par transcription *in vitro*

Deux positions sont modifiées à un degré élevé: la position correspondant à celle modifiée dans le snRNA U2 de *S. cerevisiae* et la position 49 qui n'est pas modifié *in vivo*. La modification correspondante n'est pas observée dans le snRNA U2 de *S. cerevisiae*, après incubation avec l'enzyme ScPus1p. Cette différence est sans doute liée au fait que, dans le snRNA U2 de *S. cerevisiae*, le résidu correspondant, U₅₀, n'est pas précédé par un résidu A, comme dans l'ARN humain, mais par un U. Sur la base de notre étude de la spécificité de ScPus1p, il est donc tout à fait logique d'observer la formation d'un résidu Ψ en position 49 dans le snRNA U2 humain, mais pas dans le snRNA U2 de *S. cerevisiae*, à la fois du fait de la structure 2D différente de ces ARN lorsqu'ils sont produits *in vitro* et de la différence d'identité du résidu en 5' du résidu 49 dans le snRNA U2 humain ou 50 dans le snRNA U2 de *S. cerevisiae*.

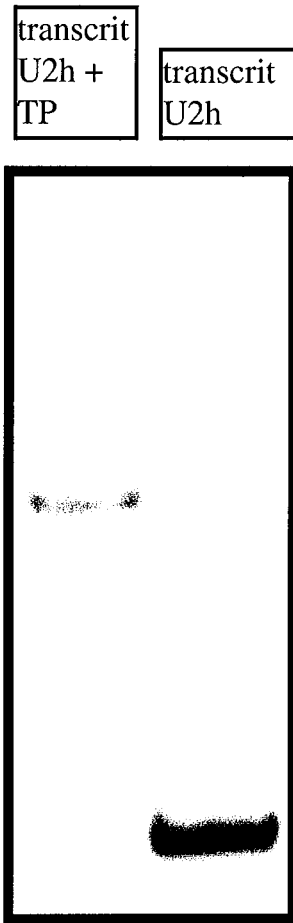


Figure 72: Analyse par électrophorèse en gel de polyacrylamide non dénaturant des complexes formés entre le snRNA U2 humain produit *in vitro*, et le mélange de protéines extraites des UsnRNP (TP).

L'enzyme ScPus1p modifie légèrement le résidu U en position 2 dans le snRNA U2 humain. La modification correspondante n'est pas observée dans le snRNA U2 de *S. cerevisiae*, car un résidu C est présent en position 2.

Comme nous l'avons déjà mentionné, la règle de reconnaissance des sites cibles par ScPus1p ne s'applique pas vraiment pour le résidu naturellement modifié par ScPus1p dans le snRNA U2 de *S. cerevisiae* et pour son homologue dans le snRNA U2 humain, mais elle s'applique bien pour les résidus additionnels formés *in vitro* dans le snRNA U2 humain. Tout ceci suggère que la Ψ -synthase ScPus1p est capable d'accommoder deux types de substrats ARN dans son site actif. Il est possible que la reconnaissance de ces deux types de substrats se fasse, au moins en partie, par l'intermédiaire d'acides aminés différents.

4.2. La présence des protéines Sm limite le taux de modification, par la Ψ -synthase ScPus1p, du snRNA U2 humain produit *in vitro*

En présence des 8 protéines Sm qui reconnaissent la séquence G₉₈AUUUUUGG₁₀₆ du snRNA U2, cet ARN semble changer de conformation. D'après des analyses structurales que K. Harthmuth avait menées au laboratoire sur les particules formées après incubation du snRNA U2 de rat avec un mélange de protéines des UsRNP, le snRNA U2 semble prendre une conformation de type B après fixation des protéines Sm. Ces 8 protéines sont supposées s'assembler dans un complexe contenant chacune des protéines Sm et, d'après des données récentes, l'ARN passe au travers d'un anneau formé par ces différentes protéines (Fig. 15 de l'introduction bibliographique) (Kambach *et al.*, 1999a). Les protéines Sm se fixent sur le site Sm et protègent ce site, ainsi qu'une partie des tige-boucles III et IV (Krämer *et al.*, 1999). Elles ne couvrent donc pas les résidus en positions 2 et 49, mais, dans nos expériences, elles peuvent avoir limité l'action de l'enzyme ScPus1p à ces deux positions par encombrement stérique.

Quoiqu'il en soit, le fait que nous trouvons, par ces expériences réalisées *in vitro*, que la présence des protéines Sm limite l'action de ScPus1p, sur le snRNA U2 humain, à la génération du résidu qu'elle forme naturellement *in vivo* dans le snRNA U2 de *S. cerevisiae*, suggère que ces protéines pourraient jouer un rôle similaire *in vivo*. Ceci soulève la question de l'étape où a lieu la modification du snRNA U2 par l'enzyme Pus1p *in vivo*. L'enzyme ScPus1p a une localisation nucléaire. Elle pourrait donc modifier le snRNA U2 immédiatement après sa transcription dans le noyau, ou après son retour dans le noyau sous forme de pré-particule, après son association avec les protéines Sm dans le cytoplasme. Sur la base des résultats obtenus avec le snRNA U2 humain et les protéines Sm, cette seconde hypothèse semble la plus vraisemblable. On peut imaginer que la fixation de ces protéines participe aussi, *in vivo*, à restreindre l'action de l'enzyme Pus1p vis-à-vis du snRNA U2. Nous

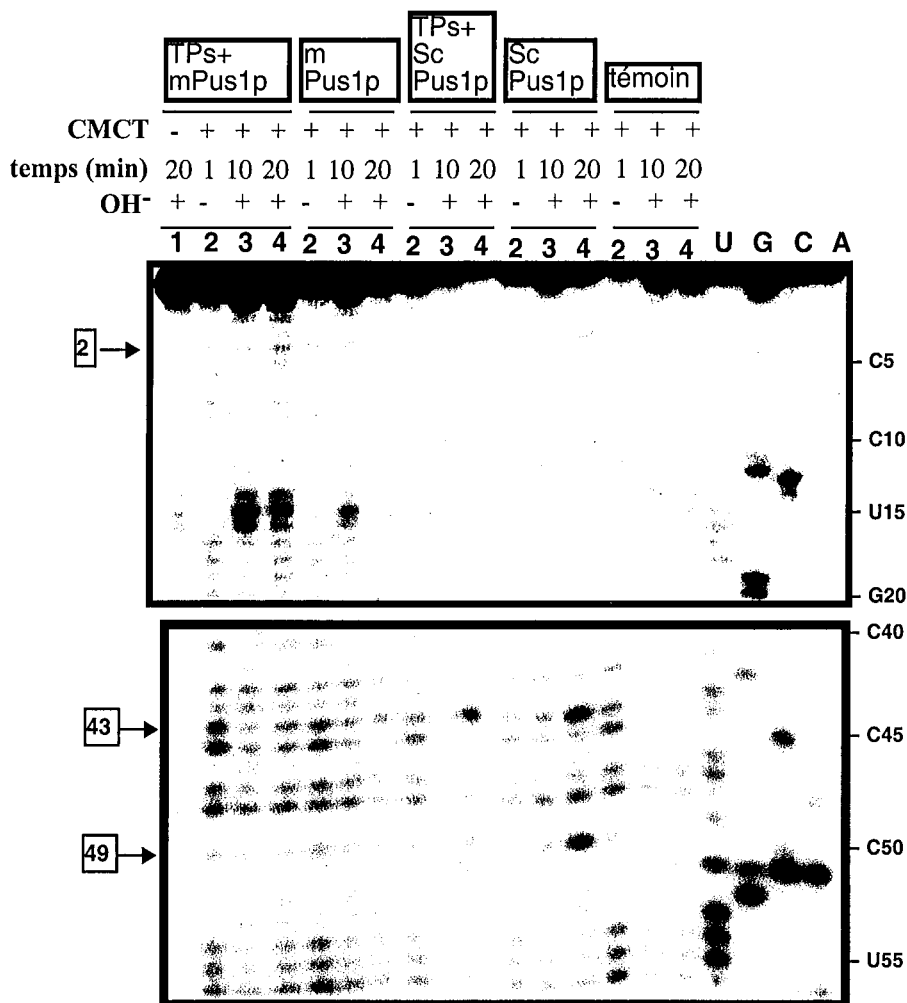


Figure 73: Identification des résidus Ψ formés *in vitro* par les enzymes mPus1p et ScPus1p dans un snRNA U2 humain produit par transcription *in vitro*, incubé au préalable en présence ou en absence des protéines des UsnRNP (TP).

Le snRNA U2 humain, produit par transcription *in vitro*, seul ou associé avec des protéines des UsnRNP (TPs) a été incubé en présence des Ψ -synthases mPus1p ou ScPus1p, puis traité par le CMCT durant 1 min (piste 2), 10 min (piste 3) ou 20 min (piste 4). Une expérience avec omission du traitement par le CMCT a été réalisée en tant que témoin (piste 1). Dans les pistes 1, 3 et 4, les ARN, modifiés ou non par le CMCT, ont été soumis à un traitement alcalin à pH 10,4 pendant 3 h. L'analyse d'un transcrit incubé en absence d'enzyme a été utilisée comme témoin (témoin).

L'oligonucléotide utilisé comme amorce est le même que dans la Figure 67. Les ADNc obtenus ont été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 7%, en parallèle avec les produits des 4 réactions de séquençage par la transcriptase inverse (pistes U, G, C et A).

n'avons malheureusement pas obtenu de données confortant cette hypothèse avec l'enzyme mPus1p, puisque cette Ψ -synthase n'a jamais été active sur le snRNA U2 humain.

En ce qui concerne le système homologue snRNA U2/ScPus1p, nous observons *in vitro* une légère modification en position 42. Il est à noter que le résidu U₄₂ est converti en résidu Ψ dans le snRNA U2 naturel de *S. cerevisiae*. Néanmoins, lorsque nous avons étudié le snRNA U2 d'une souche *pus1⁻* de *S. cerevisiae*, le résidu Ψ ₄₂ est toujours formé. La question est donc de savoir si sa formation par l'enzyme ScPus1p a lieu seulement *in vitro* ou si, *in vivo*, des enzymes distinctes peuvent avoir des activités redondantes. Il est à noter que le résidu U₄₂ est précédé par un résidu C, ce qui n'est pas favorable pour l'enzyme ScPus1p, mais ce résidu se situe dans le motif structural reconnu par l'enzyme lorsqu'elle catalyse la formation du résidu 44. Cette proximité explique sans doute le faible taux de modification en position 42 par ScPus1p *in vitro*. Mais, *in vivo*, la catalyse de la formation du résidu Ψ ₄₂ est probablement réalisée par une autre enzyme.

4.3. La Ψ -synthase mPus1p modifie le snRNA U2 de *S. cerevisiae* produit *in vitro*

Les enzymes mPus1p et ScPus1p ont des comportements très similaires vis-à-vis du snRNA U2 de *S. cerevisiae*, produit *in vitro*. La seule différence est la très faible modification du résidu U₂₄ par mPus1p. Il est à noter que ce résidu U₂₄ est bien situé dans un lien flexible entre deux structures hélicoïdales, mais qu'il est situé en 3' d'un résidu U. L'enzyme mPus1p pourrait donc être moins sensible, que l'enzyme ScPus1p, à l'identité du résidu situé en 5' du résidu cible.

Les travaux de Chen et Patton (1999) montraient que l'enzyme mPus1p a un comportement assez similaire *in vitro* à celui de l'enzyme ScPus1p vis-à-vis des ARNt de levure. Dans la plupart des cas, des modifications identiques étaient observées pour les deux enzymes, à l'exception de la position 32. Nos résultats corroborent cette observation, puisque ces deux enzymes ont des comportements très similaires, quoique non identiques, vis-à-vis du snRNA U2 de *S. cerevisiae* produit *in vitro*.

Dans la mesure où nous n'avons pas pu démontrer d'action *in vitro* de l'enzyme mPus1p sur le snRNA U2 humain, produit *in vitro*, nous n'avons pour l'instant aucun argument expérimental en faveur de la modification *in vivo*, par l'enzyme Pus1p de vertébrés, du résidu U₄₃ du snRNA U2 de vertébrés. Nous devons donc réaliser des expériences complémentaires pour essayer de répondre à cette question.

Tout d'abord, il faut noter que jusqu'à présent, comme nous disposions de moins d'enzyme mPus1p que d'enzyme ScPus1p recombinantes, nous avons utilisé un rapport enzyme/substrat

inférieur pour mPus1p, par rapport à celui utilisé pour ScPus1p (1,2 moles de mPus1p/mole de transcrit, contre 5 à 20 moles de ScPus1p/mole de transcrit). Nous devons donc, dans un premier temps, optimiser les conditions d'utilisation de mPus1p *in vitro*. Nous arriverons peut-être à obtenir une modification du snRNA U2 humain *in vitro*, en augmentant la quantité de mPus1p.

Nous ne pouvons pas non plus exclure que le tampon utilisé pour former des complexes entre le snRNA U2 humain et les protéines Sm renferme un inhibiteur de l'enzyme mPus1p. Des tests devront être réalisés en utilisant le snRNA U2 de *S. cerevisiae*, qui est substrat de mPus1p, dans le tampon de reconstitution.

Par ailleurs, afin de commencer à aborder le problème *in vivo*, la région codante de l'enzyme mPus1p a été clonée par I. Ansmant dans un vecteur de levure. Ce vecteur va être utilisé pour transformer la souche de *S. cerevisiae pus1-*, que nous avons utilisée précédemment, ceci afin de tester si l'enzyme de souris peut compléter cette souche. Cette approche a été utilisée pour montrer que la Ψ -synthase Nop60p de drosophile peut compléter une souche *cbf5-* de *S. cerevisiae* (Phillips *et al.*, 1998). Une restauration de la croissance a été observée dans ce cas. En ce qui nous concerne, il faudra tester si le résidu Ψ_{44} est formé dans le snRNA U2 de *S. cerevisiae*, après expression de l'enzyme de souris.

La seule preuve absolue du rôle de l'enzyme mPus1p de vertébrés dans la formation du résidu Ψ_{43} du snRNA U2 de vertébrés *in vivo*, serait bien sûr de bloquer l'expression de hPus1p dans des cellules HeLa, en utilisant, par exemple, des oligonucléotides anti-sens, et de rechercher s'il y a ou non formation du résidu Ψ_{43} dans le snRNA U2.

4.4. Le snRNA U2 humain renferme 14 résidus Ψ

Notre analyse des résidus Ψ contenus dans le snRNA U2 extrait de cellules HeLa confirme la présence d'un résidu Ψ en position 60, comme le suggéraient les résultats antérieurs de C. Branlant (Branlant *et al.*, 1982). Il est à noter que ce résidu Ψ appartient au segment 52 à 60 impliqué dans la formation de la structure de type pseudonoeud avec le segment 88 à 95 (Fig. 67). Cinq des 8 résidus U impliqués dans cette structure sont convertis en résidus Ψ et cela pourrait participer à stabiliser ce type de structure puisque, comme nous l'avons déjà cité plusieurs fois, différents travaux montrent que les résidus Ψ ont un effet stabilisateur des hélices d'ARN.

*PARTIE 4: RECHERCHE DES SUBSTRATS DE QUATRE Ψ-SYNTHASES
POTENTIELLES DE S. cerevisiae.*

A/

Organisme/Famille	RsuA	RluA	TruB	TruA	total
<i>E. coli</i>	4	4	1	1	10
<i>M. jannaschii</i>	0	0	1	1	2
<i>S. cerevisiae</i>	0	4	2	3	9
<i>H. sapiens</i>	0	1	3	0	/

B/

Nom de l'enzyme ou de l'ORF	Séquence du motif II	Database	Numéro d'accèsion	Substrat(s)
Famille TruB				
Pus4p	GGTL D PLASGVLVI	Swiss Prot	P48567	ARNt
Cbf5p	SGTL D PKVTGCLIV	Swiss Prot	P33322	ARNr
Famille TruA				
Pus1p	AART D KG VHAGGNL	Swiss Prot	Q12211	ARNt/snRNA U2
Pus2p	AART D KG VHAMLNL	Swiss Prot	P53167	?
Pus3p	CGRT D KGVSAMNQV	Swiss Prot	P31115	ARNt
Famille RluA				
YDL036C	CNRL D RLTSGLMFL	Swiss Prot	Q12069	?
YLR165C	VHRL D HCVTGGMLI	TREMBL	Q06244	?
YGR169C	CYRL D KITSGLLIL	Swiss Prot	P53294	?
RIB2	CNRL D KPTSGLMFL	Swiss Prot	Q12362	?

Tableau 15: La répartition des Ψ -synthases de *S. cerevisiae* dans les 4 familles RsuA, RluA, TruA et TruB.

Le Panneau A indique le nombre de Ψ -synthases de *S. cerevisiae* appartenant à chacune des familles de Ψ -synthases et le compare au nombre de Ψ -synthases appartenant à chacune de ces familles chez d'autres organismes. Le génome humain n'étant pas entièrement séquencé, le nombre total de Ψ -synthases humaines n'est pas indiqué.

Le Panneau B indique le nom, la séquence du motif II, le substrat et le numéro d'accèsion de chaque Ψ -synthase de *S. cerevisiae*. L'acide aminé en rouge est l'acide aspartique supposé être essentiel pour la catalyse.

Ces données sont tirées de Koonin (1996), Conrad *et al.* (1999) et des articles cités dans le Tableau 10 de l'Introduction.

A/ INTRODUCTION ET RESULTATS

Comme nous l'avons dit précédemment, parmi les Ψ -synthases déjà caractérisées chez *S. cerevisiae*, seule l'enzyme Pus1p agit sur les UsnRNA. Elle catalyse la formation d'un des 6 résidus Ψ que nous avons détectés dans les UsnRNA de *S. cerevisiae*, le résidu Ψ_{44} dans le snRNA U2. Il restait donc à identifier les Ψ -synthases responsables de la formation des 5 autres résidus Ψ , mis en évidence aux positions 35 et 42 du snRNA U2, 99 du snRNA U5 et 5 et 6 du snRNA U1.

Comme nous l'avons mentionné dans l'introduction bibliographique, un motif appelé II, qui constitue une signature des Ψ -synthases, a pu être établi à partir des premières séquences connues de Ψ -synthases (Fig. 36 de l'introduction bibliographique) (Koonin, 1996). La recherche de ce motif, dans le génome de 47 organismes différents, a permis récemment de recenser 111 gènes distincts codant ou susceptibles de coder des Ψ -synthases (Conrad *et al.*, 1999). Les différentes Ψ -synthases déjà caractérisées ou potentielles appartiennent à 4 familles: RluA, RsuA, TruA et TruB (Koonin, 1996; Conrad *et al.*, 1999).

Ainsi, 9 gènes ont été identifiés chez *S. cerevisiae* (Tableau 15):

⇒ *PUS1*, *PUS2*, *PUS3*, *PUS4* et *CBF5*, dont les enzymes, à l'exception de Pus2p, ont été caractérisées,

⇒ et 4 autres gènes potentiels correspondant aux ORF: YDL036C, YLR165C, YGR169C et RIB2, dont les produits n'avaient pas encore été caractérisés. D'après les comparaisons de séquences, les enzymes produites à partir de ces 4 ORF appartiendraient à la famille RluA. Elles sont caractérisées par le motif conservé HRLD dans le domaine II.

Une possibilité était donc que l'une ou plusieurs de ces ORF code(nt) pour la (ou les) Ψ -synthase(s) responsable(s) de la formation de certains des résidus Ψ des UsnRNA. Afin de tester cette hypothèse, I. Ansmant et moi-même avons entrepris la construction de 4 souches de *S. cerevisiae* portant chacune une délétion de l'une de ces 4 ORF.

1. APPROCHE UTILISEE POUR ABOLIR L'ACTIVITE DES 4 GENES CODANT DES Ψ -SYNTHASES POTENTIELLES

Différentes techniques existent pour abolir l'activité d'un gène de *S. cerevisiae*, toutes font appel à la recombinaison homologue. Nous avons utilisé la méthode de remplacement en une étape décrite par Baudin *et al.* (1993). Cette technique est basée sur le fait que, chez la levure, la recombinaison homologue est très efficace avec des fragments linéaires. En effet, l'existence d'une homologie d'environ 40 pb est suffisante pour avoir un taux de recombinaison homologue important. On remplace, par recombinaison homologue, l'ORF à éliminer par un gène corrigé

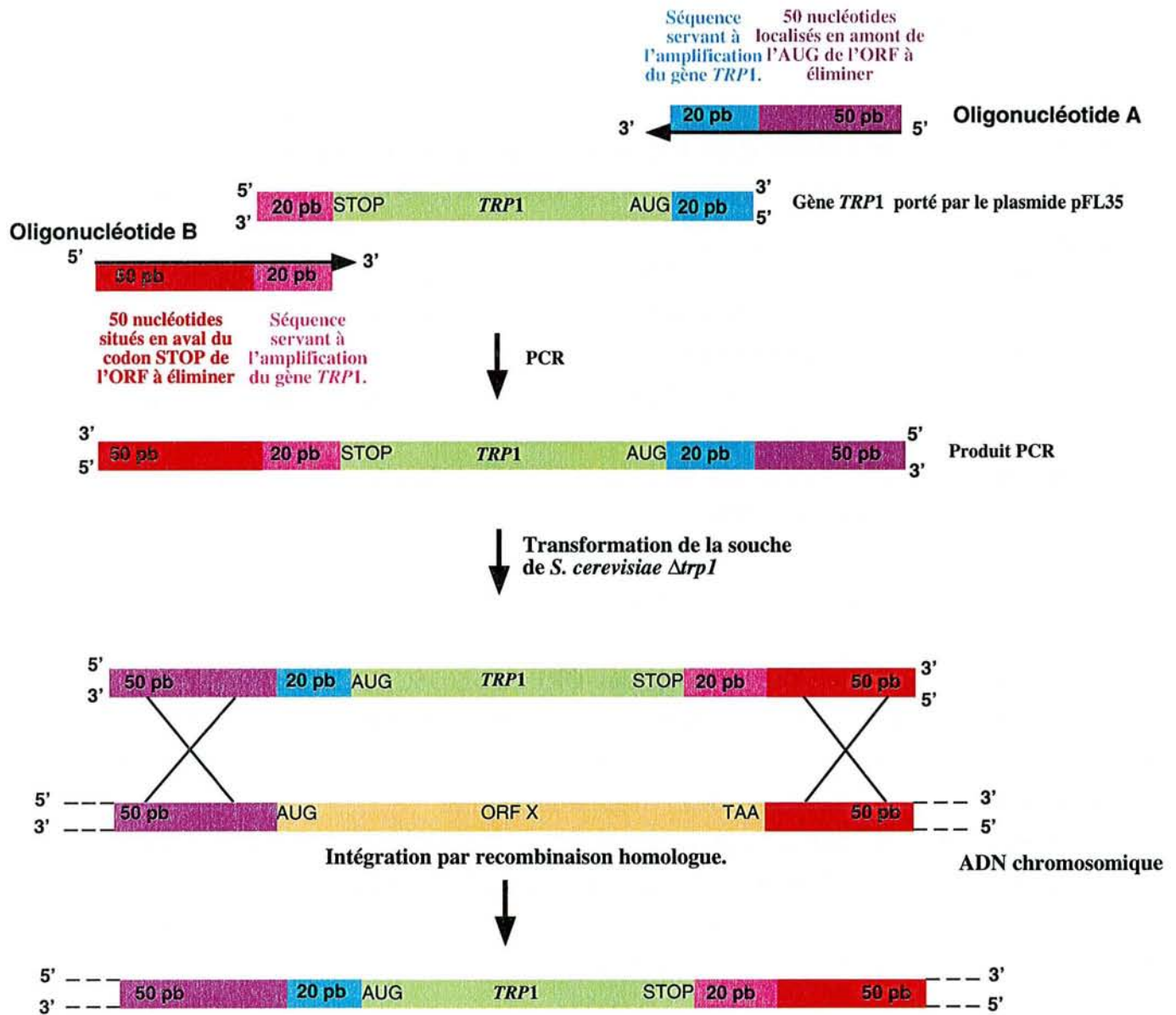


Figure 74: Principe des délétions des ORF de *S. cerevisiae* réalisées par remplacement avec le gène *TRP1* de *S. cerevisiae*.
D'après Baudin *et al.*, 1993.

une auxotrophie de la souche hôte, dans notre cas le gène *TRP1*. Pour cela, il faut générer un fragment d'ADN dans lequel le gène *TRP1* est encadré par les séquences flanquant en 5' et en 3' l'ORF cible (Baudin *et al.*, 1993) (Fig. 74). Ce fragment est produit par PCR, en utilisant des oligonucléotides contenant 2 régions distinctes: une région permettant ultérieurement la recombinaison homologue du fragment PCR au locus cible et une autre servant à l'amplification du gène *TRP1* porté par le plasmide pFL35. Une souche de *S. cerevisiae* $\Delta trp1$, diploïde ou haploïde, est ensuite transformée avec l'ADN produit par PCR et les levures ayant intégré dans leur génome le fragment d'ADN, sont sélectionnées sur un milieu sélectif des clones devenus prototrophes pour le tryptophane. Les différents oligonucléotides que nous avons utilisés sont répertoriés dans le Tableau 16 et les protocoles expérimentaux sont décrits dans la publication n°5 ci-jointe.

2. CARACTERISATION DES SOUCHES SELECTIONNEES APRES MUTATION

Nous avons vérifié systématiquement si les clones sélectionnés sur milieu sélectif sans tryptophane présentaient bien les délétions souhaitées. Pour cela, nous avons utilisé la méthode d'amplification d'ADN sur colonies. Pour chaque ORF délétée, différents couples d'oligonucléotides C/D et E/F ont été choisis, de façon à tester la présence de l'ORF non délétée (C/D) et la présence de la substitution (E/F) (Fig. 75). Les couples C/D ont été choisis de manière à amplifier des fragments d'environ 800 pb, les couples E/F de manière à amplifier des fragments d'environ 400 pb et ils sont présentés dans le Tableau 16. Les tests ont été réalisés sur les souches haploïdes et sur les souches diploïdes sélectionnées. Pour comparaison, des essais ont été réalisés sur la souche sauvage et un contrôle des amplifications a été systématiquement réalisé en absence d'ADN. Les fragments amplifiés ont été analysés par électrophorèse en gel d'agarose à 2%. L'analyse des souches sélectionnées pour les délétions des 3 ORF YDL036C, YLR165C, RIB2 est illustrée dans la Figure 76. Comme on peut le constater, aussi bien pour les souches haploïdes, que pour les souches diploïdes sélectionnées (pistes 1 et 2, respectivement, des Panneaux A, B et C), la présence des produits d'amplification attendus pour les couples d'oligonucléotides E/F (400 pb) confirmait la présence de la substitution par le gène *TRP1*. Aucune amplification n'était détectée avec ces couples d'oligonucléotides pour la souche sauvage (pistes 3 et 4 des 3 Panneaux). Une autre confirmation de la présence des délétions était apportée par le fait qu'aucune amplification n'était obtenue avec les couples C/D, permettant d'amplifier les gènes sauvages, pour les souches haploïdes (piste 5), alors que des produits d'amplification, ayant les tailles attendues, étaient obtenus avec les souches diploïdes sélectionnées (piste 6). Pour chaque délétion, ce fragment d'ADN était bien identique à celui obtenu pour la souche sauvage (pistes 7 et 8). Ces expériences ont été répétées plusieurs fois, pour chacune des souches sélectionnées, et

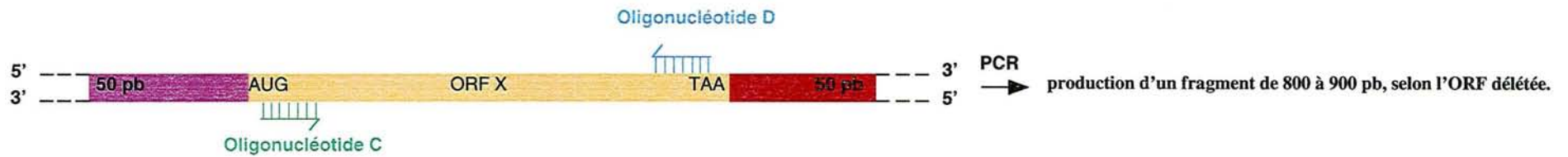
Nom	Séquence 5'→3'	Utilisation
1289	5'AAGAACATTAAAAAAAAAAACTATCAATATTAGATT AACAAACCAATCAAAGGCCAAGAGGGGAGGGC 3'	Oligonucléotide A pour la délétion de l'ORF RIB2
1290	5'TTATTGTAGATATGATGTAAGATTACATGATAATAT ATATGCTAAATAGTCTTAAATAAATACTACTC 3'	Oligonucléotide B pour la délétion de l'ORF RIB2
1295	TATTGACGATTGCTCGAGACAATCAGTGGGCAAAG GATGATTGTCAAGTGGCCAAGAGGGGAGGGC 3'	Oligonucléotide A pour la délétion de l'ORF YGR169C
1296	GGAAATGGTAACTAACTATATACAAAAAAGTTTCG CTCAAAAAAAGTTCTTAAATAAATACTACTC 3'	Oligonucléotide B pour la délétion de l'ORF YGR169C
1562	5'TTGTGAAAACTTTATTGACGATTGCTCGAGACAA TCAGTGGGCAAAGGAGGCCAAGAGGGGAGGGC 3'	Oligonucléotide A' pour la délétion de l'ORF YGR169C
1563	5'AATTTAATTTACATCCAAAGTACGGAAATGGTAA CTAACTATATACAACTTAAATAAATACTACTC 3'	Oligonucléotide B' pour la délétion de l'ORF YGR169C
1628	5'TTCGTGAAGGGGCGCTGGTTGGGTAAGTTACTAA TAGACGTTCTCGTTAGCAGCTGAAGCTTCGTACGC3'	Oligonucléotide A'' pour la délétion de l'ORF YGR169C
1631	5'TCCAAGCATGCAACCATAATTCTAGGTCTGATAGC ACAGGATCCTTTAAAGCATAGGCCACTAGTGGATCT G3'	Oligonucléotide B'' pour la délétion de l'ORF YGR169C
1291	5'GAACAATGTATGAATGGCTATAAGATCAGGATCA GTTAGGATATATTTTAGGCCAAGAGGGGAGGGC 3'	Oligonucléotide A pour la délétion de l'ORF YLR165C
1292	TATATATGTGTATTAGAAATAATTTAATAACCGATTA TTTATTGCTTATTCTTAAATAAATACTACTC 3'	Oligonucléotide B pour la délétion de l'ORF YLR165C
1293	AAAAAATTTTGAAAATGGATGAGTGCTCGTATAAT GGAATAGGAAACTTGGCCAAGAGGGGAGGGC 3'	Oligonucléotide A pour la délétion de l'ORF YDL036C
1294	TAACAGAAGACCAGTTTTCTATGTATTATTGTCCGAC AAGTCAGGAACCTCTTAAATAAATACTACTC 3'	Oligonucléotide B pour la délétion de l'ORF YDL036C
1337	5'ATGTCTCTGAAAAGCAGATACC 3'	Oligonucléotide C pour la vérification de la délétion de l'ORF YLR165C
1338	5'TTACTGGTCCCAGTTTTCCAAC 3'	Oligonucléotide D pour la vérification de la délétion de l'ORF YLR165C
1339	5'GGCTCGACAACTCAAAGGA 3'	Oligonucléotide C pour la vérification de la délétion de l'ORF YDL036C
1340	5'CTACCTTGGCCACGTATTCC 3'	Oligonucléotide D pour la vérification de la délétion de l'ORF YDL036C
1341	5'ATAATGAAGCCAGCGATGAC 3'	Oligonucléotide C pour la vérification de la délétion de l'ORF RIB2
1342	5'GTTTTTCATCGCTCATCTCGC 3'	Oligonucléotide D pour la vérification de la délétion de l'ORF RIB2
1343	5'GACTGAGAAAGGTACGGCCC 3'	Oligonucléotide C pour la vérification de la délétion de l'ORF YGR169C
1344	5'GGGATGACCGAGCCTAGCTAA 3'	Oligonucléotide D pour la vérification de la délétion de l'ORF YGR169C
1336	5'GCCGTAATCATTGACCAGAGCC 3'	Oligonucléotide F, interne au gène <i>TRP1</i> , pour la vérification des différentes délétions
1287	CCTCGACATCATCTGCCC	Oligonucléotide F', interne au gène <i>KAN</i> , pour la vérification de la délétion de l'ORF YGR169C

1368	5'ATCAGGATCAGTTAGGATATA 3'	Oligonucléotide E pour la vérification de la délétion de l'ORF YLR165C
1369	5'GAGTGCTCGTATAATGGAATAGG 3'	Oligonucléotide E pour la vérification de la délétion de l'ORF YDL036C
1564	5'TTCTTTCATTAAGGCAA 3'	Oligonucléotide E pour la vérification de la délétion de l'ORF YGR169C
1371	5'AGATTAACAAACCAATCAAA 3'	Oligonucléotide E pour la vérification de la délétion de l'ORF RIB2

Tableau 16: Description des oligonucléotides utilisés pour réaliser les délétions des ORF de *S. cerevisiae*, codant des Ψ -synthases potentielles, et pour vérifier les délétions de ces ORF.

Pour chaque ORF, un couple d'oligonucléotides A et B a été défini afin de réaliser la délétion. Pour l'ORF YGR169C, du fait d'une insertion à un site inattendu avec le premier couple d'oligonucléotides A/B, deux autres couples d'oligonucléotides A'/B' et A"/B" ont été utilisés, pour répéter l'expérience. Les oligonucléotides C, D et E ont été définis pour chaque ORF afin de vérifier la délétion par la méthode PCR sur colonies, selon le schéma représenté dans la Figure 75. Les nucléotides en bleu et en rouge correspondent aux nucléotides complémentaires aux régions codantes des gènes *TRP1* et *KAN*, respectivement.

ORF X sauvage :



Gène *TRP1* intégré par recombinaison homologue à la place de l'ORF X:



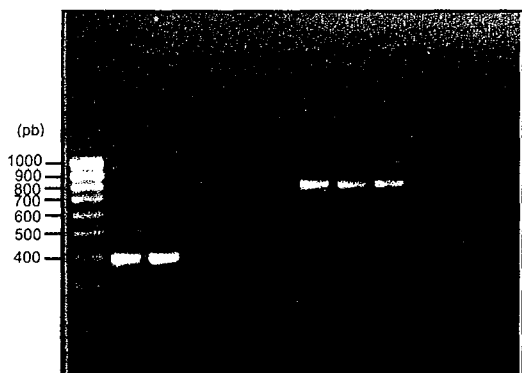
Résultats attendus lors de la vérification par PCR sur colonies de levure avec les couples d'oligonucléotides C/D et E/F:

	1. Délétion réussie	2. Délétion non réussie
Souche haploïde (n)	- couple C/D : rien - couple E/F : fragment de 400 pb	- couple C/D : fragment de 800 ou 900 pb - couple E/F : rien
Souche diploïde (nn)	- couple C/D : fragment de 800 ou 900 pb - couple E/F : fragment de 400 pb	- couple C/D : fragment de 800 ou 900 pb - couple E/F : rien

Figure 75: Principe des vérifications des délétions par approche PCR et résultats attendus dans les différents cas de figures possibles.

A. test du gène *TRP1* test du gène *YDL036C*

	délétant		WT		délétant		WT		-DNA	-DNA
	H	D	H	D	H	D	H	D		
M	1	2	3	4	5	6	7	8	9	10

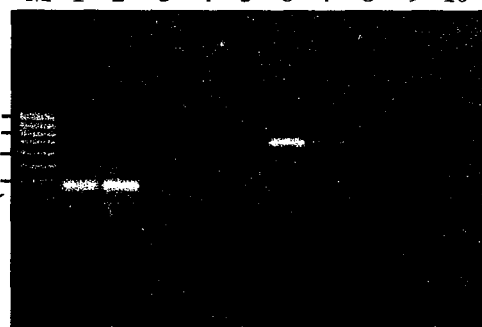


800pb
(*YDL036C*,
RIB2)

400pb (*TRP1*)

C.

	test du gène <i>TRP1</i>		test du gène <i>RIB2</i>		-DNA	-DNA				
	délétant	WT	délétant	WT						
	H	D	H	D						
M	1	2	3	4	5	6	7	8	9	10



B. test du gène *TRP1* test du gène *YLR165C*

	délétant		WT		délétant		WT		-DNA	-DNA
	H	D	H	D	H	D	H	D		
M	1	2	3	4	5	6	7	8	9	10

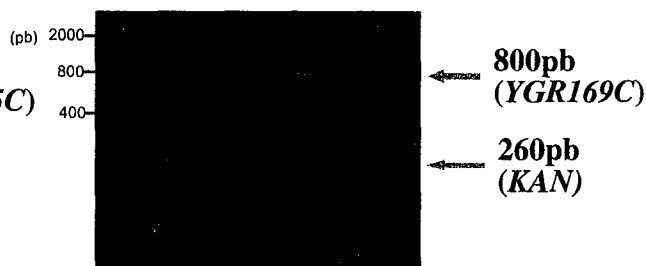


800pb
(*YLR165C*)

400pb
(*TRP1*)

D.

	test du gène <i>KAN</i>		test du gène <i>YGR169C</i>		-DNA	-DNA
	délétant	WT	délétant	WT		
M	1	2	3	4	5	6



800pb
(*YGR169C*)

260pb
(*KAN*)

Figure 76: Vérification par PCR de la délétion des ORF *YDL036C* (Panneau A), *YLR165C* (Panneau B), *RIB2* (Panneau C) et *YGR169C* (Panneau D)

Panneaux A à C: La présence du gène *TRP1* (pistes 1 à 4) et du gène de départ (pistes 5 à 8), a été testée, à la fois, sur les souches sauvages (WT) (pistes 3, 4, 7 et 8) et sur les souches haploïdes (H) et diploïdes (D) dans lesquelles la délétion a été réalisée (pistes 1, 2, 5 et 6).

Les couples d'oligonucléotides C/D et E/F (Figure 75; Tableau 16) ont été utilisés, respectivement, pour tester la présence des gènes de départ et du gène *TRP1*. Une expérience témoin, sans ajout de cellules issues d'une colonie, a été effectuée en parallèle dans le cas des PCR réalisées avec le couple d'oligonucléotides C/D (piste 10) et E/F (piste 9). Un marqueur de taille (M) a également été déposé. Les bandes d'ADN de 400 pb et 800 pb traduisent respectivement, la présence du gène *TRP1* et des gènes de départ. L'examen des différentes bandes obtenues montre que la délétion des ORF voulues a réussi dans les 3 cas.

Panneau D: La présence du gène *KAN* (pistes 1 et 2) et du gène *YGR169C* (pistes 3 et 4), a été testée, à la fois, sur les souches sauvages (WT) (pistes 2 et 4) et sur la souche haploïde sélectionnées (pistes 1 et 3). Les couples d'oligonucléotides C/D et E/F (Figure 75; Tableau 16) ont été utilisés, respectivement, pour tester la présence des gènes *YGR169C* et *KAN*. Un témoin, sans ajout de cellules issues d'une colonie a été effectué en parallèle dans le cas des PCR réalisées avec le couple d'oligonucléotides C/D (piste 5) et E/F (piste 6). Un marqueur de taille (M) a également été déposé. Les bandes d'ADN de 260 pb et 800 pb traduisent, respectivement, la présence du gène *KAN* et du gène *YGR169C*. L'examen des différentes bandes obtenues montre que la délétion du gène *YGR169C* a été réussie.

ont conduit au même résultat. Nous avons donc conclu que nous avons bien obtenu la délétion individuelle de chacune des 3 ORF.

Par contre, en ce qui concerne l'ORF YGR169C, nous n'avons jamais réussi à obtenir l'insertion du gène *TRP1* dans cette ORF. Nous sélectionnions des souches ayant acquis la capacité de croître sur milieu sans tryptophane, mais les insertions devaient avoir eu lieu ailleurs dans le génome, puisqu'en utilisant les couples d'oligonucléotides E/F, nous n'observions pas de produit d'amplification, alors qu'en utilisant le couple C/D, même pour les souches haploïdes sélectionnées, le produit d'amplification de 800 pb, correspondant à la présence du gène sauvage, était toujours obtenu. Une possibilité était que les régions utilisées pour la recombinaison homologue présentent une homologie avec une autre partie du génome. Nous avons sélectionné un autre couple d'oligonucléotides (A'/B', Tableau 16), pour produire le fragment PCR contenant le gène *TRP1*. Malheureusement, le même résultat a été obtenu. Des souches haploïdes et diploïdes ont été sélectionnées sur milieu sans tryptophane, mais d'après les analyses par PCR, la recombinaison homologue attendue n'avait encore pas lieu à l'endroit souhaité. Les deux couples d'oligonucléotides A/B et A'/B' généraient un fragment PCR destiné à déléter la totalité de l'ORF. Nous avons alors pensé que les difficultés rencontrées étaient peut-être liées à des variations des séquences flanquant l'ORF suivant les souches. Notre souche pouvait avoir une séquence légèrement différente de la souche pour laquelle la séquence complète du génome a été déterminée. Nous avons alors décidé de ne déléter qu'une partie de l'ORF (environ 80%). Pour cela, I. Ansmant a utilisé des oligonucléotides complémentaires à des régions internes de l'ORF YGR169C (A'' et B'', Tableau 16). Par ailleurs, comme une autre raison de l'échec des premiers essais de substitution pouvait être liée à l'utilisation du gène *TRP1*, nous avons réalisé un remplacement par le gène *KAN*, conférant la résistance à la kanamycine. Le fragment PCR a été produit à partir du plasmide KanMX2 (Wach *et al.*, 1994). Le principe de la délétion et les techniques utilisées étaient les mêmes que pour le remplacement par le gène *TRP1*. Une souche haploïde, présentant une grande délétion dans l'ORF a ainsi été obtenue et la présence de la délétion a été vérifiée par PCR, selon le même principe que celui utilisé pour les 3 autres ORF. Le produit d'amplification de 260 pb, obtenu avec le couple d'oligonucléotides E/F', révélait la présence du gène *KAN* à l'intérieur de l'ORF (Fig. 76, Panneau D). Cette observation, et l'absence de produit d'amplification correspondant au gène *YGR169C*, indiquait que la substitution avait bien eu lieu.

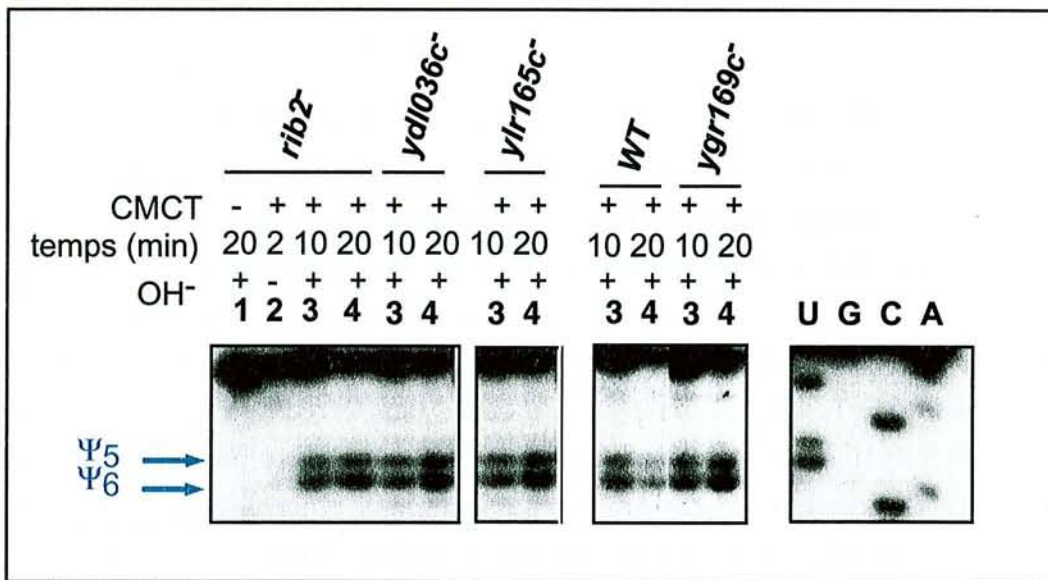
Nous avons ensuite recherché si l'une des 4 Ψ -synthases potentielles, codées par ces 4 ORF, était responsable de la formation d'un ou de plusieurs résidus Ψ dans les UsnRNA de *S. cerevisiae*. Pour chacune des souches où nous avons délété une ORF codant une Ψ -synthase potentielle, nous avons testé si un des résidus Ψ aux positions 5 et 6 du snRNA U1, 35 et 42 du

snRNA U2 et 99 du snRNA U5 disparaissait. En parallèle, l'implication possible d'une ou de plusieurs de ces enzymes dans la modification d'autres types d'ARN a également été testée. En effet, en plus des Ψ -synthases responsables de la formation des résidus Ψ dans les UsnRNA, d'autres Ψ -synthases restent à être identifiées chez *S. cerevisiae*. Chez cet organisme, la majorité des Ψ -synthases responsables de la formation des résidus Ψ dans les ARNt cytoplasmiques sont connues (Fig. 35 de l'Introduction), à l'exception de celles responsables de la formation des résidus Ψ aux positions 1, 13, 31, 32 et 35. Certaines des 4 ORF analysées pouvaient être impliquées dans la formation de ces résidus. De même, la Ψ -synthase responsable de la formation de l'unique résidu Ψ , présent dans les ARNr mitochondriaux et localisé en position 2819 de l'ARNr 21S (Bakin *et al.*, 1994), n'avait pas été identifiée. Nous avons donc débuté, en collaboration, avec l'équipe d'H. Grosjean (Gif-sur-Yvette), l'étude des effets produits par la délétion de chacune des ORF YDL036C, YLR165C et RIB2, sur la formation des résidus Ψ dans les ARNt cytoplasmiques. Nous étudions en ce moment, l'effet de chaque délétion sur les ARNt mitochondriaux. L'étude de la souche *ygr169c-*, qu' I. Ansmant vient d'obtenir, est en cours.

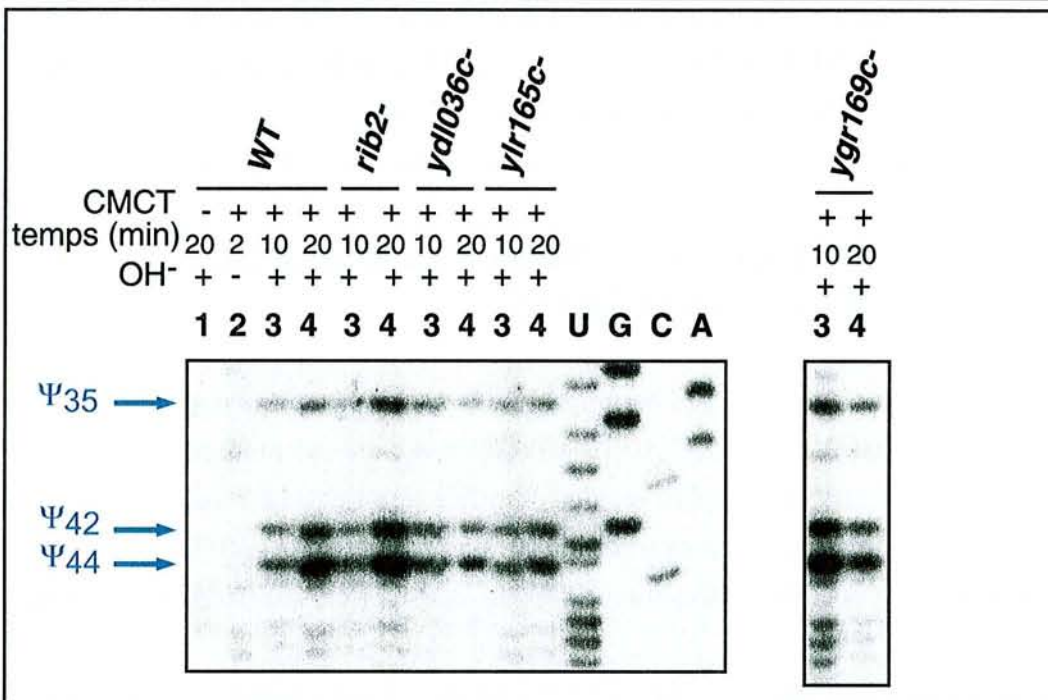
3. ANALYSE DES RESIDUS Ψ PRESENTS DANS LES UsnRNA DES SOUCHES PORTANT LES DELETIONS

Des ARN totaux, des souches haploïdes de *S. cerevisiae* présentant une délétion dans l'une des 4 ORF YDL036C, YLR165C, RIB2 et YGR169C, ont été préparés. Ces ARN ont été modifiés par le CMCT et l'analyse des résidus Ψ qu'ils portaient a été réalisée par extension d'amorce par la transcriptase inverse en utilisant des oligonucléotides spécifiques de chacun des UsnRNA, comme décrit dans la publication n°2. Cette analyse a été réalisée par I. Ansmant, dans le cadre de son stage de DEA.

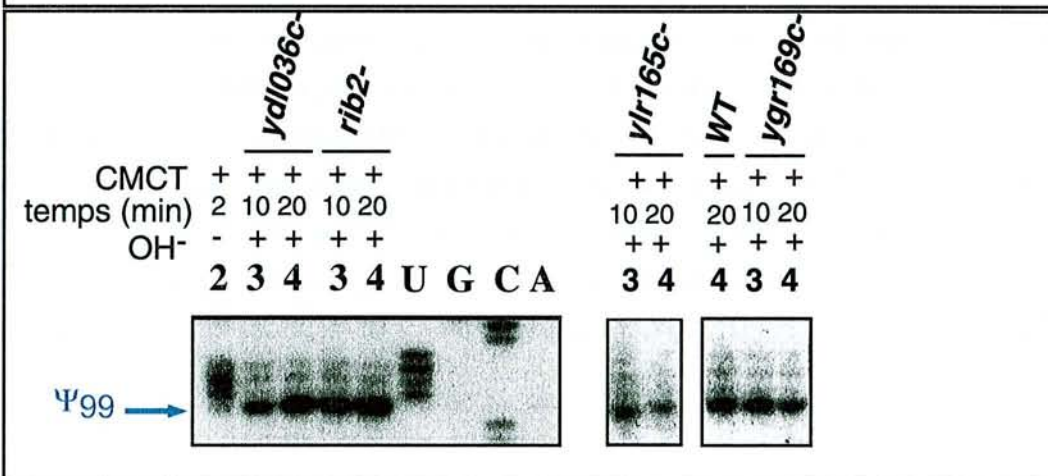
Comme on peut le constater dans la Figure 77, les 2 résidus Ψ aux positions 5 et 6 du snRNA U1 sont détectés dans toutes les souches mutantes. Il en est de même pour les résidus 35 et 42 du snRNA U2 et pour le résidu 99 du snRNA U5. Il apparaissait donc clairement que la délétion de chacune des ORF codant les 4 Ψ -synthases potentielles n'entraînait pas de changement dans la formation des résidus Ψ dans les UsnRNA de *S. cerevisiae*. Soit parmi les enzymes codées par ces 4 ORF, certaines avaient des propriétés redondantes, c'est-à-dire qu'en absence de l'une, l'autre pouvait générer la modification, soit les 4 Ψ -synthases agissaient sur d'autres ARN que les UsnRNA. Nous avons donc testé l'effet des délétions sur la formation des résidus Ψ dans d'autres ARN.



U1



U2



U5

Figure 77: Recherche des résidus Ψ présents dans les snRNA U1, U2 et U5, extraits des différentes souches où les ORF des Ψ-synthases potentielles avaient été délétées.

Les analyses ont été réalisées sur les snRNA U1, U2 et U5 contenus dans un mélange d'ARN totaux extraits des souches *rib2*⁻, *ydl036c*⁻, *ygr169c*⁻ et *ylr165c*⁻. La préparation de ces ARN est décrite dans la publication n°5. Dix µg d'ARN totaux de chacune des souches mutantes ont été modifiés par le CMCT pendant 2 min (piste 2), 10 min (pistes 3) ou 20 min (piste 4), suivi (pistes 1, 3 et 4) ou non (piste 2) d'un traitement alcalin. Des ARN non modifiés par le CMCT ont servi de témoins (piste 1). La présence de résidus Ψ a été détectée par extension d'amorce par la transcriptase inverse. Les oligonucléotides, utilisés comme amorces, sont les mêmes que dans la publication n°2.

4. ANALYSE DES RESIDUS Ψ PRESENTS AUX POSITIONS 13, 32 ET 35 DES ARNt CYTOPLASMIQUES, DES SOUCHES PORTANT UNE DELETION DANS L'UNE DES 4 ORF ETUDIEES

L'équipe d'H. Grosjean a développé un système permettant de détecter *in vitro* l'activité cellulaire capable de former des résidus Ψ aux positions 13, 32 et 35 des ARNt. La méthode repose sur l'incubation de l'ARNt^{Asp} de *S. cerevisiae* et de l'ARNt^{Tyr} de *A. thaliana*, produits par transcription *in vitro*, avec des extraits cellulaires de *S. cerevisiae*. Cette approche a été utilisée par Y. Motorin de l'équipe d'H. Grosjean, pour étudier les activités contenues dans les extraits cellulaires des 4 souches portant des délétions dans les ORF RIB2, YLR165C, YGR169C et YDL036C. Après incubation avec l'extrait, la présence des activités Ψ -synthases, responsables de la formation des résidus Ψ_{13} et Ψ_{32} dans l'ARNt^{Asp} et Ψ_{35} dans l'ARNt^{Tyr} peut être mise en évidence dans l'extrait. En effet, les résidus Ψ formés dans ces ARNt peuvent être localisés par la technique du plus proche voisin (voir Fig. 51) et ceci en choisissant des nucléotides radioactifs appropriés lors de la transcription *in vitro* des deux ARNt. Si une souche porte la délétion d'une ORF codant une des Ψ -synthases responsables de la formation d'un de ces résidus Ψ , la présence de ce résidu Ψ n'est pas attendue après incubation des transcrits ARNt avec l'extrait. Le protocole expérimental utilisé est détaillé dans la publication n°5. Comme nous l'avons déjà mentionné, la Ψ -synthase Pus1p modifie *in vitro* l'ARNt^{Tyr} de *S. cerevisiae* en position 35 (Motorin *et al.*, 1998). Mais, elle n'agit pas *in vitro* sur l'ARNt^{Tyr} de *A. thaliana*, or un résidu Ψ est formé en position 35 de cet ARN après incubation avec un extrait cellulaire de *S. cerevisiae*. L'utilisation, dans cette étude, de l'ARNt^{Tyr} de *A. thaliana* permettait donc de détecter la présence de la deuxième activité ARNt₃₅: Ψ -synthase dans les extraits.

Les résultats obtenus montraient que des résidus Ψ sont formés aux positions 13 et 32 dans l'ARNt^{Asp} et à la position 35 dans l'ARNt^{Tyr} d'*A. thaliana*, quel que soit l'extrait cellulaire utilisé. La délétion de chacune des ORF RIB2, YLR165C, YGR169C et YDL036C n'empêche donc pas la formation de résidus Ψ à ces positions.

5. RECHERCHE DE LA PRESENCE DU RESIDU Ψ LOCALISE EN POSITION 2819 DANS L'ARNr 21S MITOCHONDRIAL, DANS LES DIFFERENTES SOUCHES MUTANTES

La présence du résidu Ψ_{2819} dans l'ARNr 21S mitochondrial a été testée par modification par le CMCT, suivie d'une analyse par extension d'amorce en utilisant les conditions expérimentales et l'oligonucléotide décrits dans la publication n°5. Comme le montre la Figure 78, ce résidu Ψ est formé dans les souches de levure *rib2-*, *ygr169c-* et *ydl036c-*. Ces trois gènes

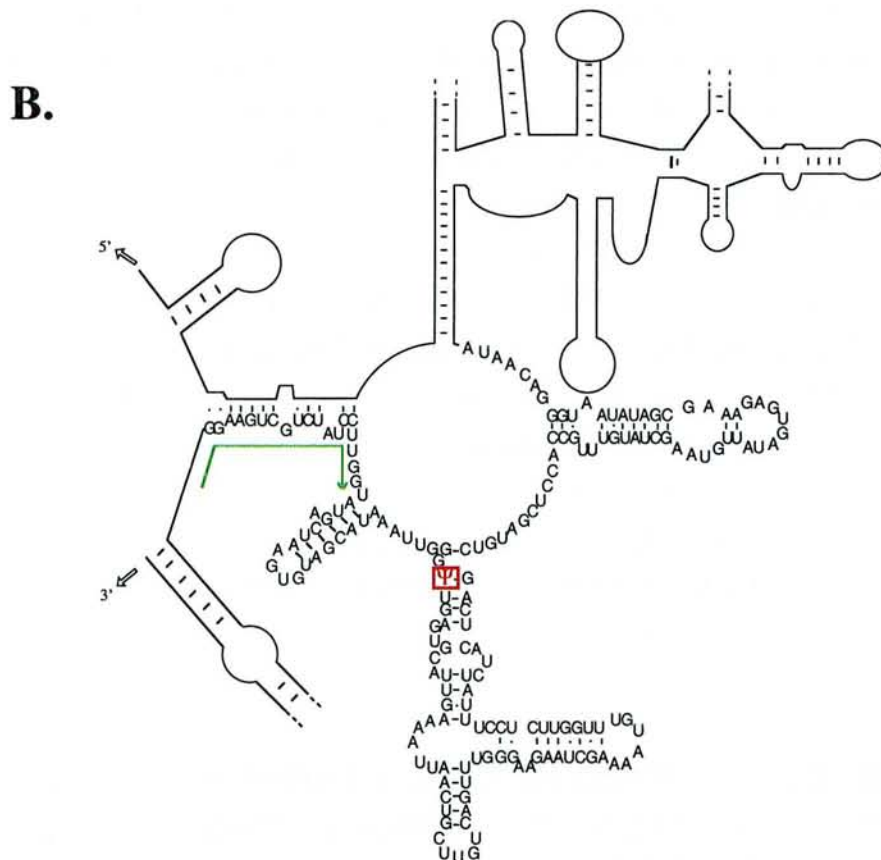
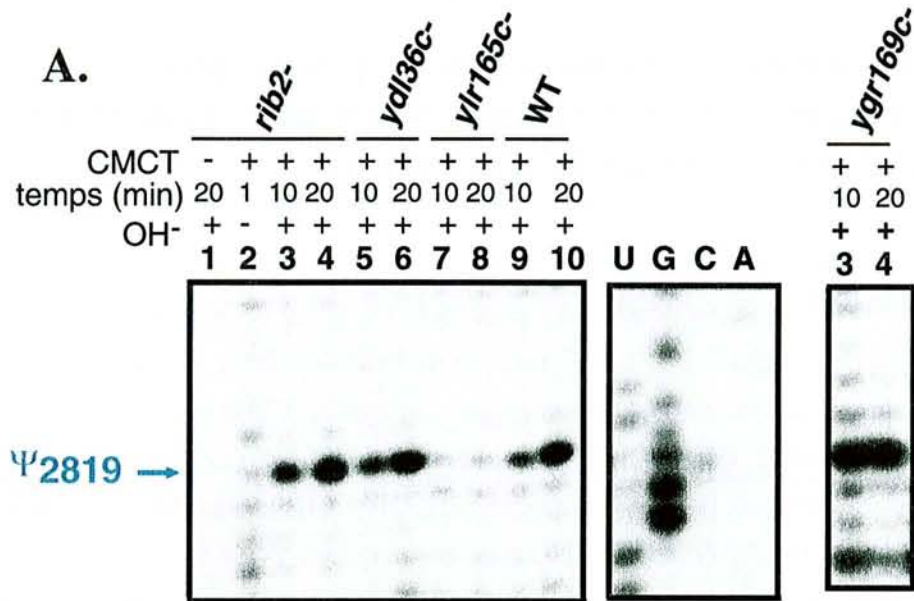


Figure 78: Recherche des résidus Ψ présents dans l'ARNr mitochondrial 21S issu des différentes souches, où les ORF de Ψ -synthases potentielles de *S. cerevisiae* étaient délétées.

Panneau A: Les analyses ont été réalisées sur l'ARNr mitochondrial 21S contenu dans un mélange d'ARN totaux des souches sauvage, *rib2-*, *ygr169c-*, *ylr165c-* et *ydl036c-*. Ces ARN totaux ont été préparés comme décrit dans la publication n°5. Dix μg d'ARN totaux de chacune des souches mutantes ont été modifiés par le CMCT pendant 1 min (piste 2), 10 min (piste 3) ou 20 min (piste 4), ce traitement étant suivi (pistes 1, 3 et 4) ou non (piste 2) d'un traitement alcalin. Des ARN non modifiés par le CMCT ont servi de témoins (piste 1). La présence de résidus Ψ a été détectée par extension d'amorce par la transcriptase inverse.

Panneau B: Structure du domaine V de l'ARNr 21S mitochondrial de *S. cerevisiae*. La région d'appariement de l'oligonucléotide utilisé est représentée, ainsi que le résidu Ψ_{2819} .

ne codent donc pas pour la Ψ -synthase responsable de la formation du résidu Ψ_{2819} . La recherche de la présence de ce résidu dans l'ARNr 21S de la souche *ylr165c*- est présentée dans la publication n°5. La disparition de ce résidu dans la souche *ylr165c*- et la restauration de sa formation, après transformation de cette souche mutante par un plasmide portant l'ORF YLR165C, suggèrent fortement que cette ORF code la Ψ -synthase responsable de la formation du résidu Ψ en position 2819 de l'ARNr 21S mitochondrial. Par homogénéité avec la nomenclature des Ψ -synthases de *S. cerevisiae*, nous avons appelé cette enzyme Pus5p.

Publication n°5

The YLR165c ORF is implicated in formation of the Ψ residue at position 2819 in the *Saccharomyces cerevisiae* 21S mitochondrial large ribosomal RNA

Ansmant I., Massenet S., Motorin, Y., Grosjean H. and Branlant C.

Soumis à NAR

Publication n°5

The YLR165c ORF is implicated in formation of the Ψ residue at position 2819 in the *Saccharomyces cerevisiae* 21S mitochondrial large ribosomal RNA.

I. Ansmant, S. Massenet, Y. Motorin, H. Grosjean and C. Branlant.

Soumis à Nucleic Acids Res.

Identification of the *Saccharomyces cerevisiae* RNA:pseudouridine synthase responsible for formation of Ψ_{2819} in 21S mitochondrial ribosomal RNA

I. Ansmant¹, S. Massenet¹, H. Grosjean², Y. Motorin², and C. Branlant¹.

¹Laboratoire de Maturation des ARN et Enzymologie Moléculaire, UMR 7567 CNRS-UHP Nancy I, Faculté des Sciences, BP 239, 54506 Vandoeuvre-les-Nancy Cedex, France

²Laboratoire d'Enzymologie et Biochimie Structurales, UPR CNRS, 1 avenue de la Terrasse, 91198 Gif-sur-Yvette, France

*Corresponding author :

Christiane Branlant

Laboratoire de Maturation des ARN et Enzymologie Moléculaire

UMR 7567 CNRS-UHP Nancy I

Faculté des Sciences, BP 239

54506 Vandoeuvre-les-Nancy Cedex, France

Tel. 33.3.83.91.20.92

Fax 33.3.83.91.20.93

e-mail cbranlant@scbim.u-nancy.fr

Abbreviations: Ψ , pseudouridine; nt, nucleotides; RT, reverse transcription; LSU rRNA, large subunit rRNA; CMCT, N-cyclohexyl-N'-(2-morpholinoethyl)-carbodiimid metho-p-toluolsulfonat; 2D-TLC, two-dimensional thin layer chromatography.

ABSTRACT

So far, four RNA:Ψ-synthases have been identified in yeast *Saccharomyces cerevisiae*. Altogether, they act on cytoplasmic and mitochondrial tRNAs, U2 snRNA and rRNAs from cytoplasmic ribosomes. However, RNA:Ψ-synthases responsible for several U to Ψ conversions in tRNAs and UsnRNAs remained to be identified. Based on conserved amino-acid motifs in already characterised RNA:Ψ-synthases, four additional ORFs encoding putative RNA:Ψ-synthases were identified in *S. cerevisiae*. Upon disruption of one of them, the YLR165c ORF, we found that the unique Ψ residue normally present in the fully matured mitochondrial rRNAs (Ψ₂₈₁₉ in 21S rRNA) was missing, while Ψ residues at all the tested pseudouridylation sites in cytoplasmic and mitochondrial tRNAs and in nuclear UsnRNAs were retained. The selective U to Ψ conversion at position 2819 in mitochondrial 21S rRNA was restored upon transformation of the deleted yeast strain with a plasmid expressing the YLR165c ORF. Hence, the nuclear YLR165c ORF, renamed *PUS5* for nomenclature homogeneity, most probably encodes the *S. cerevisiae* RNA:Ψ-synthase (Pus5p) specific for mitochondrial 21S rRNA. As already noticed for other mitochondrial RNA modification enzymes, no canonical mitochondrial targeting signal was identified in Pus5p. Our results also show that Ψ₂₈₁₉ in mitochondrial 21S rRNA is not essential for cell viability.

Key words: pseudouridine / rRNA / yeast / pseudouridine-synthase / mitochondria / maturation enzyme / gene deletion

INTRODUCTION

Pseudouridine (Ψ), an isomer of uridine with a C-C bond between C₅ of uracil and C₁ of ribose instead of the N₁-C₁ glycosidic bond in uridine, is the most abundant and universally found modified nucleoside in RNA. Indeed, numerous Ψ residues are present at specific positions in rRNAs and tRNAs from *Archaea*, *Bacteria* and *Eukarya* and in UsnRNAs from *Eukarya* (for review, (1-3). They are all formed post-transcriptionally by a family of enzymes designated as RNA: Ψ -synthases. The enzymatic reaction involves the cleavage of the N₁-C₁ glycosidic bond, rotation of the uracil ring and formation of a new C₁-C₅ bond without requiring any cofactor or energy source like ATP. The detailed mechanism of this reaction is however still unknown, yet it is thought that the carboxyl group of a conserved aspartate residue present in all RNA: Ψ -synthases identified so far plays an essential role (4-7), see below).

The first characterised RNA: Ψ -synthase was a tRNA-specific enzyme, PSU-I (recently renamed TruA), that catalyses the formation of Ψ at positions 38, 39 and/or 40 in several tRNAs from *Salmonella typhimurium* (8). Its counterpart in *Escherichia coli* was discovered and studied later (9). Since then, remarkable progress was achieved in the characterisation of many other *E. coli* RNA: Ψ -synthases and of their corresponding genes. Among them, TruB was shown to be site-specific for the formation of the conserved Ψ residue at position 55 in tRNAs (10). Five other characterised *E. coli* RNA: Ψ -synthases are responsible for the formation of nine out of ten Ψ residues present in *E. coli* ribosomal RNAs. Altogether, the two multisite specific enzymes, RluC and RluD, catalyse the formation of Ψ residues at 6 distinct sites in *E. coli* 23S rRNA (positions 955, 2504 and 2580, and positions 1911, 1915 and 1917, respectively) (5,11,12). Two other RNA: Ψ -synthases, RsuA and RluE, catalyse each an unique U to Ψ conversion at positions 516 in 16S rRNA and 2457 in 23S rRNA, respectively (6,13,14). The last enzyme, RluA, displays a dual-substrate specificity and is implicated in the formation of both Ψ_{746} in 23S rRNAs and Ψ_{32} in tRNAs (15,16).

Analysis of the sequence similarity between already characterised RNA: Ψ -synthases showed that these enzymes belong to four distinct families: TruA, TruB, RsuA and RluA

(17). Three of these families, TruB, RsuA and RluA, share two short conserved motifs (motifs I and II) (17). In contrast, the TruA family displays only little homology with proteins of the three other RNA:Ψ-synthase families. Only a GRTD motif conserved in this group is homologous to a tetra-amino acid sequence of the motif II found in the three other families (6,17). Site-directed mutagenesis of the *E. coli* RNA:Ψ-synthases RluA, TruA, TruB and RsuA demonstrated that this block of amino acids, that is conserved in all families, belongs to the enzyme active site. The aspartate residue (D) in the motif II is essential for catalysis (4-7). Extensive searches in DNA databases identified more than one hundred genes coding for putative Ψ-synthases in 47 different organisms covering altogether the three domains of living organisms (*Archaea*, *Bacteria* and *Eukarya*) (6). Based on this comparative analysis, each of these proteins can be classified in one of the already identified RNA:Ψ-synthase families.

Nine genes encoding putative RNA:Ψ-synthases have been identified in the *S. cerevisiae* genome (see (6)). Five of them correspond to the already characterised RNA:Ψ-synthases (Pus1p, Pus2p, Pus3p, Pus4p and Cbf5p) (18-22). Pus1p, Pus2p and Pus3p are members of the TruA family, whereas Pus4p and Cbf5p belong to the TruB family. The uridine targets in RNAs were identified for four of these enzymes (Pus1p, Pus3p, Pus4p and Cbf5p). Pus3p catalyses Ψ formation at positions 38 or 39 in cytoplasmic as well as mitochondrial tRNA-anticodon loops (22), while Pus1p is implicated in the Ψ formation at 8 different positions in cytoplasmic tRNAs (18,20) and also at position 44 in U2 snRNA (23). Pus4p, like TruB in *E. coli*, is responsible for the site-specific formation of the highly conserved Ψ residue at position 55 in both cytoplasmic and mitochondrial tRNAs (19). Finally, Cbf5p, a protein present in the H/ACA snoRNPs (21,24), mediates the formation of most, if not all, Ψ residues in the yeast 18S and 25S rRNAs at the precursor level (35S pre-rRNA) in the nucleolus (21,25). In the latter case, selection of the Ψ sites is mediated through direct base pairing between the snoRNAs and the nucleolar pre-rRNA and in a large number of cases the direct involvement of the snoRNA in modification at a given site was proven experimentally (26,27).

Based on the sequence homology with *E. coli* RNA:Ψ-synthases, the four other genes encoding putative RNA:Ψ-synthases in *S.cerevisiae* (ORFs YDL036c, YGR169c, YOL066c, also called RIB2, and YLR165c) all belong to the RluA family (6). The corresponding proteins have not been characterised. Their RNA substrate specificity and their cellular location have not been studied yet. In this work, we show that the deletion of the YLR165c ORF does not alter the pseudouridylation pattern of tRNAs and UsnRNAs, but causes the specific loss of the single Ψ residue present at position 2819 in the *S.cerevisiae* mitochondrial 21S LSU rRNA. The formation of Ψ₂₈₁₉ in mitochondrial 21S rRNA is restored by transformation of the disrupted strain with a plasmid expressing the YLR165c ORF. These observations strongly suggest that the YLR165c ORF codes for a novel RNA:Ψ-synthase responsible for the U to Ψ conversion at position 2819 in the yeast mitochondrial 21S rRNA. For homogeneity of nomenclature of yeast RNA:Ψ-synthases we rename this ORF as *PUS5*.

MATERIALS AND METHODS

YLR165c ORF Disruption

The YLR165c ORF (accession number in TREMBL, Q06244) was disrupted in the *S. cerevisiae* strains BMA64 (*Mat a/a ura 3-1 ade 2-1 leu 2-3, 112 his3-11, 15 trp1D can 1-100*) and BMA64-1A (*Mat a ura 3-1 ade 2-1 leu 2-3, 112 his3-11, 15 trp1D can 1-100*), by replacement of the entire coding region with the *TRP1* auxotrophic marker. Replacement was done by PCR-targeting method using the *TRP1* gene flanked by short regions homologous to the target locus of the *S. cerevisiae* DNA (28,29). The *TRP1* gene was amplified from plasmid pFL35. Both haploid and diploid strains were transformed with 3 or 6 μ g of PCR-generated DNA fragments using the lithium acetate method (30). Positive transformants (Δ YLR165c) were selected on YNB plates lacking tryptophan. The disruption was confirmed by PCR amplification on yeast colonies.

S. cerevisiae RNA preparation

Total RNA from the wild type and the Δ YLR165c yeast strains was prepared as described previously (23).

Ψ residues analysis in UsnRNAs and mitochondrial 21S LSU rRNA

Pseudouridine residues in the UsnRNAs and mitochondrial 21S LSU rRNA of the Δ YLR165c and wild type *S. cerevisiae* strains were identified by the CMCT-RT approach (31), with the modifications described previously (23). The oligonucleotides used for Ψ identification in U1, U2 and U5 snRNAs were previously described (23). The oligonucleotide complementary to nt 2846 to 2866 of the mitochondrial 21S LSU rRNA was used to detect the formation of a Ψ residue at position 2819 in this RNA.

Test for RNA: Ψ -synthase activities acting at positions 13, 31, 32 or 35 in tRNAs

The preparation of cell-free extracts from wild-type and Δ YLR165c yeast strains was done as described earlier (32). The enzymatic activity of RNA: Ψ -synthases acting at positions 13 or 32 and at position 35 in T7 transcripts of internally [32 P]labeled tRNA was tested using as substrates yeast tRNA^{Asp}(anticodon GUC) and *A. thaliana* pre-tRNA^{Tyr}

(anticodon GUA, including the intron). Quantification of pseudouridine formation in these tRNA was performed after complete T2 RNase digestion and separation of the digest by 2D-TLC, as described (20).

The presence of pseudouridine residues at positions 31 and 32 in the yeast mitochondrial tRNA^{Met} and tRNA^{Ser} was tested using the CMCT/RT approach as described for analysis of mitochondrial 21S rRNA. The oligonucleotides complementary to mitochondrial tRNA^{Met}(anticodon CAU) (nt 46-68) and to tRNA^{Ser}(anticodon GCU) (nt 44-57) were used as primers. The presence of Ψ at position 31 in cytoplasmic tRNA^{Met}(anticodon CAU) was tested by the same approach, using the oligonucleotide complementary to bases 40-57.

Complementation of the Δ YLR165c *S. cerevisiae* strain by plasmid p413TEF-YLR165c

The YLR165c ORF was amplified by PCR from the genomic DNA of strain BMA64-1A using an oligonucleotide corresponding to the 5' end of the ORF (CAAGCGGAATTCGGatgtcaaaaaagcag) and an oligonucleotide complementary to the 3' end of the ORF (GTTCGCCTTAAGTTAttactggtcccagtttc). Sense and antisense sequences of the YLR165c ORF are shown in lower-case. The 5'-oligonucleotide generated an *Eco*RI restriction site (underlined). The amplified PCR product was subcloned in the pT-Adv cloning vector (Clontech). The resulting construct was cleaved by *Eco*RI and the fragment containing ORF YLR165c was inserted at the unique *Eco*RI site of the shuttle vector p413TEF (33), downstream from the TEF promoter. The sequence of the insert was verified by DNA sequencing. The haploid BMA64-1A- Δ YLR165c strain was transformed with this construct using the standard lithium acetate procedure (34).

Protein alignment and prediction of cell sorting signals

The multiple sequence alignment of RNA: Ψ -synthases was constructed with the ClustalW (35) and MACAW (36) softwares and further refined manually. Presence of putative cellular targeting signals in the Pus5p sequence was tested with the PSORT (37), <http://psort.nibb.ac.jp>, WEB site) and NNPLS softwares (38), <http://predict.sanger.ac.uk/nnpsl/>, WEB site).

RESULTS

To test whether the protein encoded by the *S. cerevisiae* YLR165c ORF is a RNA:Ψ-synthase and to identify its natural RNA substrate, we performed a one-step gene replacement of this ORF by the gene encoding n-(5'-phosphoribosyl)-anthranilate isomerase (*TRP1* auxotrophic marker). The resulting ΔYLR165c yeast cells did not show any growth phenotype compared to the parental yeast strain. The cell-free extract and total RNA fraction obtained from the strain with the ΔYLR165c ORF were used to determine the consequences of the gene deletion on the pseudouridylation pattern of various cellular RNAs.

YLR165c ORF is not a tRNA:Ψ-synthase

Previous examination of 34 cytoplasmic and 17 mitochondrial tRNAs so far sequenced in yeast (39) revealed that pseudouridine can be found at 15 distinct positions in cytoplasmic tRNAs and at 8 positions in mitochondrial tRNAs (for review, (1). As mentioned in the introduction, three tRNA:Ψ-synthases (Pus1p, Pus3p and Pus4p) were already found to be responsible for Ψ formation at 11 out of the 15 potential Ψ positions in cytoplasmic tRNAs and at 5 out of the 8 potential Ψ positions in mitochondrial tRNAs (see (18-20,22). Only the enzymes catalysing the formation of Ψ residues at positions 1, 13, 31 and 32 in cytoplasmic tRNAs and at positions 31, 32 and 72 in mitochondrial tRNAs remain to be identified. Moreover, while yeast Pus1p is capable of forming Ψ₃₅ in tRNA^{Tyr} *in vitro*, the disruption of the corresponding *PUS1* gene does not affect the formation of Ψ₃₅ *in vivo*, attesting for the existence of another enzyme acting on U₃₅ with overlapping specificity with Pus1p (20).

Here we tested whether the protein encoded by the YLR165c ORF is involved in Ψ formation at positions 13, 31, 32 or 35 in cytoplasmic tRNAs and at positions 31 and 32 in mitochondrial tRNAs. To test the activity towards the positions 13 and 32 in cytoplasmic tRNAs, we used as substrate an *in vitro* produced wild type yeast tRNA^{Asp} that was internally radiolabeled by [α -³²P]ATP or UTP. Indeed, ATP and UTP labeling of tRNA^{Asp} transcript followed by nearest neighbour approach allows the selective detection of Ψ₁₃

and Ψ_{32} respectively (20). The putative tRNA: Ψ_{35} -synthase activity of the YLR165c ORF product was tested by comparing the action of extracts from wild-type and Δ YLR165c yeast cells on the heterologous [α - 32 P]ATP radiolabeled pre-tRNA^{Tyr} from the plant *A. thaliana*. Indeed, previous data showed that the purified yeast Pus1p does not modify *A. thaliana* tRNA^{Tyr} at position 35 *in vitro*, whereas the putative tRNA: Ψ_{35} -synthase present in the yeast cell extract does catalyse Ψ_{35} formation. After incubation of yeast tRNA^{Asp} and *A. thaliana* tRNA^{Tyr} at 30°C with yeast extracts, the presence of Ψ_{13} , Ψ_{32} or Ψ_{35} was analysed by digestion of the tRNA transcripts with RNase T2 followed by 2D TLC chromatography. The results demonstrated that deletion of the YLR165c ORF does not affect at all the enzymatic activity corresponding to Ψ formation at each of the three sites tested in cytoplasmic tRNAs (Table I).

The influence of the disruption of the YLR165c ORF on the pseudouridylation of yeast cytoplasmic tRNA at position 31 and mitochondrial tRNAs at positions 31 and 32 was tested with the CMCT/RT approach as described in the Materials and Methods. The primer extension analysis demonstrated that Ψ_{31} and Ψ_{32} in mitochondrial tRNA^{Met} and tRNA^{Ser} as well as Ψ_{31} in cytoplasmic tRNA^{Met} were retained upon gene disruption (data not shown).

YLR165c ORF is not involved in Ψ formation of UsnRNAs

Six Ψ residues have been identified in UsnRNAs from *S. cerevisiae*, two of them were found in U1 snRNA, three in U2 snRNA and one in U5 snRNA (23). So far, only one enzyme, the multisite- and dual-specific RNA: Ψ -synthase Pus1p, was shown to catalyse the formation of one of these Ψ residues (at position 44 in U2 snRNA, (23). The enzymes responsible for Ψ formation at the 5 other sites in UsnRNAs have still to be identified. To test whether the YLR165c ORF encodes such a putative snRNA: Ψ -synthase the Ψ content of UsnRNAs from the Δ YLR165c haploid strain was tested and compared with that of the isogenic yeast strain. To this goal, total RNA extracted from the wild-type and Δ YLR165c strains were treated by CMCT, followed by alkaline treatment to remove the CMC group from U. The remaining bulky CMC- Ψ were identified by reverse transcription using various oligonucleotides complementary to U1, U2 or U5 snRNA. As shown in Figure 1,

all the Ψ residues present in UsnRNAs are conserved upon deletion of the YLR165c ORF, indicating that the corresponding protein is not required for modification of UsnRNAs *in vivo*.

YLR165c ORF is implicated in formation of one Ψ residue in mitochondrial 21S rRNA

Finally, we tested the possibility that the protein encoded by the YLR165c ORF acts on the *S. cerevisiae* mitochondrial LSU rRNA, that is known to contain a single Ψ residue at position 2819 (31). Modification by CMCT was done on total RNA fractions from the Δ YLR165c haploid strain and the isogenic yeast strain, and the presence of Ψ_{2819} in the mitochondrial LSU 21S rRNA was tested as above by primer extension analysis. As shown in Figure 2A, the results obtained with the Δ YLR165c strain differed from those obtained with the wild type strain by the absence of the Ψ residue in the mitochondrial 21S rRNA.

In order to verify that the YLR165c protein is indeed involved in Ψ_{2819} formation, the Δ YLR165c strain was complemented by the p413TEF-YLR165c plasmid, containing the wild type ORF under the control of the TEF promoter. As shown in Figure 2B (lanes 3 and 4), Ψ formation at position 2819 in the mitochondrial LSU 21S rRNA was restored in the transformed yeast strain.

DISCUSSION

The protein encoded by YLR165c ORF is homologous to rRNA:Ψ-synthases from *E. coli*

Homology search using known amino acid sequences of well characterised *E. coli* RNA modification enzymes as queries has been successfully used for identifying additional putative genes belonging to the same enzyme family in *E. coli* as well as in other organisms including yeast (see, for examples (6,17,40-42)). In the present work we focused our attention on the protein product of the YLR165c ORF from yeast. This protein is homologous to the already characterised "dual specific" *E. coli* RNA:Ψ-synthase (RluA), which catalyses the formation of Ψ₇₄₆ in 23S rRNA and of Ψ₃₂ in tRNAs (15,16) and to the multi-site specific RluC and RluD enzymes, which catalyse the formation of several Ψ in 23S rRNA (5,11,12). The protein product of yeast YLR165c ORF displays the two characteristic motifs (motifs I and II) common to proteins of the three distinct RNA:Ψ-synthase families TruA, RsuA and RluA and also an additional motif III common to the RsuA and RluA families. More importantly, it bears the strictly conserved aspartic acid residue within the short sequence G/HLRD of motif II. This residue was demonstrated to be essential for RNA:Ψ-synthase activity (4,6,7,16). Altogether, this strongly argues in favour of an RNA:Ψ-synthase activity of the YLR165c encoded protein. The question was to know which cellular RNA(s) is(are) the substrate(s) of the YLR165c ORF product. To this end we tested the effect of the deletion of the YLR165c ORF on the formation of Ψ residues in RNAs at positions for which the corresponding RNA:Ψ-synthases have not been yet identified. These include positions 13, 31, 32 and 35 in cytoplasmic tRNAs, position 31 and 32 in mitochondrial tRNAs, 5 and 6 in U1 snRNA, 35 and 42 in U2 snRNA and 99 in U5 snRNA.

YLR165c ORF encodes a mitochondrial rRNA:Ψ-synthase

While disruption of the *S. cerevisiae* YLR165c ORF had no effect on the pseudouridylation pattern of UsnRNAs (U1, U2 and U5) and of tRNAs at positions 13, 31, 32 and 35, we showed that this disruption led to the disappearance of the unique residue Ψ₂₈₁₉ in the mitochondrial 21S LSU rRNA. Furthermore, plasmid p413TEF-YLR165c

expressing the YLR165c ORF restored Ψ_{2819} formation in the mitochondrial 21S rRNA of the Δ YLR165c strain. These results clearly demonstrate the implication of the YLR165c encoded protein in modification of the mitochondrial 21S rRNA. Since the YLR165c ORF product contains several motifs that are highly conserved in RNA: Ψ -synthases (see references above), we propose that it is the mitochondrial *S. cerevisiae* rRNA: Ψ -synthase converting U to Ψ at position 2819 in 21S mitochondrial rRNA. This conclusion is reinforced by the evident sequence homology that exists between the YLR165c ORF product and the three well characterised *E.coli* rRNA: Ψ -synthases RluA, RluD and especially RluC, which are all acting on 23S rRNA (> 25% of identity). Figure 3 shows details of the sequence alignment.

As the YLR165c ORF product is tightly related to the *E. coli* enzymes of the RluA family that modify rRNAs and because *E. coli* rRNA: Ψ -synthases do not require snoRNA-guides for activity (discussed in (3), it seems very unlikely that the YLR165c encoded mitochondrial rRNA: Ψ -synthase depends upon a guide-RNA located in mitochondria. The same conclusion was also reached for the specific formation of 2-O'-methyl-ribose of G₂₂₇₀ (Gm) catalysed by the Pet56 mitochondrial site-specific rRNA:ribose methylase (discussed in (43,44).

Since mitochondrial 21S rRNA is synthesised and matured exclusively in mitochondria, the corresponding rRNA: Ψ -synthase has to enter into mitochondrial compartment, by a mechanism which involves some kind of addressing signal. So far, only a few mitochondrial rRNA and tRNA modifying enzymes were characterised, all of them are encoded by nuclear DNA, synthesised by cytoplasmic ribosomes and then targeted to mitochondria (for reviews, (45,46). In all cases, the same nuclear gene encodes the protein acting in the nucleus, the cytoplasm and/or the mitochondria (19,22,47-49). In the case of tRNA-specific modification enzymes m²G₂₆-methyltransferase Trm1p and isopentenyl-transferase Mod5p, two distinct proteins are produced from the same mRNA transcript by appropriate choices between alternative initiation codons (for review, (45). It was demonstrated that the N-terminal sequence of Mod5p (amino acids 1 to 11) is

required but not sufficient for mitochondrial targeting (48). However, the amino acid composition of Mod5p N-terminal does not allow to predict its mitochondrial localisation.

The analysis of the YLR165c protein using the PSORT prediction software (37), which takes into account only the first 20 amino acids, did not revealed any plausible mitochondrial leader peptide. Another approach for determination of possible subcellular localisation is based on the global amino acid composition of proteins (NNPLS software, (38)). Using this prediction system, both Mod5p and the protein, encoded by YLR165c, can be classified as mitochondrial enzymes. Thus, despite the absence of pronounced N-terminal leader peptide, the YLR165c encoded protein is probably targeted into mitochondria.

The YLR165c encoded protein is likely dedicated only to mitochondrial rRNA modification since the residue at position 2945 in yeast cytoplasmic 26S rRNA (counterpart of position 2819 in the mitochondrial 21S rRNA) is not converted into a Ψ residue. In addition, even if a modification had to occur at this position, it would likely depend upon the Cbf5p/snoRNA-guided system. Hence, we do not expect to have two forms of enzymes produced from the YLR165c gene or two proteins with distinct localisation. This is in agreement with the fact that the second ATG codon in the YLR165c ORF is located very far from the initiator ATG and cannot be used for generation of an isoform.

Classification of the YLR165c ORF product as a mitochondrial protein by NNPLS software reinforces the idea that YLR165c ORF encoded protein is the mitochondrial rRNA: Ψ -synthase which specifically catalyses formation of the unique Ψ residue in mitochondrial LSU rRNA. For homogeneity of RNA: Ψ -synthase nomenclature in yeast, we propose to rename the YLR165c ORF to *PUS5* and the corresponding RNA: Ψ -synthase to Pus5p.

U to Ψ conversion at position 2819 of the *S. cerevisiae* mitochondrial 21S rRNA is not essential for growth

Several previous studies revealed a low content of post-transcriptional modifications in mitochondrial rRNAs. In *S. cerevisiae*, the mitochondrial SSU rRNA is apparently not

modified, but the presence of one Ψ residue has not been excluded (50). The mitochondrial *S. cerevisiae* LSU rRNA contains only three modified nucleosides: two 2'-O-methylated residues (Gm at positions 2270 and Um at position 2791) and one Ψ residue at position 2819 (51-53), reviewed in (44). The two 2'-O-methylation sites found in the *S. cerevisiae* mitochondrial 21S rRNA are strongly conserved in cytoplasmic rRNAs from *Eukarya* and *Bacteria* (54-56). In contrast, the unique pseudouridine residue present in yeast mitochondrial 21S rRNA is conserved only in mitochondrial LSU rRNAs from some other species (for example human and *M. musculus*) and in gram-negative bacteria such as *E. coli* (31,57). It is absent in the LSU rRNAs from the gram-positive bacteria *B. subtilis* (57), in 26S rRNA of eukaryotic cytoplasmic ribosomes (53,56) and also in the archaea *H. halobium* (57).

Our results demonstrate that U to Ψ conversion at position 2819 of the *S. cerevisiae* mitochondrial 21S rRNA is not essential since the disrupted Δ YLR165c yeast strain does not show any noticeable growth phenotype. Similarly, in *E. coli*, inactivation of RluC that forms residue Ψ_{2580} (the counterpart of the yeast mitochondrial Ψ_{2819} residue) and two other Ψ residues in 23S rRNA does not affect growth rate (11,12).

Nevertheless, because of its location in a 23S rRNA region considered to be located at the ribosome peptidyl transfer center, the residue Ψ_{2819} in the *S. cerevisiae* mitochondrial 21S rRNA may modulate the efficiency or the fidelity of the translation machinery. Indeed, the single-stranded sequences flanking helix 90, that contains residue Ψ_{2819} (helices are numbered according to nomenclature of (58) (Fig. 4), are the most highly conserved regions in the LSU rRNA (54,55,58). Several mutations in these single-stranded sequences confer resistance of bacterial or mitochondrial ribosomes to various antibiotics affecting the peptidyl transferase activity (for review, (59). In addition, these single-stranded regions are protected by the tRNA bound at the ribosomal P-site (60,61). Interestingly, helix 92 in the prolongation of helix 90 is terminated by the loop that was crosslinked with puromycin derivatives (54,62) and was recently shown to base-pair to the CCA terminal sequence of the tRNA located at the ribosomal A site (62). This terminal loop contains a conserved 2'-O-methylated residue (Um₂₇₉₁ in *S. cerevisiae* 21S mitochondrial rRNA).

In this study we demonstrated that one of the four putative yeast RNA:Ψ-synthases (Pus5p) of the RluA-family is implicated in the modification of yeast mitochondrial rRNA. The functional characterisation and the identification of RNA targets for the three other putative RNA:Ψ-synthases are in progress in our laboratory.

ACKNOWLEDGMENTS

This work was supported by laboratory funds from the "Ministère de la Recherche et de l'Enseignement Supérieur", the French "Centre National de la Recherche Scientifique" (Programme Physique et Chimie du Vivant 1997-99) and the "Association pour la Recherche sur le Cancer". I. Ansmant and S. Massenet are pre-doctoral fellows supported by a fellowship from the "Ministère de la Recherche et de l'Enseignement Supérieur". Y. Motorin is the recipient of an Associated researcher position at the CNRS ("Poste Rouge"). We thank M. Minet and F. Lacroute (CNRS, Gif-sur-Yvette, France) for providing us with plasmid pFL35 and the *BMA64* and *BMA64-1A* *S. cerevisiae* strains, Z. Szweykowska-Kulinska (Mickiewicz University, Poznan, Poland) for the plasmid containing the *A. thaliana* pre-tRNA^{Tyr} gene and R. Giegé (CNRS, Strasbourg, France) for the plasmid containing the yeast tRNA^{Asp} gene.

REFERENCES

1. Grosjean, H., Sprinzl, M. and Steinberg, S. (1995) *Biochimie*, **77**, 139-141.
2. Massenet, S., Mougin, A. and Branlant, C. (1998) In Grosjean, H., and Benne, R. (eds.), *The modification and editing of RNA*. ASM Press, Washington, DC, pp. 201-228.
3. Ofengand, J. and Fournier, M. J. (1998) In Grosjean, H., and Benne, R. (eds.), *The modification and editing of RNA*. ASM Press, Washington, DC, pp. 229-254.
4. Huang, L., Pookanjanatavip, M., Gu, X. and Santi, D. V. (1998) *Biochemistry*, **37**, 344-351.
5. Raychaudhuri, S., Conrad, J., Hall, B. G. and Ofengand, J. (1998) *RNA*, **4**, 1407-1417.
6. Conrad, J., Niu, L., Rudd, K., Lane, B. G. and Ofengand, J. (1999) *RNA*, **5**, 751-763.
7. Ramamurthy, V., Swann, S. L., Paulson, J. L., Spedalieri, C. J. and Mueller, E. G. (1999) *J. Biol. Chem.*, **274**, 22225-22230.
8. Arena, F., Ciliberto, G., Ciampi, S. and Cortese, R. (1978) *Nucleic Acids Res.*, **5**, 4523-4536.
9. Kammen, H. O., Marvel, C. C., Hardy, L. and Penhoet, E. E. (1988) *J. Biol. Chem.*, **263**, 2255-2263.
10. Nurse, K., Wrzesinski, J., Bakin, A., Lane, B. G. and Ofengand, J. (1995) *RNA*, **1**, 102-112.
11. Huang, L., Ku, J., Pookanjanatavip, M., Gu, X., Wang, D., Greene, P. J. and Santi, D. V. (1998) *Biochemistry*, **37**, 15951-15957.
12. Conrad, J., Sun, D., Englund, N. and Ofengand, J. (1998) *J. Biol. Chem.*, **273**, 18562-18566.
13. Wrzesinski, J., Bakin, A., Nurse, K., Lane, B. G. and Ofengand, J. (1995) *Biochemistry*, **34**, 8904-8913.
14. Niu, L. and Ofengand, J. (1999) *Biochemistry*, **38**, 629-635.
15. Wrzesinski, J., Nurse, K., Bakin, A., Lane, B. G. and Ofengand, J. (1995) *RNA*, **1**, 437-448.

16. Raychaudhuri, S., Niu, L., Conrad, J., Lane, B. G. and Ofengand, J. (1999) *J. Biol. Chem.*, **274**, 18880-18886.
17. Koonin, E. V. (1996) *Nucleic Acids Res.*, **24**, 2411-2415.
18. Simos, G., Tekotte, H., Grosjean, H., Segref, A., Sharma, K., Tollervey, D. and Hurt, E. C. (1996) *EMBO J.*, **15**, 2270-2284.
19. Becker, H. F., Motorin, Y., Planta, R. J. and Grosjean, H. (1997) *Nucleic Acids Res.*, **25**, 4493-4499.
20. Motorin, Y., Keith, G., Simon, C., Foiret, D., Simos, G., Hurt, E. and Grosjean, H. (1998) *RNA*, **4**, 856-869.
21. Lafontaine, D. L. J., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M. and Tollervey, D. (1998) *Genes Dev.*, **12**, 527-537.
22. Lecointe, F., Simos, G., Sauer, A., Hurt, E. C., Motorin, Y. and Grosjean, H. (1998) *J. Biol. Chem.*, **273**, 1316-1323.
23. Massenet, S., Motorin, Y., Lafontaine, D. L., Hurt, E. C., Grosjean, H. and Branlant, C. (1999) *Mol. Cell. Biol.*, **19**, 2142-2154.
24. Watkins, N. J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M. and Lührmann, R. (1998) *RNA*, **4**, 1549-1568.
25. Zebarjadian, Y., King, T., Fournier, M. J., Clarke, L. and Carbon, J. (1999) *Mol. Cell. Biol.*, **19**, 7461-7472.
26. Ni, J., Tien, A. L. and Fournier, M. J. (1997) *Cell*, **89**, 565-573.
27. Ganot, P., Bortolin, M. L. and Kiss, T. (1997) *Cell*, **89**, 799-809.
28. Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F. and Cullin, C. (1993) *Nucleic Acids Res.*, **21**, 3329-3330.
29. Wach, A., Brachat, A., Pohlmann, R. and Philippsen, P. (1994) *Yeast*, **10**, 1793-1808.
30. Gietz, D., St. Jean, A., Woods, R. A. and Schiestl, R. H. (1992) *Nucleic Acids Res.*, **20**, 1425.
31. Bakin, A. and Ofengand, J. (1993) *Biochemistry*, **32**, 9754-9762.
32. Jiang, H. Q., Motorin, Y., Jin, Y. X. and Grosjean, H. (1997) *Nucleic Acids Res.*, **25**, 2694-2701.

33. Mumberg, D., Muller, R. and Funk, M. (1995) *Gene*, **156**, 119-122.
34. Adams, A., Gottschling, D. E., Kaiser, C. A. and Stearns, T. (1997) *Methods in yeast genetics. A Cold Spring Harbor Laboratory Course Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
35. Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) *Nucleic Acids Res.*, **22**, 4673-4680.
36. Lawrence, C. E., Altschul, S. F., Boguski, M. S., Liu, J. S., Neuwald, A. F. and Wootton, J. C. (1993) *Science*, **262**, 208-214.
37. Nakai, K. and Kanehisa, M. (1992) *Genomics*, **14**, 897-911.
38. Reinhardt, A. and Hubbard, T. (1998) *Nucleic Acids Res.*, **26**, 2230-2236.
39. Sprinzl, S., Horn, C., Brown, M., Ioudovitch, A. and Steinberg, S. (1998) *Nucleic Acids Res.*, **26**, 148-153.
40. Gustafsson, C., Reid, R., Greene, P. J. and Santi, D. V. (1996) *Nucleic Acids Res.*, **24**, 3756-3762.
41. Reid, R., Greene, P. J. and Santi, D. V. (1999) *Nucleic Acids Res.*, **27**, 3138-3145.
42. Motorin, Y. and Grosjean, H. (1999) *RNA*, **5**, 1105-1118.
43. Bachellerie, J. P. and Cavallé, J. (1998) In Grosjean, H., and Benne, R. (eds.), *The modification and editing of RNA*. ASM Press, Washington, DC, pp. 255-272.
44. Mason, T. L. (1998) In Grosjean, H., and Benne, R. (eds.), *The modification and editing of RNA*. ASM Press, Washington, DC, pp. 273-280.
45. Martin, N. C. and Hopper, A. K. (1994) *Biochimie*, **76**, 1161-1167.
46. Maden, B. E. H. (1998) In Grosjean, H., and Benne, R. (eds.); *The modification and editing of RNA*. ASM Press, Washington, DC, pp. 421-440.
47. Hopper, A. K., Furukawa, A. H., Pham, H. D. and Martin, N. C. (1982) *Cell*, **28**, 543-550.
48. Boguta, M., Hunter, L. A., Shen, W. C., Gillman, E. C., Martin, N. C. and Hopper, A. K. (1994) *Mol. Cell. Biol.*, **14**, 2298-2306.
49. Rose, A. M., Belford, H. G., Shen, W. C., Greer, C. L., Hopper, A. K. and Martin, N. C. (1995) *Biochimie*, **77**, 45-53.

50. Klootwijk, J., Klein, I. and Grivell, L. A. (1975) *J. Mol. Biol.*, **97**, 337-350.
51. Sirum-Connolly, K. and Mason, T. L. (1993) *Science*, **262**, 1886-1889.
52. Sirum-Connolly, K., Peltier, J. M., Crain, P. F., McCloskey, J. A. and Mason, T. L. (1995) *Biochimie*, **77**, 30-39.
53. Bakin, A., Lane, B. G. and Ofengand, J. (1994) *Biochemistry*, **33**, 13475-13483.
54. Branlant, C., Krol, A., Machatt, M. A., Pouyet, J., Ebel, J. P., Edwards, K. and Kossel, H. (1981) *Nucleic Acids Res.*, **9**, 4303-4324.
55. Veldman, G. M., Klootwijk, J., de Regt, V. C., Planta, R. J., Branlant, C., Krol, A. and Ebel, J. P. (1981) *Nucleic Acids Res.*, **9**, 6935-6952.
56. Maden, B. E. (1990) *Prog. Nucleic Acid. Res. Mol. Biol*, **39**, 241-303.
57. Ofengand, J. and Bakin, A. (1997) *J. Mol. Biol.*, **266**, 246-268.
58. Leffers, H., Kjems, J., Ostergaard, L., Larsen, N. and Garrett, R. A. (1987) *J. Mol. Biol.*, **195**, 43-61.
59. Kirillov, S., Porse, B. T., Vester, B., Woolley, P. and Garrett, R. A. (1997) *FEBS Lett.*, **406**, 223-233.
60. Moazed, D. and Noller, H. F. (1989) *Cell*, **57**, 585-597.
61. Moazed, D. and Noller, H. F. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 3725-3728.
62. Kim, D. F. and Green, R. (1999) *Mol. Cell*, **4**, 859-864.

TABLES

Table I Pseudouridine formation in the extracts of wild-type and Δ YLR165c yeast strains

Position tested in tRNA	tRNA transcript/ labeling	Wild-type strain	Δ YLR165c strain
13	yeast tRNA ^{Asp} /ATP	0.75	0.65
32	yeast tRNA ^{Asp} /UTP	0.85	0.85
35	<i>A. thaliana</i> tRNA ^{Tyr} /ATP	0.55	0.65

Radiolabeled transcripts were incubated with yeast extract at 30°C for 60 min and the Ψ content was analysed by 2D TLC after RNase T2 digestion.

LEGENDS TO FIGURES

Fig. 1 UsnRNA pseudouridylation pattern in the Δ YLR165c *S. cerevisiae* strain. Primer extension analysis was performed for U1 (Panel A), U2 (Panel B) and U5 (Panel C) snRNAs from a total RNA fraction of the wild type (WT) and Δ YLR165c (*ylr165c*-) *S. cerevisiae* strains. CMCT modifications were done for 1, 10 and 20 min (lanes 2, 3 and 4, respectively), Experimental conditions are described in Materials and Methods. In lanes 3 and 4, the CMCT-modified RNA was subjected to an alkaline treatment at pH 10.4. A control extension experiment was made without addition of CMCT (lane 1). Lanes U, G, C and A correspond to the RNA sequencing ladder. Nucleotide positions, starting from the 5'-terminal nucleotide, are indicated on the right of the Panels. The reverse transcription stops, corresponding to Ψ residues in U1, U2 and U5 snRNAs, are indicated by arrows on the left of each Panel.

Fig. 2 Pseudouridine at position 2819 in yeast mitochondrial 21S rRNA disappears upon deletion of the YLR165c ORF. Panel A/ Primer extension analysis of the CMCT-modified *S. cerevisiae* mitochondrial 21S rRNA in a total RNA fraction from wild type (WT) or Δ YLR165c (*ylr165c*-) *S. cerevisiae* strains, for 2, 10 and 20 min (lanes 2, 3 and 4, respectively). The experimental conditions are the same as in Figure 1. Nucleotide positions, starting from the 5'-terminal nucleotide, are indicated on the right of the Panel. The reverse transcription stop, corresponding to residue Ψ_{2819} , is indicated by an arrow on the left of the Panel. Absence of this stop in the reverse transcription pattern of the 21S rRNA from the Δ YLR165c strain indicates the absence of residue Ψ_{2819} . Panel B/ Primer extension analysis of the CMCT-modified *S. cerevisiae* mitochondrial 21S rRNA in a total RNA fraction from the Δ YLR165c strain transformed with plasmid p413TEF-YLR165c.

Fig. 3 Sequence alignment of *E. coli* proteins from the RluA family (RluA, RluC and RluD) and *S. cerevisiae* YLR165c-encoded protein. The scale in amino acid residues is presented at the top. The three motifs (I, II and III) conserved in the RluA and RsuA

families of RNA:Ψ-synthases are shown at the bottom. Identical and conserved residues are shaded.

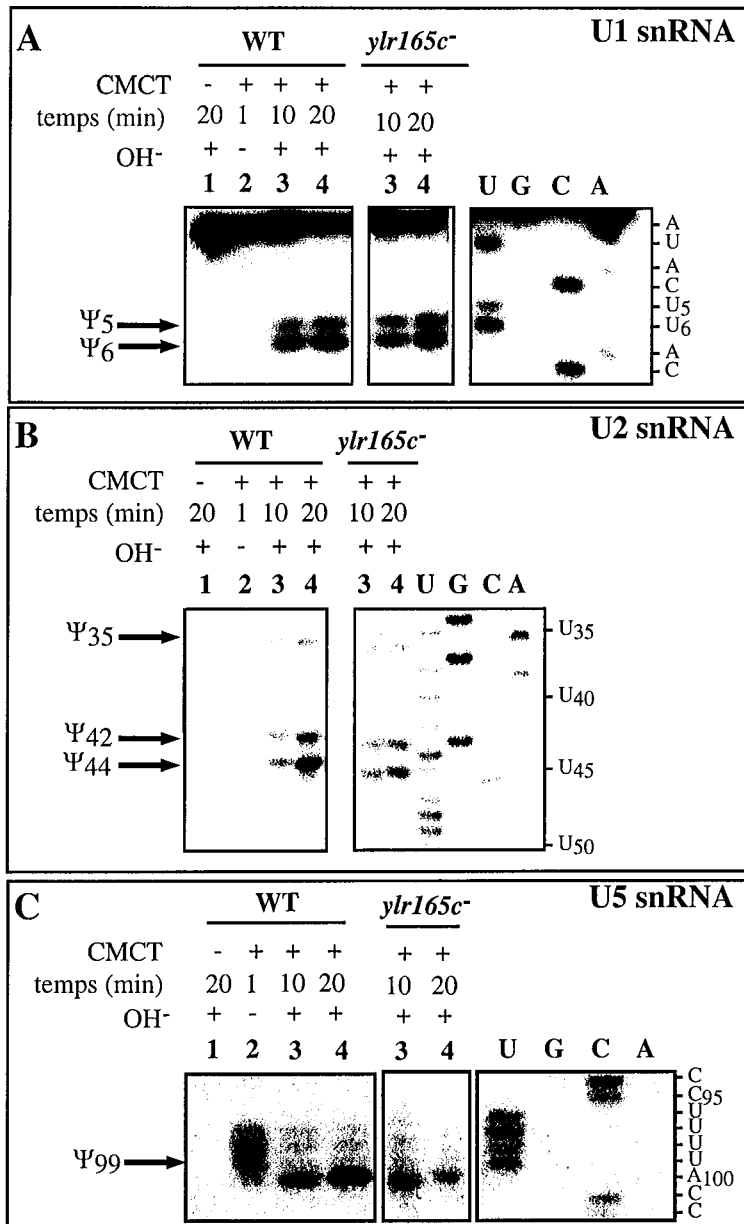


Fig. 1
Ansmant *et al.*

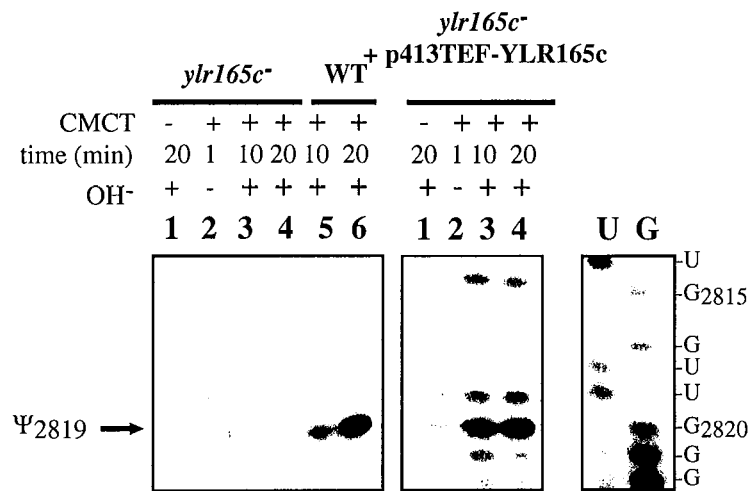


Fig. 2
Ansmant et al.

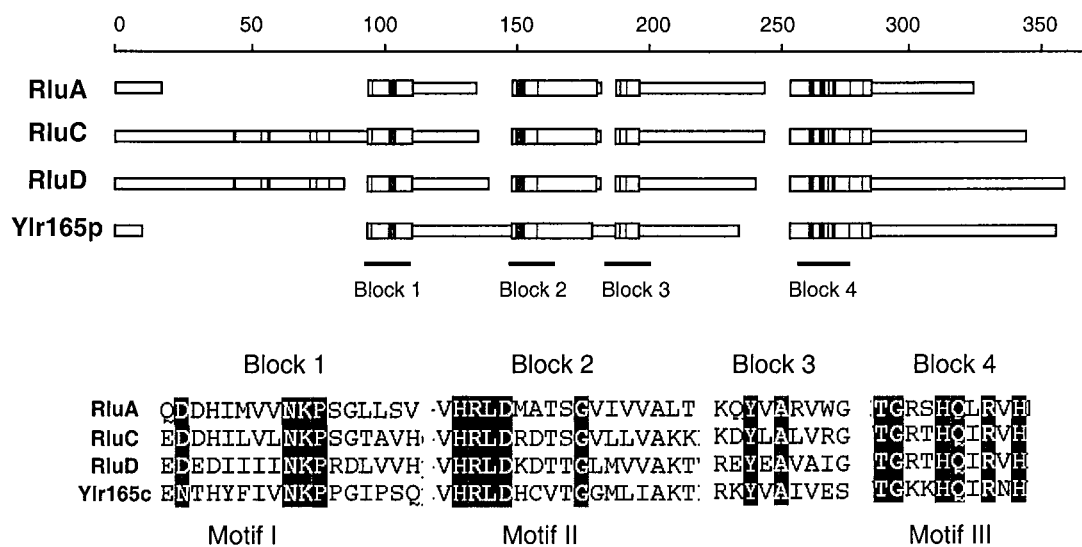


Fig. 3
Ansmant *et al*

B/ DISCUSSION ET PERSPECTIVES

1. ETUDES ENVISAGEES SUR LA Ψ -SYNTHASE Pus5p

Nous avons montré que la protéine codée par l'ORF YLR165C est nécessaire pour qu'il y ait formation du résidu Ψ_{2819} dans l'ARNr 21S mitochondrial. Etant donné que la protéine codée par cette ORF porte la signature caractéristique des Ψ -synthases, il est vraisemblable que cette protéine, que nous avons renommée Pus5p, catalyse elle-même la formation de ce résidu. Pour obtenir une vérification complète, ce gène *YLR165C* va être cloné dans un vecteur d'expression d'*E. coli*, permettant de produire, en grande quantité, l'enzyme portant une étiquette polyhistidine. Elle pourra alors être purifiée sur colonne d'agarose Ni-NTA. Nous pourrions ainsi tester son activité vis-à-vis de la position 2819 de l'ARNr mitochondrial *in vitro*, en utilisant un fragment de cet ARN produit *in vitro*. La génération de mutations permettra de rechercher les bases moléculaires de la reconnaissance de l'ARN par cette enzyme. Ce type d'expérience nous permettra de répondre définitivement à la question d'une implication éventuelle d'un ARN guide pour la formation du résidu Ψ_{2819} dans l'ARNr 21S mitochondrial. Les chances qu'un tel ARN soit impliqué sont très faibles, mais la preuve définitive sera obtenue par formation *in vitro* du résidu Ψ_{2819} par l'enzyme purifiée.

2. LA FORMATION DU RESIDU Ψ_{2819} N'EST PAS ESSENTIELLE, MAIS ELLE POURRAIT APPORTER UN AVANTAGE SELECTIF

Le résidu Ψ en position 2819 de l'ARNr 21S mitochondrial de *S. cerevisiae* est aussi présent dans de l'ARNr 23S d'*E. coli*. Or, l'ARNr 21S mitochondrial porte peu de modifications post-transcriptionnelles par rapport à l'ARNr 23S d'*E. coli* (Ψ_{2819} , Gm₂₂₇₀ et Um₂₇₉₁ contre 9 résidus Ψ , 3 résidus 2'-O-méthylés et 14 résidus méthylés sur la base dans l'ARN 23S d'*E. coli*) (Branlant *et al.*, 1981b; Bakin et Ofengand, 1993; Bakin *et al.*, 1994; Sirum-Connolly *et al.*, 1993, 1995; pour revue, Bachellerie et Cavallé, 1998). Le résidu Ψ_{2819} et les deux 2'-O-méthylations Gm₂₂₇₀ et Um₂₇₉₁ de l'ARNr 21S mitochondrial sont localisés dans trois parties du domaine V du grand ARNr, connues pour être situées dans la région catalytique de la grande sous-unité ribosomique (pour des revues récentes, Green et Noller, 1997; Noller, 1998).

Les mitochondries sont considérées comme ayant pour ancêtres des bactéries pourpres (pour revue, Gray, 1992), *E. coli* appartenant à ce groupe de bactéries. On peut donc supposer que soit les ARNr des bactéries ancestrales portaient moins de modifications post-transcriptionnelles, soit les mitochondries ont seulement retenu certaines de ces modifications, les plus importantes pour la fonction des ARNr. Les trois modifications observées dans l'ARNr 21S mitochondrial de

S. cerevisiae sont présentes dans l'ARNr 23S d'*E. coli*. Par contre, le résidu Ψ_{2819} de l'ARNr 21S mitochondrial n'a pas d'équivalent dans l'ARNr 23S des bactéries à gram + (*B. subtilis*, Ofengand et Bakin, 1997), dans l'ARN 23S de l'*Archaea H. halobium* (Ofengand et Bakin, 1997) et dans le grand ARNr cytoplasmique d'eucaryotes (Maden, 1990). Ceci explique sans doute pourquoi la génération d'un résidu Ψ en position 2819 dans l'ARNr 21S mitochondrial de *S. cerevisiae* (voir publication n°5) ou en position 2580 dans l'ARN 23S d'*E. coli* (Conrad *et al.*, 1998; Huang *et al.*, 1998a) n'est pas essentielle à la croissance. Néanmoins, la conservation de cette modification dans un système bactérien et dans un système mitochondrial, tous deux dérivés d'un ancêtre commun, suggère que la présence de ce résidu doit apporter un avantage sélectif.

3. COMMENT EXPLIQUER LA PRESENCE D'UN FAIBLE NOMBRE DE Ψ -SYNTHASES CARACTERISEES OU POTENTIELLES CHEZ *S. cerevisiae*?

Notre étude du contenu en résidus Ψ des UsnRNA dans les souches délétées spécifiquement dans l'une des quatre ORF YDL036C, YLR165C, RIB2 et YGR169C n'a pas permis d'identifier les enzymes responsables de la formation des résidus Ψ aux positions 5 et 6 du snRNA U1, 35 et 42 du snRNA U2 et 99 du snRNA U5.

De même, nous n'avons pas observé de différence, par rapport à la souche sauvage, des activités Ψ -synthases agissant aux positions 13, 32 et 35 des ARNt cytoplasmiques, après délétion des 4 ORF. Il avait été prouvé précédemment que la formation des résidus Ψ_{13} et Ψ_{32} n'était pas réalisée par Pus1p (Motorin *et al.*, 1998), Pus2p (Motorin *et al.*, résultats non publiés), Pus3p (Lecointe *et al.*, 1998) ou Pus4p (Becker *et al.*, 1997a). Pus1p pourrait participer à la formation du résidu Ψ_{35} *in vivo*, mais une deuxième activité Ψ -synthase, agissant à cette position, est également présente dans les extraits cellulaires et ne correspond pas aux enzymes codées par les ORF YDL036C, YLR165C, RIB2 et YGR169C, ni aux enzymes Pus2p, Pus3p et Pus4p. Il est à noter qu'il avait été montré précédemment que trois activités enzymatiques distinctes catalysent la formation des résidus Ψ aux positions 13, 32 et 35 des ARNt cytoplasmiques de *S. cerevisiae* (Samuelsson et Olsson, 1990).

Comment peut-on expliquer cette absence apparente d'enzymes pouvant catalyser la formation des résidus Ψ aux positions 13, 32 et 35 dans les ARNt cytoplasmiques et aux positions 5 et 6 dans le snRNA U1, 35 et 42 dans le snRNA U2 et 99 dans le snRNA U5?

⇒ Une première possibilité est que deux Ψ -synthases codées par les gènes déjà caractérisés soient capables de modifier, *in vivo*, le même site d'un même ARN. L'absence de l'une d'entre elles serait alors palliée par la présence de l'autre. Comme nous l'avons mentionné, cette hypothèse a déjà été proposée pour la formation du résidu Ψ_{35} dans des ARNt cytoplasmiques

(Motorin *et al.*, 1998). Il est tout à fait possible qu'il existe, *in vivo*, des Ψ -synthases présentant des activités redondantes. Il faut noter que ceci pourrait, en particulier, être observé pour des résidus Ψ dont la formation aurait un caractère essentiel pour la croissance cellulaire. La Ψ -synthase Ydl036c présente une forte identité de séquence avec une autre Ψ -synthase potentielle, Rib2 (62% d'identité sur 430 acides aminés comparés par la méthode de Smith-Waterman, disponible sur le site WEB <http://genome-www2.stanford.edu/Cgi-bin/SGD/SWA/alignment>). Ces deux enzymes pourraient donc avoir des propriétés communes de reconnaissance des substrats ARN. La Ψ -synthase potentielle, Ydl036c, possède un motif caractéristique d'adressage vers la mitochondrie. Les résultats préliminaires d'I. Ansmant indiquent que cette enzyme pourrait être responsable de la formation du résidu Ψ_{32} dans les ARNt mitochondriaux. On peut donc imaginer que la Ψ -synthase Rib2, qui ne possède pas ce signal d'adressage, modifie la même position dans les ARNt cytoplasmiques. Nous avons montré qu'un extrait cellulaire de la souche *rib2*- contient une activité Ψ -synthase formant *in vitro* des résidus Ψ en position 32 des ARNt cytoplasmiques. Une possibilité serait que l'enzyme mitochondriale Ydl036c (également présente dans l'extrait) soit responsable de l'activité détectée. La production de souches de *S. cerevisiae* délétées dans les deux ORF (RIB2 et YDL036C) permettra d'affirmer ou d'infirmier cette hypothèse.

D'une manière générale, la production de souches doublement délétées permettra d'avoir des informations sur la redondance des activités Ψ -synthases chez *S. cerevisiae*.

⇒ Une deuxième explication possible pour l'absence apparente de Ψ -synthases catalysant la formation des résidus Ψ aux positions 13, 32 et 35 dans les ARNt cytoplasmiques et aux positions 5 et 6 dans le snRNA U1, 35 et 42 dans le snRNA U2 et 99 dans le snRNA U5 chez *S. cerevisiae*, est l'existence d'une autre famille de Ψ -synthases, ne présentant pas d'homologie de séquence avec les Ψ -synthases déjà identifiées et ne pouvant donc pas être détectées par comparaison de séquence. L'existence de deux familles distinctes d'enzymes, responsables de la même réaction enzymatique dans une même cellule, a déjà été décrite dans plusieurs cas. Ainsi, les fructose 1,6 biphosphate aldolases de type I et II ne présentent aucune similarité de séquence entre elles. Elles procèdent par des mécanismes réactionnels différents, mais catalysent toutes deux l'hydrolyse réversible du fructose 1,6-diphosphate en glycéraldéhyde 3-phosphate et dihydroacétone-phosphate (pour revue, Fothergill-Gilmore et Michels, 1993). Ces deux classes d'enzymes coexistent chez les procaryotes et chez les eucaryotes inférieurs, tels les levures ou les algues. Chez les eucaryotes supérieurs, seules les enzymes de classe I sont présentes. Un deuxième exemple est le cas des ARNt aminoacyl synthétases. En effet, il existe deux classes différentes d'enzymes catalysant l'aminocyclation des ARNt. Ces deux classes d'enzymes diffèrent à la fois dans leurs séquences signatures et dans la structure tertiaire de leur site actif (pour revue, Meinnel *et al.*, 1995).

En ce qui concerne les Ψ -synthases, la situation chez les *Archaea* est encore plus extrême que chez *S. cerevisiae*. En effet, sur la base de la séquence signature, seules deux ORF pouvant coder des Ψ -synthases ont été détectées: l'une code une enzyme fortement homologue à l'enzyme Cbf5p de *S. cerevisiae* (famille TruB), l'autre code une enzyme pouvant être classée dans la famille TruA (Tableau 15) (Conrad *et al.*, 1999). Si la formation des résidus Ψ dans les ARNr des *Archaea* n'est pas médiée par des ARN guides, deux Ψ -synthases ne peuvent pas être suffisantes pour catalyser la formation des résidus Ψ , à la fois dans les ARNr et les ARNt et la présence de Ψ -synthases appartenant à une famille tout à fait différente est là aussi envisageable. Par contre, l'existence d'enzymes appartenant à une autre classe de Ψ -synthases chez *E. coli* est peu vraisemblable du fait que seules les Ψ -synthases responsables de la formation des résidus Ψ_{13} et Ψ_{65} dans les ARNt et du résidu Ψ_{2605} dans l'ARNr 23S ne sont pas connues et que trois gènes, codant des Ψ -synthases potentielles encore non caractérisées, ont été mis en évidence dans le génome d'*E. coli* (deux de la famille RsuA et un de la famille RluA) et ceci en plus des gènes de Ψ -synthases déjà caractérisés (Koonin, 1996; Conrad *et al.*, 1999).

La possibilité de la coexistence, dans une même cellule, d'une autre famille de Ψ -synthases, très différentes des enzymes appartenant aux familles TruA, TruB, RsuA et RluA, est donc envisageable, à la fois chez les *Archaea* et chez les eucaryotes. Comme, il ne sera pas possible de rechercher les gènes codant ces Ψ -synthases dans le génome de *S. cerevisiae* par homologie de séquence, la seule manière de procéder sera l'emploi d'une approche biochimique, calquée sur celle utilisée pour la caractérisation des premières Ψ -synthases chez *E. coli*. Cette approche, bien que longue, est pour l'instant la seule envisageable. Il faudra, en premier lieu, produire *in vitro* l'ARN, pour lequel nous voulons rechercher la Ψ -synthase responsable de sa modification, et vérifier que les modifications naturelles peuvent être obtenues avec un extrait cellulaire de *S. cerevisiae*. Ensuite, l'extrait devra être fractionné par différentes méthodes et, à chaque étape de fractionnement, la fraction contenant l'activité Ψ -synthase recherchée devra être identifiée. Ceci devrait permettre la purification de nouvelles Ψ -synthases présentes chez *S. cerevisiae*.

4. D'AUTRES SUBSTRATS ARN POSSIBLES POUR LES DIFFÉRENTES Ψ -SYNTHASES DÉJÀ CARACTÉRISÉES?

De la même manière que les Ψ -synthases responsables de la formation de certains résidus Ψ ne sont pas encore identifiées, les substrats ARN de certaines Ψ -synthases déjà caractérisées ou potentielles sont encore à rechercher. Ceci est d'autant plus vrai que certaines enzymes pourraient avoir, comme Pus1p, une multispécificité de substrats.

Certains de ces substrats pourraient être les ARNt cytoplasmiques ou mitochondriaux. En effet, les activités responsables de la formation des résidus Ψ , aux positions 1 et 31, dans les

A/

Positions des résidus Ψ dans les UsnRNA et les ARNr	Ψ-synthases responsables de la formation de ces résidus Ψ chez <i>S. cerevisiae</i>
snRNA U2 position 35	?
42	?
44	Pus1p
snRNA U5 99	?
snRNA U1 5	?
6	?
ARNr 21S mitochondrial 2819	Pus5p
ARNr cytoplasmique 44 résidus Ψ	Cbf5p

B/

Positions des résidus Ψ dans les ARNt	Ψ-synthases responsables de la formation de ces résidus Ψ dans les ARNt cytoplasmiques chez <i>S. cerevisiae</i>	Ψ-synthases responsables de la formation de ces résidus Ψ dans les ARNt mitochondriaux chez <i>S. cerevisiae</i>
1	?	/
13	?	/
26	Pus1p	/
27	Pus1p	Pus1p?
28	Pus1p	Pus1p?
31	?	?
32	Rib2?	YDLO36C
34	Pus1p	/
35	Pus1p?+?	/
36	Pus1p	/
38	Pus3p	Pus3p
39	Pus3p	Pus3p
55	Pus4p	Pus4p
65	Pus1p	/
67	Pus1p	/
72	/	?

ARNt cytoplasmiques n'ont pas été encore étudiées (Tableau 17). Il serait intéressant d'analyser la formation de ces résidus Ψ dans les différentes souches portant une délétion de l'un des 9 gènes de Ψ -synthases caractérisées ou potentielles. La formation, *in vivo*, du résidu Ψ_{31} , dans les différentes souches, peut être testée par la méthode basée sur l'emploi du CMCT. Toutefois, cette technique ne sera pas applicable pour la détection du résidu Ψ_1 . En effet, lors de l'analyse par la transcriptase inverse, les signaux correspondant à ce résidu, localisé à l'extrémité 5' des ARNt^{Arg} et ARNt^{Lys}, ne seront pas distinguables de l'arrêt d'élongation de la transcriptase inverse. Une possibilité est de réaliser le test *in vitro*, en utilisant un ARN, produit par transcription *in vitro*, correspondant à l'ARNt précurseur contenant une extension en 5'. Des extraits des souches portant les délétions des différents gènes de Ψ -synthases pourront être utilisés. Lorsque les Ψ -synthases recombinantes sont déjà purifiées, les expériences *in vitro* pourront être réalisées directement avec ces enzymes. La présence du résidu Ψ_1 dans les transcrits pourra être également testée sur l'ARNt mature, en utilisant la méthode dite "du plus proche voisin", dont le principe a été explicité précédemment.

Il est possible que les résidus Ψ_{31} et Ψ_{32} , dans les ARNt cytoplasmiques, soient synthétisés par une même Ψ -synthase, de par leur localisation contiguë. En effet, une seule Ψ -synthase (Pus3p) est responsable de la modification des résidus aux positions 38 et 39 dans les ARNt de levure. De même, l'enzyme Pus1p modifie *in vivo* les positions 26, 27 et 28 (Tableau 17). Cependant, comme nous l'avons montré, les résidus Ψ_{42} et Ψ_{44} dans le snRNA U2 de *S. cerevisiae*, bien que proches l'un de l'autre, ne sont pas modifiés par la même Ψ -synthase.

En plus des Ψ -synthases responsables de la formation des résidus Ψ aux positions 1 et 31 des ARNt cytoplasmiques, il reste encore beaucoup à faire pour identifier les différentes Ψ -synthases de *S. cerevisiae*, responsables de la formation des résidus Ψ portés par les ARNt mitochondriaux. En effet, seule l'implication respective des Ψ -synthases Pus3p et Pus4p, dans la formation des résidus $\Psi_{38/39}$ et Ψ_{55} , a été démontrée (Becker *et al.*, 1997a; Lecointe *et al.*, 1998). Ces ARNt- Ψ -synthases modifient les mêmes positions dans les ARNt cytoplasmiques et mitochondriaux (Tableau 17) (Becker *et al.*, 1997a; Lecointe *et al.*, 1998a). C'est également le cas pour trois autres enzymes de modification des ARNt: Trm1p, Mod5p et Trm2p qui sont responsables, respectivement, de la formation des résidus m^2G_{26} , i^6A_{37} et m^5U_{54} (pour revue, Maden, 1998). Sur cette base, nous pouvons également penser que la formation des résidus Ψ aux positions 27 et 28 dans les ARNt mitochondriaux soit catalysée par Pus1p. Il est également possible que la formation du résidu Ψ_{31} soit catalysée, à la fois dans le cytoplasme et dans la mitochondrie, par une même Ψ -synthase. La formation des résidus Ψ aux positions 27, 28, 31 et 72 des ARNt mitochondriaux devra donc être testée dans les différentes souches portant des délétions des Ψ -synthases démontrées ou potentielles. La présence du résidu Ψ_{72} ne pourra pas être testée par modification par le CMCT, du fait de la nécessité d'hybrider l'oligonucléotide

servant d'amorce pour la transcriptase inverse à l'extrémité 3' de la molécule. La recherche de la Ψ -synthase responsable de la formation de ce résidu pourra être réalisée *in vitro* sur un ARNt précurseur ayant une extension en 3' ou par la méthode du plus proche voisin, de la même manière que pour le résidu Ψ_1 présent dans deux ARNt cytoplasmiques.

Afin d'identifier le (ou les) substrat(s) ARN de chaque Ψ -synthase présente chez *S. cerevisiae*, il sera important de ne pas se limiter à l'étude des ARNt. En effet, d'autres ARN sont présents dans la cellule et peuvent contenir des résidus Ψ . C'est le cas:

⇒ des snoRNA: peu de données sont disponibles sur le degré de modification post-transcriptionnelle de ces ARN. Chez l'homme, les snoRNA U3 et U8 portent des résidus Ψ (Kato et Harada, 1984; Reddy *et al.*, 1985). De telles modifications pourraient également être présentes dans les ARN snR4 et snR8 de *S. cerevisiae* (Wise *et al.*, 1983). En ce qui concerne le snoRNA U3 de *S. cerevisiae*, nous n'avons pas détecté de résidu Ψ dans l'ensemble de la molécule.

⇒ La RNase P est un ribozyme universellement conservé. Elle intervient dans la maturation de l'extrémité 5' des ARNt. La RNase MRP est également un ribozyme. Elle est exclusivement présente chez les eucaryotes et est nécessaire à la maturation des pré-ARNr (pour revue, Morrissey et Tollervey, 1995). A notre connaissance, la présence de modifications post-transcriptionnelles n'a pas encore été recherchée dans les ARN portant ces activités catalytiques.

⇒ l'ARN associé à la télomérase (ARN nommé TER pour TELomerase RNA), sert de matrice à la synthèse de l'ADN télomérique par la télomérase (pour revue, O'Reilly *et al.*, 1999). La recherche des modifications post-transcriptionnelles portées par cet ARN a été débutée au laboratoire. Cet ARN contracte des appariements de bases avec l'ADN télomérique en cours d'élongation et est complexé avec de nombreuses protéines dont la télomérase elle-même (pour revue, O'Reilly *et al.*, 1999). Il ne serait donc pas étonnant que cet ARN porte un ou plusieurs résidus Ψ .

⇒ Enfin, un autre ARN, susceptible de porter des modifications post-transcriptionnelles, est l'ARN 7S faisant partie de la particule SRP impliquée dans la sécrétion des protéines (pour revues, Lutcke, 1995; Bernstein, 1998). Il a été montré récemment, chez les mammifères, que cet ARN transite dans le nucléole après synthèse (Jacobson et Pederson, 1999). Il pourrait donc être modifié par le système basé sur les snoRNA H/ACA et nous envisageons de tester cette hypothèse.

Conclusion

Mon travail de thèse a permis d'apporter des informations sur la localisation des résidus Ψ dans différents ARN issus d'*Archaea*, de *S. cerevisiae* et de cellules HeLa, ainsi que sur la synthèse de ces résidus:

⇒ La localisation des résidus Ψ , dans les UsnRNA des spliceosomes U12-dépendants de cellules HeLa et dans les spliceosomes de *S. cerevisiae*, nous a permis de montrer que certains résidus Ψ , portés par ces UsnRNA, ont été très conservés au cours de l'évolution et sont retrouvés dans les deux systèmes spliceosomaux. Cette observation suggère fortement une importance fonctionnelle de ces résidus modifiés pour l'assemblage des spliceosomes et/ou la réaction d'épissage.

⇒ Nous avons montré que la Ψ -synthase Pus1p de *S. cerevisiae* présente une double spécificité de substrats *in vivo*. A ce jour, cette enzyme est la seule Ψ -synthase identifiée ayant un UsnRNA comme substrat.

⇒ La délétion systématique des gènes codant des Ψ -synthases putatives chez *S. cerevisiae*, nous a permis de caractériser une nouvelle enzyme, Pus5p, très probablement responsable de la formation du seul résidu Ψ présent dans les ARNr 21S mitochondriaux de levure.

⇒ Enfin, la localisation des résidus Ψ dans l'ARNr 23S de l'*Archaea* hyperthermophile *S. acidocaldarius* montre que, contrairement à la situation observée pour les résidus 2'-O-méthylés, les ARN ribosomiques de ces organismes présentent, comme ceux des bactéries, un faible taux de pseudouridylatation.

Ces résultats ouvrent de larges perspectives, décrites précédemment, tant dans l'étude du rôle des modifications post-transcriptionnelles portées par les UsnRNA, que dans la recherche et l'étude des Ψ -synthases chez *S. cerevisiae*.

Les nucléotides modifiés: d'hier à aujourd'hui

Selon l'hypothèse d'un monde vivant ancien basé sur l'ARN, les molécules d'ARN auraient servi à la fois de génome et de molécules catalytiques. Selon l'observation d'une distribution universelle de certains nucléotides modifiés dans les ARN des *Archaea*, bactéries et eucaryotes, et la structure chimique simple de ces nucléotides modifiés conservés, ceux-ci auraient pu être présents, en plus des nucléotides conventionnels, dans ce qui aurait été les premières molécules du monde vivant (Cermakian et Cedergren, 1998). C'est par exemple le cas des méthylation sur le ribose et les bases, des résidus inosines ou encore des résidus Ψ . L'existence de tels nucléotides pourrait donc être très ancienne. Ils auraient pu, dans un premier temps, être incorporés dans les ARN de manière aléatoire, leur incorporation ayant ensuite pu être guidée *in cis*. Dans l'hypothèse d'un "monde des ARN", la présence de nucléotides

modifiés aurait permis l'amélioration des fonctions des ribozymes initiaux, ceci en modulant les interactions avec d'autres ARN et/ou en augmentant l'activité catalytique (Cermakian et Cedergren, 1998). Il a même été proposé que les nucléotides modifiés aient été à l'origine du monde des protéines, cette hypothèse étant basée sur le fait que les acides aminés auraient pu être originellement liés au C₅ des uraciles (Robertson et Miller, 1995).

Sur la base des données disponibles sur les ARN actuels, on observe une augmentation dans l'échelle de l'évolution du nombre de nucléotides modifiés dans les ARNt et les ARNr (pour revues, Grosjean *et al.*, 1995; Ofengand et Fournier, 1998). Nos résultats sur les résidus Ψ portés par les UsnRNA vont dans le même sens. Chez les vertébrés, le nombre de nucléotides modifiés portés par les ARN métaboliquement stables est tellement important que l'appellation de "nucléotides mineurs" pour désigner ces nucléotides est complètement inappropriée. A quoi est liée cette augmentation du nombre de résidus modifiés dans l'échelle de l'évolution: nécessité de multiplier les fonctionnalités des molécules d'ARN?

Seuls certains des nucléotides modifiés phylogénétiquement conservés pourraient avoir un rôle important dans l'activité des ARN. En effet, on remarque que la plupart du temps, l'absence de modifications post-transcriptionnelles à un site donné d'un ARN donné n'a pas d'effet clairement marqué sur l'activité de cet ARN (voir Introduction). Ceci est même vrai pour certains nucléotides modifiés phylogénétiquement conservés, comme le résidu Ψ_{55} présent dans 95% des ARNt des trois règnes du vivant. Il serait néanmoins hâtif de conclure que les modifications post-transcriptionnelles des ARN n'ont pas de rôle dans l'activité de ces molécules. La conservation phylogénétique de certains nucléotides modifiés reflète une importance biologique pour les cellules, sans doute un avantage sélectif dans la sélection naturelle, par le renforcement de la sélectivité de reconnaissance ou l'augmentation de la stabilité de la structure 3D. Ainsi, nous avons vu que les bactéries et *Archaea* thermophiles et hyperthermophiles avaient les capacités d'adapter le nombre des modifications post-transcriptionnelles, portées par leur ARNt et ARNr, à l'augmentation de la température du milieu de croissance.

Un autre point intéressant à noter est le fait que les cellules ont développé des systèmes performants qui permettent de modifier leurs ARN à de nombreux sites et ceci en utilisant un nombre minimum d'enzymes de modification. Une première solution retenue est l'existence d'enzymes de modification à spécificité de substrats multiples. La deuxième solution repose sur l'existence d'ARN guides définissant les sites d'action d'une enzyme. L'existence possible de snoRNA guides de méthylation chez les *Archaea* suggère que le système d'ARN guides, présents chez les eucaryotes, pourrait être l'héritage d'un système plus ancien présent chez les *Archaea*, qui aurait pu à l'origine être mis en place en vue d'une adaptation de la croissance cellulaire à des températures élevées.

Bibliographie

A

- Abel K, Jurnak F. 1996. A complex profile of protein elongation: translating chemical energy into molecular movement. Structure, 4:229-238.
- 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., 17:3738-3746.
- Abovich N, Legrain P, Rosbash M. 1990. The yeast *PRP6* gene encodes a U4/U6 small nuclear ribonucleoprotein particle (snRNP) protein, and the *PRP9* gene encodes a protein required for U2 snRNP binding. Mol. Cell. Biol., 10:6417-6425.
- Abovich N, Liao XC, Rosbash M. 1994. The yeast MUD2 protein: an interaction with PRP11 defines a bridge between commitment complexes and U2 snRNP addition. Genes Dev., 8:843-854.
- Abovich N, Rosbash M. 1997. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. Cell, 89:403-412.
- Achsel T, Ahrens K, Brahm H, Teigelkamp S, Lührmann R. 1998. The human U5-220kD protein (hPrp8) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase, a homologue of ribosomal elongation factor EF-2, and a novel WD-40 protein. Mol. Cell. Biol., 18:6756-6766.
- Adema GJ, van Hulst KL, Baas PD. 1990. Uridine branch acceptor is a cis-acting element involved in regulation of the alternative processing of calcitonin/CGRP-1 pre-mRNA. Nucleic Acids Res., 18:5365-5373.
- Agris PF, Koh H, Soll D. 1973. The effect of growth temperatures on the *in vivo* ribose methylation of *Bacillus stearothermophilus* transfer RNA. Arch. Biochem. Biophys., 154:277-282.
- Agris PF. 1996. The importance of being modified: roles of modified nucleosides and Mg²⁺ in RNA structure and function. Prog. Nucleic Acid. Res. Mol. Biol., 53:79-129.
- Akama K, Nass A, Junker V, Beier H. 1997. Characterization of nuclear tRNA(Tyr) introns: their evolution from red algae to higher plants. FEBS Lett., 417:213-218.
- Alvarez CJ, Romfo CM, Vanhoy RW, Porter GL, Wise JA. 1996. Mutational analysis of U1 function in *Schizosaccharomyces pombe*: pre-mRNAs differ in the extent and nature of their requirements for this snRNA *in vivo*. RNA, 2:404-418.
- Amiri KA. 1994. Fibrillarin-like proteins occur in the domain *Archaea*. J. Bacteriol., 176:2124-2127.
- Anderson GJ, Bach M, Lührmann R, Beggs JD. 1989. Conservation between yeast and man of a protein associated with U5 small nuclear ribonucleoprotein. Nature, 342:819-821.
- Ansari A, Schwer B. 1995. SLU7 and a novel activity, SSF1, act during the PRP16-dependent step of yeast pre-mRNA splicing. EMBO J., 14:4001-4009.

- 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, 3:1143-1152.
- Arena F, Ciliberto G, Ciampi S, Cortese R. 1978. Purification of pseudouridylate synthetase I from *Salmonella typhimurium*. Nucleic Acids Res., 5:4523-4536.
- Arenas JE, Abelson JN. 1993. The *Saccharomyces cerevisiae* PRP21 gene product is an integral component of the prespliceosome. Proc. Natl. Acad. Sci. USA, 90:6771-6775.
- Ares M, Jr. 1986. U2 RNA from yeast is unexpectedly large and contains homology to vertebrate U4, U5, and U6 small nuclear RNAs. Cell, 47:49-59.
- Ares M, Jr., Igel AH. 1989. Phylogenetic comparison of U2 small nuclear RNA sequences suggests a pseudoknotted structure. UCLA symp. Mol. Cell. Biol., 94:13-23.
- Ares M, Jr., Igel AH. 1990. Lethal and temperature-sensitive mutations and their suppressors identify an essential structural element in U2 small nuclear RNA. Genes Dev., 4:2132-2145.
- Arluison V, Hountondji C, Robert B, Grosjean H. 1998. Transfer RNA-pseudouridine synthetase Pus1 of *Saccharomyces cerevisiae* contains one atom of zinc essential for its native conformation and tRNA recognition. Biochemistry, 37:7268-7276.
- Arluison V, Buckle M, Grosjean H. 1999a. Pseudouridine synthetase Pus1 of *Saccharomyces cerevisiae*: kinetic characterisation, tRNA structural requirement and real-time analysis of its complex with tRNA. J. Mol. Biol., 289:491-502.
- Arluison V, Batelier G, Riès-Kautt M, Grosjean H. 1999b. RNA:pseudouridine synthase Pus1 from *Saccharomyces cerevisiae*: Oligomerization property and stoichiometry of the complex with yeast tRNA^{Phe}. Biochimie, 81:1-6.
- Arnez JG, Steitz TA. 1994. Crystal structure of unmodified tRNA(Gln) complexed with glutamyl- tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. Biochemistry, 33:7560-7567.
- Arning S, Gruter P, Bilbe G, Krämer A. 1996. Mammalian splicing factor SF1 is encoded by variant cDNAs and binds to RNA. RNA, 2:794-810.
- Ast G, Weiner AM. 1996. A U1/U4/U5 snRNP complex induced by a 2'-O-methyl-oligoribonucleotide complementary to U5 snRNA. Science, 272:881-884.
- Ast G, Weiner AM. 1997a. Antisense oligonucleotide binding to U5 snRNP induces a conformational change that exposes the conserved loop of U5 snRNA. Nucleic Acids Res., 25:3508-3513.
- Ast G, Weiner AM. 1997b. A novel U1/U5 interaction indicates proximity between U1 and U5 snRNAs during an early step of mRNA splicing. RNA, 3:371-381.
- Auffinger P, Westhof E. 1996. H-bond stability in the tRNA(Asp) anticodon hairpin: 3 ns of multiple molecular dynamics simulations. Biophys J., 71:940-954.

- Auffinger P, Westhof E. 1997. RNA hydration: three nanoseconds of multiple molecular dynamics simulations of the solvated tRNA(Asp) anticodon hairpin. J. Mol. Biol., 269:326-341.
- Auffinger P, Westhof E. 1998. Effects on pseudouridylation on tRNA hydration and dynamics: a theoretical approach. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 103-12
- Augustin S, Muller MW, Schweyen RJ. 1990. Reverse self-splicing of group II intron RNAs *in vitro*. Nature, 343:383-386.
- Auxilien S, Crain PF, Trewyn RW, Grosjean H. 1996. Mechanism, specificity and general properties of the yeast enzyme catalysing the formation of inosine 34 in the anticodon of transfer RNA. J. Mol. Biol., 262:437-458.
- Ayadi L, Miller M, Banroques J. 1997. Mutations within the yeast U4/U6 snRNP protein Prp4 affect a late stage of spliceosome assembly. RNA, 3:197-209.
- Ayadi L, Callebaut I, Saguez C, Villa T, Mornon JP, Banroques J. 1998. Functional and structural characterization of the Prp3 binding domain of the yeast Prp4 splicing factor. J. Mol. Biol., 284:673-687.

B

- ach M, Winkelmann G, Lührmann R. 1989. 20S small nuclear ribonucleoprotein U5 shows a surprisingly complex protein composition. Proc. Natl. Acad. Sci. USA, 86:6038-6042.
- Bachellerie JP, Cavaillé J. 1997. Guiding ribose methylation of rRNA. Trends Biochem. Sci., 22:257-261.
- Bachellerie JP, Cavaillé J. 1998. Small nucleolar RNAs guide the ribose methylations of eukaryotic rRNAs. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 255-272.
- Bagni C, Lapeyre B. 1998. Gar1p binds to the small nucleolar RNAs snR10 and snR30 *in vitro* through a nontypical RNA binding element. J. Biol. Chem., 273:10868-10873.
- Bakin A, Ofengand J. 1993. Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique. Biochemistry, 32:9754-9762.
- Bakin A, Kowalak J, McCloskey JA, Ofengand J. 1994. The single pseudouridine residue in *Escherichia coli* 16S ribosomal RNA is located at position 516. Nucleic Acids Res., 22:3681-3684.
- 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, 86:823-834.
- Banroques J, Abelson JN. 1989. Prp4: a protein of the yeast U4/U6 small nuclear ribonucleoprotein particle. Mol. Cell. Biol., 9:3710-3719.

- Barabino SM, Sproat BS, Lamond AI. 1992. Antisense probes targeted to an internal domain in U2 snRNP specifically inhibit the second step of pre-mRNA splicing. Nucleic Acids Res., 20:4457-4464.
- Bare LA, Uhlenbeck OC. 1986. Specific substitution into the anticodon loop of yeast tyrosine transfer RNA. Biochemistry, 25:5825-5830.
- Baudin A, Ozier-Kalogeropoulos O, Denouel A, Lacroute F, Cullin C. 1993. A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. Nucleic Acids Res., 21:3329-3330.
- Becker HF, Motorin Y, Planta RJ, Grosjean H. 1997a. The yeast gene *YNL292w* encodes a pseudouridine synthase (Pus4) catalyzing the formation of psi55 in both mitochondrial and cytoplasmic tRNAs. Nucleic Acids Res., 25:4493-4499.
- Becker HF, Motorin Y, Sissler M, Florentz C, Grosjean H. 1997b. Major identity determinants for enzymatic formation of ribothymidine and pseudouridine in the T psi-loop of yeast tRNAs. J. Mol. Biol., 274:505-518.
- Becker HF, Motorin Y, Florentz C, Giegé R, Grosjean H. 1998. Pseudouridine and ribothymidine formation in the tRNA-like domain of turnip yellow mosaic virus RNA. Nucleic Acids Res., 26:3991-3997.
- Behrens SE, Lührmann R. 1991. Immunoaffinity purification of a [U4/U6.U5] tri-snRNP from human cells. Genes Dev., 5:1439-1452.
- Behrens SE, Tyc K, Kastner B, Reichelt J, Lührmann R. 1993a. Small nuclear ribonucleoprotein (RNP) U2 contains numerous additional proteins and has a bipartite RNP structure under splicing conditions. Mol. Cell. Biol., 13:307-319.
- Behrens SE, Galisson F, Legrain P, Lührmann R. 1993b. Evidence that the 60-kDa protein of 17S U2 small nuclear ribonucleoprotein is immunologically and functionally related to the yeast PRP9 splicing factor and is required for the efficient formation of prespliceosomes. Proc. Natl. Acad. Sci. USA, 90:8229-8233.
- Beltrame M, Tollervey D. 1992. Identification and functional analysis of two U3 binding sites on yeast pre-ribosomal RNA. EMBO J., 11:1531-1542.
- Beltrame M, Tollervey D. 1995. Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis. EMBO J., 14:4350-4356.
- Benne R. 1992. RNA editing in trypanosomes. The use of guide RNAs. Mol. Biol. Rep., 16:217-227.
- Bennett M, Michaud S, Kingston J, Reed R. 1992. Protein components specifically associated with prespliceosome and spliceosome complexes. Genes Dev., 6:1986-2000.
- Bentley D. 1999. Coupling RNA polymerase II transcription with pre-mRNA processing. Curr. Opin. Cell. Biol., 11:347-351.
- Berget SM. 1995. Exon recognition in vertebrate splicing. J. Biol. Chem., 270:2411-2414.
- Berglund JA, Chua K, Abovich N, Reed R, Rosbash M. 1997. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. Cell, 89:781-787.

- Bindereif A, Wolff T, Green MR. 1990. Discrete domains of human U6 snRNA required for the assembly of U4/U6 snRNP and splicing complexes. EMBO J., 9:251-255.
- Björk GR. 1995. Biosynthesis and function of modified nucleosides. In: Söll D, RajBhandary UL, eds. tRNA: Structure, biosynthesis and Function. Washington, DC: ASM press. pp 165-205.
- 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., 9:3698-3709.
- Black DL, Chabot B, Steitz JA. 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in pre-messenger RNA splicing. Cell, 42:737-750.
- Blanton S, Srinivasan A, Rymond BC. 1992. PRP38 encodes a yeast protein required for pre-mRNA splicing and maintenance of stable U6 small nuclear RNA levels. Mol. Cell. Biol., 12:3939-3947.
- Boeck R, Lapeyre B, Brown CE, Sachs AB. 1998. Capped mRNA degradation intermediates accumulate in the yeast *spb8-2* mutant. Mol. Cell. Biol., 18:5062-5072.
- Boelens W, Scherly D, Beijer RP, Jansen EJ, Dathan NA, Mattaj IW, van Venrooij WJ. 1991. A weak interaction between the U2A' protein and U2 snRNA helps to stabilize their complex with the U2B" protein. Nucleic Acids Res., 19:455-460.
- Bordonné R, Banroques J, Abelson J, Guthrie C. 1990. Domains of yeast U4 spliceosomal RNA required for PRP4 protein binding, snRNP-snRNP interactions, and pre-mRNA splicing *in vivo*. Genes Dev., 4:1185-1196.
- Bordonné R, Tarassov I. 1996. The yeast *SME1* gene encodes the homologue of the human E core protein. Gene, 176:111-117.
- 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., 18:457-469.
- Bos JL, Heyting C, Borst P, Arnberg AC, Van Bruggen EF. 1978. An insert in the single gene for the large ribosomal RNA in yeast mitochondrial DNA. Nature, 275:336-338.
- Bossi L, Ruth JR. 1980. The influence of codon context on genetic code translation. Nature, 286:123-127.
- 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., 16:4770-4776.
- Brand RC, Klootwijk J, Sibum CP, Planta RJ. 1979. Pseudouridylation of yeast ribosomal precursor RNA. Nucleic Acids Res., 7:121-134.
- Branlant C, Krol A, Ebel JP, Lazar E, Gallinaro H, Jacob M, Sri-Widada J, Jeanteur P. 1980. Nucleotide sequences of nuclear U1A RNAs from chicken, rat and man. Nucleic Acids Res., 8:4143-4154.
- Branlant C, Krol A, Ebel JP, Gallinaro H, Lazar E, Jacob M. 1981a. The conformation of chicken, rat and human U1A RNAs in solution. Nucleic Acids Res., 9:841-858.

- Branlant C, Krol A, Machatt MA, Pouyet J, Ebel JP, Edwards K, Kossel H. 1981b. 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., 9:4303-4324.
- 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., 1:1259-1265.
- Branlant C, Krol A, Lazar E, Haendler B, Jacob M, Galego-Dias L, Pousada C. 1983. High evolutionary conservation of the secondary structure and of certain nucleotide sequences of U5 RNA. Nucleic Acids Res., 11:8359-8367.
- Bringmann P, Appel B, Rinke J, Reuter R, Theissen H, Lührmann R. 1984. Evidence for the existence of snRNAs U4 and U6 in a single ribonucleoprotein complex and for their association by intermolecular base pairing. EMBO J., 3:1357-1363.
- Brody E, Abelson J. 1985. The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. Science, 228:963-967.
- Brosi R, Hauri HP, Krämer A. 1993a. Separation of splicing factor SF3 into two components and purification of SF3a activity. J. Biol. Chem., 268:17640-17646.
- Brosi R, Groning K, Behrens SE, Lührmann R, Krämer A. 1993b. Interaction of mammalian splicing factor SF3a with U2 snRNP and relation of its 60-kD subunit to yeast PRP9. Science, 262:102-105.
- Brow DA, Guthrie C. 1988. Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. Nature, 334:213-218.
- Brow DA, Vidaver RM. 1995. An element in human U6 RNA destabilizes the U4/U6 spliceosomal RNA complex. RNA, 1:122-131.
- Brown JD, Beggs JD. 1992. Roles of PRP8 protein in the assembly of splicing complexes. EMBO J., 11:3721-3729.
- Brown JR, Doolittle WF. 1997. *Archaea* and the prokaryote-to-eukaryote transition. Microbiol. Mol. Biol. Rev., 61:456-502.
- Brulé F, Venema J, Ségault 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, 2:183-197.
- Brulé H, Holmes WM, Keith G, Giegé R, Florentz C. 1998. Effect of a mutation in the anticodon of human mitochondrial tRNA^{Pro} on its post-transcriptional modification pattern. Nucleic Acids Res., 26:537-543.
- Brys A, Schwer B. 1996. Requirement for SLU7 in yeast pre-mRNA splicing is dictated by the distance between the branchpoint and the 3' splice site. RNA, 2:707-717.

- Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NSM, Venter JC. 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science, **273**:1058-1073.
- Burge CB, Padgett RA, Sharp PA. 1998. Evolutionary fates and origins of U12-type introns. Mol. Cell., **2**:773-785.
- Burgess SM, Guthrie C. 1993. A mechanism to enhance mRNA splicing fidelity: the RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates. Cell, **73**:1377-1391.
- Burke JM, Belfort M, Cech TR, Davies RW, Schweyen RJ, Shub DA, Szostak JW, Tabak HF. 1987. Structural conventions for group I introns. Nucleic Acids Res., **15**:7217-7221.

Caffarelli E, Losito M, Giorgi C, Fatica A, Bozzoni I. 1998. *In vivo* identification of nuclear factors

interacting with the conserved elements of box C/D small nucleolar RNAs. Mol. Cell. Biol., **18**:1023-1028.

Camasses A, Bragado-Nilsson E, Martin R, Séraphin B, Bordonné R. 1998. Interactions within the yeast Sm core complex: from proteins to amino acids. Mol. Cell. Biol., **18**:1956-1966.

Carbone ML, Solinas M, Sora S, Panzeri L. 1991. A gene tightly linked to CEN6 is important for growth of *Saccharomyces cerevisiae*. Curr. Genet., **19**:1-8.

Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Kundrot CE, Cech TR, Doudna JA. 1996. Crystal structure of a group I ribozyme domain: principles of RNA packing. Science, **273**:1678-1685.

Cattaneo R. 1994. RNA editing. RNA duplexes guide base conversions. Curr. Biol., **4**:134-136.

Cavaillé J, Nicoloso M, Bachellerie JP. 1996. Targeted ribose methylation of RNA *in vivo* directed by tailored antisense RNA guides. Nature, **383**:732-735.

Cavaillé J, Bachellerie JP. 1998. SnoRNA-guided ribose methylation of rRNA: structural features of the guide RNA duplex influencing the extent of the reaction. Nucleic Acids Res., **26**:1576-1587.

Cavalier-Smith T. 1991. Intron phylogeny: a new hypothesis. Trends Genet., **7**:145-148.

Cech TR. 1985. Self-splicing RNA: implications for evolution. Int. Rev. Cytol., **93**:3-22.

Cedergren-Zeppezauer ES, Larsson G, Nyman PO, Dauter Z, Wilson KS. 1992. Crystal structure of a dUTPase. Nature, **355**:740-743.

Cellini A, Felder E, Rossi JJ. 1986. Yeast pre-messenger RNA splicing efficiency depends on critical spacing requirements between the branch point and 3' splice site. EMBO J., **5**:1023-1030.

- Cermakian N, Cedergren R. 1998. Modified nucleosides always were: an evolutionary model. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 535-542.
- Champion-Arnaud P, Reed R. 1994. The prespliceosome components SAP 49 and SAP 145 interact in a complex implicated in tethering U2 snRNP to the branch site. Genes Dev., 8:1974-1983.
- Chandrasekharappa SC, Smith JH, Eliceiri GL. 1983. Biosynthesis of small nuclear RNAs in human cells. J. Cell. Physiol., 117:169-174.
- Chanfreau G, Jacquier A. 1993. Interaction of intronic boundaries is required for the second splicing step efficiency of a group II intron. EMBO J., 12:5173-5180.
- Chanfreau G, Jacquier A. 1994. Catalytic site components common to both splicing steps of a group II intron. Science, 266:1383-1387.
- Chanfreau G, Legrain P, Dujon B, Jacquier A. 1994. Interaction between the first and last nucleotides of pre-mRNA introns is a determinant of 3' splice site selection in *S. cerevisiae*. Nucleic Acids Res., 22:1981-1987.
- Chanfreau G, Elela SA, Ares M, Jr., Guthrie C. 1997. Alternative 3'-end processing of U5 snRNA by RNase III. Genes Dev., 11:2741-2751.
- Chapon C, Legrain P. 1992. A novel gene, *spp91-1*, suppresses the splicing defect and the pre-mRNA nuclear export in the *prp9-1* mutant. EMBO J., 11:3279-3288.
- Charpentier B, Rosbash M. 1996. Intramolecular structure in yeast introns aids the early steps of *in vitro* spliceosome assembly. RNA, 2:509-522.
- Chen HR, Jan SP, Tsao TY, Sheu YJ, Banroques J, Cheng SC. 1998. Snt309p, a component of the Prp19p-associated complex that interacts with Prp19p and associates with the spliceosome simultaneously with or immediately after dissociation of U4 in the same manner as Prp19p. Mol. Cell. Biol., 18:2196-2204.
- Chen J, Patton JR. 1999. Cloning and characterization of a mammalian pseudouridine synthase. RNA, 5:409-419.
- Cheng SC, Abelson J. 1987. Spliceosome assembly in yeast. Genes Dev., 1:1014-1027.
- Chiara MD, Gozani O, Bennett M, Champion-Arnaud P, Palandjian L, Reed R. 1996. Identification of proteins that interact with exon sequences, splice sites, and the branchpoint sequence during each stage of spliceosome assembly. Mol. Cell. Biol., 16:3317-3326.
- Chiara MD, Palandjian L, Feld Krämer R, Reed R. 1997. Evidence that U5 snRNP recognizes the 3' splice site for catalytic step II in mammals. EMBO J., 16:4746-4759.
- Chu FK, Maley GF, West DK, Belfort M, Maley F. 1986. Characterization of the intron in the phage T4 thymidylate synthase gene and evidence for its self-excision from the primary transcript. Cell, 45:157-166.
- Chu JL, Elkon KB. 1991. The small nuclear ribonucleoproteins, SmB and B', are products of a single gene. Gene, 97:311-312.

- Chua K, Reed R. 1999. Human step II splicing factor hSlu7 functions in restructuring the spliceosome between the catalytic steps of splicing. Genes Dev., 13:841-850.
- Ciampi MS, Arena F, Cortese R. 1977. Biosynthesis of pseudouridine in the *in vitro* transcribed tRNA^{Tyr} precursor. FEBS Lett., 77:75-82.
- Claesson C, Lustig F, Boren T, Simonsson C, Barciszewska M, Lagerkvist U. 1995. Glycine codon discrimination and the nucleotide in position 32 of the anticodon loop. J. Mol. Biol., 247:191-196.
- Cohen JB, Snow JE, Spencer SD, Levinson AD. 1994. Suppression of mammalian 5' splice-site defects by U1 small nuclear RNAs from a distance. Proc. Natl. Acad. Sci. USA, 91:10470-10474.
- Cohn WE. 1959. 5-Ribosyl uracil, a carbon-carbon ribofuranosyl nucleoside in ribonucleic acids. Biochim. Biophys. Acta, 32:569-571.
- Cohn WE. 1960. Pseudouridine, a carbon-carbon linked ribonucleoside in ribonucleic acids: isolation, structure, and chemical characteristics. J. Biol. Chem., 235:1488-1498.
- 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., 10:1699-1708.
- Conrad J, Sun D, Englund N, Ofengand J. 1998. The *rlyC* gene of *Escherichia coli* codes for a pseudouridine synthase that is solely responsible for synthesis of pseudouridine at positions 955, 2504, and 2580 in 23 S ribosomal RNA. J. Biol. Chem., 273:18562-18566.
- Conrad J, Niu L, Rudd K, Lane BG, Ofengand J. 1999. 16S ribosomal RNA pseudouridine synthase RsuA of *Escherichia coli*: deletion, mutation of the conserved Asp102 residue, and sequence comparison among all other pseudouridine synthases. RNA, 5:751-763.
- Cooper M, Johnston LH, Beggs JD. 1995. Identification and characterization of Uss1p (Sdb23p): a novel U6 snRNA-associated protein with significant similarity to core proteins of small nuclear ribonucleoproteins. EMBO J., 14:2066-2075.
- Corollo D, Blair-Johnson M, Conrad J, Fiedler T, Sun D, Wang L, Ofengand J, Fenna R. 1999. Crystallization and characterization of a fragment of pseudouridine synthase RlyC from *Escherichia coli*. Acta Crystallogr. D. Biol. Crystallogr., 55:302-304.
- Cortes JJ, Sontheimer EJ, Seiwert SD, Steitz JA. 1993. Mutations in the conserved loop of human U5 snRNA generate use of novel cryptic 5' splice sites *in vivo*. EMBO J., 12:5181-5189.
- Cortese R, Kammen HO, Spengler SJ, Ames BN. 1974. Biosynthesis of pseudouridine in transfer ribonucleic acid. J. Biol. Chem., 249:1103-1108.
- Crick FH. 1966. Codon-anticodon pairing: the wobble hypothesis. J. Mol. Biol., 19:548-555.
- Crispino JD, Blencowe BJ, Sharp PA. 1994. Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. Science, 265:1866-1869.

Cunningham PR, Richard RB, Weitzmann CJ, Nurse K, Ofengand J. 1991. The absence of modified nucleotides affects both *in vitro* assembly and *in vitro* function of the 30S ribosomal subunit of *Escherichia coli*. Biochimie, *73*:789-796.

Curnow AW, Kung FL, Koch KA, Garcia GA. 1993. tRNA-guanine transglycosylase from *Escherichia coli*: gross tRNA structural requirements for recognition. Biochemistry, *32*:5239-5246.

Curran JF. 1998. Modified nucleosides in translation. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 493-516.

D

ahlberg JE, Lund E. 1988. The genes and transcription of the major small nuclear RNAs. In: Birnstiel

ML, eds. Structure and function of major and minor small nuclear ribonucleoprotein particles. New-York, NY: Springer-Verlag. pp 38-70.

Dalrymple MA, Petersen-Bjorn S, Friesen JD, Beggs JD. 1989. The product of the *PRP4* gene of *S. cerevisiae* shows homology to beta subunits of G proteins. Cell, *58*:811-812.

Damier L, Domenjoud L, Branlant C. 1997. The D1-A2 and D2-A2 pairs of splice sites from human immunodeficiency virus type 1 are highly efficient *in vitro*, in spite of an unusual branch site. Biochem. Biophys. Res. Commun., *237*:182-187.

Daniels CJ, Gupta R, Doolittle WF. 1985. Transcription and excision of a large intron in the tRNA^{Trp} gene of an archaeobacterium, *Halobacterium volcanii*. J. Biol. Chem., *260*:3132-3134.

Datta B, Weiner AM. 1991. Genetic evidence for base pairing between U2 and U6 snRNA in mammalian mRNA splicing. Nature, *352*:821-824.

Datta B, Weiner AM. 1992. Cross-linking of U2 snRNA using nitrogen mustard. Evidence for higher order structure. J. Biol. Chem., *267*:4497-4502.

Davis FF, Allen FW. 1957. Ribonucleic acids from yeast which contain a fifth nucleotide. J. Biol. Chem., *227*:907-915.

Davis DR, Poulter CD. 1991. 1H-15N NMR studies of *Escherichia coli* tRNA(Phe) from hisT mutants: a structural role for pseudouridine. Biochemistry, *30*:4223-4231.

Davis DR. 1995. Stabilization of RNA stacking by pseudouridine. Nucleic Acids Res., *23*:5020-5026.

Davis DR. 1998. Biophysical and conformational properties of modified nucleosides in RNA (nuclear magnetic resonance studies). In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 85-102.

- Davis DR, Veltri CA, Nielsen L. 1998. An RNA model system for investigation of pseudouridine stabilization of the codon-anticodon interaction in tRNA^{Lys}, tRNA^{His} and tRNA^{Tyr}. J. Biomol. Struct. Dyn., 15:1121-1132.
- Degoul F, Brulé H, Capanec C, Helm M, Marsac C, Leroux J, Giegé R, Florentz C. 1998. Isoleucylation properties of native human mitochondrial tRNA^{Ile} and tRNA^{Ile} transcripts. Implications for cardiomyopathy-related point mutations (4269, 4317) in the tRNA^{Ile} gene. Hum. Mol. Genet., 7:347-354.
- Deirdre A, Scadden J, Smith CW. 1995. Interactions between the terminal bases of mammalian introns are retained in inosine-containing pre-mRNAs. EMBO J., 14:3236-3246.
- Derrick WB, Horowitz J. 1993. Probing structural differences between native and *in vitro* transcribed *Escherichia coli* valine transfer RNA: evidence for stable base modification-dependent conformers. Nucleic Acids Res., 21:4948-4953.
- Deshler JO, Rossi JJ. 1991. Unexpected point mutations activate cryptic 3' splice sites by perturbing a natural secondary structure within a yeast intron. Genes Dev., 5:1252-1263.
- Dietrich RC, Inorvaia R, Padgett RA. 1997. Terminal intron dinucleotide sequences do not distinguish between U2- and U12-dependent introns. Mol. Cell., 1:151-160.
- Dix I, Russell CS, O'Keefe RT, Newman AJ, Beggs JD. 1998. Protein-RNA interactions in the U5 snRNP of *Saccharomyces cerevisiae*. RNA, 4:1675-1686.
- Dokudovskaya S, Dontsova O, Shpanchenko O, Bogdanov A, Brimacombe R. 1996. Loop IV of 5S ribosomal RNA has contacts both to domain II and to domain V of the 23S RNA. RNA, 2:146-152.
- Dreyfuss G, Matunis MJ, Pinol-Roma S, Burd CG. 1993. hnRNP proteins and the biogenesis of mRNA. Annu. Rev. Biochem., 62:289-321.
- Droogmans L, Haumont E, de Henau S, Grosjean H. 1986. Enzymatic 2'-O-methylation of the wobble nucleoside of eukaryotic tRNA^{Phe}: specificity depends on structural elements outside the anticodon loop. EMBO J., 5:1105-1109.
- Durant PC, Davis DR. 1997. The effect of pseudouridine and pH on the structure and dynamics of the anticodon stem-loop of tRNA(Lys,3). Nucleic Acids Symp. Ser., 36:56-57.
- Durant PC, Davis DR. 1999. Stabilization of the anticodon stem-loop of tRNA^{Lys,3} by an A+-C base-pair and by pseudouridine. J. Mol. Biol., 285:115-131.
- Dyhr-Mikkelsen H, Kjems J. 1995. Inefficient spliceosome assembly and abnormal branch site selection in splicing of an HIV-1 transcript *in vitro*. J. Biol. Chem., 270:24060-24066.

E

Edmonds CG, Crain PF, Gupta R, Hashizume T, Hocart CH, Kowalak JA, Pomerantz SC, Stetter KO,

McCloskey JA. 1991. Posttranscriptional modification of tRNA in thermophilic *archaea* (*Archaeobacteria*). J. Bacteriol., 173:3138-3148.

Edqvist J, Grosjean H, Straby KB. 1992. Identity elements for N2-dimethylation of guanosine-26 in yeast tRNAs. Nucleic Acids Res., 20:6575-6581.

Edqvist J, Straby KB, Grosjean H. 1993. Pleiotrophic effects of point mutations in yeast tRNA(Asp) on the base modification pattern. Nucleic Acids Res., 21:413-417.

Edqvist J, Blomqvist K, Straby KB. 1994. Structural elements in yeast tRNAs required for homologous modification of guanosine-26 into dimethylguanosine-26 by the yeast Trm1 tRNA- modifying enzyme. Biochemistry, 33:9546-9551.

Ehrenman K, Pedersen-Lane J, West D, Herman R, Maley F, Belfort M. 1986. Processing of phage T4 td-encoded RNA is analogous to the eukaryotic group I splicing pathway. Proc. Natl. Acad. Sci. USA, 83:5875-5879.

Epstein P, Reddy R, Henning D, Busch H. 1980. The nucleotide sequence of nuclear U6 (4.7 S) RNA. J. Biol. Chem., 255:8901-8906.

F

Fabrizio P, McPheeters DS, Abelson J. 1989. *In vitro* assembly of yeast U6 snRNP: a functional assay.

Genes Dev., 3:2137-2150.

Fabrizio P, Abelson J. 1990a. Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. Science, 250:404-409.

Fabrizio P, Abelson J. 1990b. Point mutations in yeast U6 snRNA can specifically block the first or second step of pre-mRNA splicing *in vitro*. Mol. Biol. Rep., 14:135.

Fabrizio P, Esser S, Kastner B, Lührmann R. 1994. Isolation of *S. cerevisiae* snRNPs: comparison of U1 and U4/U6.U5 to their human counterparts. Science, 264:261-265.

- Fabrizio P, Laggerbauer B, Lauber J, Lane WS, Lührmann R. 1997. An evolutionarily conserved U5 snRNP-specific protein is a GTP-binding factor closely related to the ribosomal translocase EF-2. EMBO J., 16:4092-4106.
- Felden B, Florentz C, Giegé R, Westhof E. 1994. Solution structure of the 3'-end of brome mosaic virus genomic RNAs. Conformational mimicry with canonical tRNAs. J. Mol. Biol., 235:508-531.
- Felden B, Florentz C, Westhof E, Giegé R. 1998. Transfer RNA identity rules and conformation of the tyrosine tRNA-like domain of BMV RNA imply additional charging by histidine and valine. Biochem. Biophys. Res. Commun., 243:426-434.
- Ferat JL, Michel F. 1993. Group II self-splicing introns in *Bacteria*. Nature, 364:358-361.
- Ferat JL, Le Gouar M, Michel F. 1994. Multiple group II self-splicing introns in mobile DNA from *Escherichia coli*. CR Acad. Sci. III, 317:141-148.
- Fetzer S, Lauber J, Will CL, Lührmann R. 1997. The [U4/U6.U5] tri-snRNP-specific 27K protein is a novel SR protein that can be phosphorylated by the snRNP-associated protein kinase. RNA, 3:344-355.
- Field DJ, Friesen JD. 1996. Functionally redundant interactions between U2 and U6 spliceosomal snRNAs. Genes Dev., 10:489-501.
- Fisher DE, Conner GE, Reeves WH, Wisniewolski R, Blobel G. 1985. Small nuclear ribonucleoprotein particle assembly *in vivo*: demonstration of a 6S RNA-free core precursor and posttranslational modification. Cell, 42:751-758.
- Fischer U, Lührmann R. 1990. An essential signaling role for the m₃G cap in the transport of U1 snRNP to the nucleus. Science, 249:786-790.
- Fischer U, Darzynkiewicz E, Tahara SM, Dathan NA, Lührmann R, Mattaj IW. 1991. Diversity in the signals required for nuclear accumulation of U snRNPs and variety in the pathways of nuclear transport. J. Cell. Biol., 113:705-714.
- Fischer U, Liu Q, Dreyfuss G. 1997. The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. Cell, 90:1023-1029.
- Florentz C, Giegé R. 1995. tRNA-like structures in plant viral RNAs. In: Söll D, RajBhandary UL, eds. tRNA: Structure, biosynthesis and Function. Washington, DC: ASM press. pp 141-163.
- Forne T, Rossi F, Labourier E, Antoine E, Cathala G, Brunel C, Tazi J. 1995. Disruption of base-paired U4.U6 small nuclear RNAs induced by mammalian heterogeneous nuclear ribonucleoprotein C protein. J. Biol. Chem., 270:16476-16481.
- Fortner DM, Troy RG, Brow DA. 1994. A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. Genes Dev., 8:221-233.
- Fothergill-Gilmore LA, Michels PA. 1993. Evolution of glycolysis. Prog. Biophys. Mol. Biol., 59:105-235.

- Fournier R. 1997. Etude de la biogenèse, de la structure, de l'évolution et du mécanisme d'action du snoRNA U3 de levures. Thèse de l'université H. Poincaré, Nancy I.
- Fouser LA, Friesen JD. 1986. Mutations in a yeast intron demonstrate the importance of specific conserved nucleotides for the two stages of nuclear mRNA splicing. Cell, 45:81-93.
- Frank D, Patterson B, Guthrie C. 1992. Synthetic lethal mutations suggest interactions between U5 small nuclear RNA and four proteins required for the second step of splicing. Mol. Cell. Biol., 12:5197-5205.
- Frank DN, Roiha H, Guthrie C. 1994. Architecture of the U5 small nuclear RNA. Mol. Cell. Biol., 14:2180-2190.
- Fresco LD, Harper DS, Keene JD. 1991. Leucine periodicity of U2 small nuclear ribonucleoprotein particle (snRNP) A' protein is implicated in snRNP assembly via protein-protein interactions. Mol. Cell. Biol., 11:1578-1589.
- Frilander MJ, Steitz JA. 1999. Initial recognition of U12-dependent introns requires both U11/5' splice-site and U12/branchpoint interactions. Genes Dev., 13:851-863.
- Fromont-Racine M, Rain JC, Legrain P. 1997. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. Nat. Genet., 16:277-282.
- Fu XD. 1995. The superfamily of arginine/serine-rich splicing factors. RNA, 1:663-680

G

alissou F, Legrain P. 1993. The biochemical defects of *prp4-1* and *prp6-1* yeast splicing mutants reveal

that the PRP6 protein is required for the accumulation of the [U4/U6.U5] tri-snRNP. Nucleic Acids Res., 21:1555-1562.

- Ganot P, Bortolin ML, Kiss T. 1997a. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. Cell, 89:799-809.
- 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., 11:941-956.
- Ganot P, Jady BE, Bortolin M, Darzacq X, Kiss T. 1999. Nucleolar factors direct the 2'-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. Mol. Cell. Biol., 19:6906-6917.
- Garcia GA, Chong S. 1997. Cysteine 265 is in the active site of, but is not essential for catalysis by tRNA-guanine transglycosylase (TGT) from *Escherichia coli*. J. Protein. Chem., 16:11-17.

- Garcia GA, Goodenough-Lashua DM. 1998. Mechanisms of RNA-Modifying and -Editing Enzymes. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 135-168.
- Gaur RK, Valcarcel J, Green MR. 1995. Sequential recognition of the pre-mRNA branch point by U2AF65 and a novel spliceosome-associated 28-kDa protein. RNA, 1:407-417.
- 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., 17:7088-7098.
- Ghetti A, Pinol-Roma S, Michael WM, Morandi C, Dreyfuss G. 1992. hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. Nucleic Acids Res., 20:3671-3678.
- Ghetti A, Company M, Abelson J. 1995. Specificity of Prp24 binding to RNA: a role for Prp24 in the dynamic interaction of U4 and U6 snRNAs. RNA, 1:132-145.
- Giegé R, Moras D, Thierry JC. 1977. Yeast transfer RNA^{Asp}: a new high-resolution x-ray diffracting crystal form of a transfer RNA. J. Mol. Biol., 115:91-96.
- Giorgianni F, Beranova S, Wesdemiotis C, Viola RE. 1995. Elimination of the sensitivity of L-aspartase to active-site-directed inactivation without alteration of catalytic activity. Biochemistry, 34:3529-3535.
- 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., 11:673-682.
- Goguel V, Rosbash M. 1993. Splice site choice and splicing efficiency are positively influenced by pre-mRNA intramolecular base pairing in yeast. Cell, 72:893-901.
- Goldschmidt-Clermont M, Choquet Y, Girard-Bascou J, Michel F, Schirmer-Rahire M, Rochaix JD. 1991. A small chloroplast RNA may be required for trans-splicing in *Chlamydomonas reinhardtii*. Cell, 65:135-143.
- Goodman HM, Olson MV, Hall BD. 1977. Nucleotide sequence of a mutant eukaryotic gene: the yeast tyrosine-inserting ochre suppressor SUP4-o. Proc. Natl. Acad. Sci. USA, 74:5453-5457.
- Goodrich-Blair H, Scarlato V, Gott JM, Xu MQ, Shub DA. 1990. A self-splicing group I intron in the DNA polymerase gene of *Bacillus subtilis* bacteriophage SPO1. Cell, 63:417-424.
- Gottschalk A, Tang J, Puig O, Salgado J, Neubauer G, Colot HV, Mann M, Séraphin B, Rosbash M, Lührmann R, Fabrizio P. 1998. A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins. RNA, 4:374-393.
- Gottschalk A, Neubauer G, Banroques J, Mann M, Lührmann R, Fabrizio P. 1999. Identification by mass spectrometry and functional analysis of novel proteins of the yeast (U4/U6.U5) tri-snRNP. EMBO, 18:4535-48
- Gozani O, Patton JG, Reed R. 1994. A novel set of spliceosome-associated proteins and the essential splicing factor PSF bind stably to pre-mRNA prior to catalytic step II of the splicing reaction. EMBO J., 13:3356-3367.

- Gozani O, Feld R, Reed R. 1996. Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. Genes Dev., 10:233-243.
- Gozani O, Potashkin J, Reed R. 1998. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. Mol. Cell. Biol., 18:4752-4760.
- Grabowski PJ, Seiler SR, Sharp PA. 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. Cell, 42:345-353.
- Grabowski PJ, Sharp PA. 1986. Affinity chromatography of splicing complexes: U2, U5, and U4 + U6 small nuclear ribonucleoprotein particles in the spliceosome. Science, 233:1294-1299.
- Green CJ, Kammen HO, Penhoet EE. 1982. Purification and properties of a mammalian tRNA pseudouridine synthase. J. Biol. Chem., 257:3045-3052.
- Green MR. 1986. Pre-mRNA splicing. Annu. Rev. Genet., 20:671-708.
- Green R, Noller HF. 1996. *In vitro* complementation analysis localizes 23S rRNA posttranscriptional modifications that are required for *Escherichia coli* 50S ribosomal subunit assembly and function. RNA, 2:1011-1021.
- Grosjean H, Sprinzl M, Steinberg S. 1995. Posttranscriptionally modified nucleosides in transfer RNA: their locations and frequencies. Biochimie, 77:139-141.
- Grosjean H, Edqvist J, Straby KB, Giegé R. 1996. Enzymatic formation of modified nucleosides in tRNA: dependence on tRNA architecture. J. Mol. Biol., 255:67-85.
- Grosjean H, Szweykowska-Kulinska Z, Motorin Y, Fasiolo F, Simos G. 1997. Intron-dependent enzymatic formation of modified nucleosides in eukaryotic tRNAs: a review. Biochimie, 79:293-302.
- Grosjean H, Houssier C, Romby P, Marquet R. 1998. Modulation role of modified nucleotides in RNA loop-loop interaction. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 113-134
- Gross T, Richert K, Mierke C, Lützelberger M, Käufer NF. Identification and characterization of *srp1*, a gene of fission yeast encoding a RNA binding domain and a RS domain typical of SR splicing factors. Nucleic Acids Res., 26:505-511.
- Gu XR, Santi DV. 1991. The T-arm of tRNA is a substrate for tRNA (m5U54)-methyltransferase. Biochemistry, 30:2999-3002.
- Gu X, Ofengand J, Santi DV. 1994. *In vitro* methylation of *Escherichia coli* 16S rRNA by tRNA (m5U54)-methyltransferase. Biochemistry, 33:2255-2261.
- Gu J, Patton JR, Shimba S, Reddy R. 1996. Localization of modified nucleotides in *Schizosaccharomyces pombe* spliceosomal small nuclear RNAs: modified nucleotides are clustered in functionally important regions. RNA, 2:909-918.

Gu X, Yu M, Ivanetich KM, Santi DV. 1998. Molecular recognition of tRNA by tRNA pseudouridine 55 synthase. Biochemistry, 37:339-343.

Gustafsson C, Reid R, Greene PJ, Santi DV. 1996. Identification of new RNA modifying enzymes by iterative genome search using known modifying enzymes as probes. Nucleic Acids Res., 24:3756-3762.

Guthrie C, Patterson B. 1988. Spliceosomal snRNAs. Annu. Rev. Genet., 22:387-419.

H

ackl W, Fischer U, Lührmann R. 1994. A 69-kD protein that associates reversibly with the Sm core

domain of several spliceosomal snRNP species. J. Cell. Biol., 124:261-272.

Hackl W, Lührmann R. 1996. Molecular cloning and subcellular localisation of the snRNP-associated protein 69KD, a structural homologue of the proto-oncoproteins TLS and EWS with RNA and DNA-binding properties. J. Mol. Biol., 264:843-851.

Hagervall TG, Ericson JU, Esberg KB, Li JN, Björk GR. 1990. Role of tRNA modification in translational fidelity. Biochim. Biophys. Acta, 1050:263-266.

Hajduk SL, Sabatini RS. 1998. Mitochondrial mRNA editing in Kinetoplastid protozoa. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 377-394.

Halfter H, Gallwitz D. 1988. Impairment of yeast pre-mRNA splicing by potential secondary structure-forming sequences near the conserved branchpoint sequence. Nucleic Acids Res., 16:10413-10423.

Hall KB, Sampson JR, Uhlenbeck OC, Redfield AG. 1989. Structure of an unmodified tRNA molecule. Biochemistry, 28:5794-5801.

Hall KB, McLaughlin LW. 1991. Properties of a U1/mRNA 5' splice site duplex containing pseudouridine as measured by thermodynamic and NMR methods. Biochemistry, 30:1795-1801.

Hall KB, McLaughlin LW. 1992. Properties of pseudouridine N1 imino protons located in the major groove of an A-form RNA duplex. Nucleic Acids Res., 20:1883-1889.

Hall SL, Padgett RA. 1994. Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. J. Mol. Biol., 239:357-365.

Hall SL, Padgett RA. 1996. Requirement of U12 snRNA for *in vivo* splicing of a minor class of eukaryotic nuclear pre-mRNA introns. Science, 271:1716-1718.

Hamm J, Mattaj IW. 1989. An abundant U6 snRNP found in germ cells and embryos of *Xenopus laevis*. EMBO J., 8:4179-4187.

Hamm J, Dathan NA, Mattaj IW. 1989. Functional analysis of mutant *Xenopus* U2 snRNAs. Cell, 59:159-169.

- Hamm J, Darzynkiewicz E, Tahara SM, Mattaj IW. 1990a. The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. Cell, 62:569-577.
- Hamm J, Dathan NA, Scherly D, Mattaj IW. 1990b. Multiple domains of U1 snRNA, including U1 specific protein binding sites, are required for splicing. EMBO J., 9:1237-1244.
- Hartmuth K, Barta A. 1988. Unusual branch point selection in processing of human growth hormone pre-mRNA. Mol. Cell. Biol., 8:2011-2020.
- Hartmuth K, Raker VA, Huber J, Branlant C, Lührmann R. 1999. An unusual chemical reactivity of Sm site adenosines strongly correlates with proper assembly of core U snRNP particles. J. Mol. Biol., 285:133-147.
- Hashimoto C, Steitz JA. 1984. U4 and U6 RNAs coexist in a single small nuclear ribonucleoprotein particle. Nucleic Acids Res., 12:3283-3293.
- Hausner TP, Giglio LM, Weiner AM. 1990. Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles. Genes Dev., 4:2146-2156.
- Hawkins JD. 1988. A survey on intron and exon lengths. Nucleic Acids Res., 16:9893-9908.
- Heinrichs V, Bach M, Winkelmann G, Lührmann R. 1990. U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5' splice site. Science, 247:69-72.
- Heinrichs V, Hackl W, Lührmann R. 1992. Direct binding of small nuclear ribonucleoprotein G to the Sm site of small nuclear RNA. Ultraviolet light cross-linking of protein G to the AAU stretch within the Sm site (AAUUUGUGG) of U1 small nuclear ribonucleoprotein reconstituted *in vitro*. J. Mol. Biol., 227:15-28.
- Heiss NS, Knight SW, Vulliamy TJ, Klauck 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., 19:32-38.
- Hellmuth K, Lau DM, Bischoff FR, Kunzler M, Hurt E, Simos G. 1998. Yeast Los1p has properties of an exportin-like nucleocytoplasmic transport factor for tRNA. Mol. Cell. Biol., 18:6374-6386.
- Hellung-Larsen P, Kulamowicz I, Frederiksen S. 1980. Synthesis of low molecular weight RNA components in cells with a temperature-sensitive polymerase II. Biochim. Biophys. Acta, 609:201-204.
- Helm M, Brulé H, Degoul F, Capanec C, Leroux JP, Giegé R, Florentz C. 1998. The presence of modified nucleotides is required for cloverleaf folding of a human mitochondrial tRNA. Nucleic Acids Res., 26:1636-1643.
- Helm M, Florentz C, Chomyn A, Attardi G. 1999. Search for differences in post-transcriptional modification patterns of mitochondrial DNA-encoded wild-type and mutant human tRNA^{Lys} and tRNA^{Leu}(UUR). Nucleic Acids Res., 27:756-763.
- 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., 17:7078-7090.

- Hermann H, Fabrizio P, Raker VA, Foulaki K, Hornig H, Brahms H, Lührmann R. 1995. snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. EMBO J., 14:2076-2088.
- Hetzer M, Wurzer G, Schweyen RJ, Mueller MW. 1997. Trans-activation of group II intron splicing by nuclear U5 snRNA. Nature, 386:417-420.
- Hinterberger M, Pettersson I, Steitz JA. 1983. Isolation of small nuclear ribonucleoproteins containing U1, U2, U4, U5, and U6 RNAs. J. Biol. Chem., 258:2604-2613.
- Hinz M, Moore MJ, Bindereif A. 1996. Domain analysis of human U5 RNA. Cap trimethylation, protein binding, and spliceosome assembly. J. Biol. Chem., 271:19001-19007.
- Hodges PE, Jackson SP, Brown JD, Beggs JD. 1995. Extraordinary sequence conservation of the PRP8 splicing factor. Yeast, 11:337-342.
- Holmes WM, Andraos-Selim C, Roberts I, Wahab SZ. 1992. Structural requirements for tRNA methylation. Action of *Escherichia coli* tRNA(guanosine-1)methyltransferase on tRNA(1Leu) structural variants. J. Biol. Chem., 267:13440-13445.
- Hong W, Bennett M, Xiao Y, Feld Krämer R, Wang C, Reed R. 1997. Association of U2 snRNP with the spliceosomal complex E. Nucleic Acids Res., 25:354-361.
- Hopper AK, Furukawa AH, Pham HD, Martin NC. 1982. Defects in modification of cytoplasmic and mitochondrial transfer RNAs are caused by single nuclear mutations. Cell, 28:543-550.
- Hornig H, Aebi M, Weissmann C. 1986. Effect of mutations at the lariat branch acceptor site on beta-globin pre-mRNA splicing *in vitro*. Nature, 324:589-591.
- Horowitz DS, Abelson J. 1993a. A U5 small nuclear ribonucleoprotein particle protein involved only in the second step of pre-mRNA splicing in *Saccharomyces cerevisiae*. Mol. Cell. Biol., 13:2959-2970.
- Horowitz DS, Abelson J. 1993b. Stages in the second reaction of pre-mRNA splicing: the final step is ATP independent. Genes Dev., 7:320-329.
- 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, 3:1374-1387.
- Horowitz DS, Krainer AR. 1997. A human protein required for the second step of pre-mRNA splicing is functionally related to a yeast splicing factor. Genes Dev., 11:139-151.
- Hu J, Xu Y, Schappert K, Harrington T, Wang A, Braga R, Mogridge J, Friesen JD. 1994. Mutational analysis of the PRP4 protein of *Saccharomyces cerevisiae* suggests domain structure and snRNP interactions. Nucleic Acids Res., 22:1724-1734.
- Huang S, Spector DL. 1997. Nuclear organization of pre-mRNA splicing factors and substrates. In: Krainer AR, eds. Eukaryotic mRNA processing. Cold Spring Harbor, N. Y.: pp 37-67.

Huang L, Ku J, Pookanjanatavip M, Gu X, Wang D, Greene PJ, Santi DV. 1998a. Identification of two *Escherichia coli* pseudouridine synthases that show multisite specificity for 23S RNA. Biochemistry, 37:15951-15957.

Huang L, Pookanjanatavip M, Gu X, Santi DV. 1998b. A conserved aspartate of tRNA pseudouridine synthase is essential for activity and a probable nucleophilic catalyst. Biochemistry, 37:344-351.

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., 259:645-654.

Hwang DY, Cohen JB. 1996. U1 snRNA promotes the selection of nearby 5' splice sites by U6 snRNA in mammalian cells. Genes Dev., 10:338-350.

Igel AH, Ares M, Jr. 1988. Internal sequences that distinguish yeast from metazoan U2 snRNA are unnecessary for pre-mRNA splicing. Nature, 334:450-453.

Igel H, Wells S, Perriman R, Ares M, Jr. 1998. Conservation of structure and subunit interactions in yeast homologues of splicing factor 3b (SF3b) subunits. RNA, 4:1-10.

Incorvaia R, Padgett RA. 1998. Base pairing with U6atac snRNA is required for 5' splice site activation of U12-dependent introns *in vivo*. RNA, 4:709-718.

Isel C, Marquet R, Keith G, Ehresmann C, Ehresmann B. 1993. Modified nucleotides of tRNA(3Lys) modulate primer/template loop-loop interaction in the initiation complex of HIV-1 reverse transcription. J. Biol. Chem., 268:25269-25272.

Izaurralde E, Lewis J, McGuigan C, Jankowska M, Darzynkiewicz E, Mattaj IW. 1994. A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell, 78:657-668.

Izaurralde E, Lewis J, Gamberi C, Jarmolowski A, McGuigan C, Mattaj IW. 1995. A cap-binding protein complex mediating U snRNA export. Nature, 376:709-712.

Jackson IJ. 1991. A reappraisal of non-consensus mRNA splice sites. Nucleic Acids Res., 19:3795-3798.

- Jackson SP, Lossky M, Beggs JD. 1988. Cloning of the *RNA8* gene of *Saccharomyces cerevisiae*, detection of the RNA8 protein, and demonstration that it is essential for nuclear pre- mRNA splicing. Mol. Cell. Biol., 8:1067-1075.
- Jacob M, Gallinaro H. 1989. The 5' splice site: phylogenetic evolution and variable geometry of association with U1RNA. Nucleic Acids Res., 17:2159-2180.
- Jacobs EY, Frey MR, Wu W, Ingledue TC, Gebuhr TC, Gao L, Marzluff WF, Matera AG. 1999. Coiled bodies preferentially associate with U4, U11, and U12 small nuclear RNA genes in interphase HeLa cells but not with U6 and U7 genes. Mol. Biol. Cell., 10:1653-1663.
- Jacobson MR, Rhoadhouse M, Pederson T. 1993. U2 small nuclear RNA 3' end formation is directed by a critical internal structure distinct from the processing site. Mol. Cell. Biol., 13:1119-1129.
- Jacquier A, Rodriguez JR, Rosbash M. 1985. A quantitative analysis of the effects of 5' junction and TACTAAC box mutants and mutant combinations on yeast mRNA splicing. Cell, 43:423-430.
- Jacquier A, Michel F. 1987. Multiple exon-binding sites in class II self-splicing introns. Cell, 50:17-29.
- Jacquier A, Michel F. 1990. Base-pairing interactions involving the 5' and 3'-terminal nucleotides of group II self-splicing introns. J. Mol. Biol., 213:437-447.
- Jacquier A. 1997. RNA-catalysed RNA processing. In: Krainer AR, eds. Eukaryotic mRNA processing. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press. pp 1-36
- Jakab G, Mougin A, Kis M, Pollak T, Antal M, Branlant C, Solymosy F. 1997. *Chlamydomonas* U2, U4 and U6 snRNAs. An evolutionary conserved putative third interaction between U4 and U6 snRNAs which has a counterpart in the U4atac-U6atac snRNA duplex. Biochimie, 79:387-395.
- Jamison SF, Pisman Z, Wang J, Will C, Lührmann R, Manley JL, Garcia-Blanco MA. 1995. U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements. Nucleic Acids Res., 23:3260-3267.
- Jarmolowski A, Mattaj IW. 1993. The determinants for Sm protein binding to *Xenopus* U1 and U5 snRNAs are complex and non-identical. EMBO J., 12:223-232.
- 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., 13:4884-4893.
- Jiang HQ, Motorin Y, Jin YX, Grosjean H. 1997. Pleiotropic effects of intron removal on base modification pattern of yeast tRNA^{Phe}: an *in vitro* study. Nucleic Acids Res., 25:2694-2701.
- Johnson PF, Abelson J. 1983. The yeast tRNA^{Tyr} gene intron is essential for correct modification of its tRNA product. Nature, 302:681-687.
- Jones MH, Guthrie C. 1990. Unexpected flexibility in an evolutionarily conserved protein-RNA interaction: genetic analysis of the Sm binding site. EMBO J., 9:2555-2561.

Jones MH, Frank DN, Guthrie C. 1995. Characterization and functional ordering of Slu7p and Prp17p during the second step of pre-mRNA splicing in yeast. Proc. Natl. Acad. Sci. USA, 92:9687-9691.

K

aine BP, Gupta R, Woese CR. 1983. Putative introns in tRNA genes of prokaryotes. Proc. Natl. Acad. Sci. USA, 80:3309-3312.

Kambach C, Walke S, Young R, Avis JM, de la Fortelle E, Raker VA, Lührmann R, Li J, Nagai K. 1999a. Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. Cell, 96:375-387.

Kambach C, Walke S, Nagai K. 1999b. Structure and assembly of the spliceosomal small nuclear ribonucleoprotein particles. Curr. Opin. Struct. Biol., 9:222-230.

Kammen HO, Marvel CC, Hardy L, Penhoet EE. 1988. Purification, structure, and properties of *Escherichia coli* tRNA pseudouridine synthase I. J. Biol. Chem., 263:2255-2263.

Kandels-Lewis S, Séraphin B. 1993. Involvement of U6 snRNA in 5' splice site selection. Science, 262:2035-2039.

Kao HY, Siliciano PG. 1992. The yeast homolog of the U1 snRNP protein 70K is encoded by the *SNP1* gene. Nucleic Acids Res., 20:4009-4013.

Kao HY, Siliciano PG. 1996. Identification of Prp40, a novel essential yeast splicing factor associated with the U1 small nuclear ribonucleoprotein particle. Mol. Cell. Biol., 16:960-967.

Kato N, Harada F. 1984. Nucleotide sequence of nuclear 5.4 S RNA of mouse cells. Biochim. Biophys. Acta, 782:127-131.

Kealey JT, Santi DV. 1995. Stereochemistry of tRNA(m⁵U54)-methyltransferase catalysis: 19F NMR spectroscopy of an enzyme-FUraRNA covalent complex. Biochemistry, 34:2441-2446.

Keith G, Roy A, Ebel JP, Dirheime rG. 1971. The nucleotide sequences of two tryptophane-tRNAs from brewer's yeast. FEBS lett., 17:306-308.

Keller EB, Noon WA. 1984. Intron splicing: a conserved internal signal in introns of animal pre- mRNAs. Proc. Natl. Acad. Sci. USA, 81:7417-7420.

Kim SH, Suddath FL, Quigley GJ, McPherson A, Sussman JL, Wang AH, Seeman NC, Rich A. 1974. Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. Science, 185:435-440.

Kim SH, Lin RJ. 1993. Pre-mRNA splicing within an assembled yeast spliceosome requires an RNA- dependent ATPase and ATP hydrolysis. Proc. Natl. Acad. Sci. USA, 90:888-892.

- Kim SH, Lin RJ. 1996. Spliceosome activation by PRP2 ATPase prior to the first transesterification reaction of pre-mRNA splicing. Mol. Cell. Biol., 16:6810-6819.
- Kim CH, Abelson J. 1996. Site-specific crosslinks of yeast U6 snRNA to the pre-mRNA near the 5' splice site. RNA, 2:995-1010.
- Kintanar A, Yue D, Horowitz J. 1994. Effect of nucleoside modifications on the structure and thermal stability of *Escherichia coli* valine tRNA. Biochimie, 76:1192-1204.
- 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, 85:1077-1088.
- 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., 17:797-807.
- Kleinschmidt AM, Patton JR, Pederson T. 1989. U2 small nuclear RNP assembly *in vitro*. Nucleic Acids Res., 17:4817-4828.
- Klenk HP, Clayton RA, Tomb JF, White O, Nelson KE, Ketchum KA, Dodson RJ, Gwinn M, Hickey EK, Peterson JD, Richardson DL, Kerlavage AR, Graham DE, Kyrpides NC, Fleischmann RD, Quackenbush J, Lee NH, Sutton GG, Gill S, Kirkness EF, Dougherty BA, McKenney K, Adams MD, Loftus B, Venter JC, *et al.* 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature, 390:364-370.
- Koch JL, Boulanger SC, Dib-Hajj SD, Hebbar SK, Perlman PS. 1992. Group II introns deleted for multiple substructures retain self-splicing activity. Mol. Cell. Biol., 12:1950-1958.
- Kohl RJ, Hall TC. 1974. Aminoacylation of RNA from several viruses: amino acid specificity and differential activity of plant, yeast and bacterial synthetases. J. Gen. Virol., 25:257-261.
- Kohtz JD, Jamison SF, Will CL, Zuo P, Lührmann R, Garcia-Blanco MA, Manley JL. 1994. Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. Nature, 368:119-124.
- Kolodrubetz D, Burgum A. 1991. Sequence and genetic analysis of NHP2: a moderately abundant high mobility group-like nuclear protein with an essential function in *Saccharomyces cerevisiae*. Yeast, 7:79-90.
- Kolossova I, Padgett RA. 1997. U11 snRNA interacts *in vivo* with the 5' splice site of U12-dependent (AU-AC) pre-mRNA introns. RNA, 3:227-233.
- Konarska MM, Grabowski PJ, Padgett RA, Sharp PA. 1985. Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors. Nature, 313:552-557.
- Konarska MM, Sharp PA. 1987. Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. Cell, 49:763-774.
- Konforti BB, Koziolkiewicz MJ, Konarska MM. 1993. Disruption of base pairing between the 5' splice site and the 5' end of U1 snRNA is required for spliceosome assembly. Cell, 75:863-873.

- 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., 22:2166-2167.
- 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., 24:2411-2415.
- Koski RA, Clarkson SG. 1982. Synthesis and maturation of *Xenopus laevis* methionine tRNA gene transcripts in homologous cell-free extracts. J. Biol. Chem., 257:4514-4521.
- Kowalak JA, Dalluge JJ, McCloskey JA, Stetter KO. 1994. The role of posttranscriptional modification in stabilization of transfer RNA from hyperthermophiles. Biochemistry, 33:7869-7876.
- Krämer A, Keller W, Appel B, Lührmann R. 1984. The 5' terminus of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA precursors. Cell, 38:299-307.
- Krämer A, Utans U. 1991. Three protein factors (SF1, SF3 and U2AF) function in pre-splicing complex formation in addition to snRNPs. EMBO J., 10:1503-1509.
- Krämer A. 1996. The structure and function of proteins involved in mammalian pre-mRNA splicing. Annu. Rev. Biochem., 65:367-409.
- Krämer A, Gruter P, Groning K, Kastner B. 1999. Combined biochemical and electron microscopic analyses reveal the architecture of the mammalian U2 snRNP. J. Cell. Biol., 145:1355-1368.
- Krecic AM, Swanson MS. 1999. hnRNP complexes: composition, structure, and function. Curr. Opin. Cell. Biol., 11:363-371.
- Kretzner L, Krol A, Rosbash M. 1990. *Saccharomyces cerevisiae* U1 small nuclear RNA secondary structure contains both universal and yeast-specific domains. Proc. Natl. Acad. Sci. USA, 87:851-855.
- Krol A, Branlant C, Lazar E, Gallinaro H, Jacob M. 1981a. Primary and secondary structures of chicken, rat and man nuclear U4 RNAs. Homologies with U1 and U5 RNAs. Nucleic Acids Res., 9:2699-2716.
- Krol A, Gallinaro H, Lazar E, Jacob M, Branlant C. 1981b. The nuclear 5S RNAs from chicken, rat and man. U5 RNAs are encoded by multiple genes. Nucleic Acids Res., 9:769-787.
- Krol A, Carbon P, Ebel JP, Appel B. 1987. *Xenopus tropicalis* U6 snRNA genes transcribed by Pol III contain the upstream promoter elements used by Pol II dependent U snRNA genes. Nucleic Acids Res., 15:2463-2478.
- Kuhn AN, Li Z, Brow DA. 1999. Splicing factor Prp8 governs U4/U6 RNA unwinding during activation of the spliceosome. Mol. Cell., 3:65-75.
- Kuhnel MG, Strickland R, Palmer JD. 1990. An ancient group I intron shared by *Eubacteria* and chloroplasts. Science, 250:1570-1573.
- Kunkel TA. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA, 82:488-492.

Kunkel GR, Maser RL, Calvet JP, Pederson T. 1986. U6 small nuclear RNA is transcribed by RNA polymerase III. Proc. Natl. Acad. Sci. USA, 83:8575-8579.

Kunkel TA, Bebenek K, McClary J. 1991. Efficient site-directed mutagenesis using uracil-containing DNA. Methods Enzymol., 204:125-139.

Kuo HC, Nasim FH, Grabowski PJ. 1991. Control of alternative splicing by the differential binding of U1 small nuclear ribonucleoprotein particle. Science, 251:1045-1050.

Kutay U, Lipowsky G, Izaurralde E, Bischoff FR, Schwarzmaier P, Hartmann E, Gorlich D. 1998. Identification of a tRNA-specific nuclear export receptor. Mol. Cell., 1:359-369.

L

afontaine 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., 241:492-497.

Lafontaine D, Vandenhaute J, Tollervey D. 1995. The 18S rRNA dimethylase Dim1p is required for pre-ribosomal RNA processing in yeast. Genes Dev., 9:2470-2481.

Lafontaine DL, Tollervey D. 1998a. Birth of the snoRNPs: the evolution of the modification-guide snoRNAs. Trends Biochem. Sci., 23:383-388.

Lafontaine DL, Tollervey D. 1998b. Regulatory aspects of rRNA modifications and pre-rRNA processing. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 281-8.

Lafontaine DLJ, Bousquet-Antonelli C, Henry Y, Caizergues-Ferrer M, Tollervey D. 1998a. The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. Genes Dev., 12:527-537.

Lafontaine DL, Preiss T, Tollervey D. 1998b. Yeast 18S rRNA dimethylase Dim1p: a quality control mechanism in ribosome synthesis? Mol. Cell. Biol., 18:2360-2370.

Lafontaine DL, Tollervey D. 1999. Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. RNA, 5:455-467.

Laggerbauer B, Lauber J, Lührmann R. 1996. Identification of an RNA-dependent ATPase activity in mammalian U5 snRNPs. Nucleic Acids Res., 24:868-875.

Laggerbauer B, Achsel T, Lührmann R. 1998. The human U5-200kD DEXH-box protein unwinds U4/U6 RNA duplexes *in vitro*. Proc. Natl. Acad. Sci. USA, 95:4188-4192.

Lamb JR, Tugendreich S, Hieter P. 1995. Tetratricopeptide repeat interactions: to TPR or not to TPR? Trends Biochem. Sci., 20:257-259.

- Lamond AI, Konarska MM, Grabowski PJ, Sharp PA. 1988. Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein. Proc. Natl. Acad. Sci. USA, **85**:411-415.
- Lane BG, Ofengand J, Gray MW. 1992. Pseudouridine in the large-subunit (23 S-like) ribosomal RNA. The site of peptidyl transfer in the ribosome? FEBS Lett., **302**:1-4.
- Lane BG, Ofengand J, Gray MW. 1995. Pseudouridine and 2'-O methylated nucleosides. Significance of their selective occurrence in rRNA domains that function in ribosome- catalyzed synthesis of the peptide bonds in proteins. Biochimie, **77**:7-15.
- Lange TS, Borovjagin A, Maxwell ES, Gerbi SA. 1998. Conserved boxes C and D are essential nucleolar localization elements of U14 and U8 snoRNAs. EMBO J., **17**:3176-3187.
- Last RL, Woolford JL, Jr. 1986. Identification and nuclear localization of yeast pre-messenger RNA processing components: RNA2 and RNA3 proteins. J. Cell. Biol., **103**:2103-2112.
- Lauber J, Fabrizio P, Teigelkamp S, Lane WS, Hartmann E, Lührmann R. 1996. The HeLa 200 kDa U5 snRNP-specific protein and its homologue in *Saccharomyces cerevisiae* are members of the DEXH-box protein family of putative RNA helicases. EMBO J., **15**:4001-4015.
- Lauber J, Plessel G, Prehn S, Will CL, Fabrizio P, Groning K, Lane WS, Lührmann R. 1997. The human U4/U6 snRNP contains 60 and 90kD proteins that are structurally homologous to the yeast splicing factors Prp4p and Prp3p. RNA, **3**:926-941.
- Lecoite F, Simos G, Sauer A, Hurt EC, Motorin Y, Grosjean H. 1998. Characterization of yeast protein Deg1 as pseudouridine synthase (Pus3) catalyzing the formation of psi 38 and psi 39 in tRNA anticodon loop. J. Biol. Chem., **273**:1316-1323.
- Lefebvre S, Burglen L, Reboullet S, Clermont O, Burlet 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**:155-165.
- Legrain P, Chapon C, Schwob E, Martin R, Rosbash M, Dujon B. 1991. Cloning of the two essential yeast genes, *PRP6* and *PRP9*, and their rapid mapping, disruption and partial sequencing using a linker insertion strategy. Mol. Gen. Genet., **225**:199-202.
- Legrain P, Chapon C. 1993. Interaction between PRP11 and SPP91 yeast splicing factors and characterization of a PRP9-PRP11-SPP91 complex. Science, **262**:108-110.
- Legrain P, Chapon C, Galisson F. 1993. Interactions between PRP9 and SPP91 splicing factors identify a protein complex required in prespliceosome assembly. Genes Dev., **7**:1390-1399.
- Legrain P, Chanfreau G. 1994. Pre-mRNA splicing: from intron recognition to catalysis. Bull. Inst. pasteur, **92**:153-179.
- Legrain P, Séraphin B, Rosbash M. 1988. Early commitment of yeast pre-mRNA to the spliceosome pathway. Mol. Cell. Biol., **8**:3755-3760.

- Lehmeier T, Raker V, Hermann H, Lührmann R. 1994. cDNA cloning of the Sm proteins D2 and D3 from human small nuclear ribonucleoproteins: evidence for a direct D1-D2 interaction. Proc. Natl. Acad. Sci. USA, 91:12317-12321.
- Lerner MR, Boyle JA, Mount SM, Wolin SL, Steitz JA. 1980. Are snRNPs involved in splicing? Nature, 283:220-224.
- Lesser CF, Guthrie C. 1993. Mutations in U6 snRNA that alter splice site specificity: implications for the active site. Science, 262:1982-1988.
- Lewis JD, Izaurralde E, Jarmolowski A, McGuigan C, Mattaj IW. 1996a. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. Genes Dev., 10:1683-1698.
- Lewis JD, Gorlich D, Mattaj IW. 1996b. A yeast cap binding protein complex (yCBC) acts at an early step in pre-mRNA splicing. Nucleic Acids Res., 24:3332-3336.
- 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., 10:1145-1152.
- Li Z, Brow DA. 1996. A spontaneous duplication in U6 spliceosomal RNA uncouples the early and late functions of the ACAGA element *in vivo*. RNA, 2:879-894.
- Li J, Esberg B, Curran JF, Björk GR. 1997. Three modified nucleosides present in the anticodon stem and loop influence the *in vivo* aa-tRNA selection in a tRNA-dependent manner. J. Mol. Biol., 271:209-221.
- Liang WQ, Fournier MJ. 1995. U14 base-pairs with 18S rRNA: a novel snoRNA interaction required for rRNA processing. Genes Dev., 9:2433-2443.
- Liao XC, Tang J, Rosbash M. 1993. An enhancer screen identifies a gene that encodes the yeast U1 snRNP A protein: implications for snRNP protein function in pre-mRNA splicing. Genes Dev., 7:419-428.
- 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., 162:623-643.
- Libri D, Stutz F, McCarthy T, Rosbash M. 1995. RNA structural patterns and splicing: molecular basis for an RNA-based enhancer. RNA, 1:425-436.
- Limbach PA, Crain PF, McCloskey JA. 1994. Summary: the modified nucleosides of RNA. Nucleic Acids Res., 22:2183-2196.
- Lin RJ, Lustig AJ, Abelson J. 1987. Splicing of yeast nuclear pre-mRNA *in vitro* requires a functional 40S spliceosome and several extrinsic factors. Genes Dev., 1:7-18.
- Lin J, Rossi JJ. 1996. Identification and characterization of yeast mutants that overcome an experimentally introduced block to splicing at the 3' splice site. RNA, 2:835-848.
- Lindsey LA, Garcia-Blanco MA. 1998. Functional conservation of the human homolog of the yeast pre-mRNA splicing factor Prp17p. J. Biol. Chem., 273:32771-32775.

- Lipowsky G, Bischoff FR, Izaurralde E, Kutay U, Schafer S, Gross HJ, Beier H, Gorlich D. 1999. Coordination of tRNA nuclear export with processing of tRNA. RNA, 5:539-549.
- Liu Q, Dreyfuss G. 1996. A novel nuclear structure containing the survival of motor neurons protein. EMBO J., 15:3555-3565.
- Liu Q, Fischer U, Wang F, Dreyfuss G. 1997a. The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. Cell, 90:1013-1021.
- Liu ZR, Laggenbauer B, Lührmann R, Smith CW. 1997b. Crosslinking of the U5 snRNP-specific 116-kDa protein to RNA hairpins that block step 2 of splicing. RNA, 3:1207-1219.
- Lockhart SR, Rymond BC. 1994. Commitment of yeast pre-mRNA to the splicing pathway requires a novel U1 small nuclear ribonucleoprotein polypeptide, Prp39p. Mol. Cell. Biol., 14:3623-3633.
- Lopato S, Mayeda A, Krainer AR, Barta A. Pre-mRNA splicing in plants: characterization of Ser/Arg splicing factors. Proc. Natl. Acad. Sci. USA, 93:3074-3079.
- Lopez AJ. 1998. Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. Annu. Rev. Genet., 32:279-305.
- Lossky M, Anderson GJ, Jackson SP, Beggs J. 1987. Identification of a yeast snRNP protein and detection of snRNP-snRNP interactions. Cell, 51:1019-1026.
- Lowe TM, Eddy SR. 1999. A computational screen for methylation guide snoRNAs in yeast. Science, 283:1168-1171.
- Lu Y, Qian XY, Krug RM. 1994. The influenza virus NS1 protein: a novel inhibitor of pre-mRNA splicing. Genes Dev., 8:1817-1828.
- Lucke S, Klockner T, Palfi Z, Boshart M, Bindereif A. 1997. Trans mRNA splicing in trypanosomes: cloning and analysis of a *PRP8*-homologous gene from *Trypanosoma brucei* provides evidence for a U5-analogous RNP. EMBO J., 16:4433-4440.
- Lührmann R, Kastner B, Bach M. 1990. Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. Biochim. Biophys. Acta, 1087:265-292.
- Lund E, Dahlberg JE. 1992. Cyclic 2',3'-phosphates and nontemplated nucleotides at the 3' end of spliceosomal U6 small nuclear RNA's. Science, 255:327-330.
- Lustig AJ, Lin RJ, Abelson J. 1986. The yeast RNA gene products are essential for mRNA splicing *in vitro*. Cell, 47:953-963.
- Lustig F, Boren T, Claesson C, Simonsson C, Barciszewska M, Lagerkvist U. 1993. The nucleotide in position 32 of the tRNA anticodon loop determines ability of anticodon UCC to discriminate among glycine codons. Proc Natl Acad Sci U S A, 90:3343-3347.
- Lützelberger M, Groß T, Käufer NF. Srp2, an SR protein family member of fission yeast: *in vivo* characterization of its modular domains. Nucleic Acids Res., 27:2618-2626.

Lutz-Freyermuth C, Query CC, Keene JD. 1990. Quantitative determination that one of two potential RNA-binding domains of the A protein component of the U1 small nuclear ribonucleoprotein complex binds with high affinity to stem-loop II of U1 RNA. Proc Natl Acad Sci U S A, 87:6393-6397.

Luukkonen BG, Séraphin B. 1997. The role of branchpoint-3' splice site spacing and interaction between intron terminal nucleotides in 3' splice site selection in *Saccharomyces cerevisiae*. EMBO J., 16:779-792.

Luukkonen BG, Séraphin B. 1998a. Genetic interaction between U6 snRNA and the first intron nucleotide in *Saccharomyces cerevisiae*. RNA, 4:167-180.

Luukkonen BG, Séraphin B. 1998b. A role for U2/U6 helix Ib in 5' splice site selection. RNA, 4:915-927.

Lybarger S, Beickman K, Brown V, Dembla-Rajpal N, Morey K, Seipelt R, Rymond B. C. 1999. Elevated levels of a U4/U6.U5 snRNP-associated protein, Spp381p, rescue a mutant defective in spliceosome maturation. Mol. Cell. Biol. 19:577-84

Lygerou Z, Christophides G, Séraphin B. 1999. A novel genetic screen for snRNP assembly factors in yeast identifies a conserved protein, Sad1p, also required for pre-mRNA splicing. Mol. Cell. Biol., 19:2008-2020.

M

acMillan AM, Query CC, Allerson CR, Chen S, Verdine GL, Sharp PA. 1994. Dynamic association

of proteins with the pre-mRNA branch region. Genes Dev., 8:3008-3020.

Maden BEH, Salim M. 1974. The methylated nucleotide sequences in HeLa cell ribosomal RNA and its precursors. J. Mol. Biol., 88:133-152.

Maden BE. 1990. The numerous modified nucleotides in eukaryotic ribosomal RNA. Prog. Nucleic Acid. Res. Mol. Biol., 39:241-303.

Maden BE, Corbett ME, Heeney PA, Pugh K, Ajuh PM. 1995. Classical and novel approaches to the detection and localization of the numerous modified nucleotides in eukaryotic ribosomal RNA. Biochimie, 77:22-29.

Maden BEH. 1998. Cellular location of modifying enzymes. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 421-440.

Madhani HD, Bordonné R, Guthrie C. 1990. Multiple roles for U6 snRNA in the splicing pathway. Genes Dev., 4:2264-2277.

Madhani HD, Guthrie C. 1992. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. Cell, 71:803-817.

Madhani HD, Guthrie C. 1994a. Randomization-selection analysis of snRNAs *in vivo*: evidence for a tertiary interaction in the spliceosome. Genes Dev., 8:1071-1086.

- Madhani HD, Guthrie C. 1994b. Dynamic RNA-RNA interactions in the spliceosome. Annu. Rev. Genet., 28:1-26.
- Manley JL, Tacke R. 1996. SR proteins and splicing control. Genes Dev., 10:1569-1579.
- Marquet R. 1998. Importance of modified nucleotides in replication of retroviruses, plant pararetroviruses, and retrotransposons. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 517-534.
- Martin R, Sibler AP, Dirheimer G. 1982. The primary structure of three yeast mitochondrial serine tRNA isoacceptors. Biochimie., 64:1073-1079.
- Massenet S, Mougin A, Branlant C. 1998. Posttranscriptional modifications in the U Small Nuclear RNAs. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 201-228.
- 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., 19:2142-2154.
- Matsumoto T, Nishikawa K, Hori H, Ohta T, Miura K, Watanabe K. 1990. Recognition sites of tRNA by a thermostable tRNA(guanosine-2'-)-methyltransferase from *Thermus thermophilus* HB27. J. Biochem., 107:331-338.
- Mattaj IW, De Robertis EM. 1985. Nuclear segregation of U2 snRNA requires binding of specific snRNP proteins. Cell, 40:111-118.
- Mattaj IW. 1986. Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. Cell, 46:905-911.
- Mattaj IW, Habets WJ, Van Venrooij WJ. 1986. Monospecific antibodies reveal details of U2 snRNP structure and interaction between U1 and U2 snRNPs. EMBO J., 5:997-1002.
- Mattaj IW, Englmeier L. 1998. Nucleocytoplasmic transport: the soluble phase. Annu. Rev. Biochem., 67:265-306.
- Maxwell ES, Fournier MJ. 1995. The small nucleolar RNAs. Annu. Rev. Biochem., 64:897-934.
- Mayeda A, Krainer AR. 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell, 68:365-375.
- Mayes AE, Verdone L, Legrain P, Beggs JD. 1999. Characterization of Sm-like proteins in yeast and their association with U6 snRNA. EMBO J., 18:4321-4331.
- Mcafee JG, Huang M, Soltaninassab S, Rech JE, Iyengar S, Lestourgeon WM. 1997. The packaging of pre-mRNA. In: Krainer AR, eds. Eukaryotic mRNA processing. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press. pp

- McLean MR, Rymond BC. 1998. Yeast pre-mRNA splicing requires a pair of U1 snRNP-associated tetratricopeptide repeat proteins. Mol. Cell. Biol., 18:353-360.
- McPheeters DS, Fabrizio P, Abelson J. 1989. *In vitro* reconstitution of functional yeast U2 snRNPs. Genes Dev., 3:2124-2136.
- McPheeters DS, Abelson J. 1992. Mutational analysis of the yeast U2 snRNA suggests a structural similarity to the catalytic core of group I introns. Cell, 71:819-831.
- McPheeters DS. 1996. Interactions of the yeast U6 RNA with the pre-mRNA branch site. RNA, 2:1110-1123.
- Meier UT, Blobel G. 1994. NAP57, a mammalian nucleolar protein with a putative homolog in yeast and *Bacteria*. J. Cell. Biol., 127:1505-1514.
- Meinzel T, Mechulam Y, Blanquet S. 1995. Aminoacyl-tRNA synthetases: occurrence, structure, and function. In: Söll D, RajBhandary UL, eds. tRNA: Structure, biosynthesis and Function. Washington, DC: ASM press. pp 251-292.
- Melton DA, De Robertis EM, Cortese R. 1980. Order and intracellular location of the events involved in the maturation of a spliced tRNA. Nature, 284:143-148.
- Méreau A, Fournier R, Grégoire A, Mouglin A, Fabrizio P, Lührmann 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., 273:552-571.
- Merryman C, Moazed D, Daubresse G, Noller HF. 1999. Nucleotides in 23S rRNA protected by the association of 30S and 50S ribosomal subunits. J. Mol. Biol., 285:107-113.
- Michaud S, Reed R. 1991. An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. Genes Dev., 5:2534-2546.
- Michel F, Jacquier A, Dujon B. 1982. Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. Biochimie, 64:867-881.
- Michel F, Jacquier A. 1987. Long-range intron-exon and intron-intron pairings involved in self-splicing of class II catalytic introns. Cold Spring Harb. Symp. Quant. Biol., 52:201-212.
- Michel F, Umesono K, Ozeki H. 1989. Comparative and functional anatomy of group II catalytic introns. Gene, 82:5-30.
- Michel F, Westhof E. 1990. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. J. Mol. Biol., 216:585-610.
- Michel F, Ferat JL. 1995. Structure and activities of group II introns. Annu. Rev. Biochem., 64:435-461.
- Miller AM. 1984. The yeast MATa1 gene contains two introns. EMBO J., 3:1061-1065.
- Miraglia L, Seiwert S, Igel AH, Ares M, Jr. 1991. Limited functional equivalence of phylogenetic variation in small nuclear RNA: yeast U2 RNA with altered branchpoint complementarity inhibits splicing and produces a dominant lethal phenotype. Proc. Natl. Acad. Sci. USA, 88:7061-7065.

- Mitchell P, Osswald M, Brimacombe R. 1992. Identification of intermolecular RNA cross-links at the subunit interface of the *Escherichia coli* ribosome. Biochemistry, 31:3004-3011.
- Moazed D, Robertson JM, Noller HF. 1988. Interaction of elongation factors EF-G and EF-Tu with a conserved loop in 23S RNA. Nature, 334:362-364.
- Moazed D, Noller HF. 1989. Interaction of tRNA with 23S rRNA in the ribosomal A, P, and E sites. Cell, 57:585-597.
- Montzka KA, Steitz JA. 1988. Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. Proc. Natl. Acad. Sci. USA, 85:8885-8889.
- Moore JM, Query CC, Sharp PA. 1993. Splicing of precursors to messenger RNAs by the spliceosome. In: Gesteland R, Athuis J, eds. The RNA world. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press. pp 303-358.
- Morin A, Auxilien S, Senger B, Tewari R, Grosjean H. 1998. Structural requirements for enzymatic formation of threonylcarbamoyladenine (t⁶A) in tRNA: an *in vivo* study with *Xenopus laevis* oocytes. RNA, 4:24-37.
- Morl M, Schmelzer C. 1990. Integration of group II intron bI1 into a foreign RNA by reversal of the self-splicing reaction *in vitro*. Cell, 60:629-636.
- Morl M, Niemer I, Schmelzer C. 1992. New reactions catalyzed by a group II intron ribozyme with RNA and DNA substrates. Cell, 70:803-810.
- Morrissey JP, Tollervey D. 1995. Birth of the snoRNPs: the evolution of RNase MRP and the eukaryotic pre-rRNA-processing system. Trends Biochem. Sci., 20:78-82.
- Motorin Y, Bec G, Tewari R, Grosjean H. 1997. Transfer RNA recognition by the *Escherichia coli* delta2-isopentenyl-pyrophosphate:tRNA delta2-isopentenyl transferase: dependence on the anticodon arm structure. RNA, 3:721-733.
- Motorin Y, Grosjean H. 1998. Chemical structures and classification of posttranscriptionally modified nucleosides in RNA. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 543-550.
- Motorin Y, Keith G, Simon C, Foiret D, Simos G, Hurt E, Grosjean H. 1998. The yeast tRNA:pseudouridine synthase Pus1p displays a multisite substrate specificity. RNA, 4:856-869.
- Motorin Y, Grosjean H. 1999. Multisite-specific tRNA:m⁵C-methyltransferase (Trm4) in yeast *Saccharomyces cerevisiae*: identification of the gene and substrate specificity of the enzyme. RNA, 5:1105-1118.
- Mougin A, Grégoire A, Banroques J, Ségault V, Fournier R, Brulé F, Chevrier-Miller M, Branlant C. 1996. Secondary structure of the yeast *Saccharomyces cerevisiae* pre-U3A snoRNA and its implication for splicing efficiency. RNA, 2:1079-1093.
- Mount SM, Steitz JA. 1981. Sequence of U1 RNA from *Drosophila melanogaster*: implications for U1 secondary structure and possible involvement in splicing. Nucleic Acids Res., 9:6351-6368.

- Mount SM. 1982. A catalogue of splice junction sequences. Nucleic Acids Res., 10:459-472.
- Mount SM, Pettersson I, Hinterberger M, Karmas A, Steitz JA. 1983. The U1 small nuclear RNA-protein complex selectively binds a 5' splice site *in vitro*. Cell, 33:509-518.
- Mueller SO, Slany RK. 1995. Structural analysis of the interaction of the tRNA modifying enzymes Tgt and QueA with a substrate tRNA. FEBS Lett., 361:259-264.
- Mullenbach GT, Kammen HO, Penhoet EE. 1976. A heterologous system for detecting eukaryotic enzymes which synthesize pseudouridine in transfer ribonucleic acids. J. Biol. Chem., 251:4570-4578.
- Muramatsu T, Nishikawa K, Nemoto F, Kuchino Y, Nishimura S, Miyazawa T, Yokoyama S. 1988. Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. Nature, 336:179-181.
- Myslinski E, Ségault V, Branlant C. 1990. An intron in the genes for U3 small nucleolar RNAs of the yeast *Saccharomyces cerevisiae*. Science, 247:1213-1216.
- Myslinski E, Branlant C. 1991. A phylogenetic study of U4 snRNA reveals the existence of an evolutionarily conserved secondary structure corresponding to 'free' U4 snRNA. Biochimie, 73:17-28.

N

agai K, Oubridge C, Jessen TH, Li J, Evans PR. 1990. Crystal structure of the RNA-binding domain

of the U1 small nuclear ribonucleoprotein A. Nature, 348:515-520.

- Nagai K. 1996. RNA-protein complexes. Curr. Opin. Struct. Biol., 6:53-61.
- Narayanan A, Lukowiak A, BE Jd, Dragon F, Kiss T, Terns RM, Terns MP. 1999. Nucleolar localization signals of Box H/ACA small nucleolar RNAs. EMBO J., 18:5120-5130.
- Neer EJ, Schmidt CJ, Nambudripad R, Smith TF. 1994. The ancient regulatory-protein family of WD-repeat proteins. Nature, 371:297-300.
- Nelissen RL, Heinrichs V, Habets WJ, Simons F, Lührmann R, van Venrooij WJ. 1991. Zinc finger-like structure in U1-specific protein C is essential for specific binding to U1 snRNP. Nucleic Acids Res., 19:449-454.
- Neubauer G, Gottschalk A, Fabrizio P, Séraphin B, Lührmann R, Mann M. 1997. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. Proc. Natl. Acad. Sci. USA, 94:385-390.
- Newman AJ, Lin RJ, Cheng SC, Abelson J. 1985. Molecular consequences of specific intron mutations on yeast mRNA splicing *in vivo* and *in vitro*. Cell, 42:335-344.

- Newman A, Norman C. 1991. Mutations in yeast U5 snRNA alter the specificity of 5' splice-site cleavage. Cell, 65:115-123.
- Newman AJ, Norman C. 1992. U5 snRNA interacts with exon sequences at 5' and 3' splice sites. Cell, 68:743-754.
- Newman AJ, Teigelkamp S, Beggs JD. 1995. snRNA interactions at 5' and 3' splice sites monitored by photoactivated crosslinking in yeast spliceosomes. RNA, 1:968-980.
- Newman AJ. 1997. The role of U5 snRNP in pre-mRNA splicing. EMBO J., 16:5797-5800.
- Ni J, Tien AL, Fournier MJ. 1997. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. Cell, 89:565-573.
- Nicoloso M, Qu LH, Michot B, Bachellerie JP. 1996. Intron-encoded, antisense small nucleolar RNAs: the characterization of nine novel species points to their direct role as guides for the 2'-O- ribose methylation of rRNAs. J. Mol. Biol., 260:178-195.
- Nilsen TW. 1998. RNA-RNA interactions in nuclear pre-mRNA splicing. In: Simos RW, Grunberg-Manago M, eds. RNA structure and fonction. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press. pp 279-308.
- Nishikura K, De Robertis EM. 1981. RNA processing in microinjected *Xenopus* oocytes. Sequential addition of base modifications in the spliced transfer RNA. J. Mol. Biol., 145:405-420.
- Niu L, Ofengand J. 1999. Cloning and characterization of the 23S RNA pseudouridine 2633 synthase from *Bacillus subtilis*. Biochemistry, 38:629-635.
- Niwa M, Berget SM. 1991. Polyadenylation precedes splicing *in vitro*. Gene Expr., 1:5-14.
- Noble SM, Guthrie C. 1996. Transcriptional pulse-chase analysis reveals a role for a novel snRNP- associated protein in the manufacture of spliceosomal snRNPs. EMBO J., 15:4368-4379.
- Noguchi S, Yamaizumi Z, Ohgi T, Goto T, Nishimura Y, Hirota Y, Nishimura S. 1978. Isolation of Q nucleoside precursor present in tRNA of an *E. coli* mutant and its characterization as 7-(cyano)-7-deazaguanosine. Nucleic Acids Res., 5:4215-4223.
- Noller HF. 1998. Ribosomal RNA. In: Simos RW, Grunberg-Manago M, eds. RNA structure and fonction. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press. pp 253-278.
- Noon KR, Bruenger E, McCloskey JA. 1998. Posttranscriptional modifications in 16S and 23S rRNAs of the archaeal hyperthermophile *Sulfolobus solfataricus*. J. Bacteriol., 180:2883-2888.
- Nurse K, Wrzesinski J, Bakin A, Lane BG, Ofengand J. 1995. Purification, cloning, and properties of the tRNA psi 55 synthase from *Escherichia coli*. RNA, 1:102-112.

O'

Connor M, Dahlberg AE. 1995. The involvement of two distinct regions of 23 S ribosomal RNA in

tRNA selection. J. Mol. Biol., 254:838-847.

O'Day CL, Dalbadie-McFarland G, Abelson J. 1996. The *Saccharomyces cerevisiae* Prp5 protein has RNA-dependent ATPase activity with specificity for U2 small nuclear RNA. J. Biol. Chem., 271:33261-33267.

O'Keefe RT, Norman C, Newman AJ. 1996. The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. Cell, 86:679-689.

O'Keefe RT, Newman AJ. 1998. Functional analysis of the U5 snRNA loop 1 in the second catalytic step of yeast pre-mRNA splicing. EMBO J., 17:565-574.

O'Reilly M, Teichmann SA, Rhodes D. 1999. Telomerases. Curr. Opin. Struct. Biol., 9:56-65.

Ochs RL, Lischwe MA, Spohn WH, Busch H. 1985. Fibrillarin: a new protein of the nucleolus identified by autoimmune sera. Biol. Cell., 54:123-133.

Ofengand J, Bakin A. 1997. Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, *archaebacteria*, mitochondria and chloroplasts. J. Mol. Biol., 266:246-268.

Ofengand J, Fournier MJ. 1998. The pseudouridine residues of rRNA: number, location, biosynthesis and function. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 229-254.

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. USA, 84:5187-5191.

Ordway RW, Pallanck L, Ganetzky B. 1994. A TPR domain in the SNAP secretory proteins. Trends Biochem. Sci., 19:530-531.

Oubridge C, Ito N, Evans PR, Teo CH, Nagai K. 1994. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. Nature, 372:432-438.

P

- adgett RA, Konarska MM, Grabowski PJ, Hardy SF, Sharp PA. 1984. Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. Science, 225:898-903.
- Palmer DT, Blum PH, Artz SW. 1983. Effects of the *hisT* mutation of *Salmonella typhimurium* on translation elongation rate. J. Bacteriol., 153:357-363.
- Pan ZQ, Prives C. 1989. U2 snRNA sequences that bind U2-specific proteins are dispensable for the function of U2 snRNP in splicing. Genes Dev., 3:1887-1898.
- Pannone BK, Xue D, Wolin SL. 1998. A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. EMBO J., 17:7442-7453.
- Parker R, Patterson B. 1987. Architecture of fungal introns: implications for spliceosome assembly. In: Inouye M, Dudock B, eds. Molecular biology of RNA: new perspectives. New York: Academic Press. pp 133-149.
- Parker R, Siliciano PG, Guthrie C. 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. Cell, 49:229-239.
- Parker R, Siliciano PG. 1993. Evidence for an essential non-Watson-Crick interaction between the first and last nucleotides of a nuclear pre-mRNA intron. Nature, 361:660-662.
- Pascolo E, Séraphin B. 1997. The branchpoint residue is recognized during commitment complex formation before being bulged out of the U2 snRNA-pre-mRNA duplex. Mol. Cell. Biol., 17:3469-3476.
- Patterson B, Guthrie C. 1987. An essential yeast snRNA with a U5-like domain is required for splicing *in vivo*. Cell, 49:613-624.
- Patterson B, Guthrie C. 1991. A U-rich tract enhances usage of an alternative 3' splice site in yeast. Cell, 64:181-187.
- Patton JR. 1991. Pseudouridine modification of U5 RNA in ribonucleoprotein particles assembled *in vitro*. Mol. Cell. Biol., 11:5998-6006.
- Patton JG, Mayer SA, Tempst P, Nadal-Ginard B. 1991. Characterization and molecular cloning of polypyrimidine tract-binding protein: a component of a complex necessary for pre-mRNA splicing. Genes Dev., 5:1237-1251.
- Patton JG, Porro EB, Galceran J, Tempst P, Nadal-Ginard B. 1993. Cloning and characterization of PSF, a novel pre-mRNA splicing factor. Genes Dev., 7:393-406.
- Patton JR. 1994a. Pseudouridine formation in small nuclear RNAs. Biochimie, 76:1129-1132.
- Patton JR. 1994b. Formation of pseudouridine in U5 small nuclear RNA. Biochemistry, 33:10423-10427.
- Peattie DA. 1979. Direct chemical method for sequencing RNA. Proc. Natl. Acad. Sci. USA, 76:1760-1764.

- Peebles CL, Gegenheimer P, Abelson J. 1983. Precise excision of intervening sequences from precursor tRNAs by a membrane-associated yeast endonuclease. Cell, 32:525-536.
- Peebles CL, Perlman PS, Mecklenburg KL, Petrillo ML, Tabor JH, Jarrell KA, Cheng HL. 1986. A self-splicing RNA excises an intron lariat. Cell, 44:213-223.
- 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, 95:615-624.
- Perret V, Garcia A, Puglisi J, Grosjean H, Ebel JP, Florentz C, Giegé R. 1990a. Conformation in solution of yeast tRNA(Asp) transcripts deprived of modified nucleotides. Biochimie, 72:735-743.
- Perret V, Garcia A, Grosjean H, Ebel JP, Florentz C, Giegé R. 1990b. Relaxation of a transfer RNA specificity by removal of modified nucleotides. Nature, 344:787-789.
- Perrot JL, Benoit F, Ségault D, Jaubert J, Guyotat D, Claudy A. 1992. Pyoderma gangrenosum aggravated by GM-CSF administration. Ann. Dermatol. Venereol., 119:846-848.
- Persson BC, Gustafsson C, Berg DE, Björk GR. 1992. The gene for a tRNA modifying enzyme, m5U54-methyltransferase, is essential for viability in *Escherichia coli*. Proc. Natl. Acad. Sci. USA, 89:3995-3998.
- Phillips B, Billin AN, Cadwell C, Buchholz R, Erickson C, Merriam JR, Carbon J, Poole SJ. 1998. The Nop60B gene of *Drosophila* encodes an essential nucleolar protein that functions in yeast. Mol. Gen. Genet., 260:20-29.
- Pikielny CW, Rymond BC, Rosbash M. 1986. Electrophoresis of ribonucleoproteins reveals an ordered assembly pathway of yeast splicing complexes. Nature, 324:341-345.
- Pikielny CW, Bindereif A, Green MR. 1989. *In vitro* reconstitution of snRNPs: a reconstituted U4/U6 snRNP participates in splicing complex formation. Genes Dev., 3:479-487.
- Plessel G, Fischer U, Lührmann R. 1994. m3G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) *in vitro*: evidence that the U1 small nuclear RNA-(guanosine-N2)-methyltransferase is a non-snRNP cytoplasmic protein that requires a binding site on the Sm core domain. Mol. Cell. Biol., 14:4160-4172.
- Potashkin J, Frendewey D. 1990. A mutation in a single gene of *Schizosaccharomyces pombe* affects the expression of several snRNAs and causes defects in RNA processing. EMBO J., 9:525-534.
- Przykorska A. 1995. Influence of modified nucleosides on tRNA structure as probed by two plant nucleases. Biochimie, 77:109-112.

Qiu Y, Nemeroff M, Krug RM. 1995. The influenza virus NS1 protein binds to a specific region in

human U6 snRNA and inhibits U6-U2 and U6-U4 snRNA interactions during splicing. RNA, 1:304-316.

Query CC, Moore MJ, Sharp PA. 1994. Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. Genes Dev., 8:587-597.

Query CC, Strobel SA, Sharp PA. 1996. Three recognition events at the branch-site adenine. EMBO J., 15:1392-1402.

Quigley GJ, Rich A. 1976. Structural domains of transfer RNA molecules. Science, 194:796-806.

Raghunathan PL, Guthrie C. 1998. A spliceosomal recycling factor that reanneals U4 and U6 small

nuclear ribonucleoprotein particles. Science, 279:857-860.

Rain JC, Tartakoff AM, Krämer A, Legrain P. 1996. Essential domains of the PRP21 splicing factor are implicated in the binding to PRP9 and PRP11 proteins and are conserved through evolution. RNA, 2:535-550.

Rain JC, Legrain P. 1997. *In vivo* commitment to splicing in yeast involves the nucleotide upstream from the branch site conserved sequence and the Mud2 protein. EMBO J., 16:1759-1771.

Raker VA, Plessel G, Lührmann R. 1996. The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle *in vitro*. EMBO J., 15:2256-2269.

Ramamurthy V, Swann SL, Paulson JL, Spedaliere CJ, Mueller EG. 1999. Critical aspartic acid residues in pseudouridine synthases. J. Biol. Chem., 274:22225-22230.

Raychaudhuri S, Conrad J, Hall BG, Ofengand J. 1998. A pseudouridine synthase required for the formation of two universally conserved pseudouridines in ribosomal RNA is essential for normal growth of *Escherichia coli*. RNA, 4:1407-1417.

Raychaudhuri S, Niu L, Conrad J, Lane BG, Ofengand J. 1999. Functional effect of deletion and mutation of the *Escherichia coli* ribosomal RNA and tRNA pseudouridine synthase RluA. J. Biol. Chem., 274:18880-18886.

Reddy R, Henning D, Epstein P, Busch H. 1981. Primary and secondary structure of U2 snRNA. Nucleic Acids Res., 9:5645-5658.

- 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., 260:5715-5719.
- Reddy R, Henning D, Das G, Harless M, Wright D. 1987. The capped U6 small nuclear RNA is transcribed by RNA polymerase III. J. Biol. Chem., 262:75-81.
- Reddy R. 1988. Compilation of small RNA sequences. Nucleic Acids Res., 16:r71-85.
- Reddy R, Singh R, Shimba S. 1992. Methylated cap structures in eukaryotic RNAs: structure, synthesis and functions. Pharmacol. Ther., 54:249-267.
- Redlak M, Andraos-Selim C, Giegé R, Florentz C, Holmes WM. 1997. Interaction of tRNA with tRNA (guanosine-1)methyltransferase: binding specificity determinants involve the dinucleotide G36pG37 and tertiary structure. Biochemistry, 36:8699-8709.
- Reed R, Maniatis T. 1985. Intron sequences involved in lariat formation during pre-mRNA splicing. Cell, 41:95-105.
- Reed R. 1989. The organization of 3' splice-site sequences in mammalian introns. Genes Dev., 3:2113-2123.
- Reed R. 1990. Protein composition of mammalian spliceosomes assembled *in vitro*. Proc. Natl. Acad. Sci. USA, 87:8031-8035.
- Reed R, Palandjian L. 1997. Spliceosome assembly. In: Krainer AR, eds. Eukaryotic mRNA processing. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press. pp 103-129.
- Reich CI, VanHoy RW, Porter GL, Wise JA. 1992. Mutations at the 3' splice site can be suppressed by compensatory base changes in U1 snRNA in fission yeast. Cell, 69:1159-1169.
- Reinhold-Hurek B, Shub DA. 1992. Self-splicing introns in tRNA genes of widely divergent *bacteria*. Nature, 357:173-176.
- Reyes JL, Kois P, Konforti BB, Konarska MM. 1996. The canonical GU dinucleotide at the 5' splice site is recognized by p220 of the U5 snRNP within the spliceosome. RNA, 2:213-225.
- Rigaut G, Shevchenko A, Rutz B, Wilm M, Séraphin B. 1999. A generic protein purification method for protein complex characterization and proteom exploration. Nature Biotech., 17:1030-1032.
- Rinke J, Appel B, Blocker H, Frank R, Lührmann R. 1984. The 5'-terminal sequence of U1 RNA complementary to the consensus 5' splice site of hnRNA is single-stranded in intact U1 snRNP particles. Nucleic Acids Res., 12:4111-4126.
- Rinke J, Appel B, Digweed M, Lührmann R. 1985. Localization of a base-paired interaction between small nuclear RNAs U4 and U6 in intact U4/U6 ribonucleoprotein particles by psoralen cross-linking. J. Mol. Biol., 185:721-731.
- Robertson MP, Miller SL. 1995. Prebiotic synthesis of 5-substituted uracils: a bridge between the RNA world and the DNA-protein world. Science, 268:702-705.

- Rochaix JD, Malnoe P. 1978. Anatomy of the chloroplast ribosomal DNA of *Chlamydomonas reinhardtii*. Cell, 15:661-670.
- Rogers J, Wall R. 1980. A mechanism for RNA splicing. Proc. Natl. Acad. Sci. USA, 77:1877-1879.
- Romby P, Carbon P, Westhof E, Ehresmann C, Ebel JP, Ehresmann B, Giegé R. 1987. Importance of conserved residues for the conformation of the T-loop in tRNAs. J. Biomol. Struct. Dyn., 5:669-687.
- Romier C, Ficner R, Suck D. 1998. Structural basis of base exchange by tRNA-guanine transglycosylases. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 169-182.
- Roscigno RF, Garcia-Blanco MA. 1995. SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. RNA, 1:692-706.
- Rotondo G, Gillespie M, Frendewey D. 1995. Rescue of the fission yeast snRNA synthesis mutant *snm1* by overexpression of the double-strand-specific Pac1 ribonuclease. Mol. Gen. Genet., 247:698-708.
- Rould MA, Perona JJ, Steitz TA. 1991. Structural basis of anticodon loop recognition by glutamyl-tRNA synthetase. Nature, 352:213-218.
- Roy J, Zheng B, Rymond BC, Woolford JL, Jr. 1995. Structurally related but functionally distinct yeast Sm D core small nuclear ribonucleoprotein particle proteins. Mol. Cell. Biol., 15:445-455.
- Ruby SW, Chang TH, Abelson J. 1993. Four yeast spliceosomal proteins (PRP5, PRP9, PRP11, and PRP21) interact to promote U2 snRNP binding to pre-mRNA. Genes Dev., 7:1909-1925.
- Rudinger J, Florentz C, Dreher T, Giegé R. Efficient mischarging of a viral tRNA-like structure and aminoacylation of a minihelix containing a pseudoknot: histidinylation of turnip yellow mosaic virus RNA. Nucleic Acids Res., 20:1865-1870.
- Rueter SM, Emeson RB. 1998. Adenosine-to-inosine conversion in mRNA. In: Grosjean H, Benne R, eds. The modification and editing of RNA. Washington, DC: ASM Press. pp 343-362.
- Ruskin B, Krainer AR, Maniatis T, Green MR. 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing *in vitro*. Cell, 38:317-331.
- Ruskin B, Green MR. 1985. Specific and stable intron-factor interactions are established early during *in vitro* pre-mRNA splicing. Cell, 43:131-142.
- Rutz B, Séraphin B. 1999. Transient interaction of BBP/ScSF1 and Mud2 with the splicing machinery affects the kinetics of spliceosome assembly. RNA, 5:819-831.
- Rymond BC, Rosbash M. 1985. Cleavage of 5' splice site and lariat formation are independent of 3' splice site in yeast mRNA splicing. Nature, 317:735-737.
- Rymond BC, Rosbash M. 1992. Yeast pre-mRNA splicing. In: Jones EW, Pringle JR, Broach JR, eds. The molecular and cellular biology of the yeast *Saccharomyces*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. pp 143-192.

Rymond BC. 1993. Convergent transcripts of the yeast *PRP38-SMD1* locus encode two essential splicing factors, including the D1 core polypeptide of small nuclear ribonucleoprotein particles. Proc. Natl. Acad. Sci. USA, 90:848-852.

Rymond BC, Rokeach LA, Hoch SO. 1993. Human snRNP polypeptide D1 promotes pre-mRNA splicing in yeast and defines nonessential yeast *Smd1p* sequences. Nucleic Acids Res., 21:3501-3505.

Salgado-Garrido J, Bragado-Nilsson E, Kandels-Lewis S, Séraphin B. 1999. Sm and Sm-like proteins

assemble in two related complexes of deep evolutionary origin. EMBO J., 18:3451-3462.

Samarsky DA, Fournier MJ. 1999. A comprehensive database for the small nucleolar RNAs from *Saccharomyces cerevisiae*. Nucleic Acids Res., 27:161-164.

Sambrook J, Fritsch FF, Maniatis T. 1989. Molecular cloning. A laboratory manual. Second edition. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sampson JR, Uhlenbeck OC. 1988. Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed *in vitro*. Proc. Natl. Acad. Sci. USA, 85:1033-1037.

Samuelsson T, Olsson M. 1990. Transfer RNA pseudouridine synthases in *Saccharomyces cerevisiae*. J. Biol. Chem., 265:8782-8787.

Santi DV, Hardy LW. 1987. Catalytic mechanism and inhibition of tRNA (uracil-5-)methyltransferase: evidence for covalent catalysis. Biochemistry, 26:8599-8606.

Sawa H, Abelson J. 1992. Evidence for a base-pairing interaction between U6 small nuclear RNA and 5' splice site during the splicing reaction in yeast. Proc. Natl. Acad. Sci. USA, 89:11269-11273.

Sawa H, Shimura Y. 1992. Association of U6 snRNA with the 5'-splice site region of pre-mRNA in the spliceosome. Genes Dev., 6:244-254.

Schaefer KP, Altman S, Soll D. 1973. Nucleotide modification *in vitro* of the precursor of transfer RNA of *Escherichia coli*. Proc. Natl. Acad. Sci. USA, 70:3626-3630.

Schatz PJ, Pillus L, Grisafi P, Solomon F, Botstein D. 1986. Two functional alpha-tubulin genes of the yeast *Saccharomyces cerevisiae* encode divergent proteins. Mol. Cell. Biol., 6:3711-3721.

Scherly D, Boelens W, van Venrooij WJ, Dathan NA, Hamm J, Mattaj IW. 1989. Identification of the RNA binding segment of human UI A protein and definition of its binding site on U1 snRNA. EMBO J., 8:4163-4170.

- Scherly D, Boelens W, Dathan NA, Van Venrooij WJ, Mattaj IW. 1990. Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B" and their cognate RNAs. Nature, 345:502-506.
- Scherly D, Kambach C, Boelens W, van Venrooij WJ, Mattaj IW. 1991. Conserved amino acid residues within and outside of the N-terminal ribonucleoprotein motif of U1A small nuclear ribonucleoprotein involved in U1 RNA binding. J. Mol. Biol., 219:577-584.
- 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., 8:4015-4024.
- Schwer B, Guthrie C. 1991. PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome. Nature, 349:494-499.
- Schwer B, Guthrie C. 1992. A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. EMBO J., 11:5033-5039.
- Schwer B, Gross CH. 1998. Prp22, a DEXH-box RNA helicase, plays two distinct roles in yeast pre- mRNA splicing. EMBO J., 17:2086-2094.
- Ségault V, Mougin A, Grégoire A, Banroques J, Branlant C. 1992. An experimental study of *Saccharomyces cerevisiae* U3 snRNA conformation in solution. Nucleic Acids Res., 20:3443-3451.
- Ségault V, Will CL, Sproat BS, Lührmann R. 1995. *In vitro* reconstitution of mammalian U2 and U5 snRNPs active in splicing: Sm proteins are functionally interchangeable and are essential for the formation of functional U2 and U5 snRNPs. EMBO J., 14:4010-4021.
- Ségault V, Will CL, Polycarpou-Schwarz M, Mattaj IW, Branlant C, Lührmann R. 1999. Conserved loop I of U5 small nuclear RNA is dispensable for both catalytic steps of pre-mRNA splicing in HeLa nuclear extracts. Mol. Cell. Biol., 19:2782-2790.
- 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., 27:587-595.
- Seiwert SD, Stuart K. 1994. RNA editing: transfer of genetic information from gRNA to precursor mRNA *in vitro*. Science, 266:114-117.
- Senger B, Auxilien S, Englisch U, Cramer F, Fasiolo F. 1997. The modified wobble base inosine in yeast tRNA^{Ile} is a positive determinant for aminoacylation by isoleucyl-tRNA synthetase. Biochemistry, 36:8269-8275.
- Séraphin B, Kretzner L, Rosbash M. 1988. A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. EMBO J., 7:2533-2538.
- Séraphin B, Rosbash M. 1989a. Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. Cell, 59:349-358.

- Séraphin B, Rosbash M. 1989b. Mutational analysis of the interactions between U1 small nuclear RNA and pre-mRNA of yeast. Gene, **82**:145-151.
- Séraphin B, Rosbash M. 1990. Exon mutations uncouple 5' splice site selection from U1 snRNA pairing. Cell, **63**:619-629.
- Séraphin B, Kandels-Lewis S. 1993. 3' splice site recognition in *S. cerevisiae* does not require base pairing with U1 snRNA. Cell, **73**:803-812.
- Séraphin 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**:2089-2098.
- Shannon KW, Guthrie C. 1991. Suppressors of a U4 snRNA mutation define a novel U6 snRNP protein with RNA-binding motifs. Genes Dev., **5**:773-785.
- Shapiro MB, Senapathy P. 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res., **15**:7155-7174.
- Sharp PA, Burge CB. 1997. Classification of introns: U2-type or U12-type. Cell, **91**:875-879.
- Shimba S, Bokar JA, Rottman F, Reddy R. Accurate and efficient N-6-adenosine methylation in spliceosomal U6 small nuclear RNA by HeLa cell extract *in vitro*. Nucleic Acids Res., **23**:2421-2426.
- Shukla GC, Padgett RA. 1999. Conservation of functional features of U6atac and U12 snRNAs between vertebrates and higher plants. RNA, **5**:525-538.
- Shuster EO, Guthrie C. 1990. Human U2 snRNA can function in pre-mRNA splicing in yeast. Nature, **345**:270-273.
- Siliciano PG, Guthrie C. 1988. 5' splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. Genes Dev., **2**:1258-1267.
- Siliciano PG, Kivens WJ, Guthrie C. 1991. More than half of yeast U1 snRNA is dispensable for growth. Nucleic Acids Res., **19**:6367-6372.
- Simos G, Tekotte H, Grosjean H, Segref A, Sharma K, Tollervey D, Hurt EC. 1996. Nuclear pore proteins are involved in the biogenesis of functional tRNA. EMBO J., **15**:2270-2284.
- Singer CE, Smith GR, Cortese R, Ames BN. 1972. Mutant tRNA His ineffective in repression and lacking two pseudouridine modifications. Nat. New. Biol., **238**:72-74.
- Singh R, Reddy R. 1989. Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. Proc. Natl. Acad. Sci. USA, **86**:8280-8283.
- Singh R, Valcarcel J, Green MR. 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science, **268**:1173-1176.
- Sirum-Connolly K, Mason TL. 1993. Functional requirement of a site-specific ribose methylation in ribosomal RNA. Science, **262**:1886-1889.

- Smith CW, Porro EB, Patton JG, Nadal-Ginard B. 1989. Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. Nature, 342:243-247.
- Smith V, Barrell BG. 1991. Cloning of a yeast U1 snRNP 70K protein homologue: functional conservation of an RNA-binding domain between humans and yeast. EMBO J., 10:2627-2634.
- Smith CW, Chu TT, Nadal-Ginard B. 1993. Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns. Mol. Cell. Biol., 13:4939-4952.
- Smith DR, Doucette-Stamm LA, Deloughery C, Lee H, Dubois J, Aldredge T, Bashirzadeh R, Blakely D, Cook R, Gilbert K, Harrison D, Hoang L, Keagle P, Lumm W, Pothier B, Qiu D, Spadafora R, Vicaire R, Wang Y, Wierzbowski J, Gibson R, Jiwani N, Caruso A, Bush D, Reeve JN, *et al.* 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* deltaH: functional analysis and comparative genomics. J. Bacteriol., 179:7135-7155.
- Solymosy F, Pollak T. 1993. Uridylate-Rich Small Nuclear RNAs (UsnRNAs), their genes and pseudogenes, and UsnRNPS in plants: structure and function. A comparative approach. Cr. Rev. in Plant Sci., 12:275-369.
- Sontheimer EJ, Steitz JA. 1992. Three novel functional variants of human U5 small nuclear RNA. Mol. Cell. Biol., 12:734-746.
- Sontheimer EJ, Steitz JA. 1993. The U5 and U6 small nuclear RNAs as active site components of the spliceosome. Science, 262:1989-1996.
- Sontheimer EJ, Sun S, Piccirilli JA. 1997. Metal ion catalysis during splicing of pre-messenger RNA. Nature, 388:801-805.
- Staknis D, Reed R. 1994a. Direct interactions between pre-mRNA and six U2 small nuclear ribonucleoproteins during spliceosome assembly. Mol. Cell. Biol., 14:2994-3005.
- Staknis D, Reed R. 1994b. SR proteins promote the first specific recognition of pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. Mol. Cell. Biol., 14:7670-7682.
- Staley JP, Guthrie C. 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell, 92:315-326.
- Staley JP, Guthrie C. 1999. An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. Mol. Cell., 3:55-64.
- Steitz TA, Steitz JA. 1993. A general two-metal-ion mechanism for catalytic RNA. Proc. Natl. Acad. Sci. USA, 90:6498-6502.
- Sterner DA, Carlo T, Berget SM. 1996. Architectural limits on split genes. Proc. Natl. Acad. Sci. USA, 93:15081-15085.
- 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. USA, 96:7226-7231.

- Strauss EJ, Guthrie C. 1991. A cold-sensitive mRNA splicing mutant is a member of the RNA helicase gene family. Genes Dev., 5:629-641.
- Strauss EJ, Guthrie C. 1994. PRP28, a 'DEAD-box' protein, is required for the first step of mRNA splicing *in vitro*. Nucleic Acids Res., 22:3187-3193.
- Strobel MC, Abelson J. 1986. Effect of intron mutations on processing and function of *Saccharomyces cerevisiae* SUP53 tRNA *in vitro* and *in vivo*. Mol. Cell. Biol., 6:2663-2673.
- Suddath FL, Quigley GJ, McPherson A, Sneden D, Kim JJ, Kim SH, Rich A. 1974. Three-dimensional structure of yeast phenylalanine transfer RNA at 3.0 angstroms resolution. Nature, 248:20-24.
- Sun JS, Manley JL. 1995. A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. Genes Dev., 9:843-854.
- Sun JS, Manley JL. 1997. The human U6 snRNA intramolecular helix: structural constraints and lack of sequence specificity. RNA, 3:514-526.
- Sylvers LA, Rogers KC, Shimizu M, Ohtsuka E, Soll D. 1993. A 2-thiouridine derivative in tRNA^{Glu} is a positive determinant for aminoacylation by *Escherichia coli* glutamyl-tRNA synthetase. Biochemistry, 32:3836-3841.
- Szkukalek A, Myslinski E, Mougin A, Lührmann R, Branlant C. 1995. Phylogenetic conservation of modified nucleotides in the terminal loop 1 of the spliceosomal U5 snRNA. Biochimie, 77:16-21.
- Szkukalek A, Mougin A, Grégoire A, Solymosy F, Branlant C. 1996. A unique U5-->A substitution in the *Physarum polycephalum* U1 snRNA: evidence at the RNA and gene levels. Biochimie, 78:425-435.
- Szweykowska-Kulinska Z, Beier H. 1992. Sequence and structure requirements for the biosynthesis of pseudouridine (psi 35) in plant pre-tRNA(Tyr). EMBO J., 11:1907-1912.
- Szweykowska-Kulinska Z, Senger B, Keith G, Fasiolo F, Grosjean H. 1994. Intron-dependent formation of pseudouridines in the anticodon of *Saccharomyces cerevisiae* minor tRNA(Ile). EMBO J., 13:4636-4644.
- Szweykowska-Kulinska Z, Krajewski J, Wypijewski K. 1995. Mutations of *Arabidopsis thaliana* pre-tRNA(Tyr) affecting pseudouridylation of U35. Biochim. Biophys. Acta, 1264:87-92.

T

- abor S, Richardson CC. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA, 84:4767-4771.
- Tacke R, Manley JL. 1999. Determinants of SR protein specificity. Curr. Opin. Cell. Biol., 11:358-362.

- Takahashi Y, Urushiyama S, Tani T, Ohshima Y. 1993. An mRNA-type intron is present in the *Rhodotorula hasegawae* U2 small nuclear RNA gene. Mol. Cell. Biol., 13:5613-5619.
- Talerico M, Berget SM. 1994. Intron definition in splicing of small *Drosophila* introns. Mol. Cell. Biol., 14:3434-3445.
- Tang J, Rosbash M. 1996. Characterization of yeast U1 snRNP A protein: identification of the N- terminal RNA binding domain (RBD) binding site and evidence that the C- terminal RBD functions in splicing. RNA, 2:1058-1070.
- Tang J, Abovich N, Rosbash M. 1996. Identification and characterization of a yeast gene encoding the U2 small nuclear ribonucleoprotein particle B" protein. Mol. Cell. Biol., 16:2787-2795.
- Tang J, Abovich N, Fleming ML, Séraphin B, Rosbash M. 1997. Identification and characterization of a yeast homolog of U1 snRNP- specific protein C. EMBO J., 16:4082-4091.
- Tani T, Ohshima Y. 1989. The gene for the U6 small nuclear RNA in fission yeast has an intron. Nature, 337:87-90.
- Tarn WY, Lee KR, Cheng SC. 1993a. The yeast PRP19 protein is not tightly associated with small nuclear RNAs, but appears to associate with the spliceosome after binding of U2 to the pre-mRNA and prior to formation of the functional spliceosome. Mol. Cell. Biol., 13:1883-1891.
- Tarn WY, Lee KR, Cheng SC. 1993b. Yeast precursor mRNA processing protein PRP19 associates with the spliceosome concomitant with or just after dissociation of U4 small nuclear RNA. Proc. Natl. Acad. Sci. USA, 90:10821-10825.
- Tarn WY, Steitz JA. 1994. SR proteins can compensate for the loss of U1 snRNP functions *in vitro*. Genes Dev., 8:2704-2717.
- Tarn WY, Steitz JA. 1995. Modulation of 5' splice site choice in pre-messenger RNA by two distinct steps. Proc. Natl. Acad. Sci. USA, 92:2504-2508.
- Tarn WY, Yario TA, Steitz JA. 1995. U12 snRNA in vertebrates: evolutionary conservation of 5' sequences implicated in splicing of pre-mRNAs containing a minor class of introns. RNA, 1:644-656.
- Tarn WY, Steitz JA. 1996a. A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron *in vitro*. Cell, 84:801-811.
- Tarn WY, Steitz JA. 1996b. Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. Science, 273:1824-1832.
- Tazi J, Kornstadt U, Rossi F, Jeanteur P, Cathala G, Brunel C, Lührmann R. 1993. Thiophosphorylation of U1-70K protein inhibits pre-mRNA splicing. Nature, 363:283-286.
- Teigelkamp S, Whittaker E, Beggs JD. 1995a. Interaction of the yeast splicing factor PRP8 with substrate RNA during both steps of splicing. Nucleic Acids Res., 23:320-326.

- Teigelkamp S, Newman AJ, Beggs JD. 1995b. Extensive interactions of PRP8 protein with the 5' and 3' splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. EMBO J., 14:2602-2612.
- Teigelkamp S, Mundt C, Achsel T, Will CL, Lührmann R. 1997. The human U5 snRNP-specific 100-kD protein is an RS domain-containing, putative RNA helicase with significant homology to the yeast splicing factor Prp28p. RNA, 3:1313-1326.
- Teigelkamp S, Achsel T, Mundt C, Gothel SF, Cronshagen U, Lane WS, Marahiel M, Lührmann 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, 4:127-141.
- Thompson-Jager S, Domdey H. 1987. Yeast pre-mRNA splicing requires a minimum distance between the 5' splice site and the internal branch acceptor site. Mol. Cell. Biol., 7:4010-4016.
- Thorsness PE, Fox TD. 1990. Escape of DNA from mitochondria to the nucleus in *Saccharomyces cerevisiae*. Nature, 346:376-379.
- 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., 10:573-583.
- 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, 72:443-457.
- Tollervey D. 1996. Trans-acting factors in ribosome synthesis. Exp. Cell. Res., 229:226-232.
- Tollervey D, Kiss T. 1997. Function and synthesis of small nucleolar RNAs. Curr. Opin. Cell. Biol., 9:337-342.
- Tronchere H, Wang J, Fu XD. 1997. A protein related to splicing factor U2AF35 that interacts with U2AF65 and SR proteins in splicing of pre-mRNA. Nature, 388:397-400.
- Tsui HC, Arps PJ, Connolly DM, Winkler ME. 1991. Absence of hisT-mediated tRNA pseudouridylation results in a uracil requirement that interferes with *Escherichia coli* K-12 cell division. J. Bacteriol., 173:7395-7400.
- 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. USA, 93:14480-14485.
- Tycowski KT, You ZH, Graham PJ, Steitz JA. 1998. Modification of U6 spliceosomal RNA is guided by other small RNAs. Mol. Cell., 2:629-638.

U

- Umen JG, Guthrie C. 1995a. Prp16p, Slu7p, and Prp8p interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. RNA, 1:584-597.
- Umen JG, Guthrie C. 1995b. The second catalytic step of pre-mRNA splicing. RNA, 1:869-885.
- 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., 25:4778-4785.

V

- Valcarcel J, Singh R, Zamore PD, Green MR. 1993. The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. Nature, 362:171-175.
- Valcarcel J, Gaur RK, Singh R, Green MR. 1996. Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA. Science, 273:1706-1709.
- Valenzuela P, Venegas A, Weinberg F, Bishop R, Rutter WJ. 1978. Structure of yeast phenylalanine-tRNA genes: an intervening DNA segment within the region coding for the tRNA. Proc. Natl. Acad. Sci. USA, 75:190-194.
- van der Veen R, Arnberg AC, van der Horst G, Bonen L, Tabak HF, Grivell LA. 1986. Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing *in vitro*. Cell, 44:225-234.
- van Tol H, Beier H. 1988. All human tRNA^{Tyr} genes contain introns as a prerequisite for pseudouridine biosynthesis in the anticodon. Nucleic Acids Res., 16:1951-1966.
- Vankan P, McGuigan C, Mattaj IW. 1990. Domains of U4 and U6 snRNAs required for snRNP assembly and splicing complementation in *Xenopus* oocytes. EMBO J., 9:3397-3404.
- 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., 9:6935-6952.
- Venema J, Tollervey D. 1995. Processing of pre-ribosomal RNA in *Saccharomyces cerevisiae*. Yeast, 11:1629-1650.
- Venema J, Bousquet-Antonelli C, Gelugne JP, Caizergues-Ferrer M, Tollervey D. 1997. Rok1p is a putative RNA helicase required for rRNA processing. Mol. Cell. Biol., 17:3398-3407.

Vijayraghavan U, Parker R, Tamm J, Iimura Y, Rossi J, Abelson J, Guthrie C. 1986. Mutations in conserved intron sequences affect multiple steps in the yeast splicing pathway, particularly assembly of the spliceosome. EMBO J., **5**:1683-1695.

Vijayraghavan U, Company M, Abelson J. 1989. Isolation and characterization of pre-mRNA splicing mutants of *Saccharomyces cerevisiae*. Genes Dev., **3**:1206-1216.

Vijayraghavan U, Abelson J. 1990. PRP18, a protein required for the second reaction in pre-mRNA splicing. Mol. Cell. Biol., **10**:324-332.

W

agner JD, Jankowsky E, Company M, Pyle AM, Abelson JN. 1998. The DEAH-box protein PRP22

is an ATPase that mediates ATP-dependent mRNA release from the spliceosome and unwinds RNA duplexes. EMBO J., **17**:2926-2937.

Wang W, Krug RM. 1998. U6atac snRNA, the highly divergent counterpart of U6 snRNA, is the specific target that mediates inhibition of AT-AC splicing by the influenza virus NS1 protein. RNA, **4**:55-64.

Wang Y, Wagner JD, Guthrie C. 1998. The DEAH-box splicing factor Prp16 unwinds RNA duplexes *in vitro*. Curr. Biol., **8**:441-451.

Wassarman KM, Steitz JA. 1992a. The low-abundance U11 and U12 small nuclear ribonucleoproteins (snRNPs) interact to form a two-snRNP complex. Mol. Cell. Biol., **12**:1276-1285.

Wassarman DA, Steitz JA. 1992b. Interactions of small nuclear RNA's with precursor messenger RNA during *in vitro* splicing. Science, **257**:1918-1925.

Watanabe K, Shinma M, Oshima T, Nishimura S. 1976. Heat-induced stability of tRNA from an extreme thermophile, *Thermus thermophilus*. Biochem. Biophys. Res. Commun., **72**:1137-1144.

Watkins NJ, Gottschalk A, Neubauer G, Kastner B, Fabrizio P, Mann M, Lührmann R. 1998a. 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**:1549-1568.

Watkins NJ, Newman DR, Kuhn JF, Maxwell ES. 1998b. *In vitro* assembly of the mouse U14 snoRNP core complex and identification of a 65-kDa box C/D-binding protein. RNA, **4**:582-593.

Weidenhammer EM, Ruiz-Noriega M, Woolford JL, Jr. 1997. Prp31p promotes the association of the U4/U6 x U5 tri-snRNP with prespliceosomes to form spliceosomes in *Saccharomyces cerevisiae*. Mol. Cell. Biol., **17**:3580-3588.

- Wells SE, Ares M, Jr. 1994. Interactions between highly conserved U2 small nuclear RNA structures and Prp5p, Prp9p, Prp11p, and Prp21p proteins are required to ensure integrity of the U2 small nuclear ribonucleoprotein in *Saccharomyces cerevisiae*. Mol. Cell. Biol., 14:6337-6349.
- Wells SE, Neville M, Haynes M, Wang J, Igel H, Ares M, Jr. 1996. *CUS1*, a suppressor of cold-sensitive U2 snRNA mutations, is a novel yeast splicing factor homologous to human SAP 145. Genes Dev., 10:220-232.
- Wersig C, Bindereif A. 1990. Conserved domains of human U4 snRNA required for snRNP and spliceosome assembly. Nucleic Acids Res., 18:6223-6229.
- Westhof E, Dumas P, Moras D. 1983. Loop stereochemistry and dynamics in transfer RNA. J. Biomol. Struct. Dyn., 1:337-355.
- Westhof E, Dumas P, Moras D. 1985. Crystallographic refinement of yeast aspartic acid transfer RNA. J. Mol. Biol., 184:119-145.
- Westhof E, Dumas P, Moras D. 1988. Hydration of transfer RNA molecules: a crystallographic study. Biochimie, 70:145-165.
- Whittaker E, Lossky M, Beggs JD. 1990. Affinity purification of spliceosomes reveals that the precursor RNA processing protein Prp8, a protein in the U5 small nuclear ribonucleoprotein particle, is a component of yeast spliceosomes. Proc. Natl. Acad. Sci. USA, 87:2216-2219.
- Wild MA, Gall JG. 1979. An intervening sequence in the gene coding for 25S ribosomal RNA of *Tetrahymena pigmentosa*. Cell, 16:565-573.
- Will CL, Rumpfer S, Klein Gunnewiek J, van Venrooij WJ, Lührmann R. 1996. *In vitro* reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E) spliceosomal complexes. Nucleic Acids Res., 24:4614-4623.
- Will LC, Lührmann R. 1997. SnRNP structure and function. In: Krainer AR, eds. Eukaryotic mRNA processing. Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory Press. pp 130-173.
- Will C, Schneider C, Reed R, Lührmann R. 1999. Identification of both shared and distinct proteins in the major and minor spliceosomes. Science, 284:2003-5
- 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, 14:37-48.
- Wise JA, Tollervey D, Maloney D, Swerdlow H, Dunn EJ, Guthrie C. 1983. Yeast contains small nuclear RNAs encoded by single copy genes. Cell, 35:743-751.
- Wise JA. 1993. Guides to the heart of the spliceosome. Science, 262:1978-1979.
- Wissinger B, Schuster W, Brennicke A. 1991. Trans splicing in *Oenothera* mitochondria: nad1 mRNAs are edited in exon and trans-splicing group II intron sequences. Cell, 65:473-482.

- Wolff T, Bindereif A. 1992. Reconstituted mammalian U4/U6 snRNP complements splicing: a mutational analysis. EMBO J., 11:345-359.
- Wolff T, Bindereif A. 1993. Conformational changes of U6 RNA during the spliceosome cycle: an intramolecular helix is essential both for initiating the U4-U6 interaction and for the first step of slicing. Genes Dev., 7:1377-1389.
- Wolff T, Menssen R, Hammel J, Bindereif A. 1994. Splicing function of mammalian U6 small nuclear RNA: conserved positions in central domain and helix I are essential during the first and second step of pre-mRNA splicing. Proc. Natl. Acad. Sci USA, 91:903-907.
- Wood DD, Pang H, Hempel A, Camerman N, Lane BG, Moscarello MA. 1995. Participation of acetylpsoudouridine in the synthesis of a peptide bond *in vitro*. J. Biol. Chem., 270:21040-21044.
- Wrzesinski J, Nurse K, Bakin A, Lane BG, Ofengand J. 1995a. A dual-specificity pseudouridine synthase: an *Escherichia coli* synthase purified and cloned on the basis of its specificity for psi 746 in 23S RNA is also specific for psi 32 in tRNA(phe). RNA, 1:437-448.
- Wrzesinski J, Bakin A, Nurse K, Lane BG, Ofengand J. 1995b. Purification, cloning, and properties of the 16S RNA pseudouridine 516 synthase from *Escherichia coli*. Biochemistry, 34:8904-8913.
- Wu J, Manley JL. 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. Genes Dev., 3:1553-1561.
- Wu JA, Manley JL. 1991. Base pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. Nature, 352:818-821.
- Wu JY, Maniatis T. 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell, 75:1061-1070.
- Wu Q, Krainer AR. 1996. U1-mediated exon definition interactions between AT-AC and GT-AG introns. Science, 274:1005-1008.
- Wu HJ, Gaubier-Comella P, Delseny M, Grellet F, Van Montagu M, Rouze R. 1996. Non-canonical introns are at least 10(9) years old. Nat. Genet., 14:383-384.
- Wu Q, Krainer AR. 1997. Splicing of a divergent subclass of AT-AC introns requires the major spliceosomal snRNAs. RNA, 3:586-601.
- Wu S, Green MR. 1997. Identification of a human protein that recognizes the 3' splice site during the second step of pre-mRNA splicing. EMBO J., 16:4421-4432.
- 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., 273:16453-16463.
- Wu Q, Krainer AR. 1998. Purine-rich enhancers function in the AT-AC pre-mRNA splicing pathway and do so independently of intact U1 snRNP. RNA, 4:1664-1673.

Wu Q, Krainer AR. 1999. AT-AC pre-mRNA splicing mechanisms and conservation of minor introns in voltage-gated ion channel genes. Mol. Cell. Biol., 19:3225-3236.

Wyatt JR, Sontheimer EJ, Steitz JA. 1992. Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. Genes Dev., 6:2542-2553.

Xie J, Beickman K, Otte E, Rymond BC. 1998. Progression through the spliceosome cycle requires

Prp38p function for U4/U6 snRNA dissociation. EMBO J., 17:2938-2946.

Xu MQ, Kathe SD, Goodrich-Blair H, Nierzwicki-Bauer SA, Shub DA. 1990a. Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in cyanobacteria. Science, 250:1566-1570.

Xu Y, Petersen-Bjorn S, Friesen JD. 1990b. The PRP4 (RNA4) protein of *Saccharomyces cerevisiae* is associated with the 5' portion of the U4 small nuclear RNA. Mol. Cell. Biol., 10:1217-1225.

Xu D, Nouraini S, Field D, Tang SJ, Friesen JD. 1996. An RNA-dependent ATPase associated with U2/U6 snRNAs in pre-mRNA splicing. Nature, 381:709-713.

Xu D, Field DJ, Tang SJ, Moris A, Bobechko BP, Friesen JD. 1998. Synthetic lethality of yeast slt mutations with U2 small nuclear RNA mutations suggests functional interactions between U2 and U5 snRNPs that are important for both steps of pre-mRNA splicing. Mol. Cell. Biol., 18:2055-2066.

Yan D, Ares M, Jr. 1996. Invariant U2 RNA sequences bordering the branchpoint recognition region are

essential for interaction with yeast SF3a and SF3b subunits. Mol. Cell. Biol., 16:818-828.

Yan D, Perriman R, Igel H, Howe KJ, Neville M, Ares M, Jr. 1998. CUS2, a yeast homolog of human Tat-SF1, rescues function of misfolded U2 through an unusual RNA recognition motif. Mol. Cell. Biol., 18:5000-5009.

Yu YT, Maroney PA, Darzynkiwicz E, Nilsen TW. 1995. U6 snRNA function in nuclear pre-mRNA splicing: a phosphorothioate interference analysis of the U6 phosphate backbone. RNA, 1:46-54.

Yu YT, Steitz JA. 1997. Site-specific crosslinking of mammalian U11 and U6atac to the 5' splice site of an AT-AC intron. Proc. Natl. Acad. Sci. USA, 94:6030-6035.

Yu YT, Shu MD, Steitz JA. 1998. Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. EMBO J., 17:5783-5795.

Yue D, Kintanar A, Horowitz J. 1994. Nucleoside modifications stabilize Mg²⁺ binding in *Escherichia coli* tRNA(Val): an imino proton NMR investigation. Biochemistry, 33:8905-8911.

Z

Zamore PD, Green MR. 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. Proc. Natl. Acad. Sci. USA, 86:9243-9247.

Zamore PD, Patton JG, Green MR. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. Nature, 355:609-614.

Zavanelli MI, Ares M, Jr. 1991. Efficient association of U2 snRNPs with pre-mRNA requires an essential U2 RNA structural element. Genes Dev., 5:2521-2533.

Zavanelli MI, Britton JS, Igel AH, Ares M, Jr. 1994. Mutations in an essential U2 small nuclear RNA structure cause cold-sensitive U2 small nuclear ribonucleoprotein function by favoring competing alternative U2 RNA structures. Mol. Cell. Biol., 14:1689-1697.

Zerby DB, Patton JR. 1996. Metabolism of pre-messenger RNA splicing cofactors: modification of U6 RNA is dependent on its interaction with U4 RNA. Nucleic Acids Res., 24:3583-3589.

Zerby DB, Patton JR. 1997. Modification of human U4 RNA requires U6 RNA and multiple pseudouridine synthases. Nucleic Acids Res., 25:4808-4815.

Zerfass K, Beier H. 1992. Pseudouridine in the anticodon G psi A of plant cytoplasmic tRNA(Tyr) is required for UAG and UAA suppression in the TMV-specific context. Nucleic Acids Res., 20:5911-5918.

Zhang X, Schwer B. 1997. Functional and physical interaction between the yeast splicing factors Slu7 and Prp18. Nucleic Acids Res., 25:2146-2152.

Zhao X, Horne DA. 1997. The role of cysteine residues in the rearrangement of uridine to pseudouridine catalyzed by pseudouridine synthase I. J. Biol. Chem., 272:1950-1955.

Zhou Z, Reed R. 1998. Human homologs of yeast Prp16 and Prp17 reveal conservation of the mechanism for catalytic step II of pre-mRNA splicing. EMBO J., 17:2095-2106.

Zhuang Y, Weiner AM. 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. Cell, 46:827-835.

Zhuang Y, Weiner AM. 1989. A compensatory base change in human U2 snRNA can suppress a branch site mutation. Genes Dev., 3:1545-1552.

Annexe

ANNEXES

*Noms, symboles et structures chimiques des nucléotides modifiés portés
par les ARN.*

D'après les revues, Limbach et al., 1994 ; Motorin et Grosjean, 1998.

pseudouridines modifiées

Symbole	nom (Code une lettre)	Source	R ₁	R ₃	R ₄
Ψ	pseudouridine (P)	A,B,E	H	H	H
Ψm	2'-O-methylpseudouridine (Z)	A,E	CH ₃	H	H
m ¹ Ψ	1-methylpseudouridine (I)	A	H	H	CH ₃
m ³ Ψ	3-methylpseudouridine	B	H	CH ₃	H
m ¹ acp ³ Ψ	1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine	E	H	acp	CH ₃

adenosines modifiées

Symbole	nom (Code une lettre)	Source	R ₁	R ₂	R ₃	R ₄
<i>Adenosine derivatives</i>						
Am	2'-O-methyladenosine (:)	A,E	CH ₃	H	H	H
m ² A	2-methyladenosine (I)	B	H	CH ₃	H	H
m ⁶ A	N ⁶ -methyladenosine (=)	A,B,E	H	H	H	CH ₃
m ⁶ ₂ A	N ⁶ ,N ⁶ -dimethyladenosine	A,B,E	H	H	CH ₃	CH ₃
m ⁶ Am	N ⁶ 2'-O-dimethyladenosine	E	CH ₃	H	H	CH ₃
m ⁶ ₂ Am	N ⁶ ,N ⁶ 2'-O-trimethyladenosine	E	CH ₃	H	CH ₃	CH ₃
ms ² m ⁶ A	2-methylthio-N ⁶ -methyladenosine	B	H	CH ₃ -S-	H	CH ₃
i ⁶ A	N ⁶ -isopentenyladenosine (+)	E	H	H	H	Dimethylallyl
ms ² i ⁶ A	2-methylthio-N ⁶ -isopentenyladenosine (*)	B	H	CH ₃ -S-	H	Dimethylallyl
io ⁶ A	N ⁶ -(<i>cis</i> -hydroxyisopentenyl)adenosine (')	E	H	H	H	<i>cis</i> -hydroxy-methyl-methylallyl
ms ² io ⁶ A	2-methylthio-N ⁶ -(<i>cis</i> -hydroxyisopentenyl)-adenosine	E	H	CH ₃ -S-	H	<i>cis</i> -hydroxy-methyl-methylallyl
g ⁶ A	N ⁶ -glycinylocarbamoyladenosine	E	H	H	H	-CONHCH ₂ COOH
t ⁶ A	N ⁶ -threonylocarbamoyladenosine (6)	A,B,E	H	H	H	threonylocarbamoyl
m ⁶ t ⁶ A	N ⁶ -methyl-N ⁶ -threonylocarbamoyladenosine (E)	B	H	H	CH ₃	threonylocarbamoyl
ms ² t ⁶ A	2-methylthio-N ⁶ -threonylocarbamoyl-adenosine (I)	A,B,E	H	CH ₃ -S-	H	threonylocarbamoyl
hn ⁶ A	N ⁶ -hydroxynorvalylcarbamoyladenosine	A,B	H	H	H	hydroxynorvalyl-carbamoyl
ms ² hn ⁶ A	2-methylthio-N ⁶ -hydroxynorvalylcarbamoyl-adenosine	A,B	H	CH ₃ -S-	H	hydroxynorvalyl-carbamoyl
Ar(p)	2'-O-ribosyladenosine (phosphate) (^)	E	2'-O-ribosyl	H	H	H
m ¹ A	1-methyladenosine (")	A,B,E	-	-	-	-
<i>Inosine derivatives</i>						
I	inosine (I)	A,B,E	H	H	-	-
Im	2'-O-methylinosine	E	CH ₃	H	-	-
m ¹ I	1-methylinosine (O)	A,E	H	CH ₃	-	-

uridines modifiées

#	Symbole	nom (code une lettre)	Source	R ₁	R ₂	R ₃	R ₄
1	Um	2'-O-methyluridine (J)	A,B,E	CH ₃	O	H	H
2	s ² U	2-thiouridine (2)	A,E	H	S	H	H
3	s ² Um	2-thio-2'-O-methyluridine	A	CH ₃	S	H	H
4	m ³ U	3-methyluridine	B	H	O	CH ₃	H
5	m ³ Um	3,2'-O-dimethyluridine	E	CH ₃	O	CH ₃	H
6	acp ³ U	3-(3-amino-3-carboxypropyl)uridine (X)	B	H	O	acp	H
7	s ⁴ U	4-thiouridine (4)	A,B	-	-	-	-
8	m ⁵ U	ribosylthymine (T)	A,B,E	H	O	H	CH ₃
9	m ⁵ Um	5, 2'-O-dimethyluridine (\\)	B,E	CH ₃	O	H	CH ₃
10	m ⁵ s ² U	5-methyl-2-thiouridine (F)	A,B	H	S	H	CH ₃
11	ho ⁵ U	5-hydroxyuridine	B,E	H	O	H	OH
12	mo ⁵ U	5-methoxyuridine (5)	B	H	O	H	-OCH ₃
13	cmo ⁵ U	uridine 5-oxyacetic acid (V)	B	H	O	H	-OCH ₂ COOH
14	mcmo ⁵ U	uridine 5-oxyacetic acid methyl ester	B	H	O	H	-OCH ₂ COOCH ₃
15	cm ⁵ U	5-carboxymethyluridine	E	H	O	H	-CH ₂ COOH
16	mcm ⁵ U	5-methoxycarbonylmethyluridine (1)	E	H	O	H	-CH ₂ COOCH ₃
17	mcm ⁵ Um	5-methoxycarbonylmethyl-2'-O-methyluridine	E	CH ₃	O	H	-CH ₂ COOCH ₃
18	mcm ⁵ s ² U	5-methoxycarbonylmethyl-2-thiouridine (3)	E	H	S	H	-CH ₂ COOCH ₃
19	ncm ⁵ U	5-carbamoylmethyluridine (&)	E	H	O	H	-CH ₂ CONH ₂
20	ncm ⁵ Um	5-carbamoylmethyl-2'-O-methyluridine (-)	E	CH ₃	O	H	-CH ₂ CONH ₂
21	chm ⁵ U	5-(carboxyhydroxymethyl)uridine	E	H	O	H	-CH(OH)COOH
22	mchm ⁵ U	5-(carboxyhydroxymethyl)uridinemethyl ester (,)	E	H	O	H	-CH(OH)COOCH ₃
23	nm ⁵ s ² U	5-aminomethyl-2-thiouridine	B	H	S	H	-CH ₂ NH ₂
24	mnm ⁵ U	5-methylaminomethyluridine(!)	B	H	O	H	-CH ₂ NHCH ₃
25	mnm ⁵ s ² U	5-methylaminomethyl-2-thiouridine (S)	B	H	S	H	-CH ₂ NHCH ₃
26	mnm ⁵ se ² U	5-methylaminomethyl-2-selenouridine	A,B	H	Se	H	-CH ₂ NHCH ₃
27	cmnm ⁵ U	5-carboxymethylaminomethyluridine (!)	B,E	H	O	H	-CH ₂ NHCH ₂ COOH
28	cmnm ⁵ Um	5-carboxymethylaminomethyl-2'-O-methyluridine (!)	B	CH ₃	O	H	-CH ₂ NHCH ₂ COOH
29	cmnm ⁵ s ² U	5-carboxymethylaminomethyl-2-thiouridine (\$)	B	H	S	H	-CH ₂ NHCH ₂ COOH
30	D	dihydrouridine (D)	A,B,E	H	O	H	H
31	m ⁵ D	dihydroribosylthymine	E	H	O	H	CH ₃

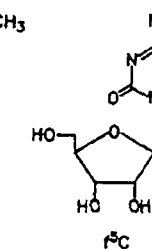
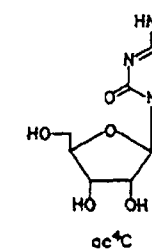
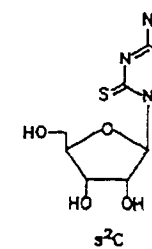
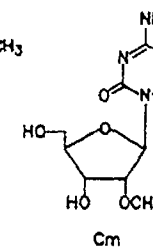
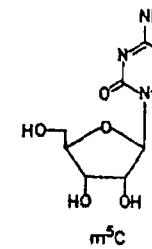
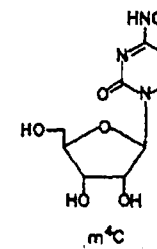
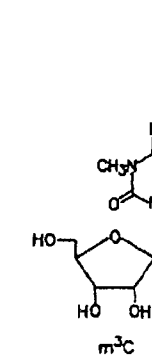
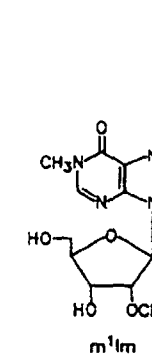
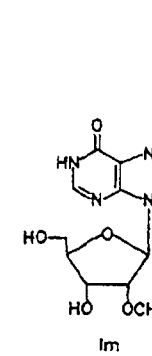
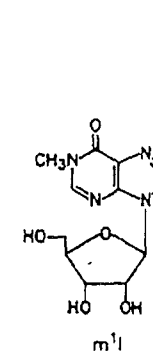
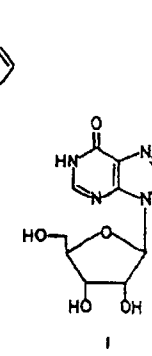
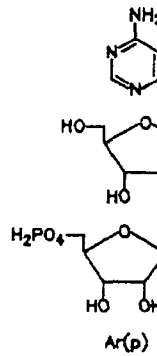
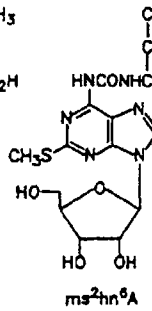
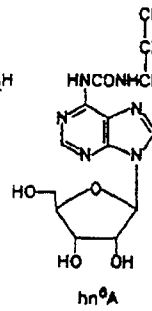
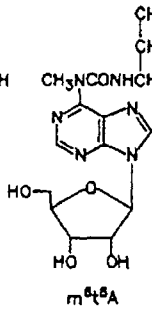
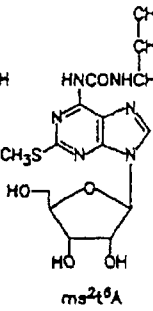
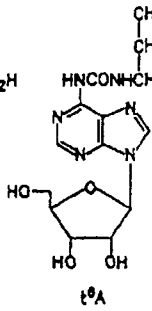
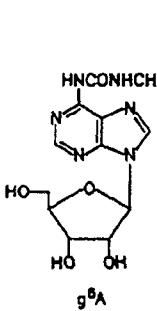
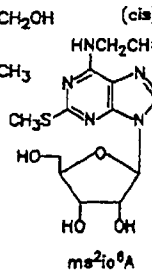
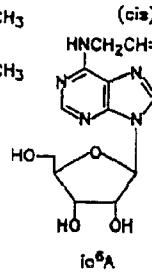
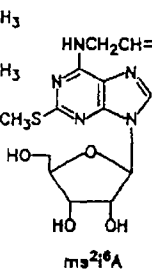
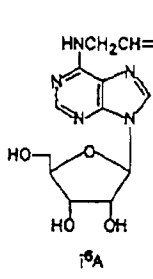
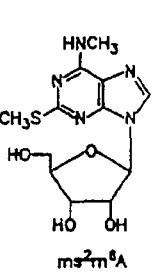
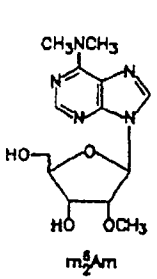
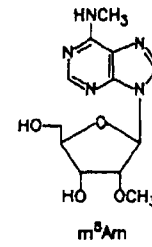
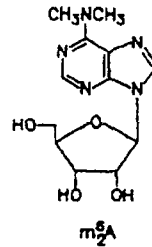
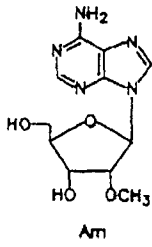
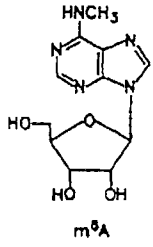
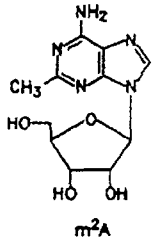
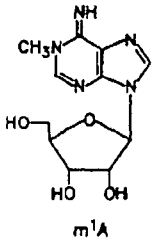
guanosines modifiées

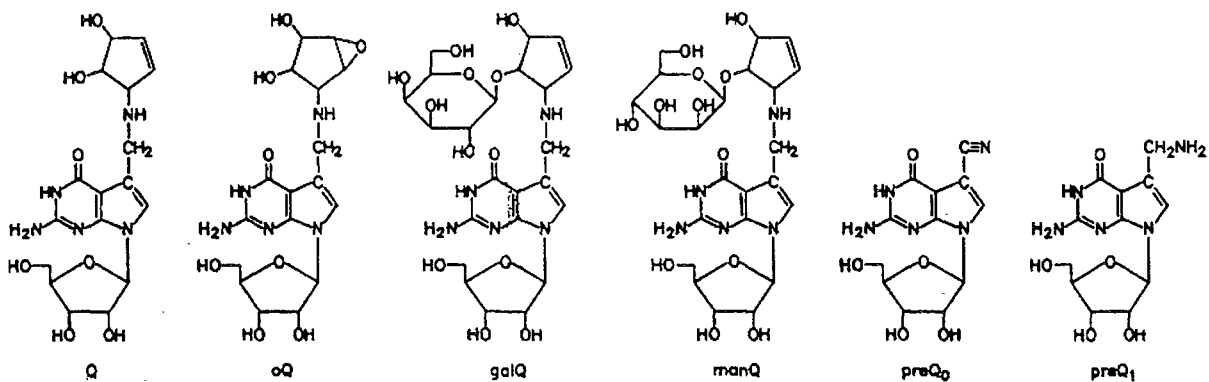
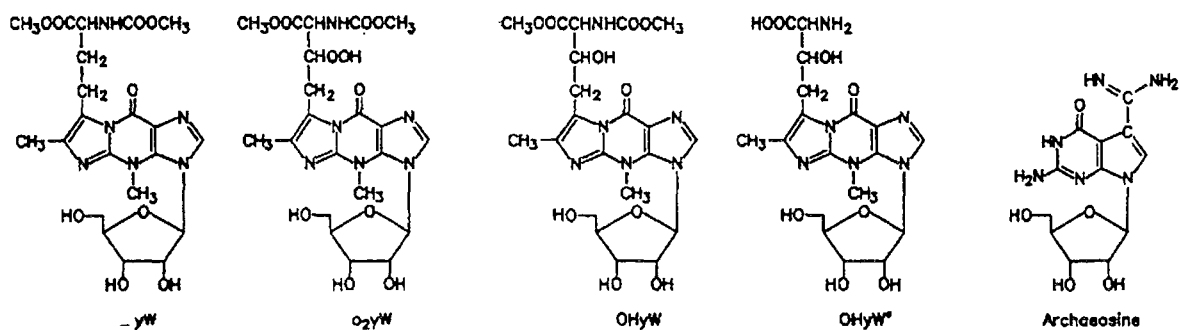
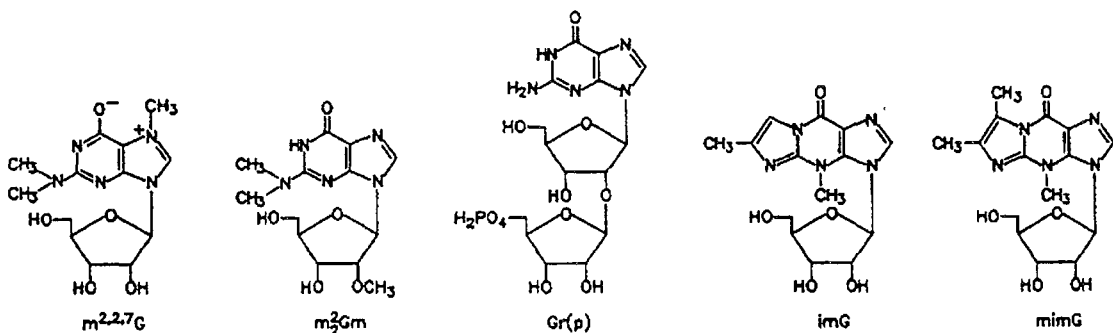
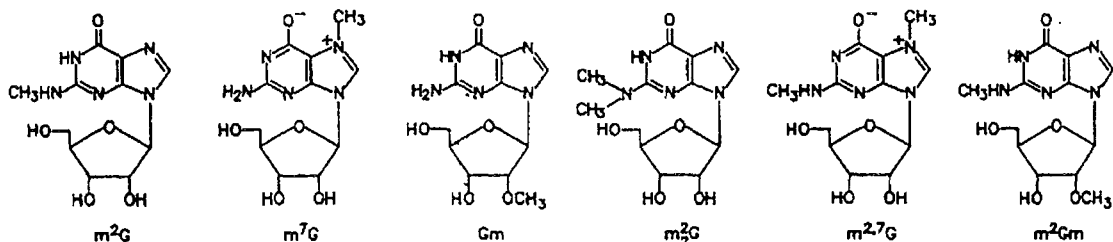
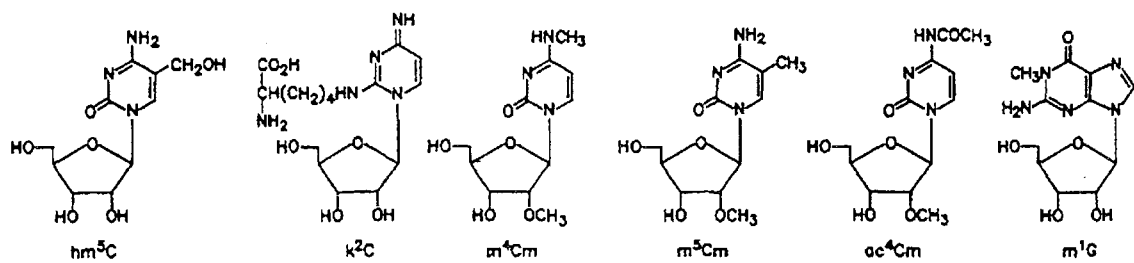
Symbole	nom (code une lettre)	Source	R ₁	R ₂	R ₃	R ₄
<i>Guanosine derivatives</i>						
Gm	2'-O-methylguanosine (#)	A,B,E	CH ₃	H	H	H
m ¹ G	1-methylguanosine (K)	A,B,E	H	H	H	CH ₃
m ² G	N ² -methylguanosine (L)	A,E	H	CH ₃	H	H
m ² ₂ G	N ² ,N ² -dimethylguanosine (R)	A,E	H	CH ₃	CH ₃	H
m ² Gm	N ² ,2'-O-dimethylguanosine	A	CH ₃	CH ₃	H	H
m ² ₂ Gm	N ² ,N ² ,2'-O-trimethylguanosine (I)	A	CH ₃	CH ₃	CH ₃	H
Gr(p)	2'-O-ribosylguanosine (phosphate)	E	2-O- ribosyl	H	H	H
<i>m⁷-Guanosine derivatives</i>						
m ⁷ G	7-methylguanosine (7)	A,B,E	-	H	H	-
m ^{2,7} G	N ² ,7-dimethylguanosine	E	-	CH ₃	H	-
m ^{2,2,7} G	N ² ,N ² ,7-trimethylguanosine	E	-	CH ₃	CH ₃	-
<i>Wyosine derivatives</i>						
imG	wyosine	E	-	-	-	H
mimG	methylwyosine	A	-	-	-	CH ₃
OHyW*	undermodified hydroxywybutosine	E	-	-	-	S1
yW	wybutosine (Y)	E	-	-	-	S2
OHyW	hydroxywybutosine	E	-	-	-	β-hydroxy-S2
o ₂ yW	peroxywybutosine (W)	E	-	-	-	β-peroxy-S2
<i>7-deazaguanosine derivatives</i>						
Q	queuosine (Q)	B,E	-	-	-	S3
oQ	epoxyqueuosine	B	-	-	-	epoxy-S3
galQ	galactosyl-queuosine (9)	E	-	-	-	gal-S3
manQ	mannosyl-queuosine (8)	E	-	-	-	man-S3
preQ ₀	7-cyano-7-deazaguanosine	B	-	-	-	-CN
gQ (G ⁺)	Archaeosine (alternate name 7-formamidino-7-deazaguanosine) (I)	A	-	-	-	-
preQ ₁	7-aminomethyl-7-deazaguanosine	B	-	-	-	-CH ₂ NH ₂

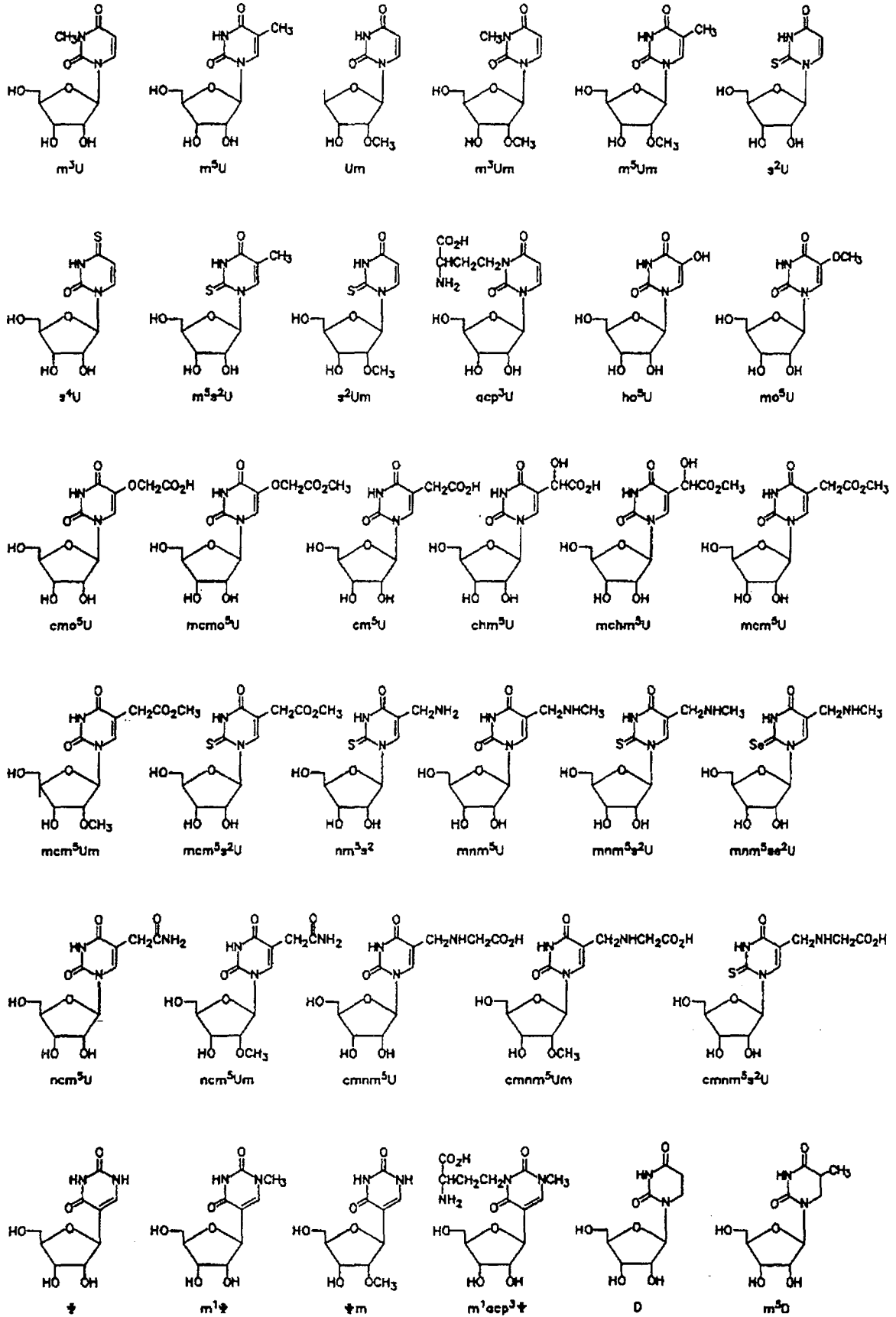
cytidines modifiées

#	Symbole	nom (Code une lettre)	Source	R ₁	R ₂	R ₃
1	Cm	2'- <i>O</i> -methylcytidine (B)	A,B,E	CH ₃	H	H
2	m ⁴ C	N ⁴ -methylcytidine	E	H	CH ₃	H
3	m ⁴ Cm	N ⁴ ,2'- <i>O</i> -dimethylcytidine	B	CH ₃	CH ₃	H
4	ac ⁴ C	N ⁴ -acetylcytidine (M)	A,B,E	H	-COCH ₃	H
5	ac ⁴ Cm	N ⁴ -acetyl-2'- <i>O</i> -methylcytidine	A	CH ₃	-COCH ₃	H
6	m ⁵ C	5-methylcytidine (?)	A,E	H	H	CH ₃
7	m ⁵ Cm	5,2'- <i>O</i> -dimethylcytidine	A	CH ₃	H	CH ₃
8	hm ⁵ C	5-hydroxymethylcytidine	E	H	H	-CH ₂ OH
9	f ⁵ C	5-formylcytidine (>)	E	H	H	-CHO
10	f ⁵ Cm	2'- <i>O</i> -methyl-5-formylcytidine (°)	E	CH ₃	H	-CHO
11	m ³ C	3-methylcytidine (')	B,E	-	-	-
12	s ² C	2-thiocytidine (%)	A,B	-	-	-
13	k ² C	lysidine (j)	B,E	-	-	-

structures chimiques des différents nucléotides modifiés

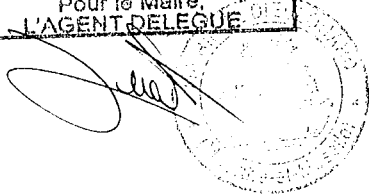
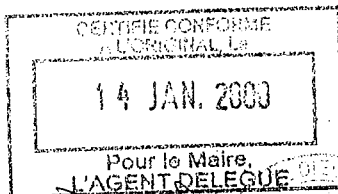






Mademoiselle MASSENET Séverine

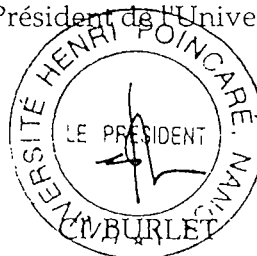
DOCTORAT de l'UNIVERSITÉ HENRI POINCARÉ, NANCY-I
en BIOLOGIE STRUCTURALE, MOLECULAIRE & CELLULAIRE



VU, APPROUVÉ ET PERMIS D'IMPRIMER

Nancy, le 10 JAN 2000 ~ 277

Le Président de l'Université



Tous les ARN métaboliquement stables, ARNr, ARNt et UsnRNA, portent des modifications post-transcriptionnelles. Les nucléotides modifiés portés par les UsnRNA sont essentiellement des résidus Ψ et 2'-O-méthylés. Quand j'ai débuté ma thèse, la localisation de ces résidus modifiés n'avait été réalisée que dans les UsnRNA des spliceosomes U2-dépendants de plantes et de quelques vertébrés. Ces modifications étaient localisées dans des régions des UsnRNA, impliquées dans des interactions ARN-ARN et ARN-protéines, ce qui pouvait suggérer un rôle important des nucléotides modifiés dans la réaction d'épissage. Afin d'identifier les résidus Ψ phylogénétiquement conservés dans les UsnRNA, nous avons recherché la présence de ces modifications dans les UsnRNA des spliceosomes de *S. cerevisiae* et des spliceosomes U12-dépendants de cellules HeLa. L'analyse montre que ces UsnRNA portent beaucoup moins de résidus Ψ que les UsnRNA des spliceosomes U2-dépendants de plantes et de vertébrés, mais les résidus détectés sont, pour la plupart, conservés phylogénétiquement et localisés dans des régions des UsnRNA jouant un rôle majeur dans l'assemblage du spliceosome ou dans son mécanisme d'action.

L'identification des résidus Ψ des UsnRNA de *S. cerevisiae* nous a permis de rechercher, par voie génétique, les Ψ-synthases responsables de leur synthèse. L'utilisation de souches de levure, délétées spécifiquement dans chacun des gènes codant les Ψ-synthases de levure déjà caractérisées, nous a permis de mettre en évidence la première Ψ-synthase agissant sur les UsnRNA, Pus1p, qui est responsable de la modification du résidu U₄₄ dans le snRNA U12. Cette enzyme catalyse également l'isomérisation de plusieurs résidus U en résidus Ψ dans les ARNt (Motorin et al., 1998). A l'heure actuelle, c'est la seule Ψ-synthase connue qui possède, *in vivo*, à la fois une multispécificité d'action au sein d'un même ARN et une action vis-à-vis de deux substrats distincts. Cette caractéristique surprenante posait la question de la spécificité de reconnaissance de Pus1p vis-à-vis de ces ARN substrats. En utilisant, *in vitro*, des transcrits correspondant à des domaines de l'ARNr 26S de *S. cerevisiae*, au domaine 3'-terminal de l'ARN génomique du virus BMV et à des ARNt tronqués, nous avons montré que Pus1p reconnaît un motif structural particulier. En effet, l'uridine modifiée se trouve toujours à proximité d'une jonction flexible liant deux structures tige-boucles. L'enzyme Pus1p de mammifère, récemment identifiée, modifie *in vitro* les mêmes sites dans les ARNt que son homologue chez *S. cerevisiae*. Nous avons débuté l'étude de la spécificité de cette enzyme vis-à-vis du snRNA U12 et nous avons pu montrer qu'elle modifie un snRNA U12 de *S. cerevisiae*, produit par transcription *in vitro*.

Il restait à identifier la (ou les) Ψ-synthase(s) responsable(s) de l'isomérisation des autres résidus Ψ dans les UsnRNA de *S. cerevisiae*. Quatre ORF, codant des Ψ-synthases potentielles, avaient été détectées chez *S. cerevisiae*, sur la base de la recherche d'une séquence signature établie pour les Ψ-synthases. Nous avons délété ces 4 ORF, de manière individuelle, et montré qu'aucune des enzymes codées par ces ORF n'agit sur les UsnRNA. Nous avons alors étudié l'action possible de ces Ψ-synthases sur d'autres ARN et nous avons déjà montré que l'une d'entre elles est responsable de la formation de l'unique résidu Ψ de l'ARNr mitochondrial 21S de *S. cerevisiae*.

Tous les résidus Ψ dans les ARNr eucaryotes ont la particularité d'être formés par une Ψ-synthase unique, dont la spécificité est assurée par un ensemble d'ARN guides. Par contre, chez les bactéries, plusieurs Ψ-synthases de spécificité différente assurent la formation des quelques résidus Ψ portés par les ARNr. Quel est le système utilisé chez les Archaea? Pour répondre à cette question, il était nécessaire de localiser les résidus Ψ dans les ARNr d'Archaea. C'est pourquoi nous avons entrepris l'identification des résidus Ψ portés par l'ARNr 23S de *Sulfolobus acidocaldarius*. Le nombre de résidus Ψ détectés est comparable à celui précédemment observé dans l'ARNr 23S de bactéries, mais la localisation de ces résidus est différente.