



**HAL**  
open science

# Rôle de la $\beta$ 1,3-Galactosyltransférase 6 ( $\beta$ 3GalT6) dans la pathogénie d'une maladie génétique rare, les syndromes d'Ehlers-Danlos (SED)

Benjamin Jolivet

## ► To cite this version:

Benjamin Jolivet. Rôle de la  $\beta$ 1,3-Galactosyltransférase 6 ( $\beta$ 3GalT6) dans la pathogénie d'une maladie génétique rare, les syndromes d'Ehlers-Danlos (SED). Biochimie, Biologie Moléculaire. Université de Lorraine, 2020. Français. NNT : 2020LORR0085 . tel-02964603

**HAL Id: tel-02964603**

**<https://hal.univ-lorraine.fr/tel-02964603v1>**

Submitted on 12 Oct 2020

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

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



## AVERTISSEMENT

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

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

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

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

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

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

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

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



**UNIVERSITÉ  
DE LORRAINE**

BioSE



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

**Thèse**

**Présentée et soutenue publiquement pour l'obtention du titre de**

**DOCTEUR DE L'UNIVERSITE DE LORRAINE**

**Mention : « Sciences de la Vie et de la Santé »**

**Par Benjamin JOLIVET**

**Rôle de la  $\beta$ 1,3-Galactosyltransférase 6 ( $\beta$ 3GalT6) dans la pathogénie d'une maladie  
génétique rare, les syndromes d'Ehlers-Danlos (SED)**

**Le 16 Juillet 2020**

**Membre du jury :**

**Rapporteurs :**

Madame Mirjam CZJZEK, Directeur de Recherche CNRS, Laboratoire de Biologie Intégrative des modèle Marins (LBI2M), UMR 8227 – CNRS et Sorbonne Université, Roscoff, France

Madame Sylvie BRASSART-PASCO, Chargé de Recherche CNRS, Laboratoire de Biochimie Médicale et de Biologie Moléculaire, UMR 7365 CNRS et Université de Reims Champagne Ardenne, UFR Médecine, Reims, France

**Examineurs :**

Monsieur Bruno LEHEUP, Professeur de Génétique Médicale, Chef du Service de Génétique Clinique, Faculté de Médecine, Université de Lorraine, Vandœuvre-lès-Nancy, France

Monsieur Régis DANIEL, Directeur de Recherche, Laboratoire Analyse et Modélisation pour la Biologie et l'Environnement, CNRS UMR 8587, Evry, France

Madame Sylvie FOURNEL-GIGLEUX, Directeur de Recherche INSERM, UMR 7365 CNRS-Université de Lorraine, Vandoeuvre-lès- Nancy, France (Directeur de Thèse)

Madame Sandrine GULBERTI, Maître de Conférences, UMR 7365 CNRS-Université de Lorraine, Vandoeuvre-lès-Nancy, France (Co-Directeur de Thèse)

---

Laboratoire IMoPA (UMR 7365 CNRS – Université de Lorraine) – Equipe Ingénierie Moléculaire, Cellulaire, Thérapeutique & Glycosyltransférases (MoCelTEG), Campus Biologie-Santé, Faculté de Médecine, Biopôle de l'Université de Lorraine, 9 Avenue de la Forêt de Haye, BP 50184, 54505 Vandoeuvre-lès-Nancy







## Remerciements

Je tiens tout d'abord à remercier Madame le Docteur Sylvie FOURNEL-GIGLEUX pour m'avoir accueilli au sein de son équipe et pour avoir dirigé ce travail de thèse. Je la remercie de sa compétence et de sa rigueur scientifique qui m'ont beaucoup appris. Qu'elle trouve ici l'expression de mes sincères remerciements.

Mes remerciements s'adressent également à Madame le Docteur Sandrine GULBERTI qui m'a co-encadré ce travail de thèse. Je la remercie d'avoir transmis son enthousiasme, de m'avoir appris la patience et la rigueur. Merci beaucoup pour son soutien durant les moments difficiles au cours de ce travail de thèse et le temps consacré à la correction de ce manuscrit.

Mes remerciements s'adressent également à Monsieur le Docteur Nick RAMALANJONA, qui m'a beaucoup appris et conseillé durant ces années de thèses. Je le remercie aussi pour soutien et le temps qu'il a accordé pour la correction de ce manuscrit.

Je tiens également à remercier le Professeur Jean-Yves JOUZEAU de m'avoir accueilli au sein de son laboratoire, qu'il trouve ici l'expression de mon profond respect.

Je voudrais remercier Madame le Docteur Mirjam CZJZEK, Madame le Docteur Sylvie BRASSART-PASCO pour avoir accepté de juger ce travail de thèse en qualité de rapporteur.

Je voudrais remercier Monsieur le Professeur Bruno LEHEUP et Monsieur le Docteur Régis DANIEL d'avoir accepté de juger ce travail de thèse en qualité d'examineur.

Dans un registre plus personnel, je tiens à remercier mes compagnons de thèse. Je remercie Paul, Madhia, Dima, Elise, Dafné, Sajida pour leur écoute, leur soutien et leur précieux conseils. Merci également à Isabelle, Anne et Jean-Baptiste qui m'ont beaucoup appris durant mes années de Master et mes années de thèse. Je remercie Anne et Roméo qui m'ont apporté leur aide pour réaliser des manipulations présentés dans ce manuscrit. Je remercie également toutes ces personnes pour tous les moments qu'on a partagé ensemble.

Je remercie également toutes les autres personnes croisées au sein du laboratoire pour leur disponibilités et leur bonne humeur.

Mes remerciements s'adressent également au Docteur Frédérique BERNARD pour son aide durant mes années de thèses pour m'avoir assisté à surmonter mes soucis avec la langue française. Je la remercie également pour son aide sur la correction de ce manuscrit.





Enfin mes plus profonds remerciements vont à mes parents sans qui je ne serai pas là aujourd'hui ! Les mots ne suffiront pas pour vous remercier de tout ce que vous avez pu faire pour moi. Merci à mes frères d'avoir été à mon écoute durant les moments difficiles.

Et toi, Mylène, je t'adresse un remerciement qui est particulier et unique. Merci de m'avoir soutenu durant ces années qui ont été parfois difficiles pour moi, mais également pour toi. Merci d'avoir été là quand j'en avais besoin et de croire en moi. Merci pour ton amour qui a été essentiel durant ces années !







<b>REMERCIEMENTS</b> .....	<b>3</b>
<b>LISTE DES TABLEAUX</b> .....	<b>23</b>
<b>LISTE DES FIGURES</b> .....	<b>25</b>
<b>LISTE DES ABREVIATIONS</b> .....	<b>14</b>
<b><u>AVANT-PROPOS</u></b> .....	<b><u>19</u></b>
<b><u>INTRODUCTION BIBLIOGRAPHIQUE</u></b> .....	<b><u>23</u></b>
<b>I. LES SYNDROMES D'EHLERS-DANLOS :</b> .....	<b>24</b>
I.1. HISTORIQUE DES SED :	24
I.2. CLASSIFICATION DES SED :	25
I.3. SYMPTOMES DES SED :	30
I.4. MODES DE TRANSMISSION DES SED :	30
I.5. DIAGNOSTIC DES DIFFERENTS TYPES DE SED :	31
I.6. TRAITEMENTS DES SED :	33
I.6.1. LES TRAITEMENTS PREVENTIFS :	33
I.6.2. LES TRAITEMENTS MEDICAMENTEUX :	34
I.6.3. LA PHYSIOTHERAPIE ET LES TRAITEMENTS PSYCHOLOGIQUES :	34
I.7. LES SED ASSOCIES A DES DEFAUTS DE SYNTHESE DES GLYCOSAMINOGLYCANES :	35
<b>II. LES PROTEOGLYCANES : STRUCTURE, FONCTIONS BIOLOGIQUES ET BIOSYNTHESE :</b> .....	<b>39</b>
II.1. STRUCTURE DES PROTEOGLYCANES ET DES GLYCOSAMINOGLYCANES :	39
II.1.1. LA STRUCTURE DE LA PROTEINE CORE :	39
II.1.2. LA STRUCTURE DES GLYCOSAMINOGLYCANES :	39
II.2. LES FONCTIONS BIOLOGIQUES DES PROTEOGLYCANES :	40
II.2.1. GENERALITES :	40
II.2.2. FONCTIONS DES PROTEOGLYCANES DANS LES TISSUS CONJONCTIFS :	41
II.2.2.1. PROTEOGLYCANES MEMBRANAIRES PORTANT DES CHAINES D'HEPARANES SULFATES, LES SYNDECANS :	41
II.2.2.2. PETIT PROTEOGLYCANE MATRICIEL PORTANT UNE CHAINE DE CHONDROÏTINE/DERMATANE SULFATE, LA DECORINE :	42
II.3. LA BIOSYNTHESE DES PROTEOGLYCANES :	43
II.3.1. SYNTHESE DE L'AMORCE TETRASACCHARIDIQUE :	44
II.3.2. L'ELONGATION ET LA MATURATION DES CHAINES DE GAGS :	45
<b>III. LES GLYCOSYLTRANSFERASES ET LA B1,3-GALACTOSYLTRANSFERASE 6 :</b> .....	<b>48</b>
III.1. LES GLYCOSYLTRANSFERASES, DES ENZYMES UBIQUITAIRES :	48
III.2. LA FAMILLE DES B1,3-GALACTOSYLTRANSFERASES (B3GALTS) :	49



III.3. LA B1,3-GALACTOSYLTRANSFERASE 6 (B3GALT6) :.....	50
III.3.1. ORGANISATION STRUCTURALE DE LA B3GALT6 :.....	50
III.3.2. REACTION ENZYMATIQUE CATALYSEE PAR LA B3GALT6 :.....	51
III.4. IMPLICATION DU GENE B3GALT6 DANS LA PATHOGENIE DU SED SPONDYLO-DYSPLASTIQUE : 52	
III.4.1. DES MUTATIONS DU GENE B3GALT6 IMPLIQUEES DANS DEUX TYPES DE PATHOLOGIES : .....	52
III.4.2. IMPACT DES MUTATIONS SUR LA B3GALT6 SUR L'ACTIVITE ET LA LOCALISATION DE L'ENZYME :.....	53
III.4.3. IMPACT DES MUTATIONS DE LA B3GALT6 SUR LA PATHOGENIE DES SED : .....	54

**PROBLEMATIQUE ET OBJECTIFS..... 59**

**MATERIELS ET METHODES..... 62**

<b>I. MATERIELS :.....</b>	<b>63</b>
I.1. LES REACTIFS CHIMIQUES ET BIOCHIMIQUES :.....	63
I.2. LES LIGNEES CELLULAIRES ET LE MILIEU UTILISE EN CULTURE CELLULAIRE : .....	63
I.3. LES REACTIFS DE BIOLOGIE MOLECULAIRE :.....	63
<b>II. METHODES EXPERIMENTALES : .....</b>	<b>64</b>
II.1. CARACTERISATION FONCTIONNELLE DE LA $\beta$ 3GALT6 RECOMBINANTE HUMAINE :.....	64
II.1.1. CLONAGE DES SEQUENCES CODANT LES DIFFERENTES TRONCATIONS DE LA $\beta$ 3GALT6 RECOMBINANTE HUMAINE DANS LES VECTEURS D'EXPRESSION PETM41, PET41A ET PETM30 :.....	64
II.1.2. PRODUCTION DE LA $\beta$ 3GALT6 RECOMBINANTE HUMAINE EN SYSTEME BACTERIEN :.....	65
II.1.2.1. ETAPE D'OPTIMISATION DE LA PRODUCTION DE LA h $\beta$ 3GALT6 POUR LES DIFFERENTES TRONCATIONS DE LA PROTEINE :.....	65
II.1.2.2. PRODUCTION DE LA $\beta$ 3GALT6 RECOMBINANTE HUMAINE : 6His-MBP-h $\beta$ 3GALT6 $\Delta$ NTER30 PAR LA SOUCHE E. COLI BL21 DE3 :.....	66
II.1.2.3. EXTRACTION DE LA PROTEINE RECOMBINANTE HUMAINE : .....	66
II.1.3. PURIFICATION DE LA PROTEINE DE FUSION 6His-MBP-h $\beta$ 3GALT6 $\Delta$ NTER30 : .....	67
II.1.4. ANALYSE DE L'EXPRESSION ET DE LA SOLUBILITE DES PROTEINES DE FUSION 6His-MBP-h $\beta$ 3GALT6 $\Delta$ NTER30, NTER34, NTER43, NTER50 ET NTER55 :.....	67
II.1.4.1. DOSAGE DES PROTEINES TOTALES :.....	67
II.1.4.2. ANALYSE DE L'EXPRESSION DE LA HB3GALT6 $\Delta$ NTER30, 34, 43, 50 ET 55 PAR SDS-PAGE : 67	
II.1.4.2.1. ANALYSE DE L'EXPRESSION DES PROTEINES APRES COLORATION DU GEL PAR LE BLEU DE COOMASSIE : .....	68
II.1.4.2.2. ANALYSE DE L'EXPRESSION DES PROTEINES PAR WESTERN BLOT :.....	68





II.1.5. EVALUATION DE L'ACTIVITE GALACTOSYLTRANSFERASE : .....	69
II.2. CLONAGE DES SEQUENCES CODANT LES DIFFERENTES TRONCATIONS DE LA β3GALT6 RECOMBINANTE HUMAINE DANS LES VECTEURS DE TRANSFERT, PFASTBAC-SS : .....	69
II.3. GENERATION D'UN MODELE CELLULAIRE DEFICIENT EN B3GALT6 PAR LA TECHNIQUE CRISPR/CAS9 : .....	70
II.3.1. ENTRETIEN ET CONGELATION DES CELLULES HeLA : .....	70
II.3.2. TRANSFECTION DES CELLULES HeLA SAUVAGE PAR LE PLASMIDE PX458_SGHB3GALT6-2 : ..	71
II.3.3. SELECTION DES CELLULES RECOMBINANTES ET ISOLEMENT DES CLONES CELLULAIRES DEFICIENTS .....	71
II.3.4. CARACTERISATION GENOMIQUE DES CELLULES DEFICIENTE EN hβ3GALT6 : .....	72
II.3.5. ANALYSE DE L'EXPRESSION DES ARNM DE LA hβ3GALT6 PAR RT-PCR QUANTITATIVE : .....	72
II.3.5.1. EXTRACTION DES ARN TOTAUX : .....	72
II.3.5.2. LA RT-PCR QUANTITATIVE : .....	73
II.3.6. OBTENTION DES PLASMIDES CONTENANT LES SEQUENCES NUCLEOTIDIQUES CODANT LES MUTANTS DE LA hβ3GALT6 PAR MUTAGENESE DIRIGEE : .....	73
II.3.6.1. LA MUTAGENESE DIRIGEE : .....	73
II.3.6.2. AMPLIFICATION ET EXTRACTION D'ADN PLASMIDIQUE : .....	74
II.3.7. ANALYSE DE LA SYNTHESE DES GAGs A PARTIR D'UN XYLOSIDE EXOGENE, LE 4- METHYLOMBELLIFERYL-B-D-XYLOPYRANOSIDE (4-MOX) : .....	74
II.3.8. ANALYSE DE L'ACTIVITE GALACTOSYLTRANSFERASE ET DE LA GLYCOSYLATION DE LA DECORINE DANS LES CELLULES DEFICIENTES : .....	75
II.3.8.1. TRANSFECTION DES CELLULES DEFICIENTES AVEC LE PLASMIDE pCDNA_DECORINE ET/OU pCDNA_β3GALT6 : .....	75
II.3.8.2. PURIFICATION DE LA DECORINE, PREPARATION DES HOMOGENATS CELLULAIRES ET DOSAGES DES PROTEINES TOTALES : .....	75
II.3.8.3. ANALYSE DE L'EXPRESSION DE LA hβ3GALT6 SAUVAGE OU MUTEE ET DE LA GLYCOSYLATION DE LA DECORINE PAR WESTERN BLOT : .....	76
II.3.8.4. ÉVALUATION DE L'ACTIVITE GALACTOSYLTRANSFERASE : .....	76
<b><u>RESULTATS ET DISCUSSION .....</u></b>	<b>77</b>
<b>I. CARACTERISATION FONCTIONNELLE DE LA β3GALT6 HUMAINE RECOMBINANTE : .....</b>	<b>78</b>
I.1. TRAVAUX PRELIMINAIRES (RAPPELS DES RESULTATS DE MASTER 2) : .....	78
I.2. OPTIMISATION DES CONDITIONS DE PRODUCTION POUR L'ETUDE DE LA SOLUBILITE DE LA B3GALT6 RECOMBINANTE HUMAINE : .....	79



I.2.1. ETUDE DE LA SOLUBILITE DE LA B3GALT6 RECOMBINANTE HUMAINE EN FONCTION DU PARTENAIRE DE FUSION ET DE LA SOUCHE BACTERIENNE UTILISEE POUR LA PRODUCTION: .....	80
I.2.2. ETUDE DE LA SOLUBILITE DE LA B3GALT6 RECOMBINANTE HUMAINE EN FONCTION DES CONDITIONS DE CULTURE :.....	82
I.2.3. ETUDE DE LA SOLUBILITE DE LA B3GALT6 RECOMBINANTE HUMAINE EN FONCTION DES TRONCATIONS DE LA PROTEINE :.....	82
I.3. PURIFICATION DE LA 6HIS-MBP-HB3GALT6ΔNTER30 : .....	83
I.4. ANALYSE DE L'ACTIVITE ENZYMATIQUE DE LA 6HIS-MBP-B3GALT6 ΔNTER30 RECOMBINANTE HUMAINE :.....	85
I.4.1. MISE AU POINT DU TEST ENZYMATIQUE DE LA 6HIS-MBP-B3GALT6ΔNTER30 :.....	85
I.4.2. ANALYSE DE L'ACTIVITE GALACTOSYLTRANSFERASE EN FONCTION DE LA QUANTITE DE SUBSTRAT ACCEPTEUR : .....	86
I.5. PREPARATION D'UN VECTEUR D'EXPRESSION RECOMBINANT CONTENANT LA SEQUENCE CODANT LA HB3GALT6 POUR LA PRODUCTION EN CELLULES D'INSECTES : .....	86
I.6. PRINCIPAUX RESULTATS ET DISCUSSION DE L'AXE 1 : .....	87
<b>II. GENERATION ET CARACTERISATION D'UN MODELE CELLULAIRE DEFICIENT EN B3GALT6 :..</b>	<b>92</b>
II.1. ANALYSE DE L'EXPRESSION DE L'ARNM DE LA β3GALT6 DES CLONES RECOMBINANTS DE CELLULES HeLa : .....	94
II.2. CARACTERISATION GENOMIQUE DES CELLULES HeLa B3GALT6 DEFICIENTES :.....	94
II.3. ANALYSE DE L'ACTIVITE ENZYMATIQUE DE LA β3GALT6 DANS LES HOMOGENATS CELLULAIRES DES CLONES DEFICIENTS :.....	95
II.4. ANALYSE DE LA SYNTHESE DES GAGS DANS LES LIGNEES HeLa B3GALT6 DEFICIENTES :.....	95
II.4.1. ANALYSE DE LA SYNTHESE DES GAGS A PARTIR D'UN XYLOSIDE EXOGENE, LE 4-METHYLOMBELLIFERYL-B-D-XYLOPYRANOSIDE (4-MOX) :.....	95
II.4.2. ANALYSE DE LA GLYCOSYLATION DE LA PROTEINE CORE DE LA DECORINE : .....	96
II.5. ANALYSE DE LA SUREXPRESSION DES MUTANTS DE LA B3GALT6 DANS LA LIGNEE HeLa B3GALT6 DEFICIENTE :.....	97
II.5.1. ANALYSE DE L'ACTIVITE GALACTOSYLTRANSFERASE DES PROTEINES MUTANTES :.....	97
II.5.2. ANALYSE DE LA GLYCOSYLATION DE LA PROTEINE CORE DE LA DECORINE PAR LES MUTANTS DE LA B3GALT6 :.....	98
II.6. DISCUSSION :.....	99
<b><u>CONCLUSIONS ET PERSPECTIVES .....</u></b>	<b>108</b>
<b><u>BIBLIOGRAPHIE .....</u></b>	<b>112</b>







**LISTES DES  
ILLUSTRATIONS**





## Liste des tableaux

**Tableau 1** – Classification des syndromes d’Ehlers-Danlos selon la classification internationale de 2017

**Tableau 2** – Grille clinique pour l’établissement du score d’hypermobilité articulaire de Beighton

**Tableau 3** – Classification des protéoglycanes en fonction de leur localisation cellulaire

**Tableau 4** – Composition des motifs dissaccharidiques des glycosaminoglycanes

**Tableau 5** – Exemples de ligands et fonctions biologiques des syndécans

**Tableau 6** – Listes des fonctions physiologiques et pathologiques associées à la décorine

**Tableau 7** – Liste des modifications et des enzymes intervenant durant l’élongation et la maturation des chaînes de glycosaminoglycanes

**Tableau 8** – Caractéristiques fonctionnelles des membres de la famille des  $\beta$ 1,3-Galactosyltransférases humaines

**Tableau 9** – Mutations du gène *B3GALT6* et de la protéine décrite chez les différents patients

**Tableau 10** – Fournisseurs et sources biologiques des anticorps utilisés

**Tableau 11** – Séquences nucléotidiques des amorces utilisées pour l’étude de la  $\beta$ 3GalT6 recombinante humaine (h $\beta$ 3GalT6)

**Tableau 12** – Durée et température des différentes étapes du protocole de PCR utilisées pour amplifier la séquence de la h $\beta$ 3GalT6.

**Tableau 13** – Liste des différents paramètres étudiés lors de l’optimisation de la production de la h $\beta$ 3GalT6

**Tableau 14** – Durée et température des différentes étapes du protocole de PCRq utilisées pour amplifier la séquence de la h $\beta$ 3GalT6.

**Tableau 15** – Résumé des caractéristiques des patients atteints de SEDsp *B3GALT6* et des études des mutations de la  $\beta$ 3GalT6 dans la lignée déficiente (clone H7).



## Liste des figures

**Figure 1** – Représentation schématique de l'organisation des composants de la matrice extracellulaire.

**Figure 2** – Photographies présentant les principaux signes cliniques des SED.

**Figure 3** – Analyse biochimique des collagènes dans des fibroblastes de patients atteints du SED de type vasculaire.

**Figure 4** – Modèle de la structure de la  $\beta 4\text{GalT}7$  et position des mutations décrites dans le cas de SEDsp.

**Figure 5** – Représentation schématique des protéoglycanes et des glycosaminoglycanes.

**Figure 6** – Représentation schématique des différents membres de la famille des syndécans et fonctions des différents domaines du protéoglycane.

**Figure 7** – Modèle putatif de l'assemblage des fibres de collagène en présence ou absence de décorine.

**Figure 8** – Représentation schématique de la voie de biosynthèse des glycosaminoglycanes.

**Figure 9** – Représentation schématique de la biosynthèse de l'amorce tétrasaccharidique dans les cellules.

**Figure 10** – Représentation schématique de la biosynthèse et de la maturation des chaînes de CS/DS et Hep/HS.

**Figure 11** – Représentation des différents types de repliement des glycosyltransférases.

**Figure 12** – Représentation schématique de l'organisation modulaire de la  $\beta 3\text{GalT}6$ .

**Figure 13** – Représentation schématique de la réaction catalysée par la  $\beta 3\text{GalT}6$ .

**Figure 14** – Effets de la déficience en  $\beta 3\text{GalT}6$  sur l'organisation des fibres de collagène.

**Figure 15** – Carte du plasmide pET41a.

**Figure 16** – Carte du plasmide pETM41.

**Figure 17** – Carte du plasmide pETM30.

**Figure 18** – Carte du plasmide pFastBac-SS.

**Figure 19** – Carte du plasmide pX458.

**Figure 20** – Représentation schématique des différentes troncations de la h $\beta 3\text{GalT}6$  utilisées pour optimiser la solubilité et la production de la protéine.

**Figure 21** – Analyse de la solubilité de la h $\beta 3\text{GalT}6\Delta\text{Nter}30$  en fonction du partenaire de fusion.

**Figure 22** – Vérification de la solubilité de la 6His-MBP-h $\beta 3\text{GalT}6\Delta\text{Nter}30$  en fonction de la température de production.



**Figure 23** – Analyse de la solubilité des différentes formes tronquées de la 6His-MBP-h $\beta$ 3GalT6.

**Figure 24** – Analyse de la purification de la 6His-MBP- $\beta$ 3GalT6 $\Delta$ Nter30 par chromatographie d'affinité.

**Figure 25** – Analyse de la distribution de la taille des macromolécules présentes dans la fraction d'éluion de la 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 par DLS.

**Figure 26** – Analyse de l'activité enzymatique de la 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 purifiée en fonction de la quantité de substrat accepteur.

**Figure 27** – Chromatogramme de l'activité de la 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30.

**Figure 28** – Modèle de la structure de la h $\beta$ 3GalT6.

**Figure 29** – Modèle moléculaire de la h $\beta$ 3GalT6 sauvage et positions des mutations de la protéine décrites dans les syndromes d'Ehlers-Danlos.

**Figure 30** - Analyse du taux d'expression des ARNm de la  $\beta$ 3GalT6 dans les cellules HeLa suite à l'étape d'inactivation.

**Figure 31** – Analyse de l'activité enzymatique de la  $\beta$ 3GalT6 dans les homogénats cellulaires issus des clones déficients.

**Figure 32** – Analyse de la conséquence de la déficience en  $\beta$ 3GalT6 dans les lignées HeLa KO sur l'anabolisme des GAGs.

**Figure 33** – Analyse de la glycosylation de la protéine *core* de la décorine dans les cellules HeLa contrôle et déficientes, E2 et H7.

**Figure 34** – Analyse de l'expression et de l'activité enzymatique des mutants de la  $\beta$ 3GalT6 exprimés dans les cellules HeLa B3GALT6 déficientes (clone H7).

**Figure 35** – Analyse de la glycosylation de la protéine *core* de la décorine dans les cellules HeLa déficientes transfectées avec la  $\beta$ 3GalT6 sauvage ou mutée.

**Figure 36** – Impact potentiel des mutations de la  $\beta$ 3GalT6 présentes chez les patients sur le GAGosome et la synthèse des GAGs.



## Liste des abréviations

A	Adénine
AD	Autosomique dominante
ADAMTS	<i>A disintegrin and metalloproteinase with thrombospondin motifs</i>
ADN	Acide désoxyribonucléique
ADNc	Acide désoxyribonucléique complémentaire
AH	Acide hyaluronique
AINS	Anti-inflammatoires non stéroïdiens
Ala	Alanine
APS	Ammonium persulfate
AR	Autosomique récessif
Arg	Arginine
ARN	Acide ribonucléique
ARNg	Acide ribonucléique guide
ARNm	Acide ribonucléique messenger
Asp	Aspartate
Asn	Asparagine
β3GalT	β1,3-Galactosyltransférase
β3GalT6	β1,3-Galactosyltransférase 6
β4GalT7	β1,4-Galactosyltransférase 7
C	Cytosine
C4ST	<i>Chondroitin 4-O-Sulfotransferase</i>
C6ST	<i>Chondroitin 6-O-Sulfotransferase</i>
Cas9	<i>CRISPR associated protein 9</i>
CAZy	<i>Carbohydrate Active Enzyme</i>
CHO	Chinese Hamster Ovary
CHST14	<i>Carbohydrate Sulfotransferase 14</i>
ChSy	<i>Chondroitin Synthases</i>
ChGn	<i>Chondroitin GalNAc transferases</i>
ChPF	<i>Chondroitin Polymerizing factor</i>
CLHP	Chromatographie Liquide à Haute Performance
CS	Chondroïtine Sulfate





CSS	<i>Chondroitin Sulfates Synthase</i>
Cys	Cystéine
D4ST	<i>Dermatan 4-O-sulfotransferase</i>
DBSD2	Dysplasie squelettique de Desbuquois de type 2
DLS	<i>Dynamic Light Scattering</i>
DMEM	<i>Dulbecco's Modified Eagle Medium</i>
DMSO	Diméthylsulfoxyde
DO	Densité optique
DES	<i>Dermatan Sulfate epimerase</i>
DS	Dermatane Sulfate
ECCC	<i>European Collection of Cell Cultures</i>
EGFP	<i>Enhanced Green Fluorescent Protein</i>
EXT	Exostosine
EXTL	<i>Exostosin-like glycosyltransferases</i>
Fam20B	<i>Family with sequence similarity 20 member B</i>
FGF	<i>Fibroblast Growth Factor</i>
FGFR	<i>FGF Receptor</i>
G	Guanine
GAG	Glycosaminoglycane
Gal	Galactose
GalNAc	N-acétyl-D-galactosamine
GalNAcT-I	N-acétylgalactosaminyltransférase-I
GalNAcT-II	N-acétylgalactosaminyltransférase-II
GlcA	Acide glucuronique
GlcT-I	Glucuronyltransférase I
GlcAT-II	Acétylglucosyltransférase-II
GlcNAc	N-acétyl-D-glucosamine
GlcNAcT	N-acétylglucaminyltransférase
Glu	Glutamate
Gly	Glycine
GPI	Glycosylphosphatidylinositol
GT	Glycosyltransférase
HA	<i>Human influenza hemagglutinin</i>
HEK293	<i>Human Embryonic Kidney 293</i>



Hep	Héparine
His	Histidine
HP	Hydroxylsypyrindoline
HPRT	Hypoxanthine-guanine phosphoribosyltransférase
HRP	<i>Horseradish Peroxidase</i>
HS	Héparane sulfate
IdoA	Acide Iduronique
IL-1	Interleukine 1
Ile	Isoleucine
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
KS	Kératane sulfates
LB	<i>Luria Bertani</i>
LBNB	<i>Luria Bertani</i> NaCl Bétaine
Lc <sub>3</sub> Cer	Lactosylcéramide
Leu	Leucine
LG	<i>Large globular</i>
LH1	Lysyl hydroxylase 1
LOX	Lysyl Oxidase
LP	Lysylpyridinoline
LRR	<i>Leucine Rich Repeats</i>
MBP	<i>Maltose Binding Protein</i>
MEC	Matrice extracellulaire
Met	Méthionine
OMN	<i>O-Méthoxynaphtyle</i>
PBS	<i>Phosphate-buffered saline</i>
PCR	<i>Polymerase chain reaction</i>
PDGF	<i>Platelet Derived Growth Factor</i>
PDZ	<i>Post synaptic density protein, Drosophila discs large tumor suppressor, Zonula occludens-1 protein</i>
PEI	Polyéthylénimine
PG	Protéoglycane
Phe	Phénylalanine
PLOD1	<i>Procollagen lysine 2-oxoglutarate 5-dioxygénase 1</i>
Pro	Proline



PVDF	PolyFluorure de vinylidene
RE	Réticulum endoplasmique
RPS29	<i>Ribosomal Protein S29</i>
RT-PCR	<i>Reverse transcription Polymerase chain reaction</i>
SAB	<i>Serum albumine bovine</i>
SDS-PAGE	<i>Sodium Dodecyl Sulfate - PolyAcrylamide Gel Electrophoresis</i>
SED	Syndromes d'Ehlers-Danlos
SEDsp	Syndromes d'Ehlers-Danlos spondylo-dysplastique
SEMD-JL1	Dysplasie épimétaphysaire avec hyperlaxité ligamentaire de type I
Ser	Sérine
SLRP	<i>Small Leucine Rich Proteoglycan</i>
SVF	Sérum de veau fœtal
T	Thymine
TBST	<i>Tris Buffer Saline Tween</i>
TEMED	<i>Tetramethylethylenediamine</i>
TFA	Acide trifluoroacétique
TGF- $\beta$	<i>Transforming growth Factor beta</i>
Thr	Thréonine
Trp	Tryptophane
Tyr	Tyrosine
UDP-Gal	Uridine-diphosphate-Galactose
Val	Valine
VEGF	<i>Vascular endothelial growth factor</i>
Xyl	Xylose
XylP	Xylose Phosphate
XylT-I/II	Xylosyltransférase I et -II
2OST	2-O-Sulfotransférase
6OST	6-O-Sulfotransférase
3OST	3-O-Sulfotransférase









# AVANT-PROPOS



Les syndromes d'Ehlers-Danlos (SED) sont un groupe de maladies génétiques rares des tissus conjonctifs touchant principalement les constituants des matrices extracellulaires (MEC) comme le(s) collagène(s) et les protéoglycanes (PGs). Les SED conduisent à des symptômes très divers, plus ou moins sévères, mais associés à des signes cliniques communs comme une peau fragile et élastique, des articulations hypermobiles et une fragilité généralisée des tissus. Dans les premières formes de SED décrites, les patients présentaient des mutations des gènes codant différents types de collagène ou codant les enzymes impliquées dans leur biosynthèse et leur maturation. Plus récemment, d'autres molécules matricielles (en particuliers les PGs) ont été impliquées dans des formes plus rares et sévères. Par exemple, des mutations du gène codant la  $\beta$ 1,3-Galactosyltransférase 6 ( $\beta$ 3GalT6), une glycosyltransférase (GT) impliquée dans l'initiation de la biosynthèse des PGs, ont été décrites et associées à la forme spondylo-dysplastique des SED (SEDsp) et à une autre maladie génétique du tissu conjonctif, la dysplasie spondylo-épimétaphysaire avec hyperlaxité ligamentaire de type 1 (SEMD-JL1), faisant des PGs des molécules pertinentes à étudier dans le contexte de ces pathologies.

Les PGs sont des macromolécules principalement situées dans les MEC et au niveau de la membrane plasmique des cellules. Grâce à leur position stratégique à l'interface cellule-matrice, ils jouent des rôles majeurs dans le maintien de la structure tridimensionnelle et l'organisation fonctionnelle des tissus conjonctifs. Ils sont par exemple particulièrement abondants dans la peau, le cartilage, les os, ou encore la paroi des vaisseaux sanguins. Les PGs sont également impliqués dans de multiples processus physiopathologiques comme la prolifération, la différenciation, la migration cellulaire ou encore la réparation tissulaire grâce à leur importante capacité d'interaction avec des effecteurs solubles comme les facteurs de croissance dont ils peuvent contrôler la biodisponibilité et l'activité biologique. Les PGs sont composés d'un squelette peptidique appelé protéine *core* sur lequel sont fixées de longues chaînes linéaires polysaccharidiques appelées glycosaminoglycanes (GAGs). La biosynthèse des GAGs implique de nombreuses enzymes de la famille des GTs et débute par la synthèse d'une amorce tétrasaccharidique commune aux différents types de GAGs [Acide glucuronique- $\beta$ 1,3-Galactose- $\beta$ 1,3-Galactose- $\beta$ 1,4-Xylose- $\beta$ 1-O-]. Cette dernière servira d'amorce à la polymérisation des différents types de chaînes de GAGs, qui seront simultanément maturés par diverses réactions d'épimérisation et de sulfatation, conduisant à une importante diversité structurale et fonctionnelle de PGs produits et expliquant leurs multiples fonctions cellulaires.

Dans le cadre de ce travail de thèse, je me suis focalisé sur l'étude de la  $\beta$ 3GalT6, l'enzyme responsable du transfert du troisième résidu (galactose) de l'amorce tétrasaccharidique sur la protéine *core* des PGs. Des précédents travaux réalisés au sein de l'équipe, en collaboration



avec le Centre de Génétique Médicale de Gent (Belgique), ont montré que des fibroblastes de derme de peau issus de patients porteurs de mutations du gène *B3GALT6* présentaient un défaut de synthèse des GAGs associé à un retard de migration par rapport à des cellules non porteuses de la déficience génétique (Malfait et al., 2013; Van Damme et al., 2018). Sur la base de ces résultats et des récentes publications impliquant la  $\beta$ 3GalT6 dans des maladies génétiques rares des tissus conjonctifs (comme le SEDsp et la SEMD-JL1), l'objectif de ma thèse a été de mieux comprendre l'implication de cette enzyme dans la pathogénie des SED et d'évaluer les conséquences fonctionnelles de cette déficience au niveau cellulaire.

Dans la première partie de ce travail, j'ai mis en place une étude fonctionnelle de la  $\beta$ 3GalT6 recombinante humaine sauvage après optimisation de la production et de la purification d'une forme tronquée soluble de l'enzyme en système bactérien. En parallèle, une collaboration avec le Dr. François Hoh (UMR 5048 CNRS Centre de Biochimie Structurale, Montpellier) a conduit à la réalisation d'un modèle structural de la protéine qui a permis d'illustrer et de mieux comprendre l'impact des mutations observées chez les patients sur la structure et l'activité de l'enzyme.

Dans un deuxième temps, j'ai généré à l'aide de la technologie CRISPR/Cas9 un modèle cellulaire déficient en  $\beta$ 3GalT6 afin de mieux comprendre l'impact de l'extinction du gène *B3GALT6* dans des cellules humaines, les cellules HeLa, choisies comme modèle de travail. Ce modèle cellulaire *B3GALT6* déficient a été caractérisé sur le plan moléculaire et utilisé pour étudier les conséquences de l'absence d'expression de la  $\beta$ 3GalT6 sur la biosynthèse des GAGs en particulier. De façon complémentaire, ce modèle déficient a permis d'initier l'étude des conséquences de la restauration de l'expression du gène codant la  $\beta$ 3GalT6 (enzyme sauvage et mutants décrits chez les patients) pour analyser l'éventuelle restauration de la biosynthèse des GAGs dans ces cellules génétiquement modifiées. L'objectif de cette partie du projet est de déterminer les conséquences des mutations de l'enzyme décrites chez les patients en termes d'activité enzymatique et d'anabolisme des GAGs.

Ce manuscrit de thèse débutera par une introduction bibliographique présentant la classification et les symptômes cliniques des SED, ainsi que les différents traitements actuellement proposés aux patients. Les mutations des gènes responsables des différents types de SED seront présentées, en particulier celles décrites au niveau des enzymes codant la biosynthèse des GAGs, comme la  $\beta$ 3GalT6, sur laquelle se focalise ce travail. Une partie consacrée plus précisément à la structure et aux fonctions biologiques des PGs (et des GAGs) sera présentée avec une description des processus de biosynthèse et de maturation de ces chaînes polysaccharidiques. La famille des  $\beta$ 1,3-Galactosyltransférases ( $\beta$ 3GalTs), en



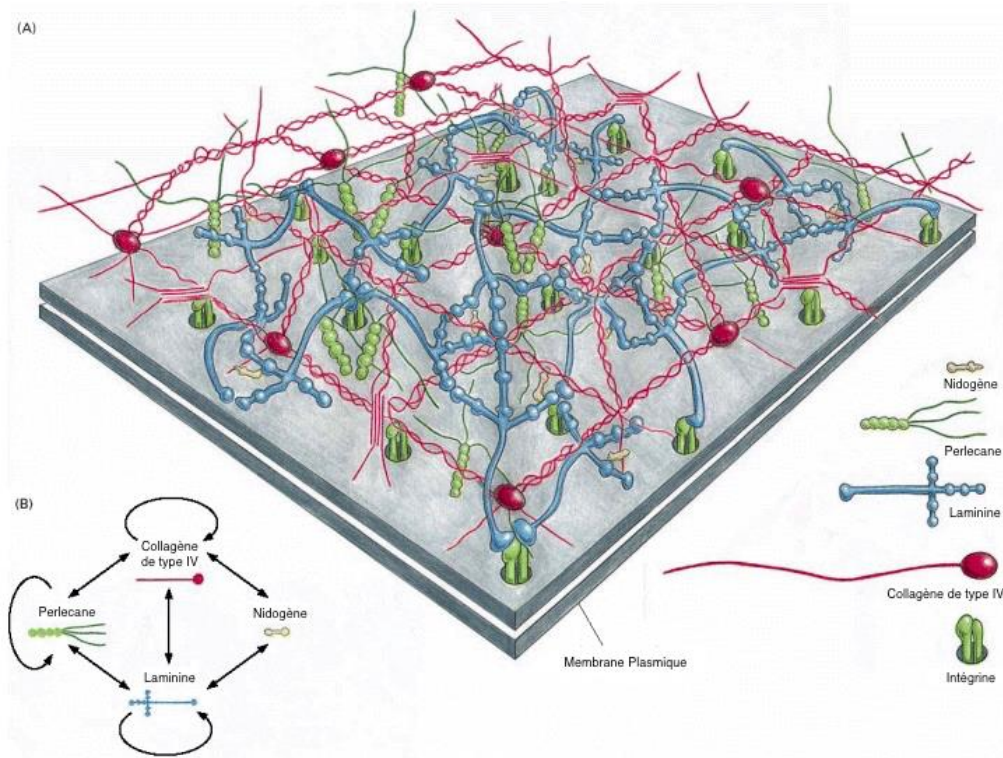
particulier la  $\beta$ 3GalT6, sera également décrite d'un point de vue fonctionnel dans la mesure où aucune information relative à la structure tridimensionnelle de cette enzyme n'est actuellement disponible. La problématique et les objectifs de ce travail seront exposés. Après une description détaillée des protocoles expérimentaux dans la section « Matériels et Méthodes », les résultats obtenus seront présentés en deux parties : une première partie dédiée à l'analyse fonctionnelle de l'enzyme recombinante humaine et une seconde partie présentant l'établissement du modèle cellulaire *B3GALT6* déficient et l'étude de la conséquence de la déficience en  $\beta$ 3GalT6 sur la synthèse des GAGs. Les résultats de chaque partie seront discutés avant de conclure et de présenter les perspectives de ce projet à court, moyen et plus long terme.

L'ensemble des résultats obtenus et à venir devrait permettre de mieux comprendre les relations génotype-phénotype de la pathologie par l'établissement d'un lien entre la nature de la mutation décrite, la gravité des symptômes des patients et l'évolution de la pathologie. Les réponses fournies pourraient permettre également une meilleure compréhension et prédiction de l'évolution de la pathologie, des points qui font défaut aux cliniciens à l'heure actuelle, pour améliorer la prise en charge des patients. A terme, ce travail devrait aussi contribuer au développement de stratégies thérapeutiques visant à s'opposer à la réduction d'anabolisme des GAGs et aux altérations fonctionnelles qui en résultent.





**INTRODUCTION**  
**BIBLIOGRAPHIQUE**



**Figure 1 – Représentation schématique de l’organisation des composants de la matrice extracellulaire (Alberts et al., 2002).**

La MEC est principalement formée de collagènes (protéines fibrillaires), de glycoprotéines (comme la laminine) et de PGs (comme le perlécan) (A). Tous ces constituants vont interagir entre eux afin de donner une structure fonctionnelle au tissu (B). Les constituants de la matrice vont aussi interagir avec d’autres protéines de la membrane plasmique telles que les intégrines, expliquant par exemple l’implication des molécules matricielles dans les phénomènes d’adhésion et de communication intercellulaire.

Dans tous les tissus et les organes, les cellules sont très intimement associées à des composants non cellulaires qui sont principalement synthétisés et sécrétés par les cellules, et qui interagissent les uns avec les autres dans un microenvironnement cellulaire pour former un réseau tridimensionnel, fonctionnel et organisé nommé MEC. La matrice est responsable du maintien de l'homéostasie cellulaire et de sa structure tridimensionnelle en créant un échafaudage moléculaire dynamique pour les cellules. Ce réseau est composé de nombreuses macromolécules différentes qui peuvent être classées en quatre groupes majeurs : (1) les collagènes, (2) les protéines élastiques, (3) les glycoprotéines et (4) les PGs (Figure 1). Une perturbation du métabolisme (biosynthèse ou dégradation) de ces macromolécules peut être à l'origine de situations pathologiques, comme dans le cas des SED dans lesquels peuvent être observés des défauts de synthèse ou de maturation des collagènes ou des PGs.

## *I. Les Syndromes d'Ehlers-Danlos :*

Les SED sont un groupe hétérogène de maladies génétiques touchant principalement les constituants matriciels des tissus conjonctifs, comme la peau, les ligaments, les articulations, la paroi des vaisseaux sanguins ou encore des organes internes. La première description de ce syndrome a été réalisée par deux dermatologues, les docteurs Edward Ehlers et Alexandre Danlos. Edward Ehlers a montré, chez un patient de 21 ans, la présence d'un relâchement articulaire anormal et d'hémorragies cutanées à la suite de traumatismes (Ehlers, 1900). Alexandre Danlos a découvert, chez un jeune patient, la présence de cicatrices atrophiques sur le front, les mains et les jambes issues de traumatismes et suivies de suppuration (Danlos, 1908). Ces premières descriptions de la pathologie illustrent encore aujourd'hui des symptômes cliniques caractéristiques de différents types de SED.

### *I.1. Historique des SED :*

Les premières formes de SED décrites dans la littérature ont montré des mutations sur les gènes codant les collagènes fibrillaires de type I, III et V, ainsi que sur les gènes codant les enzymes impliquées dans la synthèse et la maturation de ces importants constituants matriciels (Tableau 1) (Byers and Murray, 2014; Paepe and Malfait, 2012). A titre d'exemple, des mutations (ponctuelles ou des délétions avec décalage du cadre de lecture) au niveau du gène *COL5A1*, codant la chaîne  $\alpha 1$  du collagène de type V, ont été décrites et conduisent à des symptômes cliniques identiques à ceux décrits pour la forme classique des SED, tels qu'une hyperextensibilité cutanée, des cicatrices atrophiques, une hyperlaxité ligamentaire et une peau

**Tableau 1 – Classification des syndromes d’ Ehlers-Danlos selon la classification internationale de 2017 (Malfait et al., 2017)**

Type des SED	Nomenclature de Villefranche	Principaux symptômes cliniques	Cause génétiques	Mode de transmission	Références
Classique	<i>Classical type</i>	Hyperextensibilité de la peau Cicatrices hypertrophiques Hypermobilité des articulations	<i>COL5A1</i> <i>COL5A2</i> <i>COL1A1</i>	AD	(Symoens et al., 2012) (Ritelli et al., 2013) (Malfait et al., 2007)
Classique-Like	-	Hyperextensibilité de la peau Hypermobilité des articulations	<i>TNXB</i>	AR	(Schalkwijk et al., 2009) (O’ Connell et al., 2010) (Burch et al., 1997)
Cardio-vasculaire	-	Atteintes cardio-vasculaires sévères Hyperextensibilité de la peau Cicatrices hypertrophiques Hypermobilité des articulations	<i>COL1A2</i>	AR	(Guarnieri et al., 2019) (Hata et al., 1988) (Malfait et al., 2006)
Vasculaire	<i>Vascular type</i>	Peau mince et translucide Fragilité ou rupture des artères, de l’ intestin ou de l’ , utérus Contusions étendues Aspect caractéristique de la face	<i>COL3A1</i>	AD	(Germain and Herrera-Guzman, 2004) (Pepin et al., 2009)
Hypermobile	<i>Hypermobility type</i>	Hyperextensibilité variable de la peau Pas de cicatrice hypertrophique Hypermobilité généralisée des articulations	Inconnue	AD	(Levy, 1993) (Syx et al., 2017) (Tinkle et al., 2017)

<b>Types de SED</b>	<b>Nomenclature de Villefranche</b>	<b>Principaux symptômes cliniques</b>	<b>Cause génétiques</b>	<b>Mode de transmission</b>	<b>Références</b>
Arthrochalasique	<i>Arthrochalasia type</i>	Hypermobilité sévère des articulations avec subluxations récurrentes Luxation congénitale bilatérale des hanches	<i>COL1A1</i> <i>COL1A2</i>	AD	(Giunta et al., 2008a) (Klaassens et al., 2012)
Dermatosparaxis	<i>Dermatosparaxis type</i>	Fragilité cutanée grave Ecchymoses étendues Peau laxo et redondante	<i>ADAMTS2</i>	AR	(Colige et al., 1999) (Colige et al., 2004) (Malfait et al., 2004)
Cypho-scoliotique	<i>Kyphoscoliotic type</i>	Articulations lâches Hypotonie musculaire sévère à la naissance Scoliose à la naissance Rupture du globe oculaire	<i>PLOD1</i>	AR	(Rohrbach et al., 2011) (Yeowell and Steinmann, 1993)
<i>Brittle Cornea syndrome</i>	-	Cornée fine avec ou sans rupture Kératocône progressif précoce Kératoglobule progressif précoce	<i>ZNF469</i> <i>PRDM5</i>	AR	(Burkitt Wright et al., 2011) (Davidson et al., 2015) (Wan et al., 2018)
Spondylo-dysplasique	-	Peau hyperextensibilité et fine Petite taille Hypotonie des muscles Scoliose	<i>B4GALT7</i> <i>B3GALT6</i> <i>SLC39A13</i>	AR	(Fayyaz-Ul-Haque et al., 2004) (Van Damme et al., 2018) (Giunta et al., 2008b)

<b>Types de SED</b>	<b>Nomenclature de Villefranche</b>	<b>Principaux symptômes cliniques</b>	<b>Cause génétiques</b>	<b>Mode de transmission</b>	<b>Références</b>
<i>Musculo-contractural</i>	-	<p>Multiplés contractions congénitales</p> <p>Aspect caractéristique de la face</p>	<p><i>CHST14</i></p> <p><i>DES</i></p>	AR	<p>(Janecke et al., 2016)</p> <p>(Malfait et al., 2010)</p> <p>(Syx et al., 2015)</p>
<i>Myopathic</i>	-	<p>Hypotonie congénitale des muscles</p> <p>Contraction des articulations proximales</p>	<i>COL12A1</i>	AD ou AR	<p>(Hicks et al., 2014)</p> <p>(Zou et al., 2014)</p> <p>(Punetha et al., 2017)</p>
<i>Peridontal</i>	-	<p>Plaques pré-tibiales</p> <p>Récession gingivale</p>	<p><i>CIR</i></p> <p><i>CIS</i></p>	AD	<p>(Kapferer-Seebacher et al., 2016)</p> <p>(Wu et al., 2018)</p>

duveteuse (De Paepe et al., 1997; Malfait et al., 2005; Toriello et al., 1996; Wenstrup et al., 1996) (Voir partie I.3). Des observations identiques ont pu être faites dans le contexte des mutations ponctuelles du gène *COL1A1* (codant la chaîne  $\alpha 1$  du collagène de type I), comme la mutation c.934 C>T, p.(Arg312Cys) induisant l'introduction d'une cystéine en position 312 au lieu d'une arginine (Nuytinck et al., 2000). Des mutations de ce même gène peuvent par ailleurs conduire à des formes plus sévères de la maladie et induire des risques vasculaires avec rupture artérielle (Malfait et al., 2007). Par la suite, d'autres mutations ont été décrites au niveau du collagène de type I, plus précisément au niveau du gène *COL1A2* (codant la chaîne  $\alpha 2$  du collagène de type I), et ont été mises en évidence dans le cas de la forme vasculaire des SED. Ces patients présentent une hypermobilité articulaire, une hyperextensibilité de la peau et des atteintes cardio-vasculaires (Malfait et al., 2006; Schwarze et al., 2004).

Des gènes codant des protéines impliquées dans la maturation des collagènes sont également associés aux SED. C'est le cas de la protéine ADAMTS-II, une métalloprotéase impliquée dans le clivage du propeptide N-terminal des procollagènes de type I, II et III, dont la mutation est associée à la forme *dermatosparaxis*, une pathologie rare caractérisée par une importante fragilité cutanée, un aspect hyperlaxe de la peau et des ecchymoses (Colige et al., 1999, 2004; Malfait et al., 2004).

Plus rarement, des mutations au niveau d'autres composants matriciels, comme les PGs, ont été décrites dans des formes rares de SED et seront détaillées dans la deuxième partie. Des mutations au niveau des enzymes intervenant dans les premières étapes de la biosynthèse des GAGs (voir Partie II), comme les gènes *B4GALT7* et *B3GALT6* (détaillées dans la partie I.7 et III.4), ont été associées à la forme SEDsp dans la classification des SED révisée en 2017 (Brady et al., 2017; Malfait et al., 2017).

## **I.2. Classification des SED :**

Les SED sont considérés comme appartenant aux maladies génétiques dites rares, pour lesquelles la pathologie touche moins d'une personne sur 5000 (1/5 000) (Pyeritz, 2000). Selon la classification internationale de 2017, ils sont classés en treize types qui se distinguent par la nature, le mode de transmission, les symptômes cliniques et les défauts génétiques à l'origine de la pathologie lorsqu'ils sont identifiés (Tableau 1) (Malfait et al., 2017).

- (i) Le type classique (prévalence 1/20 000) est caractérisé par une hyperextensibilité cutanée, des cicatrices atrophiques, une hyperlaxité ligamentaire et une peau duveteuse. Les causes génétiques de ce type de SED sont des mutations des gènes codant les collagènes (majoritairement les collagènes de type V et plus rarement





ceux du type I) (Ritelli et al., 2013; Symoens et al., 2012). Les collagènes de type V jouent un rôle important dans la nucléation des fibrilles de collagène de type I (collagènes majeurs de la peau). Les mutations observées chez les patients induisent une diminution de la synthèse du collagène de type V et provoquent une diminution de la formation des fibrilles de collagène de type I. Cette altération de la nucléation des fibrilles de collagène impacte directement la structure de la MEC et la fonctionnalité des tissus.

- (ii) Le type classique-*like* (prévalence inférieure à 1/1 000 000) présente des symptômes proches du type classique avec comme particularités une hyperextensibilité cutanée avec absence de cicatrices atrophiques et une hyperlaxité ligamentaire avec des dislocations récurrentes (Burch et al., 1997; O'Connell et al., 2010; Schalkwijk et al., 2009). La maladie est due à un défaut en ténascine X, une glycoprotéine qui joue un rôle important dans la fibrillogenèse des collagènes dans les tissus conjonctifs. La ténascine X est capable de ponter les fibrilles de collagène en interagissant avec les fibres de collagène par l'intermédiaire de la décorine, un petit PG matriciel. Chez les patients atteints de ce type de SED, les fibrilles de collagène ne s'assemblent plus en fibres contrairement aux personnes saines.
- (iii) Le type cardio-valvulaire (prévalence inférieure à 1/1 000 000) est caractérisé par des risques cardio-vasculaires sévères, une peau hyperextensible et des articulations hypermobiles. Des études ont montré que des mutations au niveau du gène *COL1A2*, codant la chaîne  $\alpha 2$  du collagène de type I entraînent une réduction ou une absence totale de la chaîne  $\alpha 2$  du collagène de type I chez les patients atteints de cette forme de SED (Guarnieri et al., 2019; Hata et al., 1988; Malfait et al., 2006). Cette absence induit la formation d'un homotrimère de collagène de type I (composé de trois chaînes  $\alpha 1$ ) au lieu d'un hétérotrimère (deux chaînes  $\alpha 1$  et une chaîne  $\alpha 2$ ). Toutefois il n'y aurait pas d'impact direct sur la structure de l'homotrimère et donc sur la structure des fibres de collagène (Sharma et al., 2017).
- (iv) Le type vasculaire (prévalence estimée entre 1/100 000 – 1/250 000) se caractérise par une fragilité des parois artérielles, du tube digestif et de l'utérus, de multiples ecchymoses et une peau fine et translucide. Chez les patients, des mutations du gène *COL3A1*, codant la chaîne  $\alpha 1$  du collagène de type III, ont été décrites. Il a pu être observé chez certains patients l'absence de sécrétion du procollagène de type III, ce dernier s'accumulant au niveau du réticulum endoplasmique (RE) rugueux. Actuellement, le mécanisme qui induit cette accumulation n'est pas encore connu.



Au final, chez ces patients, une absence de sécrétion du procollagène de type III induit une diminution de la quantité de fibres de collagène III dans la MEC (Byers and Murray, 2014; Germain and Herrera-Guzman, 2004; Pepin et al., 2009).

- (v) Le type hypermobile (prévalence 1 à 5/10 000) est caractérisé par une hypermobilité articulaire associée à des luxations, une hyperextensibilité cutanée, l'absence de cicatrices hypertrophiques et de multiples autres symptômes, tels que la fatigue, des troubles de la proprioception, des douleurs chroniques (Levy, 1993; Syx et al., 2017; Tinkle et al., 2017). A l'heure actuelle, les bases moléculaires de cette forme de SED ne sont pas connues et aucun gène cible n'a pu être directement associé à cette forme.
- (vi) Le type arthrochalasique (très rare, seulement 30 cas dans le monde) présente une hypermobilité articulaire généralisée sévère, des subluxations (dislocations partielles) et des luxations récidivantes, une hyperextensibilité cutanée et une fragilité tissulaire. Cette forme est caractérisée par la présence de mutations au niveau du gène *COL1A1* ou *COL1A2* (Giunta et al., 2008a; Klaassens et al., 2012) qui induisent un défaut de synthèse de la chaîne de collagène mutée avec d'importantes répercussions sur la formation, la structure et la fonction du collagène.
- (vii) Le type *dermatosparaxis* (prévalence inférieure à 1/1 000 000) est caractérisé par une fragilité cutanée sévère, des ecchymoses étendues et une peau laxo et redondante. Dans cette forme, des mutations du gène *ADAMTS2* codant l'enzyme *procollagen I N-proteinase* (Colige et al., 1999, 2004; Malfait et al., 2004) entraînent l'accumulation d'une forme immature du collagène de type I (un collagène présentant encore son domaine N-terminal) qui induit une désorganisation des fibres de collagène.
- (viii) Les patients atteints du type cypho-scoliotique (prévalence inconnue) présentent des articulations lâches, une hypotonie musculaire sévère, une scoliose dès la naissance, s'aggravant au cours de la vie et dans certains cas une rupture du globe oculaire. Les mutations décrites chez les patients sont présentes sur le gène *PLOD1* (*procollagen lysine 2-oxoglutarate 5-dioxygenase 1*) qui code une enzyme catalysant l'hydroxylation des lysines des chaînes de collagène, modification post-traductionnelle importante pour l'établissement des liaisons de pontage entre les fibrilles de collagène (Rohrbach et al., 2011; Yeowell and Steinmann, 1993). Chez ces patients, ces mutations induisent une réduction ou une absence de l'hydroxylation des résidus lysine présents dans les collagènes, qui conduit à une



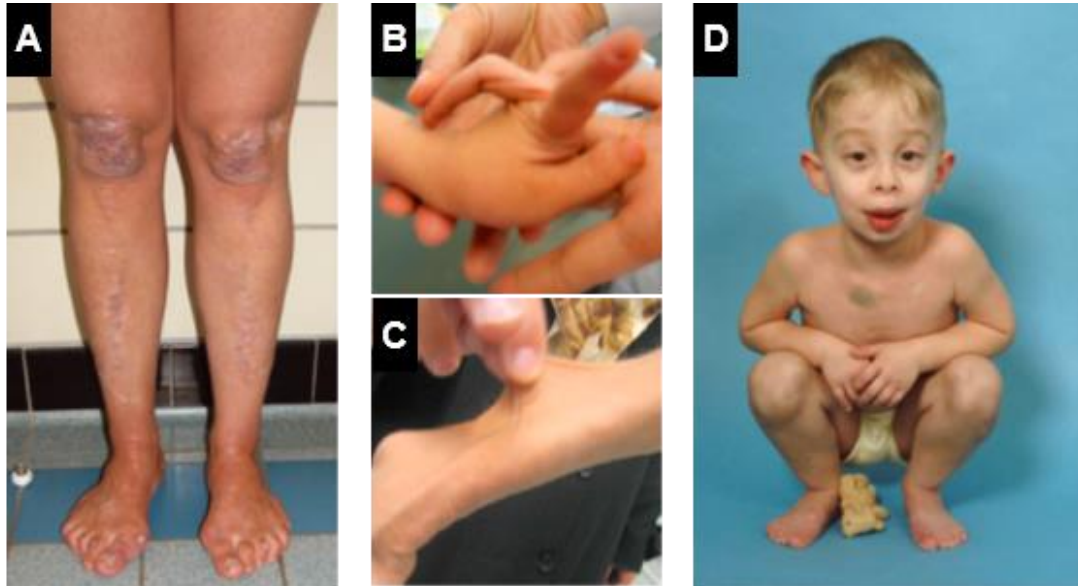
augmentation du diamètre des fibrilles et à une perte de leur forme circulaire qui impactent l'architecture des tissus.

- (ix) Le *Brittle Cornea Syndrome* (prévalence inférieure à 1/1 000 000) est un syndrome caractérisé par une cornée fine présentant ou non des ruptures, un kératocône (perte de la sphéricité de la cornée) et un kératoglobus (amincissement de la cornée dans la zone périphérique ou centrale) qui sont progressifs et précoces (normalement détectés avant 40 ans). Les patients atteints par cette forme de SED présentent des mutations des gènes *ZNF469* et *PRMD5* (Burkitt Wright et al., 2011; Davidson et al., 2015; Micheal et al., 2019; Wan et al., 2018). Ces deux gènes codent pour des facteurs de transcription régulant la synthèse des collagènes ou leur structuration. La régulation de ces gènes se fait par le recrutement d'histones méthyltransférases et/ou d'histones déacétylases qui induisent la répression de l'expression transcriptionnelle de certains gènes, parmi lesquels ceux codant les chaînes des collagènes. La diminution de l'expression de certains collagènes peut induire une fragilité de la cornée et l'apparition des symptômes observés chez ces patients.
- (x) Le type SEDsp (très rare, environ 70-100 patients dans le monde) présente des symptômes divers parmi lesquels une peau duveteuse, fine et hyperextensible, une hypermobilité articulaire généralisée, une petite taille, une hypotonie des muscles, les pieds plats, une dysplasie squelettique, un retard du développement moteur et parfois mental et de l'ostéopénie. On retrouve également des caractéristiques faciales telles qu'une sclérotique bleue (blanc de l'œil bleuté), des fentes palpébrales inclinées, une hypoplasie médio-faciale ou encore des oreilles basses. Il a été montré que les gènes *B4GALT7* (codant la  $\beta$ 1,4-Galactosyltransférase 7,  $\beta$ 4GalT7), *B3GALT6* (codant la  $\beta$ 3GalT6) et *SLC39A13* (codant un transporteur de zinc) sont impliqués dans ce sous-type de SED (Faiyaz-Ul-Haque et al., 2004; Giunta et al., 2008b; Van Damme et al., 2018). Les implications de ces mutations dans la pathogénie des SED seront détaillées dans les parties I.7 et III.4).
- (xi) Le type musculo-contractural (prévalence inférieure à 1/1 000 000) est caractérisé par de multiples contractures et un aspect caractéristique de la face (large fontanelle, hypertélorisme, fente palpébrale courte et inclinée, blanc de l'œil de couleur bleue au lieu de blanche, petite bouche). La présence de mutations au niveau des gènes *CHST14* (codant la *Carbohydrate Sulfotransferase 14*, une enzyme de maturation des GAGs responsable du transfert d'un groupement sulfate en position 4 sur les résidus N-acétylgalactosamine (GalNAc) des chaînes de dermatane sulfate



(DS)) et *DES* (codant la *Dermatane-Sulfate Epimerase*, une autre enzyme de maturation des GAGs responsable de l'épimérisation de l'acide glucuronique en acide iduronique au niveau des chaînes de DS) a été décrite chez ces patients. Elle induit une altération de la composition et de la structure fine des chaînes de DS qui ne sont pas maturées de façon correcte (Janecke et al., 2016; Malfait et al., 2010; Syx et al., 2015). L'absence de maturation au niveau de la chaîne de GAG de la décorine, un petit PG contenant une chaîne de chondroïtine sulfate (CS) ou de DS (voir partie II.2.2.2), a un impact direct sur l'ultrastructure des fibres de collagène.

- (xii) Le type myopathique (prévalence inférieure à 1/1 000 000) présente une hypotonie congénitale des muscles et une contraction des articulations proximales. Des mutations au niveau du gène de la chaîne  $\alpha 1$  du collagène XII (*COL12A1*) ont été décrites (Hicks et al., 2014; Punetha et al., 2017; Zou et al., 2014). Le collagène de type XII est un collagène fibrillaire impliqué dans les interactions avec le collagène de type I et d'autres molécules matricielles, telles que la ténascine X. Ces interactions avec le collagène de type XII permettent de réguler l'organisation et les propriétés mécaniques des fibres de collagène (Chiquet et al., 2014). Chez les patients atteints de SED myopathique, une réduction de l'expression des collagènes XII a été observée, induisant une altération de l'organisation des fibres de collagène et de leurs propriétés mécaniques.
- (xiii) Le type péri-odontal (prévalence inférieure à 1/1 000 000) est caractérisé par une périodontite sévère et incurable, un défaut de l'adhésion des gencives, des ecchymoses fréquentes, une hypermobilité articulaire, une hyperextensibilité de la peau, une augmentation du taux d'infections ou encore des hernies. Des études ont montré l'implication de mutations au niveau des gènes *C1R* et *C1S*, des gènes codant des enzymes protéolytiques impliquées dans les premières étapes de l'activation de la voie classique du système du complément, un composant du système immunitaire inné (Kapferer-Seebacher et al., 2016 ; Wu et al., 2018). Les protéines C1r et C1s forment un hétérotétramère (2C1s/2C1r) capable de se lier au domaine N-terminal des chaînes  $\alpha 1$  et  $\alpha 2$  des procollagènes de type I et des chaînes  $\alpha 2$  des procollagènes de type III. L'établissement de modèles structuraux des protéines C1r et C1s a permis de déterminer les régions interagissant avec les chaînes de procollagènes (Kapferer-Seebacher et al., 2016). Grâce à ces modèles, les mutations présentes chez les patients ont été localisées et sont retrouvées dans les régions impliquées dans les interactions avec les chaînes de procollagène. Même si le rôle de ces interactions



**Figure 2 – Photographies présentant les principaux signes cliniques des SED.**

(A) Patient atteint de la forme classique de la maladie présentant des cicatrices larges et atrophiques sur les genoux et les tibias, des pieds plats avec *hallux valgus*. (B et C) Patient atteint du SEDsp présentant une hyperlaxité articulaire (B) et une peau fine et hyperextensible (C) (Malfait et al., 2013). (D) Patient atteint de la forme *dermatosparaxis* qui présente un aspect typique du visage avec épicanthus (un aspect bridé des yeux), des fentes palpébrales inclinées, une sclérotique bleue et des cicatrices du visage en particulier autour de la bouche (Paepe and Malfait, 2012).



n'est pas encore établi à l'heure actuelle, il est possible que leur altération soit à l'origine des effets délétères observés sur l'architecture du collagène.

### **1.3. Symptômes des SED :**

Les SED conduisent à des symptômes très divers, de sévérité plus ou moins importante mais présentant des signes cliniques communs (Figure 2). Une des manifestations les plus typiques des SED est une peau hyperextensible dotée d'une élasticité importante (elle s'étend aisément et reprend sa forme initiale après relâchement) (Figure 2B). La peau des patients est fine avec des vaisseaux apparents et présente un aspect lisse et velouté au toucher. Dans le cas du sous-type vasculaire, la peau n'est pas hyperextensible, mais mince et transparente avec des veines apparentes (Byers, 1993). La peau des patients présente une fragilité et une facilité à se déchirer après un traumatisme même léger, ce qui conduit à la formation d'ecchymoses et de cicatrices atrophiques (Xu et al., 2019).





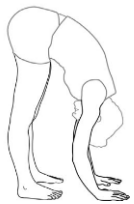
L'hypermobilité des articulations est généralisée et variable en intensité et en fonction de l'âge des patients et du type de SED (Figure 2B). Elle est souvent asymptomatique dans l'enfance et l'adolescence, mais peut devenir sévère avec le temps. Elle se caractérise par des complications articulaires associées à des subluxations répétitives, des entorses et des douleurs articulaires chroniques. L'hypermobilité articulaire est estimée au moyen du score de Beighton, un système d'évaluation semi-quantitatif (Tableau 2) basé sur une série de neuf manipulations cliniques. Le diagnostic d'un SED est retenu pour un score supérieur ou égal à 5 points (Beighton et al., 1998). Lors de ce diagnostic, il est important de tenir compte de l'âge des patients, dû au fait que la souplesse des enfants est plus importante que celle des adultes.

L'hypotonie musculaire peut contribuer à un retard du développement moteur, avec des difficultés à l'apprentissage de la marche et une perturbation de la motricité. Les patients sont fréquemment sujets à des ecchymoses, aboutissant spontanément à la formation d'hématomes et conduisant souvent au dépôt d'hémosidérine (un pigment insoluble présent dans le sang et contenant de l'hydroxyde ferrique) dans les zones exposées telles que les tibias et les genoux (Paepe and Malfait, 2004). Ces patients ont tendance à présenter des saignements prononcés, par exemple suite à un simple brossage de dents, en dépit d'une coagulation normale (Paepe and Malfait, 2004, 2012).

### **1.4. Modes de transmission des SED :**

Les SED possèdent différents modes de transmission dépendant de la forme de la maladie. Trois types de transmission sont répertoriés :

**Tableau 2 – Grille clinique pour l'établissement du score d'hypermobilité articulaire de Beighton (Beighton et al., 1998; Malfait et al., 2017).**

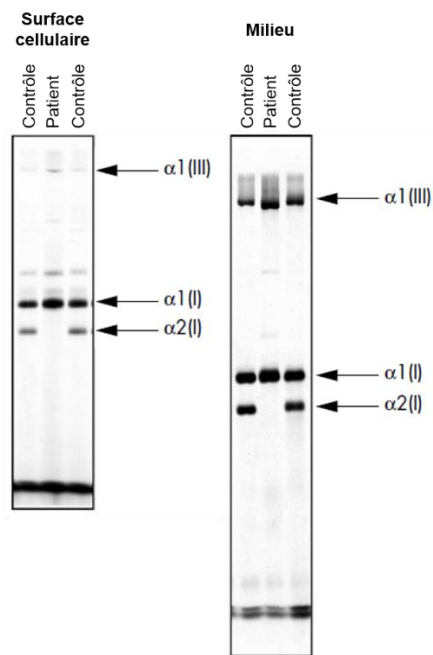
Articulation/Signe	Illustration	Négatif	Positif	
			Unilatéral	Bilatéral
Dorsiflexion passive du 5 <sup>ème</sup> doigt > 90 °		0	1	2
Flexion passive du pouce vers l'avant-bras		0	1	2
Hyperextension du coude > 10 °		0	1	2
Hyperextension des genoux > 10 °		0	1	2
Flexion complète du tronc, genoux complètement étendus, avec les paumes complètement sur le sol		0	1	
Score positif si ≥ 5				

- La transmission selon un mode autosomique dominant (AD) est une transmission de la maladie où le gène déficient est porté par un chromosome autosome de l'un ou l'autre des parents. Dans ce cas, la présence d'un seul allèle muté suffit pour que la maladie se développe. Cette forme de transmission est majoritaire parmi les SED (comme par exemple pour le type classique) (Malfait et al., 1993).
- Si le SED suit un mode de transmission autosomique récessif (AR), les deux allèles déficients du gène sont transmis l'un par le père, l'autre par la mère pour qu'il y ait une expression clinique de la maladie (comme par exemple pour le type spondylo-dysplastique). Dans ce cas, les parents sont eux-mêmes des porteurs sains (hétérozygotes) (Hanset and Lapiere, 1974; Van Damme et al., 2016). **Dans certains cas, un des deux parents peut transmettre un allèle sain qui va subir une mutation. L'apparition de cette néomutation peut rendre l'allèle déficient et la présence d'un autre allèle déficient issu du second parent va alors induire l'expression clinique de la maladie sans que l'un des deux parents soit porteur.**
- Plus rarement, le gène déficient peut être porté par un chromosome sexuel, comme c'est le cas, par exemple, du SED associé à des mutations du gène *FLNA* (codant la filamine A, une protéine qui se fixe sur l'actine et stabilise sa structure) avec une hétérotopie nodulaire périventriculaire, une pathologie rare qui se manifeste dans l'enfance par des troubles neurologiques importants liés à la perturbation de la migration des neurones lors de la formation du système nerveux du fœtus. Cette pathologie présente un mode de transmission dominant lié au chromosome X maternel. La probabilité d'avoir des enfants atteints de cette maladie dès la naissance est de 50% (Gómez-Garre et al., 2006). Dans cette forme de SED, la mutation du gène *FLNA* induit une mortalité prénatale chez l'homme et avec un risque de 50% chez la femme (Guerrini and Carrozzo, 2001; Guerrini and Dobyns, 1998).

### **1.5. Diagnostic des différents types de SED :**

En fonction du sous-type de SED, le diagnostic s'appuiera sur toutes, ou une partie, des analyses suivantes : (i) l'analyse des signes cliniques, en particulier l'hypermobilité articulaire, (ii) l'imagerie médicale, (iii) l'analyse structurale de la peau, (iv) l'analyse moléculaire des collagènes (par SDS-PAGE), (v) le séquençage du génome.

Le diagnostic initial repose principalement sur une évaluation clinique. Le médecin établira son diagnostic en analysant les symptômes du patient et ses antécédents familiaux et pathologiques. En sachant que les symptômes des différentes ostéochondrodysplasies sont souvent similaires,



**Figure 3 – Analyse biochimique des collagènes dans des fibroblastes de patients atteints du SED de type vasculaire (Malfait et al., 2006).**

Les analyses des collagènes de la surface cellulaire et du milieu de culture sont réalisées par SDS-PAGE (*Sodium DodecylSulfate PolyAcrylamide Gel Electrophoresis*). Chez un patient atteint de SED, l'absence de la chaîne  $\alpha 2$  du collagène de type I indique un SED de type vasculaire.

ces premières étapes sont cruciales pour établir un diagnostic correct du type de pathologie et du sous-type de SED.

Parfois utilisées pour aider au diagnostic différentiel des sous-types de SED, des analyses structurales de la peau et les analyses biochimiques et moléculaires des tissus sont réalisées. L'analyse de la peau peut être effectuée par microscopie électronique afin de vérifier la structure et l'organisation des fibrilles de collagène. Ces dernières prennent, dans le cas des SED, un aspect irrégulier et désordonné avec une variation de leur diamètre (augmentation dans le cas de SED cypho-scoliotique) (Rohrbach et al., 2011). Cependant, ces anomalies sont assez fréquentes et communes aux différents sous-types de SED et ne permettent pas de distinguer de façon précise un sous-type de SED par rapport à un autre. Il faut rester conscient que ces analyses ne sont pas des analyses de routine. L'analyse des collagènes de type I, III et V par SDS-PAGE permet de montrer des anomalies qualitatives ou quantitatives des collagènes dans les fibroblastes de derme issus de patients (Figure 3) (Malfait et al., 2006). Ces analyses sur le collagène III couplées avec le génotypage du gène *COL3A1* peuvent par exemple confirmer le diagnostic d'un SED de type vasculaire (Jørgensen et al., 2015).

Dans le cas du SED de type cypho-scoliotique, qui est caractérisé par des mutations du gène *PLOD1* (Yeowell and Steinmann, 1993), une déficience en *lysyl hydroxylase 1* (LH1) est observée (Voir partie I.2). Chez les patients atteints de ce sous-type de SED, une augmentation du rapport lysylpyridinoline (LP) / hydroxylysylpyridinoline (HP), qui sont les produits de la protéolyse du collagène et qui sont excrétés dans les urines, est décrite. La mesure de ce rapport (par chromatographie liquide à haute performance (CLHP)) est un test sensible pour le diagnostic de cette forme de SED (Rohrbach et al., 2011). L'imagerie médicale est une technique complémentaire importante dans le diagnostic. Dans le cas des SED classiques, une échocardiographie permet de déceler des troubles cardiaques. Dans la forme vasculaire, il est souvent proposé de faire une échographie doppler artérielle ou une angio-IRM cervico-encéphalique de l'aorte abdominale et de ses branches pour vérifier l'absence de formation d'anévrisme (Eagleton, 2016).

En fonction des risques liés aux antécédents familiaux, il peut être conseillé de réaliser un diagnostic prénatal sur l'ADN des cellules fœtales afin de rechercher une éventuelle anomalie génétique.

Actuellement, le seul moyen de diagnostiquer de manière définitive et précise un sous-type de SED est de procéder à l'analyse d'un panel de gènes du patient. Ces analyses permettent d'identifier le (les) gène(s) muté(s) et de déterminer la forme de SED dont le patient est atteint.

Pour déterminer les différents gènes impliqués dans les SED, il est tout d'abord réalisé un



séquençage d'un panel de gènes sélectionnés sur la base d'un tableau clinique. Si les résultats ne sont pas concluants, des analyses par séquençage d'exome sont menées. Les techniques actuelles d'analyse haut débit permettent de réaliser le séquençage simultané de plusieurs gènes (environ 20 000 gènes) dans un délai assez court. Ainsi, de nouvelles mutations de gènes ont pu être identifiées et associées à différents sous-types de SED de façon spécifique (Delbaere et al., 2019a; Syx et al., 2019). A l'heure actuelle, la recherche des gènes impliqués dans les SED est en plein développement afin d'améliorer de façon continue le diagnostic moléculaire de la maladie et, le cas échéant, la prise en charge du patient. Les médecins espèrent en particulier pouvoir identifier le (les) gène(s) impliqué(s) dans la forme hypermobile des SED grâce aux dernières évolutions des techniques de séquençage haut débit et aux approches de type « omiques » telles que la transcriptomique.

### **I.6. Traitements des SED :**

Actuellement, il n'existe aucun traitement curatif pour les SED. Ceux proposés aux patients sont uniquement symptomatiques et préventifs et comprennent la physiothérapie, l'administration d'antalgiques et d'anti-inflammatoires, la prévention des troubles cardiovasculaires, l'adaptation du mode de vie ainsi qu'un suivi psychologique.

#### **I.6.1. Les traitements préventifs :**

Les patients qui présentent des ecchymoses multiples sont encouragés à éviter les sports violents et de contact. Ils sont également incités à éliminer les médicaments interférant avec les mécanismes de l'hémostase, comme l'aspirine. Il a été suggéré que des compléments alimentaires, tels que l'acide ascorbique (vitamine C), puissent être administrés à certains patients atteints de SED (Mantle et al., 2005). La vitamine C est connue pour jouer un rôle important dans la maturation des collagènes. Il a pu être mis en évidence que la supplémentation en vitamine C améliore la cicatrisation, en plus de la synthèse des collagènes (Ringsdorf and Cheraskin, 1982; Shashikiran et al., 1999). Cet effet au niveau tissulaire de la vitamine C provient de son rôle en tant que cofacteur de la prolyl-4-hydroxylase, une enzyme impliquée dans l'hydroxylation des résidus proline du collagène et assurant la formation des liaisons croisées entre les fibrilles de collagène (Jong and Kemp, 1984; Vasta and Raines, 2016). Ainsi, la supplémentation en vitamine C pourrait améliorer la fibrillogénèse des collagènes.

Selon la gravité des symptômes, des orthèses (mitaines, shorts, genouillères de contention) peuvent être utilisées, pour limiter les pressions trop fortes sur la peau et prévenir les entorses, les luxations et les ecchymoses (Castori, 2012). Dans les cas les plus sévères, une adaptation de





l'environnement quotidien s'avère utile pour améliorer la qualité de vie. Pour tous les patients atteints de SED, il est important que le médecin traitant soit sensibilisé aux risques éventuels des gestes invasifs et aux possibilités de complication, liée à la fragilité tissulaire, à la diversité et la complexité des symptômes des patients.

Dans le cas des femmes enceintes atteintes de SED, en particulier celles atteintes de la forme vasculaire, il existe un risque plus élevé de rupture utérine et/ou de rupture artérielle pouvant entraîner le décès de la patiente (risque d'environ 11%). Il est important que le suivi clinique soit très étroit et réalisé dans des centres spécialisés (Pepin et al., 2009; Rudd et al., 1983).

### *1.6.2. Les traitements médicamenteux :*

Les traitements pharmacologiques proposés aux patients atteints de SED sont uniquement symptomatiques, visant à soulager le patient et améliorer sa qualité de vie. Les douleurs associées aux SED sont principalement des douleurs articulaires, qui peuvent être traitées par des anti-inflammatoires non stéroïdiens (AINS, comme l'ibuprofène) et/ou antalgiques (paracétamol) (Castori, 2012). D'autres traitements médicamenteux peuvent être employés en complément ou non des précédents incluant par exemple d'autres AINS (comme le célécoxib) ou des opiacés (tramadol, codéine). Ces derniers sont efficaces contre les douleurs chroniques importantes que subissent certains patients atteints de SED, mais ces traitements ne sont pas efficaces chez tous les patients.

### *1.6.3. La physiothérapie et les traitements psychologiques :*

Dans les cas de patients atteints d'hypotonie musculaire associée à un retard du développement moteur, un suivi par physiothérapie est indispensable. Le traitement comprend une rééducation douce, passive et qui implique, de préférence, de la natation. Ces traitements passifs ont pour principaux rôles d'aider à soulager la douleur et à lutter contre la perte de sommeil associée à la douleur.

En plus de ces traitements, des soutiens psychologiques et comportementaux peuvent être envisagés pour accompagner tous les patients atteints de SED. Un accompagnement avec une mise en relation avec des associations de malades, comme l'Union Nationale des Syndromes d'Ehlers-Danlos (UNSED) ou l'Association Française des Syndromes d'Ehlers-Danlos (AFSED), peut également être proposé. Ces associations assurent un soutien psychologique au patient, et accompagnent également les familles en charge d'un ou plusieurs enfants atteints par la maladie. Elles interviennent également dans l'aide à la prise en charge de certains soins et organisent de nombreuses actions (organisation de rencontres avec des spécialistes de la



maladie) afin d'aider à financer l'aide aux malades.

### **I.7. Le SEDsp associé à une déficience en B4GALT7 :**

Outre les mutations affectant les enzymes de la biosynthèse et de la maturation des collagènes, d'autres formes de SED ont été décrites. Les gènes impliqués dans ces formes plus rares et souvent sévères, touchent d'autres composants matriciels, comme par exemple les GAGs dont les enzymes de biosynthèse ou de maturation sont mutées et exprimées sous des formes peu ou non fonctionnelles. Les patients atteints de ces formes de SED présentent les symptômes typiques des SED (voir partie I.3), mais également des caractéristiques cliniques différentes et variables selon les formes décrites comme une petite taille, des malformations squelettiques, une hypotonie musculaire, des contractures bilatérales ou des mouvements limités du coude, une hypermobilité articulaire généralisée et un aspect caractéristique de la face (Kresse et al., 1987; Malfait et al., 2017).

En 1987, Kresse et al. (1987) décrivent une forme de SED dite « progéroïde » en raison de l'aspect vieillissant de leur peau (renommée SEDsp dans la nouvelle classification proposée en 2017 - Tableau 1 - Brady et al., 2017). Ils observent, dans les fibroblastes de derme de ces patients, un défaut de synthèse et de glycanation de la protéine *core* de la décorine, un SLRP (*Small Leucine Rich Proteoglycan*) qui est un petit PG portant une seule chaîne de CS/DS (Kresse et al., 1987).

Plus tard, deux mutations ponctuelles du gène *B4GALT7*, localisé sur le chromosome 5 au locus 5q35.2 et codant la  $\beta$ 4GalT7 (une GT qui intervient dans les premières étapes de la biosynthèse des GAGs (voir partie II.3)), ont été identifiées chez ces patients. Il s'agit de la substitution d'une alanine par un aspartate en position 186 (Ala186Asp) et d'une leucine par une proline en position 206 (Leu206Pro) (Götte and Kresse, 2005 ; Salter et al., 2016). Une autre mutation ponctuelle (homozygote) de la protéine avec le remplacement d'une arginine par une cystéine en position 270 (Arg270Cys) a été rapportée chez des patients atteints de cette même forme de SED (Faiyaz-UI-Haque et al., 2004). Plus récemment, d'autres mutations du gène *B4GALT7* ont été également décrites comme la substitution d'une leucine en proline en position 41 (Leu41Pro) et une mutation avec décalage du cadre de lecture à partir de l'acide glutamique en position 277 (p.Glu277\*) associé à un épissage aberrant de l'ARNm (Guo et al., 2013; Ritelli et al., 2017; Salter et al., 2016).

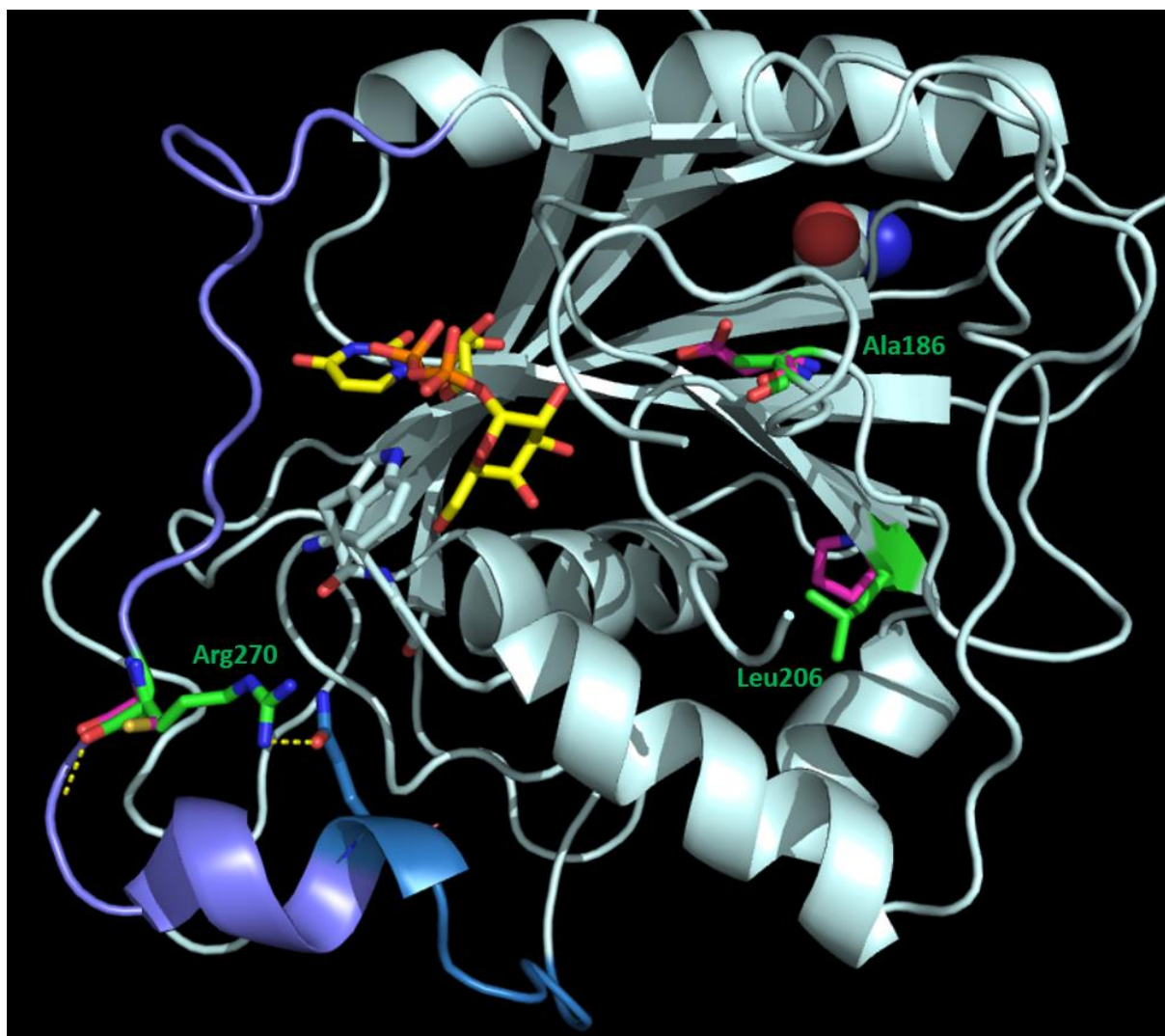
La mutation Arg270Cys a été également associée au syndrome de Larsen chez des patients habitant l'île de la Réunion (Cartault et al., 2015). Ce syndrome est caractérisé par un nanisme, une hyperlaxité, de multiples dislocations et une apparence crânio-faciale caractéristique (front



large, proptosis et bouche étroite). Plus récemment, des mutations du gène *B4GALT7* ont également été décrites chez deux autres patients (sans lien de parenté) présentant une variante du syndrome de Larsen (Salter et al., 2016). Pour le premier patient, la présence d'une mutation faux-sens induit le changement d'une cystéine en tyrosine à la position 214 (Cys214Tyr) et d'une mutation de l'histidine 93 en proline avec l'apparition d'un décalage du cadre de lecture (His93Pro *fs\*73*). Le second patient présente par ailleurs la mutation Arg270Cys, précédemment décrite par Seidler et al. (2006) et une mutation ponctuelle qui induit la substitution du résidu arginine en tryptophane en position 141 (Arg141Trp). Ces deux patients ont montré des symptômes très proches des patients atteints du syndrome de Larsen et du SEDsp.

Des études récentes ont essayé de comprendre l'implication du gène *B4GALT7* dans le développement des SED. Dans ce contexte, les mutations faux-sens observées sont principalement localisées dans ou à proximité de domaines conservés de l'enzyme. Ces motifs conservés sont impliqués soit dans le maintien de la structure tridimensionnelle de la protéine, soit dans la fixation de l'un des substrats et/ou dans la catalyse (Bui et al., 2010; Rahuel-Clermont et al., 2010). Ces mutations conduisent ainsi à la production d'une protéine non fonctionnelle dans la majorité des cas. La mutation Leu206Pro conduit à une perte totale d'activité galactosyltransférase qui a été associée à sa délocalisation dans le cytoplasme (au lieu de l'appareil de Golgi, localisation normale de la protéine) (Okajima et al., 1999). Des études phylogénétiques et structurales ont montré que le résidu Leu206 est un résidu très conservé au sein de toutes les  $\beta$ 4GalT7s de différentes espèces, laissant supposer un rôle important de ce résidu et attestant de sa conservation au cours de l'évolution. La mutation est localisée dans un feuillet  $\beta$ , suggérant qu'elle pourrait induire un changement de conformation avec un mauvais repliement de la protéine qui pourrait être à l'origine de sa localisation défectueuse (Bui et al., 2010). Des analyses de la glycosylation de la protéine *core* de la décorine dans des cellules déficientes en  $\beta$ 4GalT7 (CHOpGs-618), surexprimant la forme mutée de l'enzyme Leu206Pro, ont montré que la protéine était dans l'incapacité d'initier la biosynthèse des GAGs *in cellulose* (Götte and Kresse, 2005). Le défaut de la biosynthèse des GAGs, plus particulièrement une perte de la glycanation de la protéine *core* de la décorine, expliquent les altérations structurales de la MEC et les perturbations fonctionnelles telles que la mauvaise cicatrisation des tissus et de la peau observée chez les patients.

Dans le cas du mutant Ala186Asp, la protéine est correctement localisée dans l'appareil de Golgi, mais présente une perte d'activité galactosyltransférase de l'ordre de 50% (Bui et al., 2010; Okajima et al., 1999). Il a été montré que l'alanine 186 est localisée dans le site de fixation



**Figure 4 – Modèle de la structure de la  $\beta$ 4GalT7 et de la position des mutations décrites dans le cas de SEDsp** (d'après Isaure Chauvot de Beauchène, communication personnelle).

Le modèle est issu de la structure cristallographique de la  $\beta$ 4GalT7 humaine (référéncée 4IRP) (Tsutsui et al., 2013) avec en bleu/violet la boucle flexible décrite dans l'enzyme. Le substrat donneur (UDP-Gal) de l'enzyme est présent dans la structure (en jaune) ainsi que la position des résidus qui sont mutés chez les patients (résidus de la protéine sauvage en vert et résidus mutés en magenta). Travail réalisé dans le cadre du projet Action de site Mirabelle CNRS- Université de Lorraine, Projet DynaCriGT - *Criblage virtuel et dynamique moléculaire pour la recherche de bioactifs ciblant la  $\beta$ 4GalT7, une enzyme de biosynthèse des GAGs*, 2017-2018, en collaboration avec le LORIA, équipe CAPSID.

du substrat donneur, l'uridine-diphosphate-galactose (UDP-Gal), lui-même situé dans un domaine très conservé du site actif de l'enzyme (Rahuel-Clermont et al., 2010). Il est intéressant de noter ici que ce résidu ne semble pas important pour la structure ni pour la localisation de la protéine, mais qu'il semble plutôt crucial pour la reconnaissance du substrat et le déroulement du mécanisme catalytique en impactant l'interaction de l'enzyme avec l'un de ses substrats.

La mutation Arg270Cys entraîne une perte de deux tiers de l'activité enzymatique dans les fibroblastes de patients portant cette mutation (Seidler et al., 2006). Des analyses structurales ont montré que ce résidu est localisé dans une boucle flexible qui se replie sur le site actif une fois le substrat donneur fixé (Bui et al., 2010; Rahuel-Clermont et al., 2010). La structure cristallographique de la protéine décrite par la suite va dans le même sens et propose que la mutation induirait une modification de la flexibilité de la boucle à la suite de la fixation du substrat donneur, ce qui modifierait la capacité du substrat accepteur à se fixer ensuite correctement dans le site actif (Figure 4) (Tsutsui et al., 2013). Dans les fibroblastes de patients présentant la mutation Arg270Cys, la glycosylation de la protéine *core* de la décorine, mais également du biglycane, est aberrante et incomplète. La fibrillogenèse du collagène et la prolifération cellulaire des fibroblastes sont également affectées par la mutation Arg270Cys (Seidler et al., 2006). La maturation des PGs, et particulièrement la sulfatation des héparanes sulfates (HS), est également perturbée dans ce contexte. Une réduction de la sulfatation des HS a été observée dans les fibroblastes de patients possédant cette mutation. Cette altération entraîne un retard de cicatrisation, une augmentation de l'adhésion cellulaire et une difficulté des cellules à organiser un réseau de collagène structuré et fonctionnel (Götte et al., 2008).

En conclusion, les différentes mutations du gène *B4GALT7* qui ont été décrites peuvent conduire à (i) une modification de l'affinité de l'enzyme mutée vis-à-vis de ses substrats, (ii) une altération de la structure de l'enzyme, (iii) une mauvaise localisation et/ou (iv) une baisse ou une perte de l'activité enzymatique qui impacte la biosynthèse des GAGs, la fibrillogenèse du collagène et plus généralement l'architecture de la MEC et les propriétés fonctionnelles des tissus conjonctifs.

Des mutations du gène *B3GALT6*, codant la  $\beta$ 3GalT6, une GT qui catalyse l'étape suivant celle catalysée par la  $\beta$ 4GalT7 dans la biosynthèse des GAGs (figure 8), ont également été associées à la forme SEDsp et seront détaillées dans la partie III.

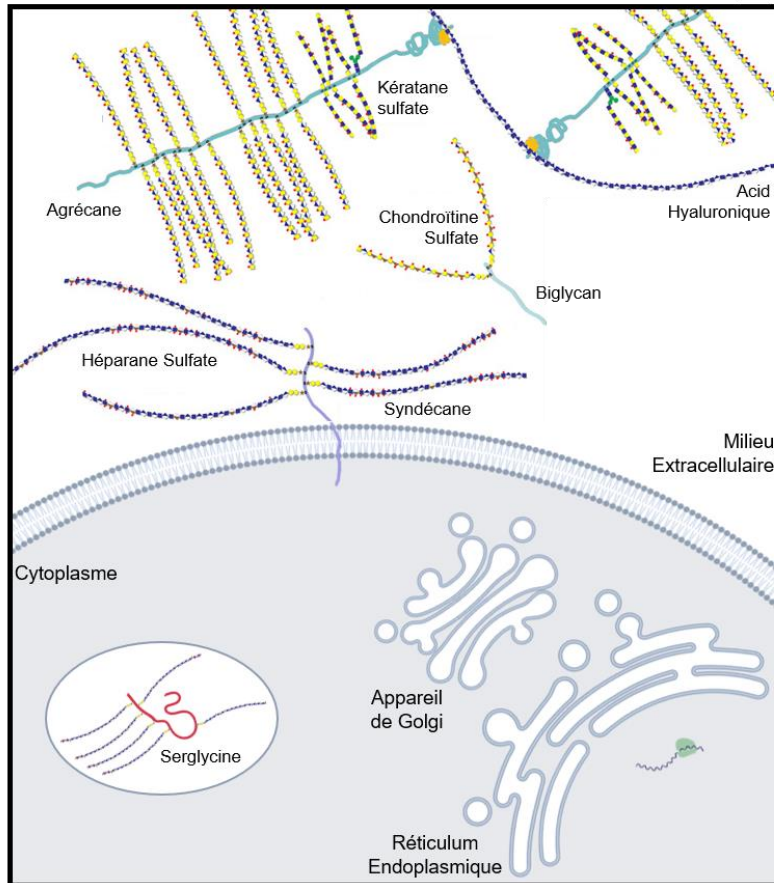
D'autres formes de SED et d'ostéochondrodysplasies ont également été mises en relation avec d'autres enzymes impliquées dans la synthèse et la maturation des GAGs. Plusieurs mutations au niveau des enzymes impliquées dans la synthèse de l'amorce tétrasaccharidique





induisent diverses pathologies regroupées sous le terme générique de « *linkeropathies* ». Des mutations de la Xylosyltransférase I (XylT-I) provoquent la dysplasie squelettique de Desbuquois de type 2 (DBSD2) (Bui et al., 2014; Jamsheer et al., 2016; Koningsbruggen et al., 2016). Dans cette pathologie, les patients présentent une dislocation des articulations, un sévère retard de croissance pré- et post-natal et une hyperlaxité des articulations. Ces patients présentent ainsi des symptômes assez proches de ceux qui sont décrits chez les patients atteints de SEDsp. Chez ces patients, une forte diminution du contenu global en PGs cellulaires est observée, montrant l'importance de cette enzyme dans le processus de biosynthèse (voir partie II.3.1) (Bui et al., 2014). Plusieurs mutations de la Glucuronyltransférase I (GlcAT-I), codée par le gène *B3GAT3*, sont impliquées dans des pathologies comme le syndrome de Larsen-like. Les patients présentent des malformations cardiaques congénitales, une petite taille avec des caractéristiques crânio-faciales variables qui incluent une brachycéphalie, des sourcils épais, de grands yeux avec une descente des fissures palpébrales, un creux nasal déprimé et une bouche étroite (Baasanjav et al., 2011; Budde et al., 2015; Jones et al., 2015; von Oettingen et al., 2014). Plus récemment, une patiente présentant des mutations de cette enzyme a montré des symptômes cliniques correspondant à une forme proche du SEDsp (Ritelli et al., 2019). Dans ce cas, la protéine mutée est nettement moins active que la protéine sauvage, ce qui induit une diminution de la synthèse des GAGs (Héparine (Hep) /HS et CS/DS), montrant que le défaut de la synthèse des GAGs affecte tous les types de PGs (Baasanjav et al., 2011). Toutes ces pathologies présentent des symptômes assez proches, et montrent bien l'importance des enzymes impliquées dans les premières étapes de la biosynthèse des GAGs.

L'ensemble de ces études a permis de faire le lien entre les déficiences génétiques affectant l'activité des GTs, la biosynthèse des chaînes de GAGs et celle des PGs, et les symptômes cliniques observés chez les patients atteints de la forme de SEDsp. La description du SEDsp a d'ailleurs permis de mettre en évidence l'importance des PGs et de leurs chaînes de GAGs dans la fibrillogenèse des collagènes, dans l'architecture matricielle et dans la fonctionnalité des tissus conjonctifs. C'est pourquoi un défaut de leur biosynthèse peut conduire à des effets délétères chez les patients porteurs de mutation(s) au niveau des gènes codant les enzymes de biosynthèse des GAGs. La structure, la fonction et la biosynthèse de ces molécules va maintenant être abordée plus en détail dans la partie II.



**Figure 5 - Représentation schématique des protéoglycannes et des glycosaminoglycannes (Li et al., 2012).**

Les PGs sont des macromolécules situées à la surface des cellules (syndécane), à l'intérieur des cellules (serglycine) et dans les MEC des tissus conjonctifs (agrécane, biglycane). Ils sont constitués d'une partie protéique nommée protéine *core*, sur laquelle sont fixées de manière covalente une ou plusieurs chaînes hétéropolysaccharidiques linéaires appelées GAGs.

## II. Les protéoglycanes : structure, fonctions biologiques et biosynthèse :

### II.1. Structure des protéoglycanes et des glycosaminoglycanes :

Les PGs sont des macromolécules biologiques composées d'un squelette peptidique, nommé protéine *core*, sur lequel sont fixées de manière covalente des chaînes hétéropolysaccharidiques linéaires appelées GAGs.

#### II.1.1. La structure de la protéine core :

La protéine *core* est spécifique de chacune des principales familles de PGs. Elle détermine la classification des PGs selon leur localisation cellulaire, membranaire ou matricielle (Tableau 3 et Figure 5). Dans le cas des PGs membranaires, la protéine *core* possède un domaine hydrophobe dans la région C-terminale ou un domaine GPI (glycosylphosphatidylinositol, comme observé dans le glypican) qui permet aux PGs de s'ancrer à la membrane plasmique des cellules (Theocharis et al., 2016).

La protéine *core* est constituée de divers motifs structuraux et fonctionnels, conservés et spécifiques dont certains sont des potentiels sites d'ancrage pour les chaînes de GAGs. Ces motifs, fréquemment localisés dans la région N-terminale de la protéine *core*, sont constitués d'un motif Sérine-Glycine qui peut être répété plusieurs fois (n) (Ser-Gly)<sub>n</sub> et au niveau duquel le groupement hydroxyle de la chaîne latérale de la sérine est impliqué dans une liaison O-glycosidique avec le xylose des chaînes de GAGs (Iozzo and Schaefer, 2015).

Il a été suggéré que certains motifs de la protéine *core* pourraient jouer un rôle dans la sélection du type de GAGs qui sera attaché sur la protéine porteuse. En effet, la présence de certains motifs peptidiques spécifiques semble favoriser la synthèse d'un type particulier de GAGs (Zhang et al., 1995). Par ailleurs, la présence de résidus acides et/ou de tryptophane à proximité du motif d'ancrage Ser-Gly, semble être un signal favorisant la synthèse des chaînes de HS. Les travaux réalisés sur le sujet indiquent que l'attachement des chaînes de GAGs au niveau du site d'ancrage de la protéine *core* pourrait être régulé par la composition et la structure de la protéine *core* elle-même (Esko and Zhang, 1996; Zhang and Esko, 1994).

#### II.1.2. La structure des glycosaminoglycanes :

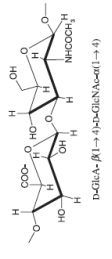
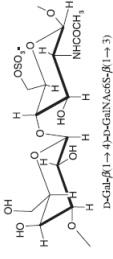
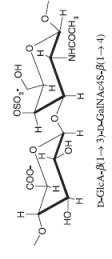
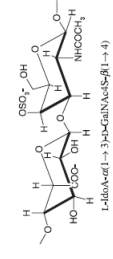
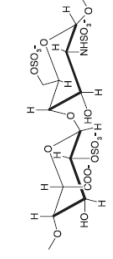
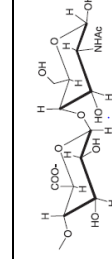
Les GAGs sont composés d'une zone de liaison (nommée amorce tétrasaccharidique) entre la protéine *core* et la chaîne de GAG qui initie la polymérisation des chaînes polysaccharidiques et qui est commune aux principaux types de GAGs (CS/DS et HS). Cette amorce est composée de quatre sucres [Acide Glucuronique-β1,3-Galactose-β1,3-Galactose-β1,4-Xylose-β1-O-]



**Tableau 3 – Classification des protéoglycanes en fonction de leur localisation cellulaire (Iozzo and Schaefer, 2015).**

<b>Localisation</b>	<b>Protéoglycanes</b>	<b>Gène</b>	<b>Glycosaminoglycanes</b>
Intracellulaire	Serglycine	<i>SRGN</i>	Héparine
Surface cellulaire	Syndécan (1 à 4)	<i>SDC</i>	Héparane sulfate
	Betaglycan	<i>TGFBR3</i>	Chondroïtine sulfate / Dermatane sulfate
	Phosphacan	<i>PTPRZ1</i>	Chondroïtine sulfate
	Glypican (1 à 6)	<i>GPC</i>	Héparane sulfate
Péricellulaire	Perlécane	<i>HSPG2</i>	Héparane sulfate
	Agrine	<i>AGRN</i>	Héparane sulfate
Extracellulaire	Aggrécane	<i>ACAN</i>	Chondroïtine sulfate / Kératane sulfate
	Versican	<i>VCAN</i>	Chondroïtine sulfate
	Biglycan	<i>BGN</i>	Chondroïtine sulfate
	Décorine	<i>DCN</i>	Dermatane sulfate
	Lumican	<i>LUM</i>	Kératane sulfate

**Tableau 4 – Composition des motifs dissaccharidiques des glycosaminoglycans (Gandhi and Mancera, 2008).**

Glycosaminoglycans	Motifs dissaccharidiques	Acide uronique	Hexosamine	Autres oses	Caractéristiques	Références
Acide Hyaluronique (AH)		GlcA	GlcNAc	-	Poids moléculaire : 0,8 – 2 000 kDa Non sulfaté et non lié à une protéine <i>core</i>	(Jiang et al., 2011)
Kératine sulfate (KS)		-	GlcNAc	Gal	Poids moléculaire : 4 – 19 kDa Sulfaté sur le galactose et/ou sur la GlcNAc en position 6	(Funderburgh, 2002) (Uchimura, 2015)
Chondroïtine sulfate (CS)		GlcA	GalNAc	-	Poids moléculaire : 5 – 50 kDa Sulfaté en position 4 et/ou 6 du GalNAc	(Lauder, 2009) (Wang et al., 2016)
Dermatane sulfate (DS)		GlcA ou IdoA	GalNAc	-	Poids moléculaire : 15 – 45 kDa Sulfaté en position 4 et/ou 6 de la GalNAc et/ou en position 2 du GlcA/IdoA	
Héparine (Hep)		GlcA ou IdoA	GlcNAc	-	Poids moléculaire : 3 – 30 kDa Sulfaté en position 2 du GlcA/IdoA et/ou en position 2, 3 et/ou 6 de la GlcNAc	(Li and Kusche-Gullberg, 2016) (Meneghetti et al., 2015)
Héparane sulfate (HS)		GlcA ou IdoA	GlcNAc	-	Poids moléculaire : 10 – 70 kDa Sulfaté en position 2 du GlcA/IdoA et/ou en position 2, 3 et/ou 6 de la GlcNAc	

(GlcA- $\beta$ 1,3-Gal- $\beta$ 1,3-Gal- $\beta$ 1,4-Xyl- $\beta$ 1-O-). C'est à partir de cette amorce qu'aura lieu l'allongement des différentes chaînes hétéropolysaccharidiques qui formeront les GAGs. L'élongation de ces chaînes s'effectue par ajout successif, séquentiel et coordonné d'un motif disaccharidique composé (i) d'un sucre aminé, la N-acétyl-D-glucosamine (GlcNAc) pour les GAGs de type HS ou la N-acétyl-D-Galactosamine (GalNAc) dans le cas des CS/DS et (ii) d'un acide uronique, soit l'acide D-glucuronique (GlcA) ou son épimère en position 5, l'acide L-iduronique (IdoA), prédominant dans les DS.

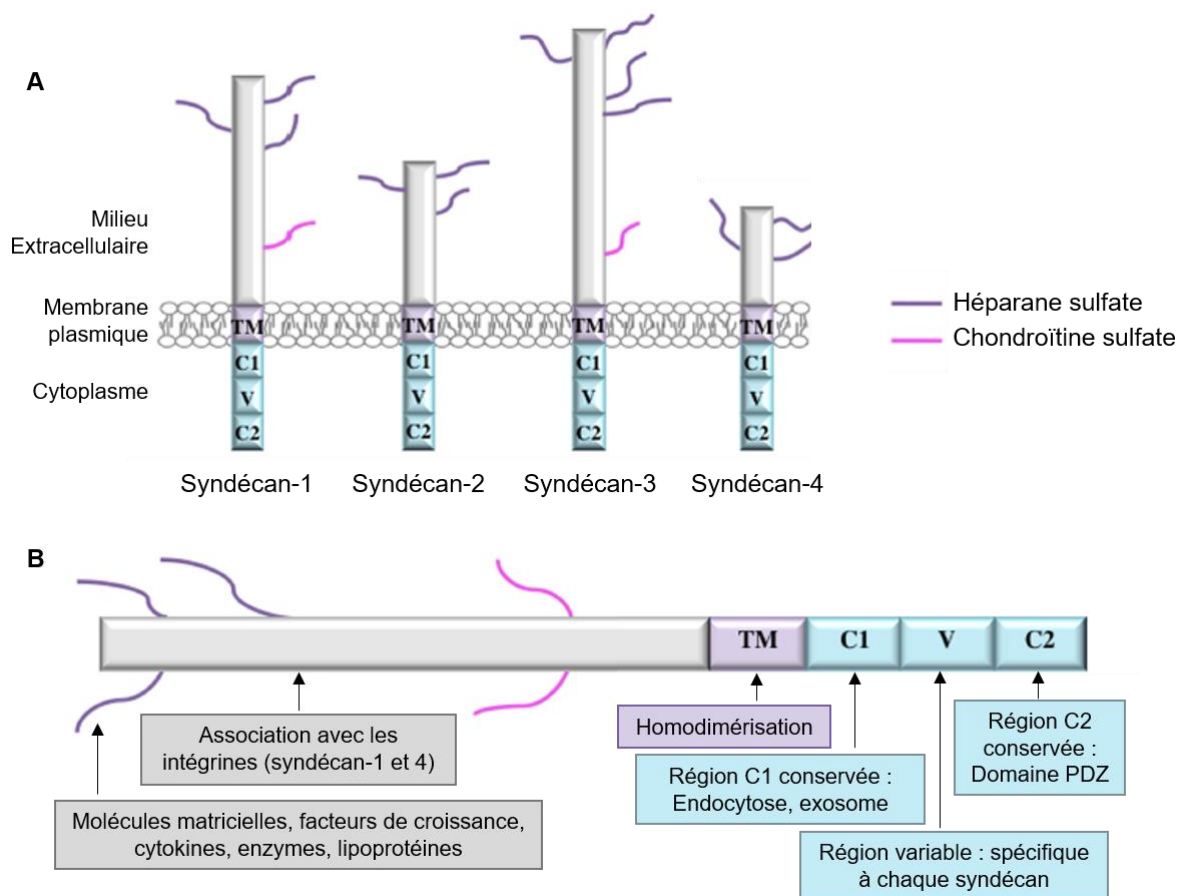
La composition des unités disaccharidiques et la nature de la liaison entre les deux oses déterminent le type de GAG synthétisé. Quatre grands types de chaînes de GAGs peuvent être distingués sur cette base (Tableau 4) : (i) l'acide hyaluronique (AH), (ii) les Hep/HS, (iii) les CS/DS et (iv) les kératanes sulfates (KS). A l'exception de l'AH qui n'est pas sulfaté et non lié à une protéine *core*, les GAGs présentent différents profils de sulfatation qui vont contribuer à la complexité structurale et fonctionnelle de ces chaînes et donc des PGs qui en sont constitués.

## **II.2. Les fonctions biologiques des protéoglycanes :**

### **II.2.1. Généralités :**

Comme nous avons pu le voir, les PGs présentent une grande variabilité structurale et fonctionnelle du fait de leur composition biochimique et de leur localisation. Cette diversité leur permet d'assurer de nombreux rôles biologiques, tels que le maintien de la structure et l'assemblage de la MEC des tissus conjonctifs. Leur implication dans divers processus physiologiques et pathologiques est également associée à leur capacité d'interagir avec un grand nombre d'effecteurs solubles, comme des protéases matricielles, des facteurs de croissance et/ou des récepteurs situés à la surface des cellules. C'est le cas du perlécan qui peut interagir avec un grand nombre de facteurs de croissance comme le FGF-2 (*Fibroblast Growth Factor 2*), le FGF-7, le VEGF (*Vascular Endothelial Growth Factor*) et le PDGF (*Platelet-Derived Growth Factor*) (Iozzo, 2005; Mongiat et al., 2000, 2001). Il a été montré que le FGF-2 pouvait se lier aux chaînes de HS du perlécan, permettre l'activation du récepteur FGFR (*FGF Receptor*) et induire l'activation des voies de signalisation à l'origine de l'angiogenèse, la cicatrisation et l'invasion tumorale. Dans ce contexte, le perlécan joue le rôle de corécepteur en permettant la présentation du ligand au récepteur et l'induction d'une signalisation efficace (Gubbiotti et al., 2017; Nugent and Iozzo, 2000; Whitelock et al., 1996).

Dans un autre contexte, l'agrécane, un PG de haut poids moléculaire, s'associe à l'AH et conduit à la formation de complexes matriciels supramoléculaires qui peuvent atteindre une masse moléculaire supérieure à 200 000 kDa. Ces complexes vont contribuer à l'organisation



**Figure 6 – Représentation schématique des différents membres de la famille des syndécans et fonctions des différents domaines (Cheng et al., 2016).**

Les syndécans sont des protéines transmembranaires portant majoritairement des chaînes de HS. La famille des syndécans compte quatre membres chez les vertébrés qui possèdent tous un domaine extracellulaire portant les chaînes de GAGs, un seul domaine transmembranaire et une extrémité C-terminale cytoplasmique. (A) Représentation schématique des membres de la famille des syndécans (1-4). (B) Représentation schématique montrant les rôles des différents domaines des syndécans avec des exemples de leurs fonctions et de leurs partenaires d'interaction (C : domaine conservé et V : domaine variable).



tridimensionnelle de la matrice des tissus conjonctifs (Kiani et al., 2002; Wight et al., 2011). Dans le cas du cartilage, ces larges agrégats présentent une densité élevée en chaînes de CS, ce qui explique leur grande densité en charges négatives. La présence d'un nombre élevé de charges négatives permet de maintenir une grande quantité d'eau au niveau de la matrice, conférant ainsi une élasticité et une forte résistance des tissus aux stress mécaniques (Aspberg, 2012; Heinegård, 2009).

## II.2.2. Fonctions des protéoglycanes dans les tissus conjonctifs :

### II.2.2.1. Protéoglycanes membranaires portant des chaînes d'héparanes sulfates, les syndécans :

Les syndécans sont des PGs transmembranaires présents à la membrane plasmique des cellules et jouant des rôles fonctionnels importants en raison de leur localisation à l'interface cellule-matrice. Il s'agit d'une famille de PGs comptant quatre gènes distincts, *SDC1* (Saunders et al., 1989), *SDC2* (Marynen et al., 1989; Pierce et al., 1992), *SDC3* (Carey et al., 1992; Gould et al., 1992) et *SDC4* (David et al., 1992; Kojima et al., 1992), qui codent des protéines *core* avec un seul domaine transmembranaire, un domaine N-terminal extracellulaire et un domaine C-terminal intracellulaire (Figure 6A) (Couchman, 2010).

Le domaine extracellulaire, ou ectodomaine, est un domaine à faible homologie de séquence au sein d'une même famille de syndécans (entre 10 et 20%), ce qui explique que les syndécans ont la propriété d'interagir avec une grande diversité de protéines effectrices ou de ligands (Figure 6B) (Choi et al., 2011). Les interactions des syndécans avec les protéines effectrices impliquent dans de nombreux cas les chaînes de HS du PG (entre 2 à 5 chaînes) et dans certains cas des chaînes de CS. Le domaine extracellulaire contient également un site spécifique de clivage protéolytique, permettant sa libération dans la matrice suite à l'action de différentes protéases. Dans la matrice, la forme soluble du PG joue un rôle régulateur important dans divers processus physiopathologiques, comme c'est le cas dans certains cancers. Par exemple, la présence de syndécan 1 soluble favorise la croissance tumorale du myélome (Yang et al., 2002). Dans ces cellules, le syndécan 1 soluble est capable d'interagir avec le VEGF par l'intermédiaire des chaînes de HS. Ce complexe est reconnu par l'intégrine  $\alpha\beta3$  (connue pour son rôle dans l'angiogenèse et l'activation endothéliale) et par le récepteur VEGFR2, favorisant la croissance tumorale (Purushothaman et al., 2010). Dans le cas du syndécan 2, la forme soluble du PG interagit avec le récepteur CD148 (un récepteur à activité tyrosine phosphatase) qui est responsable de la déphosphorylation de l'intégrine  $\beta1$ . L'enlèvement du groupement

**Tableau 5 – Exemples de ligands et fonctions biologiques des syndécans.**

<b>Ligands</b>	<b>Fonctions biologiques</b>	<b>Références</b>
<b>Molécules matricielles</b>		
Fibronectine	Adhésion cellulaire	(Xian et al., 2010) (Klass et al., 2000)
Laminine-5	Adhésion et migration cellulaire	(Bachy et al., 2008) (Ogawa et al., 2007)
Vitronectine	Adhésion cellulaire	(Wilkins-Port et al., 2003)
<b>Cytokines</b>		
Ostéopontine	Prolifération des chondrocytes	(Shimo et al., 2004)
IL-1	Arthrite inflammatoire	(Wang et al., 2011) (Korb-Pap et al., 2012)
<b>Facteur de croissance</b>		
FGF-2	Prolifération et cicatrisation	(Clasper et al., 1999) (Mukhopadhyay et al., 2010)
TGF- $\beta$ -1	Cicatrisation, régénération tissulaire	(Sebestyén et al., 2000) (Worapamorn et al., 2001)

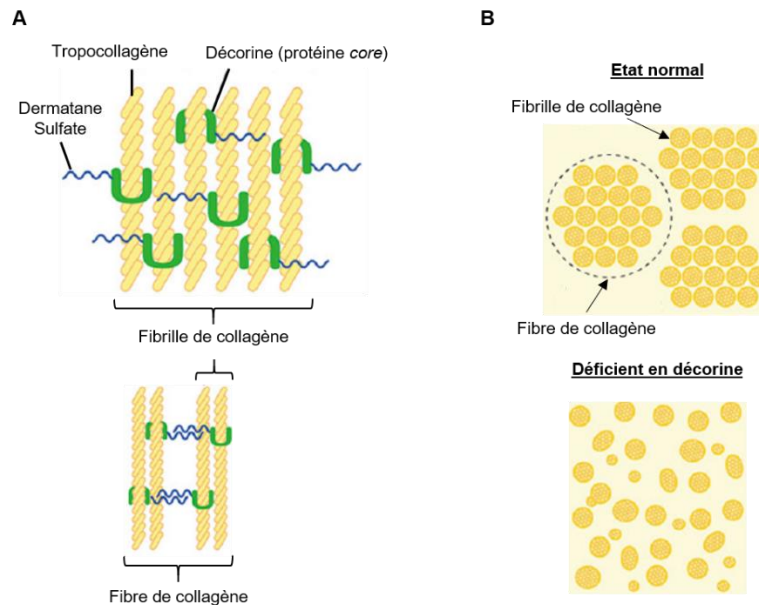
phosphate provoque l'inactivation de l'intégrine  $\beta 1$  et induit l'inhibition de l'angiogénèse (Rossi et al., 2014). En revanche, le syndécan 2 membranaire favorise l'angiogénèse en jouant le rôle de corécepteur pour des facteurs de croissance comme le VEGF (Chen et al., 2004a). Le clivage du domaine extracellulaire de ce syndécan aurait donc un rôle régulateur dans le processus angiogénique.

Le domaine cytoplasmique du PG membranaire est composé de deux régions conservées (appelées C1 et C2) séparées par une séquence variable (notée V) qui est propre à chaque famille de syndécans (Figure 6) (Iozzo and Schaefer, 2015). Le domaine cytoplasmique comprend également une séquence térapeptidique, nommée EFYA (Glutamate-Phénylalanine-Tyrosine-Alanine), qui peut fixer des protéines ayant un domaine PDZ (*Post synaptic density protein*), comme la synténine-1 (Grootjans et al., 1997) ou la synectine (Tkachenko Eugene et al., 2006). Ces interactions, entre les syndécans et les protéines PDZ, vont intervenir dans les mécanismes d'adhésion et de migration cellulaire. Par exemple, l'interaction entre le syndécan 4 et la synectine permet d'initier une cascade de signalisation qui conduit à l'activation de Rac1 (une Rho-GTPase). L'activation de cette protéine favorise la migration cellulaire des cellules endothéliales (Cheng et al., 2016; Tkachenko et al., 2006).

Les chaînes de GAGs portées par les syndécans jouent un rôle important dans l'activité biologique de ces PGs. Comme nous l'avons dit précédemment, les HS interagissent avec de nombreux ligands tels que des facteurs de croissance, comme le FGF-2 ou le TGF- $\beta$  (*Transforming Growth Factor  $\beta$* ), des cytokines (ostéopontine, interleukine-1 (IL-1)) ou encore des molécules matricielles (fibronectine, laminine, etc). Les principaux ligands des syndécans et leurs fonctions biologiques sont présentés dans le tableau 5. Des études ont montré des interactions entre les syndécans et les laminines, dont la laminine-5 (ou laminine-332) qui est une isoforme majeure au niveau de la membrane basale de la peau (Sugawara et al., 2008). La laminine-5 est composée de trois chaînes protéiques ( $\alpha 3$ ,  $\beta 3$  et  $\gamma 2$ ). Un domaine de la chaîne  $\alpha 3$  (domaine LG1-3 ; LG : *Large globular*) interagit avec les syndécans 2 et 4 pour favoriser l'adhésion cellulaire (Utani et al., 2001). Il a également été montré que l'ectodomaine du syndécan 2 interagissait avec le TGF- $\beta$ . Ces interactions permettent de réguler certaines fonctions biologiques du TGF- $\beta$ , comme son implication dans la fibrose (Chen et al., 2004b; Lichtman et al., 2016).

#### II.2.2.2. Petit protéoglycane matriciel portant une chaîne de chondroïtine/dermatane sulfate, la décorine :

La décorine est un PG de la famille des SLRPs. C'est un petit PG présentant un domaine



**Figure 7 – Modèle putatif de l’assemblage des fibres de collagène en présence ou absence de décorine (Miyake et al., 2014).**

(A) La décorine fixe directement le tropocollagène et aide à former les fibrilles de collagène. Avec les chaînes de CS/DS, les fibrilles pourront former les fibres de collagène. (B) Représentation schématique de la structure de la fibre de collagène chez une souris saine (état physiologique) et chez une souris déficiente en décorine (Danielson et al., 1997).

**Tableau 6 – Listes des fonctions physiologiques et pathologiques associées à la décorine.**

Mécanismes et pathologies impliquant la décorine	Références
Adhésion et migration cellulaire	(Ferdous et al., 2010)
Asthme	(Marchica et al., 2011)
Cicatrisation et angiogenèse	(Järveläinen et al., 2006) (Neill et al., 2012) (Neill et al., 2013) (Järveläinen et al., 2015)
Fibrillogenèse des collagènes	(Rühland et al., 2007) (Seidler et al., 2005) (Ferdous et al., 2008)
Infarctus du myocarde	(Weis et al., 2005)
Initiation de l’immunité et de l’inflammation	(Merline et al., 2011) (Seidler et al., 2011) (Bocian et al., 2013)
Maladie de Lyme	(Brown et al., 2001)

protéique central riche en leucine, nommée LRRs (*Leucine Rich Repeats*). La décorine possède une unique chaîne de GAG de type CS/DS au niveau de l'extrémité N-terminale de la protéine *core*. La protéine *core* de la décorine est une métalloprotéine à zinc de 40 kDa qui interagit avec les fibrilles de collagène dans les MEC (Figure 7A) (Keene et al., 2000; Scott, 1988). Cette interaction permet de maintenir une organisation structurale fonctionnelle du réseau de collagène dans les matrices (Raspanti et al., 2008; Rühland et al., 2007). Chez des souris déficientes en décorine, les fibrilles de collagène présentent une organisation anormale dans la peau et les tendons. Ces phénotypes sont proches de ceux qui ont pu être observés chez les patients atteints de SED (Figure 7B) (Byers and Murray, 2014; Danielson et al., 1997).

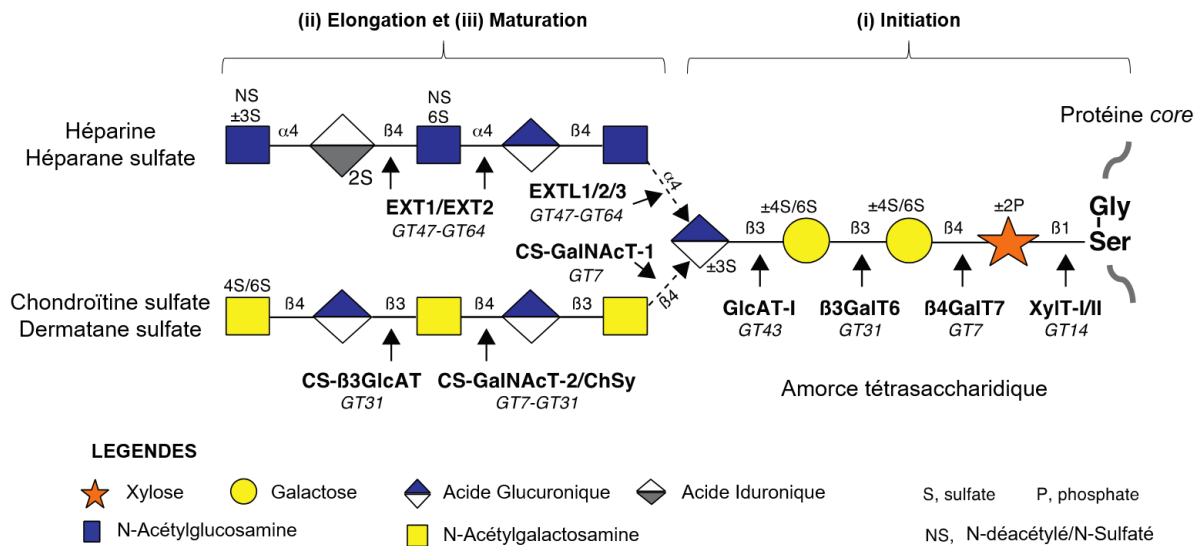
La pertinence du modèle de souris déficiente en décorine ne s'est pas limitée au cas des SED. Des liens entre la décorine et d'autres pathologies pouvant être associées à la mécanique cellulaire (exemple : l'adhésion et la migration cellulaire, la cicatrisation et l'angiogenèse) ont pu être établis dans la maladie de Lyme, l'asthme, les infarctus du myocarde (Brown et al., 2001; Marchica et al., 2011; Weis et al., 2005). Une liste non exhaustive des implications physiopathologiques de la décorine est présentée dans le tableau 6.

Pour conclure sur cette partie, nous voyons par ces deux exemples de PG membranaire et matriciel que les PGs jouent des rôles structuraux et fonctionnels très importants dans le maintien de l'architecture tridimensionnelle de la matrice des tissus conjonctifs et dans la régulation des mécanismes de signalisation intra- et intercellulaire. Pour preuve, des modifications de leurs structures (et donc de leurs fonctions) sont associées à l'apparition et au développement de pathologies.

Nous proposons de nous focaliser sur l'étude des voies de biosynthèses des PGs et de leurs chaînes de GAGs qui vont être présentées plus en détail dans la partie qui suit avec un focus sur l'une des GTs impliquées dans les étapes précoces de biosynthèse des GAGs, la  $\beta$ 3GalT6, qui fait l'objet de ce travail.

### **II.3. La biosynthèse des protéoglycanes :**

La biosynthèse des PGs débute par la formation de la protéine *core* dans le RE et se poursuit par la synthèse des GAGs dans les différents compartiments tardifs du RE et de l'appareil de Golgi. Elle s'effectue en trois étapes (Figure 8) : (i) la synthèse d'une séquence tétrasaccharidique sur la protéine *core* commune aux GAGs de type CS/DS et Hep/HS, (ii) la polymérisation des chaînes sur cette amorce, et finalement (iii) la modification des chaînes (épimérisation, sulfatation, acétylation) par différentes enzymes (épimérase, N- et O-



**Figure 8 – Représentation schématique de la voie de biosynthèse des glycosaminoglycans (Breton et al., 2012).**

La biosynthèse des PGs est initiée par la synthèse de la protéine *core* dans le RE. Elle se poursuit dans l'appareil de Golgi par (i) la synthèse d'une amorce tétrasaccharidique (GlcA-β1,3-Gal-β1,3-Gal-β1,4-Xyl-β1-O-Ser) réalisée séquentiellement par l'ajout d'un résidu xylose, de deux galactoses et d'un acide glucuronique. (ii) Les motifs disaccharidiques seront ensuite transférés par les GTs de la famille des CS-synthases (CSS) pour l'élongation des CS/DS et de la famille des exostosines (EXT) pour les Hep/HS qui définiront la nature des chaînes de GAGs produites. Durant la biosynthèse des GAGs, différentes modifications auront lieu telles que des sulfatations en positions 4 (4S) et/ou 6 (6S) sur les résidus N-acétylgalactosamine ou des N-sulfatations (NS) sur les résidus N-acétylglucosamine par exemple.

sulfotransférases, N-acétyltransférases respectivement) pour produire des chaînes de GAGs matures et fonctionnelles.

### II.3.1. Synthèse de l'amorce tétrasaccharidique :

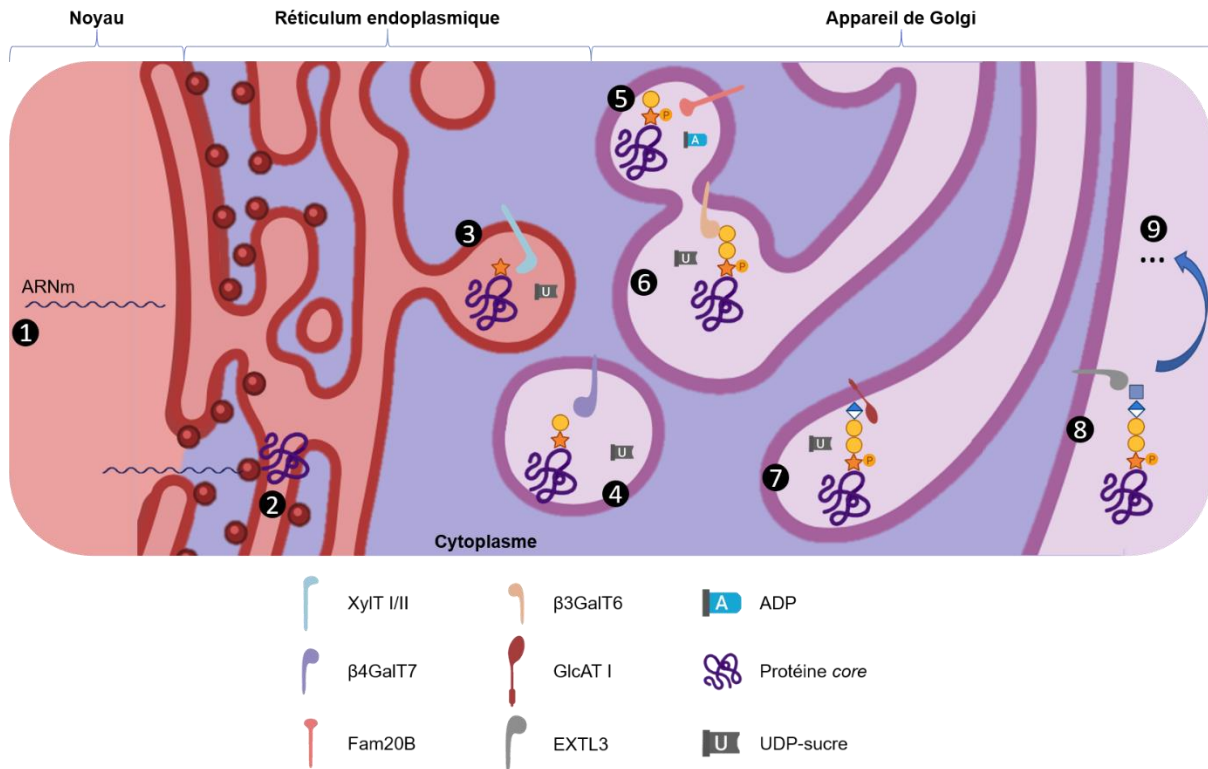
La première étape de la biosynthèse des GAGs consiste à former une amorce tétrasaccharidique composée d'un GlcA, de deux galactoses (Gal) et d'un xylose (Xyl), attachée à un résidu sérine de la protéine *core* du PG (GlcA- $\beta$ 1,3-Gal- $\beta$ 1,3-Gal- $\beta$ 1,4-Xyl- $\beta$ 1-O-Ser) (Figure 8). Cette région charnière est commune aux CS/DS et aux Hep/HS. Le transfert de chaque ose est effectué par des GTs spécifiques qui vont agir de manière séquentielle et coordonnée (Figures 8 et 9).

La fixation du xylose sur un résidu sérine spécifique de la protéine *core* est catalysée par les XylT-I/II. Ces enzymes catalysent la formation d'une liaison glycosidique  $\beta$ 1-O- entre le carbone 1 (C1) du xylose et le groupement hydroxyle de la chaîne latérale de la sérine (Götting et al., 2000). Le premier galactose est transféré sur le résidu xylose par la  $\beta$ 4GalT7 en formant une liaison glycosidique  $\beta$ 1,4 (Almeida et al., 1999). La  $\beta$ 3GalT6 ajoute le second galactose grâce à une liaison glycosidique  $\beta$ 1,3 (Bai et al., 2001). La synthèse de l'amorce est finalisée par l'action de la GlcAT-I (Kitagawa et al., 1998) qui ajoute un résidu d'acide glucuronique sur l'amorce [Gal- $\beta$ 1,3-Gal- $\beta$ 1,4-Xyl- $\beta$ 1-O-Ser].

Des modifications de l'amorce tétrasaccharidique ont été décrites comme la phosphorylation du xylose ou la sulfatation des galactoses (Sugahara et al., 1988, 1992; Tone et al., 2008). Ces modifications peuvent moduler l'activité des GTs et réguler la synthèse des Hep/HS ou CS/DS, comme cela a été montré pour la GlcAT-I, dont l'activité est impactée par la sulfatation des deux résidus galactose (Gulberti et al., 2005). Dans le même sens, la phosphorylation du groupement hydroxyle du xylose en position 2, catalysée par la kinase Fam20B (*Family with sequence similarity 20 member B*) (Koike et al., 2009) impacte fortement l'activité de la  $\beta$ 4GalT7 et de la  $\beta$ 3GalT6. En effet, l'accepteur Xyl(2P) n'est pas un substrat de la  $\beta$ 4GalT7 (Gulberti et al., 2005; Talhaoui et al., 2010) alors qu'un accepteur de type Gal-Xyl(2P) est un bien meilleur substrat pour la  $\beta$ 3GalT6 comparé au substrat non modifié (Wen et al., 2014). En parallèle, il a été montré que la xylose phosphatase (XylP) qui enlève le groupement phosphate du xylose, fonctionnerait de façon conjointe avec la kinase Fam20B et serait impliquée dans la régulation de l'élongation de la chaîne de GAG en Hep/HS ou CS/DS (Izumikawa et al., 2015; Koike et al., 2014). Cette modification semble donc cruciale pour arrêter ou amorcer de façon efficace la biosynthèse des GAGs et semble agir de manière très

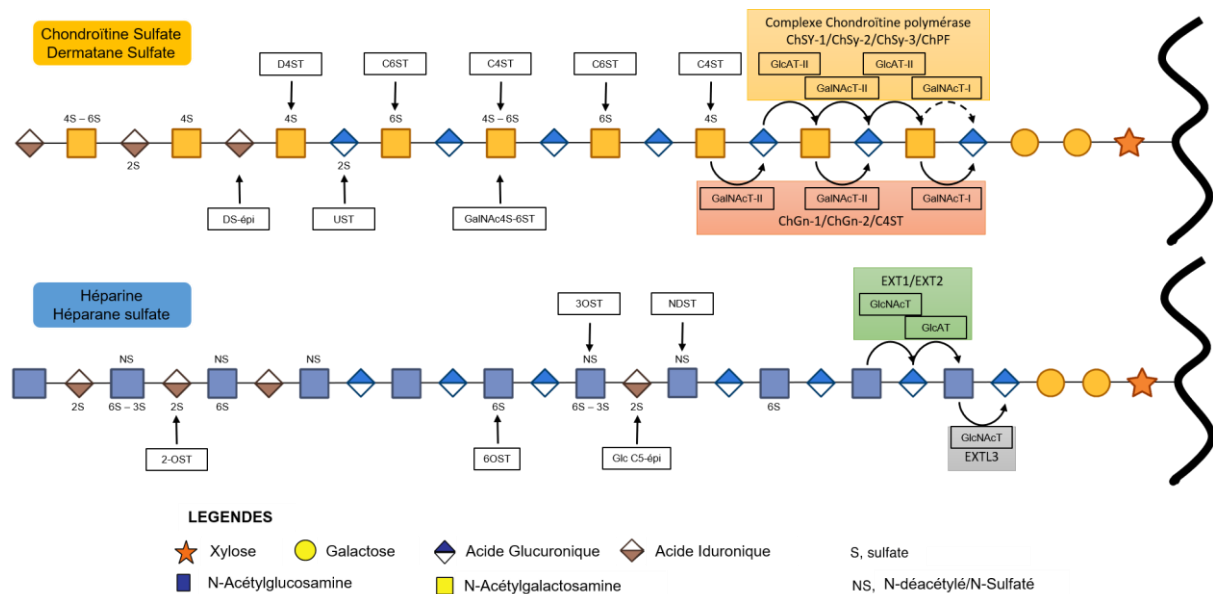






**Figure 9 – Représentation schématique de la biosynthèse de l'amorce tétrasaccharidique dans les cellules.**

La biosynthèse des PGs commence par (1) et (2) l'expression de l'ARNm d'une protéine *core* qui est traduit en protéine au niveau du RE. (3) Dans les compartiments tardifs du RE et/ou dans le *cis*-Golgi, l'ajout du xylose par la XylT-I ou II est effectué. (4) L'ajout du premier galactose par la β4GalT7 est ensuite réalisé dans le cis-Golgi. Une étape de maturation de l'amorce peut avoir lieu avec (5) l'ajout d'un groupement phosphate sur le xylose par la Fam20B (Koike et al., 2009). Ensuite, l'amorce modifiée est utilisée par la β3GalT6 (6) pour l'ajout du second galactose. Enfin, (7) la synthèse de l'amorce est finalisée par l'ajout de l'acide glucuronique par la GlcAT-I. A partir de cette amorce, il va y avoir élongation des chaînes de GAGs (Hep/HS ou CS/DS) (8 et 9).



**Figure 10 – Représentation schématique de la biosynthèse et de la maturation des chaînes de CS/DS et Hep/HS (Mikami and Kitagawa, 2013; Poulain and Yost, 2015).**

Les motifs disaccharidiques qui vont définir la nature des chaînes de GAGs sont ajoutés par des systèmes enzymatiques de type CSS pour l'élongation des CS/DS et des GTs de la famille des EXT pour celle des Hep/HS. Au cours de leur biosynthèse, différentes étapes de maturation vont avoir lieu telles que qu'une N-déacétylation, des sulfatations (N-, 2-O-, 3-O-, 4-O- et 6-O-sulfatations) ou encore une épimérisation du résidu GlcA en IdoA.

fine car elle impacte différemment deux enzymes qui interviennent de façon consécutive dans la formation de la zone de liaison. Par conséquent, les modifications du tétrasaccharide semblent agir en tant que régulateurs de l'activité des GTs impliquées dans la biosynthèse des GAGs et donc contrôler l'initiation et l'élongation des chaînes polysaccharidiques. Elles pourraient également intervenir dans le choix du type de GAGs ainsi que le niveau de production du GAG sur les PGs et avoir ainsi des fonctions biologiques importantes (Gulberti et al., 2005; Izumikawa et al., 2011; Koike et al., 2014).

### II.3.2. L'élongation et la maturation des chaînes de GAGs :

L'élongation des chaînes de GAGs est réalisée à partir de l'amorce tétrasaccharidique par l'ajout successif d'un motif disaccharidique composé : (i) d'un sucre aminé, le résidu GlcNAc pour l'Hep et les HS ou le résidu GalNAc pour les CS/DS et (ii) d'un acide uronique, GlcA pour les CS/HS ou IdoA pour l'Hep/DS. Dans tous les cas, le transfert des oses constituant le motif disaccharidique est catalysé par des GTs spécifiques, les enzymes de la famille des exostosines (EXT) pour les Hep/HS (Kim et al., 2001; Kitagawa et al., 1999) et les *chondroïtin sulfate synthases* (CSS) pour les CS/DS (Yada et al., 2003a, 2003b). De façon concomitante à leur élongation, différentes modifications, comprenant une épimérisation du GlcA en IdoA dans les DS et des O- et N-sulfatations par des sulfotransférases spécifiques, vont être réalisées au niveau des chaînes de GAGs (Figure 10).

Six GTs, comprenant les CSS (ChSy-1, ChSy-2 et ChSy-3) (Kitagawa et al., 2001; Yada et al., 2003a, 2003b), les *Chondroïtin GalNAc transferases* (ChGn-1, ChGn-2) (Uyama et al., 2002, 2003) et le *Chondroïtin polymerizing factor* (ChPF) (Kitagawa et al., 2003), vont intervenir dans la biosynthèse des chaînes de CS/DS.

Les enzymes, ChSy-1, ChSy-2 et ChSy-3, présentent des activités (i) N-acétylgalactosaminyltransférase-II (GalNAcT-II) pour le transfert des résidus GalNAc sur la chaîne en cours de synthèse, et (ii) glucuronosyltransférase-II (GlcAT-II) pour le transfert des résidus GlcA sur les résidus GalNAc (Figure 10). Seules, ces enzymes ne peuvent pas polymériser les chaînes de GAGs. Elles nécessitent la co-expression d'au moins une autre des enzymes de ce groupe (et/ou du ChPF) pour être actives. Les enzymes ChGn-1 et ChGn-2, présentent seulement les activités N-acétylgalactosaminyltransférase-I et N-acétylgalactosaminyltransférase-II. L'activité N-acétylgalactosaminyltransférase-I permet d'initier la synthèse des chaînes de CS/DS en ajoutant le premier résidu GalNAc sur l'amorce tétrasaccharidique. Durant la synthèse des chaînes de CS/DS, différentes modifications peuvent avoir lieu comme une épimérisation des résidus GlcA en IdoA pour former les DS et/ou des N-

**Tableau 7 – Liste des modifications et des enzymes intervenant durant l'élongation et la maturation des chaînes de glycosaminoglycanes.**

<b>Modifications</b>	<b>Enzymes</b>	<b>Références</b>
<b>Modifications des Chondroïtines sulfates/Dermatanes sulfates</b>		
Epimérisation	Dermatane Sulfate Epimérase (DES)	(Malmström et al., 2012)
4-O-Sulfatation	Chondroïtine 4-O-Sulfotransférase 1 (C4ST-1)	(Hiraoka et al., 2000)
	Chondroïtine 4-O-Sulfotransférase 2 (C4ST-2)	(Kang et al., 2002)
	Chondroïtine 4-O-Sulfotransférase 3 (C4ST-3)	(Mikami et al., 2003)
	Dermatane 4-O-sulfotransférase (D4ST)	(Izumikawa et al., 2011)
6-O-Sulfatation	Chondroïtine 6-O-Sulfotransférase (C6ST)	(Fukuta et al., 1998) (Uchimura et al., 2002)
2-O-Sulfatation	Uronyl 2-O-sulfotransférase	(Kobayashi et al., 1999)
<b>Modifications des Héparine/Héparanes sulfates</b>		
N-déacétylation/ N-sulfatation	N-déacétylases/ N-sulfotransférases (NDST)	(Bame et al., 1994)
2-O-Sulfatation	2-O-Sulfotransférases (2OST)	(Liu et al., 2014)
6-O-Sulfatation	6-O-Sulfotransférases (6OST)	(Liu et al., 2014)
3-O-Sulfatation	3-O-Sulfotransférases (3OST)	(Mao et al., 2016)
		(Vanpouille et al., 2007)
Epimérisation	C5-Epimérase	(Qin et al., 2015)

déacétylation et N- et O-sulfatations (Tableau 7).

Dans le cas des Hep/HS, l'élongation commence par l'ajout du premier résidu GlcNAc par l'action de la GT EXTL3, de la famille des EXTL (*Exostosin-like glycosyltransferases*), qui possède une activité  $\alpha$ -GlcNAc transférase (GlcNAcT) (Figure 10) (Busse et al., 2007; Han et al., 2004; Holmborn et al., 2012). Une fois le premier résidu lié à l'amorce, les chaînes de Hep/HS sont polymérisées par l'addition successive des résidus GlcA et GlcNAc par les GTs EXT1 et EXT2 (Kreuger and Kjellén, 2012). De façon complémentaire à la biosynthèse, les chaînes sont maturées par l'ajout de groupements sulfates ou de l'épimérisation d'un résidu GlcA en IdoA grâce à l'activité de sulfotransférases et d'une épimérase spécifiques présentées dans le Tableau 7.

La biosynthèse des CS/DS et des Hep/HS est un processus complexe et coordonné faisant intervenir de nombreuses GTs. A l'heure actuelle, le mécanisme déterminant l'orientation préférentielle de la biosynthèse des GAGs vers celle des Hep/HS ou CS/DS à partir de l'amorce n'est pas encore connu. Il a été proposé que la protéine *core* pourrait jouer un rôle dans la sélection de la chaîne de GAGs (voir partie II.1.1) (Bame et al., 1994; Esko and Zhang, 1996), mais plus récemment, il a été montré que des modifications de l'amorce tétrasaccharidique, comme la phosphorylation du xylose, pourraient jouer également un rôle dans cette orientation (Izumikawa et al., 2015; Koike et al., 2014). Des études cinétiques sur les enzymes impliquées dans la biosynthèse des CS/DS ont montré que l'amorce tétrasaccharidique phosphorylée ne semble pas être un substrat adapté pour certaines enzymes impliquées dans la synthèse de la chaîne de GAGs. Dans ce cas, la déphosphorylation du xylose serait une étape préalable à l'initiation de la biosynthèse des CS et à la formation de l'amorce tétrasaccharidique (Koike et al., 2014). En effet, il a été montré que l'enzyme EXTL2 n'est pas capable de fixer le résidu GlcNAc sur une amorce phosphorylée (Koike et al., 2014). Par contre, l'enzyme ChGn-1 est capable de catalyser le transfert d'un résidu GalNAc sur cette amorce de façon préférentielle si le xylose est phosphorylé (Izumikawa et al., 2015). Par contre, l'élongation ultérieure de la chaîne de GAGs semble moins efficace en absence de la phosphorylation. Il semble donc que, comme pour la  $\beta$ 3GalT6 au niveau de la biosynthèse de l'amorce, une étape de phosphorylation-déphosphorylation du xylose pourrait jouer un rôle crucial pour la détermination de la nature de la chaîne de GAGs synthétisée par les cellules.

Pour résumer, nous avons vu que la biosynthèse des GAGs est un processus complexe qui se déroule en deux étapes : l'initiation et l'élongation. Elle consiste d'abord en la synthèse d'une amorce tétrasaccharidique commune aux Hep/HS et CS/DS. Le fait que cette amorce soit commune à la biosynthèse des principaux types de GAGs implique une régulation étroite des



étapes de biosynthèse et renforce le rôle crucial des enzymes impliquées dans la formation de cette amorce. En effet, des mutations au niveau de ces enzymes conduisent à l'apparition de maladies génétiques rares telles que le SEDsp dans lequel la  $\beta$ 4GalT7 ou la  $\beta$ 3GalT6 sont déficientes ou encore dans le DBSD2 et le syndrome de Larsen-like impliquant des mutations de la XylT-I ou de la GlcAT-I, respectivement. Il est donc important de mieux comprendre l'implication de ces enzymes dans la pathogénie des SED, en particulier les SEDsp, objectif visé par ce travail de thèse.

Mon travail se consacre à l'étude de la  $\beta$ 3GalT6, la galactosyltransférase impliquée dans la fixation du second galactose de l'amorce tétrasaccharidique, et son implication dans le SEDsp. Nous allons dans un premier temps rappeler la structure et la fonction des GTs, avant de nous concentrer sur la famille des  $\beta$ 3GalTs et plus précisément sur la  $\beta$ 3GalT6.





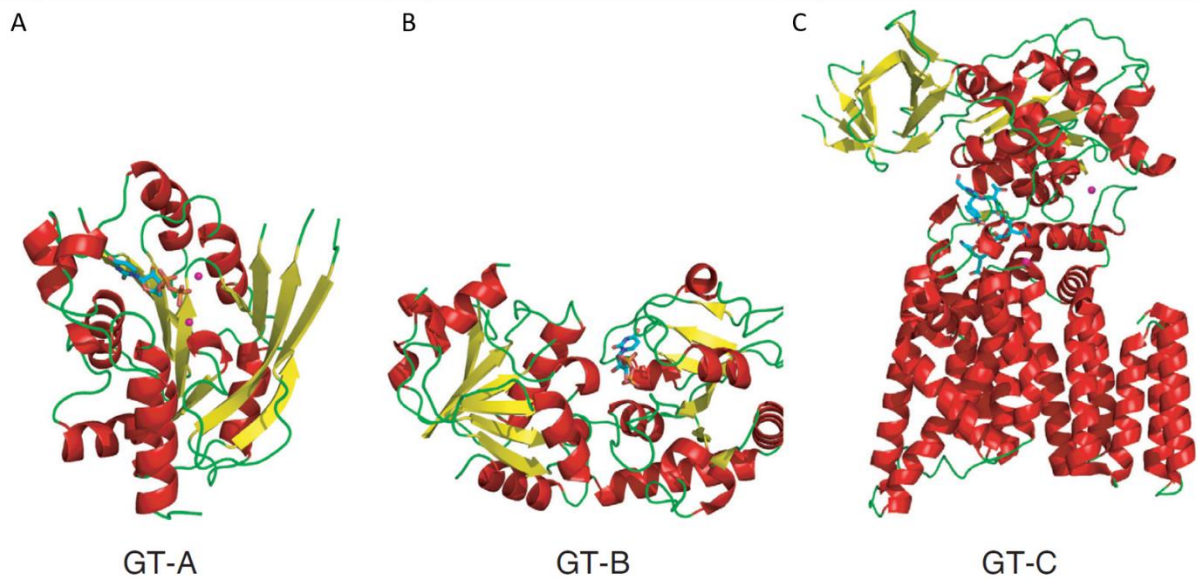
### III. Les glycosyltransférases et la $\beta$ 1,3-Galactosyltransférase 6 :

#### III.1. Les glycosyltransférases, des enzymes ubiquitaires :

Les GTs sont des enzymes ubiquitaires responsables de la création d'une gamme complexe et diversifiée d'oligosaccharides et de glycoconjugués. La diversité chimique et structurale de ces molécules (épimères, configuration anomérique, position et nature de la liaison osidique et ramification) s'accompagne d'une forte complexité et nécessite que les enzymes qui catalysent leur synthèse, leur dégradation et leurs modifications soient hautement spécifiques. Les GTs catalysent la formation d'une liaison osidique entre un sucre et un substrat accepteur, qui peut être une large gamme de biomolécules, telles que d'autres sucres, des protéines, des lipides et d'autres petites molécules. Les donneurs de sucre sont le plus souvent des dérivés nucléotidiques (comme les UDP-sucres), mais ils peuvent également être des lipides et/ou des phosphates non substitués.

Les GTs ont été classées par homologie de séquences en acides aminés en 96 familles au sein d'une banque de données, la *Carbohydrate Active enZyme database* (CAZy) (Lombard et al., 2014). Cette base de données est un outil prédictif très puissant car le repliement structural et le mécanisme catalytique des GTs sont similaires dans la plupart des familles. Par conséquent, lorsque la structure et le mécanisme d'une GT pour une famille donnée ont été décrits, certaines hypothèses concernant d'autres membres de la même famille peuvent être émises. Cependant, la spécificité de substrat (donneur et accepteur) est plus difficile à prévoir et nécessite une caractérisation expérimentale. Déterminer la spécificité de substrat donneur et/ou accepteur pour une GT de fonction inconnue peut être difficile, et c'est l'une des raisons pour lesquelles relativement peu de GTs sont bien caractérisées.

Les GTs catalysent la formation d'une liaison glycosidique avec soit une rétention, soit une inversion de la configuration anomérique du sucre donneur. Les GTs qui catalysent une glycosylation avec inversion de configuration réalisent leur réaction par un déplacement unique selon un mécanisme de substitution nucléophile de type 2 ( $S_N2$ ) (Lairson et al., 2008). Dans la plupart des réactions enzymatiques de ce type, un résidu du site actif agit en tant que catalyseur (basique) de la réaction. Malgré tout, des données structurales ont montré que plusieurs enzymes, induisant une inversion de configuration, ne présentent pas de catalyseur basique évident, indiquant que ces enzymes doivent utiliser un mécanisme alternatif (Gloster, 2014; Lazarus et al., 2011). Le mécanisme catalytique utilisé par les GTs ayant une rétention de la conformation anomérique du sucre ne semble pas être le même pour toutes les enzymes. Le premier mécanisme proposé est un mécanisme à double déplacement selon un mécanisme



**Figure 11 – Représentation des différents types de repliements des glycosyltransférases (Gloster, 2014).**

Les GTs présentent trois types de repliements structuraux : les repliements de type GT-A, GT-B et GT-C. (A) Le repliement de type GT-A se base sur un repliement de type *Rossmann-like* ( $\alpha/\beta/\alpha$ , hélices  $\alpha$  entourant un feuillet  $\beta$ ) représenté ici par la structure cristallographique de la SpsA issue de *Bacillus subtilis* en complexe avec l'UDP et le magnésium (code PDB : 1QGS) (Charnock and Davies, 1999). (B) La  $\beta$ -Glucosyltransférase du phage T4 est un bon modèle de la structure des GTs de type GT-B (code PDB : 2BGU) avec deux domaines *Rossmann-like* qui se font face (Vrieling et al., 1994). (C) La structure de type GT-C est une structure présentant un grand domaine transmembranaire composé uniquement d'hélices  $\alpha$  comme illustré ici avec la structure de l'oligosaccharyltransférase, issue de *Campylobacter lari* (code PDB : 3RCE) (Lizak et al., 2011).

impliquant la formation d'un intermédiaire réactionnel covalent, analogue à celui utilisé par les glycosidases (Soya et al., 2011). Le mécanisme le plus probable dans ce contexte est un mécanisme de type substitution nucléophile intramoléculaire ( $S_{Ni}$ ) ou  $S_{Ni}$ -like, qui implique une interaction entre le groupe partant et le nucléophile attaquant sur la même face (Lira-Navarrete et al., 2011).

Actuellement, toutes les structures des GTs décrites adoptent trois types de repliements appelés GT-A, GT-B et GT-C (Figure 11). Les enzymes adoptant une conformation de type GT-A comprennent deux domaines structuraux juxtaposés de type *Rossmann-like* ( $\alpha/\beta/\alpha$ , hélices  $\alpha$  entourant un feuillet  $\beta$ ). Le maintien de la conformation active de ces enzymes dépend généralement de la présence d'un cofacteur, en général un cation métallique divalent (le manganèse dans le cas de la  $\beta$ 3GalT6) (Figure 11A). Au niveau du site actif de l'enzyme, l'ion métallique est coordonné par un motif appelé DXD (Asp-acide aminé quelconque-Asp) hautement conservé dans le site actif de ces GTs. Ce motif facilite la réaction catalytique en stabilisant les groupements phosphates (chargés) présents au niveau du substrat donneur (un sucre-nucléotide comme l'UDP-galactose dans le cas des galactosyltransférases par exemple) (Charnock and Davies, 1999; Tsutsui et al., 2013). Les enzymes possédant une conformation de type GT-B possèdent deux domaines *Rossmann-like* qui se font face et dont le site actif est généralement localisé dans la région située entre ces deux domaines structuraux (Figure 11B). Ces enzymes ne dépendent pas de la présence d'un ion métallique, dans la mesure où seuls les résidus du site actif sont impliqués dans la fixation du substrat donneur et dans la catalyse (Albesa-Jové et al., 2014). Plus récemment, une troisième famille structurale de GTs a été identifiée, les enzymes possédant un repliement de type GT-C. Les représentants structuraux de ces GTs sont des protéines membranaires hydrophobes. Ces enzymes présentent un motif DXD modifié dans la première boucle extracellulaire et utilisent des donneurs de sucre liés à un lipide (Liu and Mushegian, 2003; Lizak et al., 2011; Schmid et al., 2016) (Figure 11C).

Après avoir introduit les GTs, nous allons maintenant nous concentrer sur la famille des  $\beta$ 1,3-Galactosyltransférases ( $\beta$ 3GalTs) à laquelle appartient notre enzyme d'intérêt.

### **III.2. La famille des $\beta$ 1,3-Galactosyltransférases ( $\beta$ 3GalTs) :**

La famille des  $\beta$ 3GalTs compte six membres chez l'Homme, notés h $\beta$ 3GalT1, 2, 4, 5, 6, 7. Ces enzymes sont des protéines transmembranaires présentant une topologie de type II, c'est-à-dire possédant un court domaine N-terminal cytoplasmique, une seule région transmembranaire composée d'une vingtaine d'acides aminés (hydrophobes) organisés en hélice, une région *stem* qui relie le segment transmembranaire à un large domaine

**Tableau 8 – Caractéristiques fonctionnelles des membres de la famille des  $\beta$ 1,3-Galactosyltransférases humaines.**

Enzymes	Localisation chromosomique	Taille de la protéine (acides aminés)	Homologie (par rapport à la $\beta$ 3GalT6)	Substrats accepteurs	Références
$\beta$ 3GalT1	2p14	326	23,7%	GlcNAc- $\beta$ 1 $\rightarrow$ glycoprotéine	(Amado et al., 1998) (Hennet et al., 1998)
$\beta$ 3GalT2	1q31	422	28,4%	GlcNAc- $\beta$ 1 $\rightarrow$ ovalbumine GlcNAc- $\beta$ 1,3-Gal- $\beta$ 1,4-Glc- $\beta$ 1-Ceramide (Lc <sub>3</sub> , lactosylcéramide)	(Amado et al., 1998) (Hennet et al., 1998) (Kolbinger et al., 1998)
$\beta$ 3GalT4	6p21	378	31,3%	GalNAc- $\beta$ 1,4-(NeuAc- $\alpha$ 2,3)-Gal- $\beta$ 1,4-Glc- $\beta$ 1-Ceramide (GM <sub>2</sub> ) GalNAc- $\beta$ 1,4-Gal- $\beta$ 1,4-Glc- $\beta$ 1-Ceramide (AsialoGM <sub>2</sub> )	(Amado et al., 1998) (Miyazaki et al., 1997)
$\beta$ 3GalT5	21q22.3	310	27,9%	GalNAc- $\beta$ 1,3-Gal- $\alpha$ 1,4-Gal- $\beta$ 1,4-Glc- $\beta$ 1-Cer (Gb <sub>4</sub> ) O-linked core3 Antigène de Lewis A	(Isshiki et al., 1999) (Zhou et al., 1999)
$\beta$ 3GalT6	1p36.33	329	-	Gal- $\beta$ 1,4-Xyl(2P)-protéine <i>core</i> (GAG)	(Bai et al., 2001)
$\beta$ 3GalT7	7	363	26,4%	GalNAc- $\alpha$ 1-O-phenylglycopeptide PSGL-1 ( <i>P-selectin glycopeptide ligand</i> )	(Ju et al., 2002)

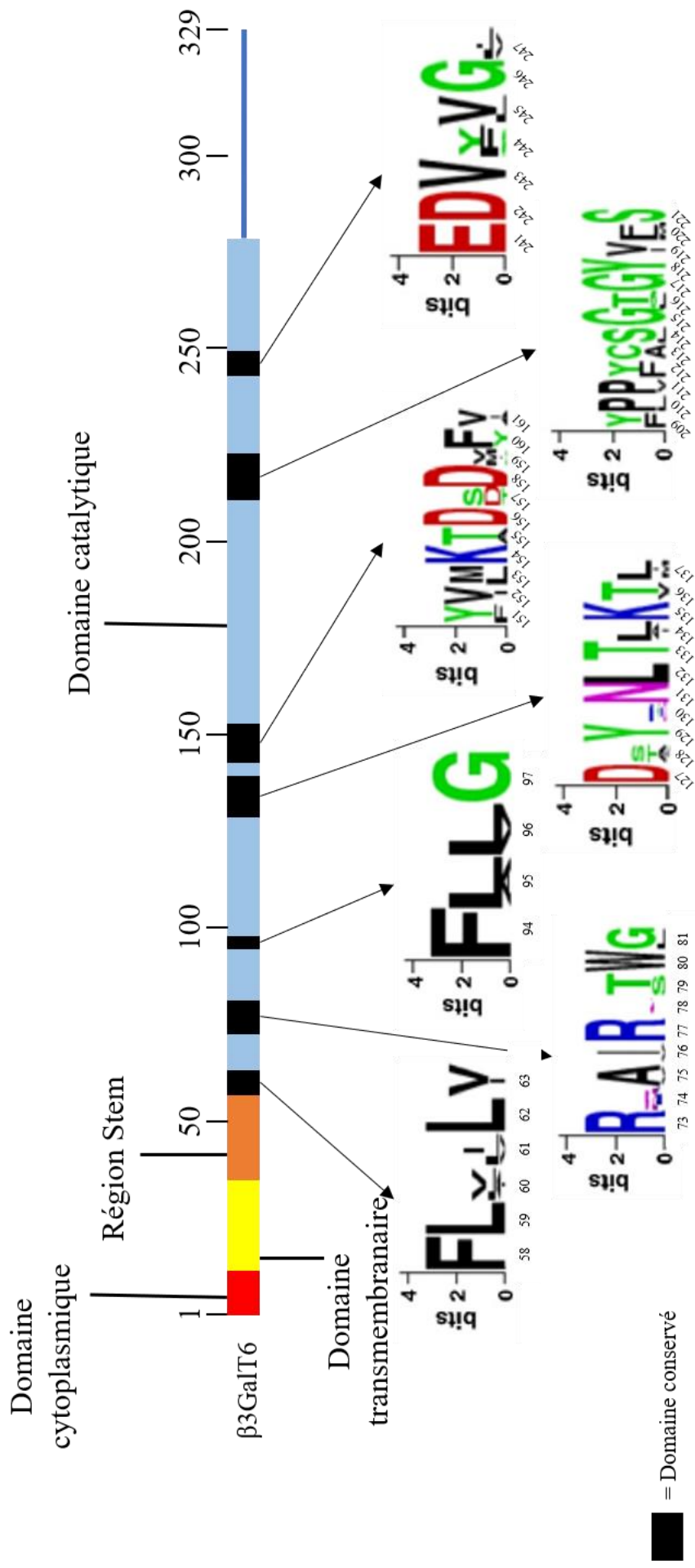
C-terminal contenant le centre catalytique de l'enzyme. Les  $\beta$ 3GalTs transfèrent un Gal à partir d'un nucléotide sucre donneur, l'UDP-Gal, sur des substrats accepteurs présentant des résidus GlcNAc, GalNAc ou Gal terminaux en formant une liaison  $\beta$ 1,3 (avec inversion de configuration du Gal).

Les propriétés fonctionnelles des différents membres de la famille des  $\beta$ 3GalTs sont présentées dans le Tableau 8. Les  $\beta$ 3GalT1 et -2 présentent une spécificité de substrat similaire en transférant le Gal sur un résidu GlcNAc terminal d'accepteurs de nature lipidique, comme par exemple le Lc<sub>3</sub>Cer (lactosylcéramide), pour former un motif [Gal- $\beta$ 1,3-GlcNAc] (Amado et al., 1998; Hennet et al., 1998). Les  $\beta$ 3GalT4, -5 et -7 catalysent la formation d'un motif [Gal- $\beta$ 1,3-GalNAc] et interviennent dans la synthèse des dérivés gangliosides (Miyazaki et al., 1997) ou encore dans la synthèse de l'antigène de Lewis de type 1 [Gal- $\beta$ 1,3-(Fuc- $\alpha$ 1,4)-GlcNAc] (Isshiki et al., 1999). La  $\beta$ 3GalT6 est la seule galactosyltransférase qui réalise le transfert d'un Gal sur un autre résidu Gal terminal et la seule qui intervienne dans la biosynthèse des GAGs en catalysant l'addition du deuxième Gal de l'amorce tétrasaccharidique [GlcA- $\beta$ 1,3-Gal- $\beta$ 1,3-Gal- $\beta$ 1,4-Xyl- $\beta$ 1-O-Ser] sur le motif disaccharidique fixé sur la protéine *core* des PGs [Gal- $\beta$ 1,4-Xyl-O-Ser] (Amado et al., 1998, 1999; Bai et al., 2001; Hennet, 2002).

### **III.3. La $\beta$ 1,3-Galactosyltransférase 6 ( $\beta$ 3GalT6) :**

#### **III.3.1. Organisation structurale de la $\beta$ 3GalT6 :**

La  $\beta$ 3GalT6 est une GT membranaire présente dans les compartiments médians de l'appareil de Golgi. Elle catalyse l'une des étapes précoces de la biosynthèse des GAGs en transférant un résidu Gal sur la zone de liaison protéine *core*-GAG en cours de synthèse (Bai et al., 2001). Le gène codant la protéine humaine (*hB3GALT6*) est localisé sur le chromosome 1 au locus 1p36.33 et est composé d'un unique exon d'environ 2.8 kb. L'ADNc a été cloné pour la première fois par Bai et al. (2001) et code une protéine de 329 acides aminés présentant une masse moléculaire apparente de 37 kDa (Figure 12). Comme nous l'avons précédemment décrit pour les autres membres de la famille des  $\beta$ 3GalTs, la  $\beta$ 3GalT6 présente une topologie de type II avec un court domaine cytoplasmique N-terminal de 11 acides aminés et un unique domaine transmembranaire composé de 19 acides aminés majoritairement hydrophobes. Il est intéressant de noter que ce domaine transmembranaire contient un résidu cystéine conservé chez l'Homme, la souris et le nématode *Caenorhabditis elegans* qui pourrait être impliqué dans l'oligomérisation de l'enzyme, ce qui n'a pas été clairement montré à ce jour (Bai et al., 2001; Qian et al., 2001). Le domaine transmembranaire est suivi d'une région charnière appelée *stem*



**Figure 12 – Représentation schématique de l’organisation modulaire de la  $\beta 3\text{GalT6}$  (Patel and Balaji, 2007).**

La  $\beta 3\text{GalT6}$  humaine est une enzyme de 329 acides aminés localisée dans les compartiments médians de l’appareil de Golgi. Elle présente une topologie de type II avec un court domaine N-terminal cytoplasmique (en rouge), un unique domaine transmembranaire (en jaune) et le domaine catalytique (en bleu) composant la majeure partie de la protéine. Ce dernier est relié au domaine transmembranaire par une courte région nommée *stem* (en orange).

Sept domaines conservés ont pu être mis en évidence au sein de la famille des  $\beta 3\text{GalTs}$  sur la base de leur homologie de séquence primaire (en noir). Ces domaines sont localisés dans le domaine catalytique et sont potentiellement impliqués dans le maintien de la structure tridimensionnelle et la fonction catalytique de l’enzyme.

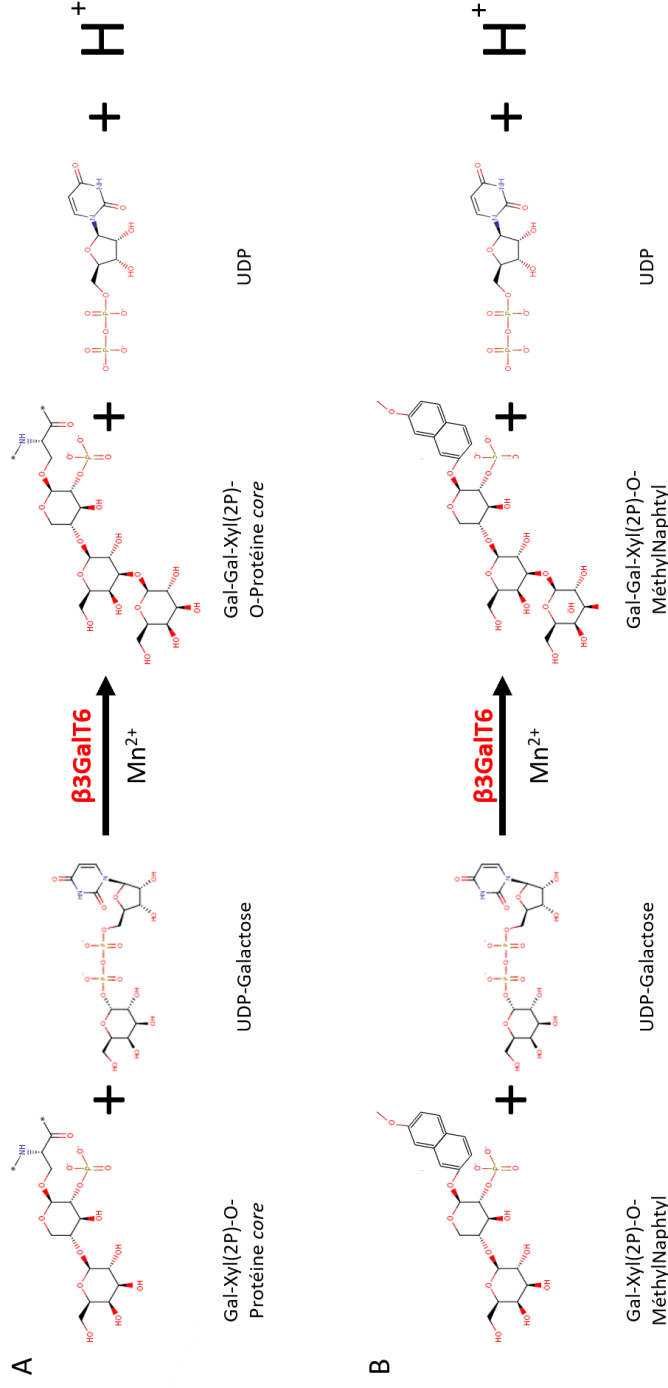
qui fait le lien entre l'hélice du domaine transmembranaire et le domaine catalytique. Cette région a la particularité de contenir un segment riche en Ser, Pro, Ala et Gly, appelé domaine « SPLAG » présent également dans d'autres GTs (comme les autres  $\beta$ 3GalTs ou la  $\beta$ 4GalT7 par exemple) (Almeida et al., 1999; Patel and Balaji, 2007). L'ensemble de la région *stem* correspond à un domaine structural important dont la présence serait nécessaire pour maintenir une conformation tridimensionnelle active de l'enzyme au contact de la membrane et/ou être impliquée dans la formation de complexes multimériques.

Patel and Balaji (2007) ont mis en évidence la présence de sept domaines conservés à travers la famille des  $\beta$ 3GalTs (Figure 12). Ces éléments conservés sont localisés dans le domaine catalytique et sont potentiellement impliqués dans le repliement du site actif et/ou dans la fixation des substrats et/ou dans le mécanisme catalytique de l'enzyme.

La structure cristallographique de l'enzyme n'est pas résolue à ce jour. Cependant, l'établissement d'un modèle moléculaire de la  $\beta$ 3GalT6 et des enzymes de la famille des  $\beta$ 3GalTs a permis de proposer que ces enzymes appartiennent à la famille structurale GT31 (selon la nomenclature CAZy, *Carbohydrate Active Enzyme*) avec un repliement de type GT-A (Bai et al., 2001; Patel and Balaji, 2007). Ce repliement est caractéristique des GTs utilisant un nucléotide diphosphate-sucre comme substrat donneur (voir partie III.2.2). La structure de la  $\beta$ 3GalT6 présente par ailleurs un feuillet  $\beta$  central entouré d'hélices  $\alpha$  appelé *Rossmann-like fold* contenant un motif DXD. Dans le cas de la  $\beta$ 3GalT6, ce motif est composé de trois aspartates des positions 156 à 158 (Asp156-Asp157-Asp158) et intervient dans la fixation du substrat donneur, l'UDP-Gal.

### III.3.2. Réaction enzymatique catalysée par la $\beta$ 3GalT6 :

La  $\beta$ 3GalT6 transfère un résidu Gal à partir du donneur UDP-Gal, sur un  $\beta$ -galactoside ou sur la zone de liaison protéine *core*-GAG possédant un Gal terminal impliqué dans une liaison  $\beta$ -O-glycosidique (substrat accepteur) pour former un motif (Gal- $\beta$ 1,3-Gal) (Figure 13). La réaction catalysée par l'enzyme nécessite la présence d'un cofacteur métallique, le cation divalent manganèse ( $Mn^{2+}$ ), comme cela a été décrit pour d'autres galactosyltransférases (Robinson and Robinson, 1985). Les données cinétiques obtenues pour d'autres GTs appartenant à la même famille structurale (repliement GT-A, famille GT31, voir partie III.2.1.) suggèrent que la réaction enzymatique se déroule selon un mécanisme séquentiel ordonné en suivant une réaction de substitution nucléophile de type 2 ( $S_N2$ ): le premier substrat interagissant avec la  $\beta$ 3GalT6 serait le substrat donneur (UDP-Gal), suivi par le substrat



**Figure 13 – Représentation schématique de la réaction catalysée par la  $\beta 3\text{GalT6}$ .**

La réaction catalysée par la  $\beta 3\text{GalT6}$  est l'addition d'un galactose sur le début de l'amorce tétrasaccharidique (Gal- $\beta 1,4$ -Xyl(2P)- $\beta 1$ -O-Proteine core) (A). Le résidu galactose est transféré depuis le nucléotide diphosphate donneur, l'UDP-Gal. Le galactose est fixé au substrat accepteur par l'intermédiaire d'une liaison O-glycosidique de type  $\beta 1,3$  en présence d'un cofacteur métallique, le cation divalent manganèse ( $\text{Mn}^{2+}$ ), indispensable à l'activité enzymatique. La réaction enzymatique peut être également réalisée à partir de substrat accepteur synthétique comme le Gal- $\beta 1,4$ -Xyl(2P)- $\beta 1$ -O-Méthoxy-naphtyl (B).



accepteur (galactoside) après un changement de conformation de l'enzyme induit par la fixation du premier substrat (Breton et al., 2012). La réaction se termine par la libération des deux produits de réaction : le trisaccharide, Gal- $\beta$ 1,3-Gal- $\beta$ 1,4-Xyl-Ser (protéine *core*) et l'UDP.

Intervenant dans la biosynthèse de l'amorce tétrasaccharidique, la  $\beta$ 3GalT6 est une enzyme importante pour l'initiation de la biosynthèse des GAGs de type CS/DS et HS, comme le montre la production de GAGs fortement réduite dans les cellules HeLa après extinction de l'expression du gène *B3GALT6* suite à l'utilisation d'un siRNA ciblant spécifiquement le gène d'intérêt (Bai et al., 2001). Ces résultats ont été également observés dans des cellules CHO (*Chinese Hamster Ovary*) après une extinction du gène *B3GALT6* par la technologie CRISPR/Cas9 (Chen et al., 2018).

Récemment, des mutations du gène *B3GALT6* ont été identifiées et associées à un groupe de maladies génétiques rares des tissus conjonctifs : le SEDsp et le SEMD-JL1. La compréhension du fonctionnement de l'enzyme ainsi que l'analyse des défauts engendrés par ces mutations chez les patients pourront permettre d'apporter des informations sur le rôle de certains acides aminés ou motifs peptidiques dans la fonctionnalité de la  $\beta$ 3GalT6.

### **III.4. Implication du gène *B3GALT6* dans la pathogénie du SED spondylo-dysplastique :**

#### **III.4.1. Des mutations du gène *B3GALT6* impliquées dans deux types de pathologies :**

Des mutations du gène *B3GALT6* ont été récemment identifiées et associées à deux ostéochondrodysplasies proches, le SEDsp et la SEMD-JL1 (Voir partie I, Tableau 9) (Ben-Mahmoud et al., 2018; Brady et al., 2017; Malfait et al., 2013; Nakajima et al., 2013; Ritelli et al., 2015; Van Damme et al., 2018; Vorster et al., 2015).

Le SEDsp est une maladie autosomale récessive impliquant des mutations génétiques au niveau des gènes *B4GALT7*, *B3GALT6* et *SLC39A13* (voir partie I.2). Les patients atteints de ce sous-type de SED présentent une petite taille, une hypotonie musculaire, une hyperlaxité et une fragilité de la peau, des contractures et une hypermobilité des articulations, une cyphoscoliose, un trouble du développement moteur et parfois une déficience intellectuelle (Tableau 1) (Malfait et al., 2013; Van Damme et al., 2016). La sévérité du tableau clinique est cependant variable et dépend de la nature des mutations (mutation ponctuelle ou délétion, mutation homozygote ou hétérozygote composite). Les études moléculaires réalisées à partir

**Tableau 9 - Mutations du gène *B3GALT6* et de la protéine décrite chez les différents patients.**

A) Forme spondylo-dysplastique du Syndrome d'Ehlers-Danlos (SEDsp) d'après Malfait et al., 2013 et (B) Familles 1-7 : Dysplasie spondylo-épimétaphysaire avec hyperlaxité ligamentaire type 1 (SEMD-JL1) et famille 8-10 : SED d'après Nakajima et al., 2013. (C) SEDsp d'après Vorster et al., 2015. (D) Patient 1 et 2 : SEMD-JL1 d'après Ritelli et al., 2015. (E) SEDsp d'après Van Damme et al., 2018.

A)

Dénomination des patients	Mutations au niveau du gène <b>B3GALT6</b>	Mutations au niveau de la protéine <b>β3GalT6</b>
Patient 2	c.619G>C	p.Asp207His
Patient 3	c.323_344del c.619G>C	Délétion p.Ala108Glyfs*163 p.Asp207His
Patient 5	c.649G>A	p.Gly217Ser

B)

Dénomination des familles de patients	Mutations au niveau du gène <b>B3GALT6</b>	Mutations au niveau de la protéine <b>β3GalT6</b>
Famille 1	c.1A>G c.694C>T	p.Met1 ? p.Arg232Cys
Famille 2	c.1A>G c.466G>A	p.Met1 ? p.Asp156Asn
Famille 3	c.1A>G	p.Met1 ?
Famille 4	c.1A>G c.694C>T	p.Met1 ? p.Arg232Cys
Famille 5	c.694C>T c.899G>C	p.Arg232Cys p.Cys300Ser
Famille 6	c.1A>G c.193A>G	p.Met1 ? p.Ser65Gly
Famille 7	c.200C>T c.694C>T	p.Pro67Leu p.Arg232Cys
Famille 8	c.353delA c.925T>A	Délétion p.Ala118Alafs*160 p.Ser309Thr
Famille 9	c.588delG c.925T>A	Délétion p.Arg197Alafs*81 p.Ser309Thr
Famille 10	c.16C>T c.415_423del	p.Arg6Trp Délétion p.Met139Ala141

C)

Dénomination des familles de patients	Mutations au niveau du gène <b>B3GALT6</b>	Mutations au niveau de la protéine <b>β3GalT6</b>
Famille 1	c.16C>T c.200C>T	p.Arg6Trp p.Pro67Leu
Famille 2	c.235A>G	p.Thr79Ala
Famille 3	c.235A>G	p.Thr79Ala
Famille 4	c.235A>G	p.Thr79Ala
	c.200C>T	p.Pro67Leu
	c.235A>G	p.Thr79Ala
Famille 5	c.200C>T	p.Pro67Leu
	c.235A>G	p.Thr79Ala
Famille 8	c.235A>G	p.Thr79Ala
	c.235A>G	p.Thr79Ala
Famille 9	c.235A>G	p.Thr79Ala
	c.200C>T	p.Pro67Leu
Famille 10	c.235A>G	p.Thr79Ala

D)

Dénomination des patients	Mutations au niveau du gène <b>B3GALT6</b>	Mutations au niveau de la protéine <b>β3GalT6</b>
Patient 1	c.227del c.766C>T	Délétion p.Ile76Thrfs*202 p.Arg256Trp
Patient 2	c.227del c.766C>T	Délétion p.Ile76Thrfs*202 p.Arg256Trp

E)

Dénomination des patients	Mutations au niveau du gène <b>B3GALT6</b>	Mutations au niveau de la protéine <b>β3GalT6</b>
PI:1	c.795A>C c.929A>G	p.Glu265Asp p.Tyr310Cys
PII :1	c.513_520del c.925T>A	Délétion p.Glu174Alafs*266 p.Ser309Thr
PIII :1	c.477del c.782G>A	Délétion p.Phe160Serfs*118 p.Arg261His
PIV :1 PIV :2 PIV :3	c.197_253del c.953C>T	p.Ala66_Arg84del p.Pro318Leu
PV :1	c.3G>T c.631C>T	p. ? p.Pro211Ser
PVI :1	c.513_520del	Délétion p.Glu174Alafs*266

	c.430G>A	p.Asp144Asn
PVII :1	c.545A>G	p.Tyr182Cys
PVIII :1	c.77T>C c.808G>A	p.Leu26Pro p.Gly270Ser
PIX :1	c.513_520del	Délétion p.Glu174Alafs*266
PIX :2	c.181G>C	p.Val61Leu

d'échantillons de tissus de ces patients ont permis de montrer la présence de mutations chez 14 patients issus de 11 familles indépendantes (Malfait et al., 2013; Van Damme et al., 2018). Une autre équipe a également identifié quatre patients (issus de 3 familles indépendantes) atteints de la même forme de la maladie (Nakajima et al., 2013). Ces patients présentent des symptômes proches de ceux décrits précédemment. Pour le moment, il n'est pas possible de relier la sévérité de ces symptômes à la nature de la mutation de la protéine, ni de prévoir l'évolution de la pathologie.

L'autre forme de pathologie présentant des mutations du gène *B3GALT6* est la SEMD-JL-1. Cette pathologie est une maladie autosomale récessive qui a été décrite en 1984 chez des patients qui présentaient des symptômes similaires au SEDsp (Beighton et al., 1984). Plus récemment, un certain nombre de patients atteints de SEMD-JL1 ont été décrits comme porteurs de mutations du gène *B3GALT6*, dont huit patients provenant de sept familles indépendantes (Nakajima et al., 2013) et deux patientes provenant d'une même famille (Ritelli et al., 2015) (Tableau 9B et 9D). Ces patients présentent un tableau clinique affectant principalement les tissus squelettiques. Les symptômes retrouvés chez les patients sont similaires à la forme SEDsp avec également un mauvais alignement spinal, une hypermobilité articulaire (en particulier au niveau des mains), une asymétrie de la cage thoracique, une dysplasie généralisée et sévère du squelette et des pieds bots.

#### III.4.2. Impact des mutations sur la protéine :

Les mutations de la  $\beta$ GalT6 identifiées correspondent (i) à des mutations faux-sens de résidus localisés dans ou à proximité de motifs conservés entre les  $\beta$ GalTs ou entre les  $\beta$ GalT6 issues de différentes espèces ou (ii) à des délétions entraînant des décalages du cadre de lecture avec apparition d'un codon stop prématuré et donc conduisant à la production d'une protéine tronquée non fonctionnelle (Tableau 9). A l'heure actuelle, il a pu être répertorié une trentaine de mutations au niveau du gène *B3GALT6*. La protéine mutée présente une perte d'activité galactosyltransférase variable selon les mutations, mais le plus souvent sévère. Une réduction significative ou complète de l'activité enzymatique de la  $\beta$ GalT6 a été observée dans des homogénats cellulaires issus de fibroblastes de peau de patients porteurs de mutations faux sens et/ou de délétions, qu'elles soient homozygotes ou hétérozygotes (Malfait, 2013 et Van Damme, 2018). Des résultats comparables ont été obtenus dans des cellules HeLa surexprimant les mutants Ser65Gly, Pro67Leu, Asp156Asn, Arg232Cys et Cys300Ser attestant du caractère



non fonctionnel des enzymes mutées (Nakajima et al., 2013). Les autres mutations décrites par Ritelli et al. (2015) et Vorster et al. (2015) n'ont pas été étudiées d'un point de vue fonctionnel et leurs conséquences sur l'activité enzymatique ne sont pas connues. Pour certaines mutations (Ser65Gly, Pro67Leu, Thr79Ala, Phe186Leu, Arg232Cys, Arg256Trp), la protéine est localisée dans le RE au lieu de l'appareil de Golgi (Ben-Mahmoud et al., 2018; Nakajima et al., 2013), expliquant l'absence d'activité des protéines et/ou de la diminution de la synthèse des GAGs (Nakajima et al., 2013; Ritelli et al., 2015). Dans le cas des mutations de la  $\beta$ 3GalT6 associées à la forme SEDsp, les protéines mutées Glu265Asp, Tyr310Cys, Arg261His et Tyr182Cys sont plus faiblement exprimées dans les fibroblastes de patients, ce qui pourrait être dû à une instabilité de la protéine produite (Van Damme et al., 2018).

### III.4.3. Impact des mutations de la $\beta$ 3GalT6 sur la pathogénie des SED :

A partir de cellules lymphoblastoïdes de patients (patients 1 et 2 de la famille 1 et patient 3 de la famille 2) (Nakajima et al., 2013), des analyses de la composition en disaccharides des chaînes de GAGs ont été réalisées par chromatographie échangeuse d'anions après digestion enzymatique des GAGs par la chondroïtinase ou l'héparinase. Ces analyses ont montré une réduction du taux de HS synthétisés par les cellules mais également une forte augmentation de la quantité de CS/DS. Comme il a été discuté précédemment, la  $\beta$ 3GalT6 est impliquée dans la formation de l'amorce tétrasaccharidique, qui est commune aux deux types de GAGs (HS et CS/DS). Donc, il est surprenant de voir une augmentation et non une diminution de la quantité des chaînes de CS/DS. Pour que les cellules puissent augmenter la synthèse des CS/DS, une hypothèse pourrait être que leur synthèse puisse faire appel à un mécanisme alternatif ne faisant pas intervenir la  $\beta$ 3GalT6. Il peut également être supposé que la diminution de la synthèse des HS entraîne l'activation de la voie de synthèse d'un autre type de GAGs, dans un effet compensatoire. Des analyses par immunofluorescence dans les fibroblastes de peau de patients atteints de SEDsp (patients 2, 3 et 5) ont montré une forte diminution de l'expression des chaînes de HS membranaires (Malfait et al., 2013). De la même manière, Ritelli et al. (2015) ont décrit une forte diminution de la synthèse des HS dans des fibroblastes de derme de patients. Par immunofluorescence, ils ont également montré une diminution concomitante de la protéine *core* du perlécan, une diminution des chaînes de CS sans impact sur l'expression des protéines *core* de la décorine et du versican. Chez ces patients, la synthèse des CS est également impactée mais seuls les HSPG voient l'expression de leur protéine *core* diminuer. Dans ces conditions, il est possible que cette diminution permette à la cellule de prendre en charge les protéines *core*



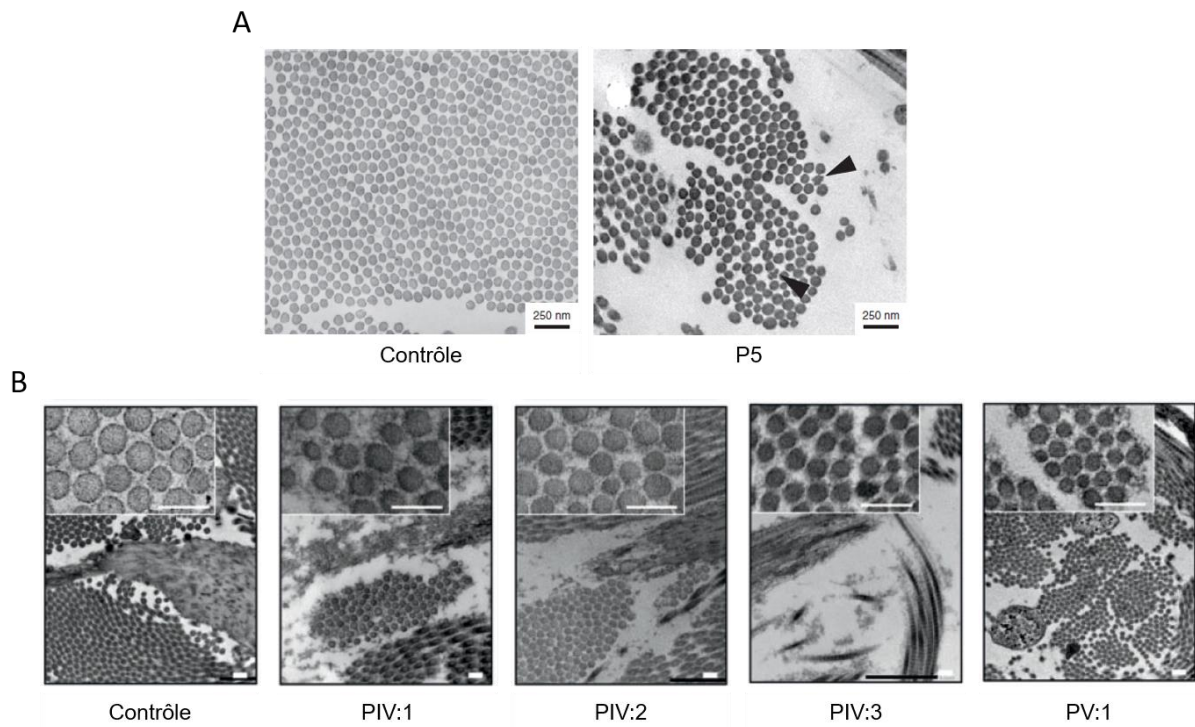


des CSPGs et favoriser leur glycanation.

Dans le cadre des études menées par Malfait et al. (2013) et Van Damme, Pang et al., (2018), les mutations du gène *B3GALT6* conduisent à une diminution plus ou moins importante de la glycosylation de la protéine *core* de la décorine *in cellulo* (variable en fonction des patients). Il est donc possible que l'impact des mutations sur le niveau d'expression et/ou sur le niveau d'activité enzymatique de la  $\beta 3\text{GalT6}$ , et donc sur le niveau d'expression des chaînes de GAGs de type Hep/HS et CS/DS, explique la variabilité des symptômes et leur sévérité différente d'un patient à l'autre. En revanche, le mécanisme par lequel ces déficits génétiques en  $\beta 3\text{GalT6}$  modifient de façon différente les voies de synthèse des HS et des CS/DS reste difficile à comprendre. Récemment, l'existence d'une voie de biosynthèse non canonique pour les CS a été proposée suite à l'étude de la bikunine (un PG soluble portant des chaînes de CS) présente dans les urines de volontaires sains. Cette étude montre l'existence d'une amorce protéine *core*-GAG trisaccharidique [GlcA-Gal-Xyl], dépourvue d'un résidu Gal en lieu et place du tétrasaccharide classique [GlcA-Gal-Gal-Xyl] (Persson et al., 2019). La présence de chaînes de GAGs observées chez les patients porteurs d'une déficience génétique pour l'une des galactosyltransférases impliquées dans la biosynthèse de cette amorce tétrasaccharidique, pourrait s'expliquer par une voie alternative utilisant cette forme d'amorce trisaccharidique. Des études sont en cours pour répondre à cette question, en collaboration avec le Dr Frederik Noborn et le Dr Göran Larson (*Institute of Biomedicine, Sahlgrenska Academy, Université de Göteborg, Sweden*).

L'impact au niveau moléculaire et cellulaire des mutations du gène *B3GALT6* et ses répercussions au niveau du tissu et de leurs propriétés physiologiques et/ou biomécaniques sont en cours d'étude en collaboration avec les collègues généticiens de l'Université de Gent (Belgique). Les travaux réalisés précédemment ont permis de mettre en évidence un retard de migration des fibroblastes issus de patients porteurs de mutations sur le gène *B3GALT6* par rapport à des fibroblastes de patients non déficients. Sur la base de ces résultats, un lien peut être potentiellement établi entre les mutations présentes sur la séquence de l'enzyme et le ralentissement du processus de cicatrisation observé chez les patients porteurs des mutations (Malfait et al., 2013; Van Damme et al., 2018).

Ces défauts au niveau de la synthèse des HS et des CS/DS sont également associés à des défauts d'assemblage des collagènes, mis en évidence par l'absence des collagènes de type III et V dans la MEC de fibroblastes de peau de patients (Ritelli et al., 2015). Des analyses par microscopie électronique à transmission sur les fibroblastes de derme de patients ont par ailleurs montré une organisation anormale du réseau de collagène fibrillaire chez les patients porteurs



**Figure 14 – Effets de la déficience en  $\beta$ 3GalT6 sur l’organisation des fibres de collagène** (Malfait et al., 2013; Van Damme et al., 2018).

Des analyses par microscopie électronique à transmission ont permis d’analyser l’organisation des fibres de collagène et l’architecture matricielle au niveau des tissus de patients atteints de SEDsp. Les patients présentent des fibres de collagène plus petites avec un contour irrégulier. (A) Patient 5 issu de Malfait et al. (2013). (B) Patients IV: 1, 2, 3 et V:1 issus de Van Damme, Pang et al., (2018).

de mutations du gène *B3GALT6*, avec des fibrilles de collagène de diamètre plus petit et présentant un aspect irrégulier (Figure 14) (Malfait et al., 2013; Van Damme et al., 2018). Le lien entre la déficience génétique en  $\beta 3\text{GalT6}$  et la fibrillogenèse anormale du collagène chez ces patients n'est pas établi à ce jour.

Nous avons vu précédemment que la décorine était impliquée dans la fibrillogenèse du collagène en s'associant aux fibres de collagène dont elle peut réguler le diamètre et la taille (voir partie II.2.2.2 et Figure 7), mais d'autres PGs peuvent également intervenir dans ce phénomène. Les petits PGs de la famille des SLRPs, comme la décorine, la fibromoduline ou encore le biglycan, sont impliqués dans la fibrillogenèse des collagènes matriciels. Des études ont mis en évidence une augmentation de la formation de liaisons croisées (*cross-link*) entre les chaînes de collagène dans des fibroblastes de souris déficients en décorine après ajout de décorine recombinante dans le milieu de culture (Seidler et al., 2005). Une autre étude sur la décorine a permis de montrer que ce petit PG est capable de se lier à la région C-terminale d'un monomère de collagène et plus précisément au niveau d'un site de liaisons croisées (Keene et al., 2000). Il a ainsi été montré l'importance de la décorine dans la formation des liaisons croisées et donc dans la fibrillogenèse du collagène. Par ailleurs, un autre PG, la fibromoduline (un PG pouvant porter jusqu'à quatre chaînes de KS), a été également impliqué dans la formation de ces liaisons croisées au cours de la maturation des collagène. Des études sur des tendons de souris déficientes en fibromoduline (*Fmod* *-/-*) ont montré un aspect irrégulier et une diminution de la taille des fibres de collagène, ainsi que l'augmentation d'une forme aberrante de liaisons croisées pour la chaîne  $\alpha 2(I)$  (Jepsen et al., 2002; Kalamajski et al., 2014; Svensson et al., 1999). D'autres études ont permis de démontrer que la fibromoduline interagit avec les sites de liaisons croisées au niveau du collagène et avec la *Lysyl Oxidase* (LOX). Ces interactions permettent d'augmenter l'activité de la LOX et donc de favoriser la fibrillogenèse (Kalamajski et al., 2016; Tillgren et al., 2016). Ces études ont permis de montrer l'importance des PGs, dont la décorine et la fibromoduline, dans le processus de fibrillogenèse du collagène et expliquent qu'un défaut d'expression de ces composants peut induire des altérations de la formation des fibres de collagène. Chez les patients atteints de SEDsp, les mutations du gène *B3GALT6* induisent une diminution de la synthèse des chaînes de GAGs affectant le niveau de glycosylation de la décorine. Si ces PGs présentent une mauvaise ou une absence de glycanation de leur protéine *core*, leurs rôles dans la fibrillogenèse vont être affectés et donc potentiellement conduire à une mauvaise organisation structurale des fibres de collagène. Il est donc possible que les mutations du gène *B3GALT6* conduisent à des perturbations de la fibrillogenèse des collagènes fibrillaires. Ritelli et al. (2015) ont observé le même aspect irrégulier des fibres de



collagène par immunofluorescence, avec une destruction des fibres de collagène de type I, III et V dans les fibroblastes de patients atteints de SEDsp. Les mutations du gène codant la  $\beta$ 3GalT6 impactent donc de façon négative la biosynthèse des GAGs et les fonctions biologiques qui leur sont associées, comme ici l'organisation matricielle et la migration cellulaire, fonctions qui peuvent expliquer un certain nombre de symptômes observés chez les patients comme l'hyperlaxité articulaire, une peau élastique et un retard de cicatrisation.



# PROBLEMATIQUE ET OBJECTIFS





Le SED est un groupe de maladies génétiques rares qui conduit à des manifestations cliniques très diverses et de sévérité variable, touchant principalement les tissus conjonctifs comme la peau, les os, les articulations et la paroi des vaisseaux sanguins. Les premières formes de SED décrites présentaient des mutations au niveau des gènes codant les collagènes fibrillaires ou des enzymes intervenant dans leur biosynthèse ou leur maturation. Récemment, les GTs qui catalysent la biosynthèse des chaînes de GAGs sur les PGs ont également été impliquées dans ces syndromes. Malgré la différence d'origine de la déficience génétique, les symptômes observés chez les patients présentent des caractéristiques communes comme une hypermobilité articulaire, une hyperlaxité de la peau et un défaut de cicatrisation associés à des altérations multiples des tissus conjonctifs.

L'équipe MolCelTEG (*Molecular, Cellular, Therapeutic Engineering and Glycosyltransferases*) s'est, depuis de nombreuses années, intéressée à la caractérisation structurale et fonctionnelle des GTs impliquées dans la biosynthèse des GAGs, en particulier aux enzymes responsables de la formation de l'amorce tétrasaccharidique (Bui et al., 2010; Gulberti et al., 2005; Talhaoui et al., 2010). Dans des travaux réalisés en collaboration avec le Centre de Génétique Médicale de Gent (Belgique), la présence de mutations au niveau du gène *B3GALT6* a été identifiée (Malfait et al., 2013; Van Damme et al., 2018). Les études biochimiques ont montré que les fibroblastes de derme de peau des patients atteints de SED présentaient une perte de l'activité de l'enzyme. Cet impact a pour conséquence d'induire un défaut de synthèse des GAGs membranaires et matriciels, associé à un retard de migration cellulaire par rapport aux cellules non déficientes. La diminution de la synthèse des GAGs va induire une altération de la structure et de la fonctionnalité des tissus qui, par extension, pourraient expliquer les symptômes observés chez ces patients atteints de SEDsp.

Néanmoins, les relations génotype-phénotype n'ont pu être complètement établies pour ce sous-type de SED, ainsi il n'est pas possible de prédire la sévérité des symptômes observés chez les patients au regard des différentes mutations du gène *B3GALT6*. L'impact de ces mutations sur la structure et la fonction de la protéine n'est pas non plus connu à l'heure actuelle et sera abordé dans ce travail. Par conséquent, il est impossible de savoir quelle sera l'importance du défaut de synthèse des PGs et donc de la désorganisation de la MEC, et par extension, de présager de la sévérité et de l'évolution des symptômes. Les cliniciens manquant d'informations sur ces points, la prise en charge des patients est compliquée à mettre en place du fait de l'évolution de la pathologie qui ne peut être prédite.

Sur la base des travaux impliquant la  $\beta$ 3GalT6 dans la forme SEDsp (Malfait et al., 2017), nous cherchons à comprendre l'impact des mutations de la  $\beta$ 3GalT6 dans la pathogénie de la



maladie en essayant de comprendre par quel(s) mécanisme(s) celles-ci affectent-elles la fonctionnalité et l'activité biologique de cette enzyme. Nous espérons ainsi mieux comprendre les relations existant entre la nature de la mutation de l'enzyme, les répercussions au niveau cellulaire et la sévérité des symptômes observés chez les patients.

Pour ce faire, nous proposons deux axes de travail :

(1) Le premier axe sera consacré à l'étude fonctionnelle et cinétique de la  $\beta$ 3GalT6 recombinante humaine sauvage (et mutée le cas échéant) produite en système bactérien sous forme tronquée soluble. En parallèle, nous avons envisagé de réaliser la production de ces mêmes protéines en cellules d'insectes en collaboration avec le CERMAV à Grenoble. Mon travail a consisté à construire les vecteurs d'expression recombinants et à les fournir à nos collègues pour les premiers tests de production en cellules d'insectes. A l'issue de la production, l'enzyme recombinante produite sera purifiée et caractérisée d'un point de vue fonctionnel, afin d'apporter des informations sur l'impact des mutations sur la fonction de l'enzyme. La caractérisation fonctionnelle de la  $\beta$ 3GalT6 sauvage débutera par la détermination de ses paramètres cinétiques vis-à-vis d'un substrat accepteur synthétique (un disaccharide fluorescent [Gal-Xyl-naphtyle] phosphorylé). La caractérisation structurale sera initiée par des études de DLS (*Dynamic Light Scattering*) pour vérifier la stabilité et l'homogénéité de la préparation de protéine purifiée et par la réalisation d'un modèle structural de l'enzyme pour comprendre l'impact des mutations sur la structure et potentiellement sur la fonction de la protéine. Ce travail permettra d'envisager ensuite une étude des mutants de la  $\beta$ 3GalT6 en nous focalisant sur les mutations ponctuelles décrites chez les patients afin de fournir des informations sur l'impact des modifications de l'enzyme sur sa structure et sa fonction.

(2) Le second axe de l'étude vise au développement d'un modèle cellulaire déficient en  $\beta$ 3GalT6 par invalidation du gène par la technique de CRISPR/Cas9. Ce modèle nous permettra d'étudier les conséquences de l'extinction du gène *B3GALT6* dans des cellules HeLa (lignée humaine de cancer du col de l'utérus, exprimant la  $\beta$ 3GalT6 et produisant des GAGs) en culture. Après validation du modèle développé, les cellules déficientes pourront servir à évaluer le métabolisme cellulaire, en particulier les conséquences sur la production de GAGs. Par la suite, une surexpression de la protéine sauvage et des mutants de la  $\beta$ 3GalT6 sera réalisée pour déterminer si les protéines mutées présentent ou non une activité enzymatique, et si elles conservent (ou non) la capacité à initier la biosynthèse des GAGs à partir d'un substrat xylosidique exogène ou sur la protéine *core* d'un PG.



Ce projet permettra de fournir des informations sur les bases moléculaires de la déficience génétique en  $\beta$ 3GalT6 ainsi que de comprendre ses répercussions cellulaires. Une meilleure connaissance du fonctionnement de cette enzyme dans le contexte des SED pourrait permettre une amélioration de la prise en charge des patients si la sévérité des symptômes et l'évolution de la pathologie peuvent être mieux prédites à partir d'une mutation donnée sur l'enzyme. A plus long terme, ces travaux pourraient permettre de proposer des stratégies thérapeutiques visant à s'opposer aux déficits observés dans les cellules déficientes et potentiellement à restaurer les fonctions biologiques associées.



**MATERIELS ET  
METHODES**

**Tableau 10 – Fournisseurs et sources biologiques des anticorps utilisés.**

<b>Anticorps</b>	<b>Source biologique</b>	<b>Dilution</b>	<b>Fournisseur</b>	<b>Référence</b>
Anti-MBP	Rat	1/10 000	Sigma-Aldrich, France	(Rodríguez-Herva et al., 2012)
Anti-HA	Lapin	1/10 000	Sigma-Aldrich, France	(Skehel and Wiley, 2000)
Anti-hDécorine	Souris	1/5 000	R&D System, USA	(Malfait et al., 2013)
Anti-Souris conjugué à la HRP	Cheval	1/5 000	Cell Signaling Technology USA	(Ma et al., 2020)
Anti-Rat conjugué à la HRP	Chèvre	1/5 000	Cell Signaling Technology USA	(Wang et al., 2020)
Anti-Lapin conjugué à la HRP	Chèvre	1/5 000	Cell Signaling Technology USA	(Shao et al., 2020)



## I. Matériels :

### I.1. Les réactifs chimiques et biochimiques :

L'UDP-Gal, l'ampiciline, la kanamycine et le milieu Luria Bertani (LB) ont été fournis par Sigma Aldrich (St Quentin Fallavier, France). Le disaccharide Gal- $\beta$ 1,4-Xyl(2-*O*-Phosphate)-*O*-MéthoxyNaphthyle (Gal-Xyl(2P)-OMN) a été synthétisé et gracieusement fourni par le Pr Chrystel Lopin-Bon (Institut de chimie organique et analytique, ICOA-UMR7311, Université d'Orléans, France).

Les colonnes de purification MBPTrap<sup>TM</sup>HHP proviennent de GE Healthcare (Velizy-Villacoublay, France). L'albumine sérique bovine (*serum albumin bovine*, SAB) est fournie par Euromedex (Souffelweyersheim, France). L'anticorps polyclonal couplé à la HRP (*Horseradish Peroxidase*) a été fourni par Sigma-Aldrich (Saint-Quentin, Fallavier, France). Le sulfate de sodium radiomarqué (Na<sub>2</sub><sup>35</sup>SO<sub>4</sub>) a été fourni par Perkin Elmer Life Sciences (Courtaboeuf, France). La résine Sephadex G50 fine provient d'Amersham Biosciences (GE Healthcare, Velizy-Villacoublay, France). Les anticorps utilisés pour les analyses par western blot et pour les analyses par immunofluorescence sont listés dans le Tableau 10.

### I.2. Les lignées cellulaires et le milieu utilisé en culture cellulaire :

Les cellules HeLa (cellules épithéliales humaines de carcinome de col utérin) sont obtenues auprès de la banque de cellules *European Collection of Cell Culture* (ECCC, Sigma-Aldrich, Saint-Quentin, Fallavier, France).

Le milieu de culture DMEM (*Dulbecco's Modified Eagle Medium*) provient de chez Sigma-Aldrich (Saint-Quentin, Fallavier, France). La glutamine et le mélange antibiotique pénicilline-streptomycine proviennent de Gibco (ThermoFischer Scientific, Waltham, MA, USA). Le sérum de vœu fœtal (SVF) est fourni par Dominique Dutscher (Brumath, France).

### I.3. Les réactifs de biologie moléculaire :

Le vecteur d'extinction eucaryote pX458 utilisé pour générer des cellules déficientes en *B3GALT6* par la technologie CRISPR/Cas9 a été fourni par TACGENE (Sorbonne Université, Paris, France). Le vecteur d'expression eucaryote pcDNA3.1a (+) et les vecteurs d'expression bactérien pETM41, pET41a et pETM30 ont été fournis par Invitrogen (Thermo Fisher Scientific, Waltham, MA, USA) et la plasmide pFastBac-SS nous a été fourni par le CERMAV (Grenoble). L'agent de transfection PEI (PolyEthylenimine) provient de Polyscience (Washington, PA, USA).

**Tableau 11 – Séquences nucléotidiques des amorces utilisées pour l'étude de la  $\beta$ 3GalT6 recombinante humaine (h $\beta$ 3GalT6).**

Noms de la séquence	Séquences nucléotidiques	
<b>Réalisation des troncations de la h<math>\beta</math>3GalT6 en baculovirus</b>		
B3GALT6 $\Delta$ Ter30	Sens	5'-CGCCATGGCAGCTGCGCGGCCGAG-3'
	Antisens	5'-TGCTCGAGTCAGGGGATGCCCTCCCTTCTC-3'
B3GALT6 $\Delta$ Ter55	Sens	5'-CGCCATGGCAGCCGCCTTCCTGGCAGTGC-3'
	Antisens	5'-TGCTCGAGTCAGGGGATGCCCTCCCTTCTC-3'
<b>Réalisation des troncations de la h<math>\beta</math>3GalT6 en bactérie</b>		
B3GALT6 $\Delta$ Ter30	Sens	5'-CGCCATGGCAGCTGCGCGGCCGAGCCC-3'
	Antisens	5'-TGCTCGAGTCAGGGGATGCCCTCCCTTCTCTGGCA-3'
B3GALT6 $\Delta$ Ter34	Sens	5'-CGCCATGGCAGAGCCCGGGACCCCAGGGC-3'
	Antisens	5'-TGCTCGAGTCAGGGGATGCCCTCCCTTCTCTGGCA-3'
B3GALT6 $\Delta$ Ter43	Sens	5'-CGCCATGGCAGGCCGCAGCCCGCCTCCCC-3'
	Antisens	5'-TGCTCGAGTCAGGGGATGCCCTCCCTTCTCTGGCA-3'
B3GALT6 $\Delta$ Ter50	Sens	5'-CGCCATGGCAGCGCCCGCGCGCGCC-3'
	Antisens	5'-TGCTCGAGTCAGGGGATGCCCTCCCTTCTCTGGCA-3'
B3GALT6 $\Delta$ Ter55	Sens	5'-CGCCATGGCAGCCGCCTTCCTGGCAGTGTGGTG-3'
	Antisens	5'-TGCTCGAGTCAGGGGATGCCCTCCCTTCTCTGGCA-3'
<b>Mutagenèse dirigée</b>		
Asp207His (1)	Sens	5'-GGCAACTCTGCC <u>ACT</u> ACTACCT-3'
	Antisens	5'-AGGTAGTAGTGGCAGAGTTGCC-3'
Gly217Ser (2)	Sens	5'-CACGTAGGAG <u>CC</u> GCCC-3'
	Antisens	5'-GGGCGG <u>CT</u> CCTACGTGC-3'
Val61Leu (3)	Sens	5'-CTTCCTGGC <u>ACT</u> GCTGGTGG-3'
	Antisens	5'-CCACCAGCAGTGCCAGGAAG-3'
Arg261His (4)	Sens	5'-CACGACCCGC <u>ACT</u> TCGACA-3'
	Antisens	5'-TGTCGAAGTGC <u>GG</u> TCTGTG-3'
Tyr310Cys (5)	Sens	5'-GCCTGTCC <u>TGC</u> GTGTACGA-3'
	Antisens	5'-TCGTACACGCAGGACAGGC-3'
Glu265Asp (6)	Sens	5'-TCGACACCG <u>ACT</u> ACCGGTCC-3'
	Antisens	5'-GGACCGGTAG <u>TC</u> GGTGTCTGA-3'
Tyr182Cys (7)	Sens	5'-CCGCCTCT <u>GCT</u> GGGGC-3'
	Antisens	5'-GCCCCAGCAGAGGCGG-3'
<b>Séquençage de l'ADN génomique des cellules HeLa</b>		
PCR 1	Sens	5'-ATGAAGCTGCTGCGGCGG-3'
	Antisens	5'-CCGCCTCTACTGGGGCTTCTT-3'
PCR 2	Sens	5'-CTAGGCCTGGGCACGCTGG-3'
	Antisens	5'-CCTCACGGCCAAGGTGCT-3'
<b>Analyse de l'expression de l'ARNm dans les cellules HeLa</b>		
B3GALT6	Sens	5'-TCACGTACTTTAACACATCCTTGAA-3'
	Antisens	5'-CGGAGCTCGAACTCAAC-3'
RPS 29	Sens	5'-TTCAAACCGGCACGGTCTGA-3'
	Antisens	5'-TGCCGTA <u>CTG</u> ACGGAAACAC-3'
18S	Sens	5'-GTAACCCGTTGAACCCATT-3'
	Antisens	5'-CCATCCAATCGGTAGTAGCG-3'
HPRT	Sens	5'-GAGGAGTCCCTGTTGATGTTGCCAG-3'
	Antisens	5'-GGCTGGCCTATAGGCTCATAGTGC-3'

Le kit d'extraction d'ARN (*RNeasy Mini Kit*) provient de Qiagen (Courtaboeuf, France). La réaction de transcription inverse (*reverse transcription*, RT) est réalisée avec le kit SuperScript VILO (Invitrogen). L'agent intercalant utilisé pour les réactions de PCRq (*Quantitative Polymerase Chain Reaction*), le SYBR Green, a été fourni par Bio-Rad (Marnes-la-Coquette, France).

La séquence des amorces utilisées pour l'amplification du gène codant pour la version recombinante humaine de la  $\beta$ GalT6 délétée de ces 30, 34, 43, 50 ou 55 premiers acides aminés sont indiquées dans le tableau 11. Les amorces utilisées pour la réalisation de la mutagenèse ainsi que la séquence des amorces de la PCRq sont présentées dans ce tableau.

## II. Méthodes Expérimentales :

### II.1. Caractérisation fonctionnelle de la $\beta$ GalT6 recombinante humaine :

#### II.1.1. Clonage des séquences codant les différentes troncations de la $\beta$ GalT6 recombinante humaine dans les vecteurs d'expression pETM41, pET41a et pETM30 :

Les ADNc codant les différentes troncations de la  $\beta$ GalT6 humaine (NM\_080605.3), délétée de ses 30, 34, 43, 50 ou 55 premiers acides aminés N-terminaux, ont été amplifiés par PCR à partir de l'ADNc de la  $\beta$ GalT6 humaine en utilisant des amorces (i) sens contenant le site de restriction NcoI et (ii) anti-sens contenant le site de restriction XhoI après le codon stop (Tableau 11).

La réaction de PCR a été effectuée en utilisant la polymérase *Phusion* (ThermoFischer Scientific) durant 40 cycles de PCR en suivant les étapes présentées dans le tableau 12.

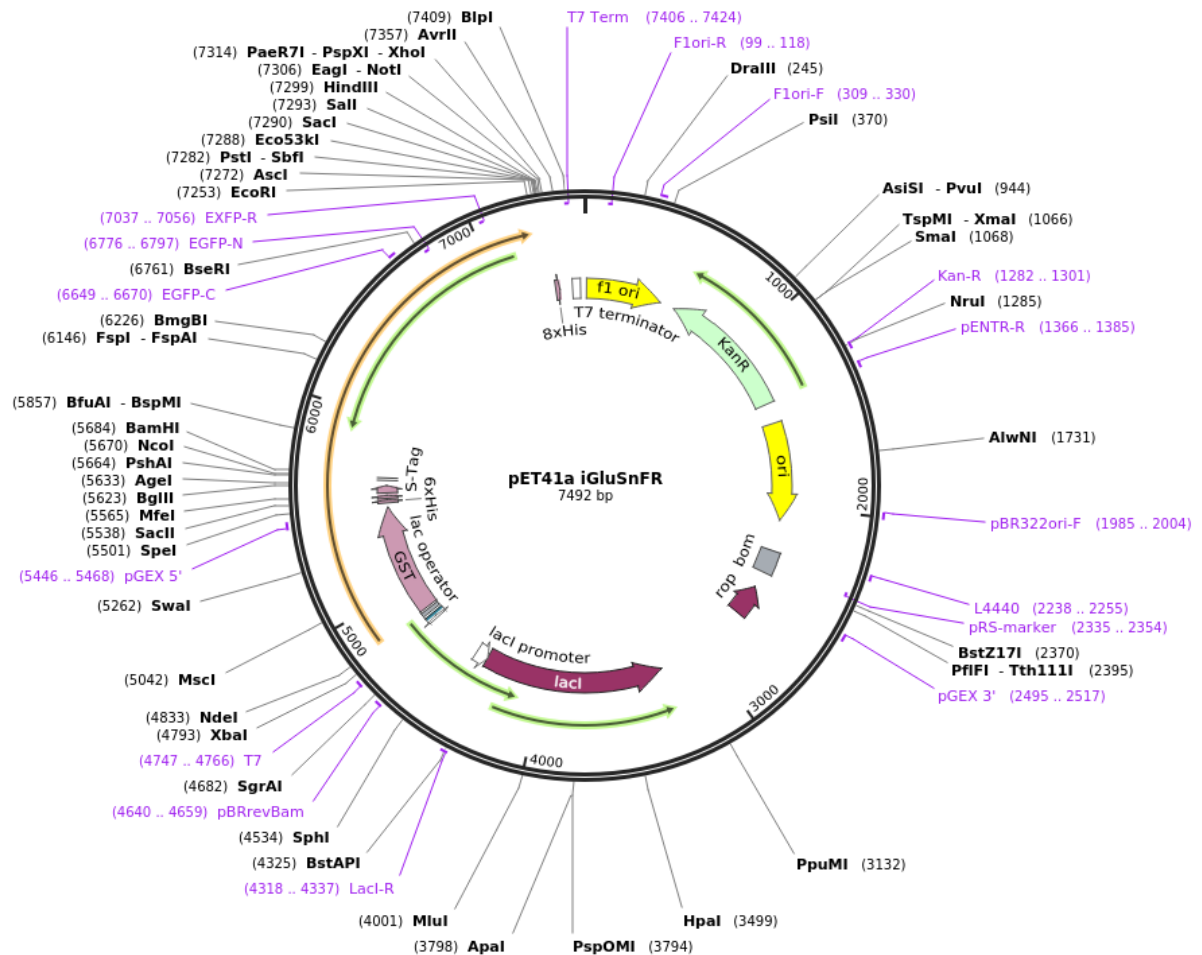
Les fragments de PCR amplifiés et les plasmides pET41a, pETM30 et pETM41 ont été digérés individuellement par les enzymes de restriction NcoI et XhoI (Figure 15, 16 et 17). La digestion a été réalisée en utilisant 500 ng d'ADN (fragment de PCR ou plasmide) dans le tampon de digestion *Green FastDigest* (ThermoFischer Scientific). Le clonage dans les vecteurs d'expression a été effectué avec le kit *DNA Ligation Kit Mighty Mix* (Takara, Paris, France). Le mélange de ligation a été incubé à 16°C durant la nuit et le produit de ces ligations a été utilisé pour transformer des bactéries compétentes Top10 (Invitrogen) pour la production et la purification des plasmides.

Pour réaliser la transformation, 20  $\mu$ L de bactéries compétentes sont mises en présence de 0.2 ng de plasmide dans la glace pendant 30 min. Un choc thermique est réalisé en plaçant le mélange bactéries-plasmide à 42°C pendant 30 secondes puis immédiatement refroidi dans la

**Tableau 12 – Durées et températures des différentes étapes du protocole de PCR utilisées pour amplifier la séquence de la h $\beta$ 3GalT6.**

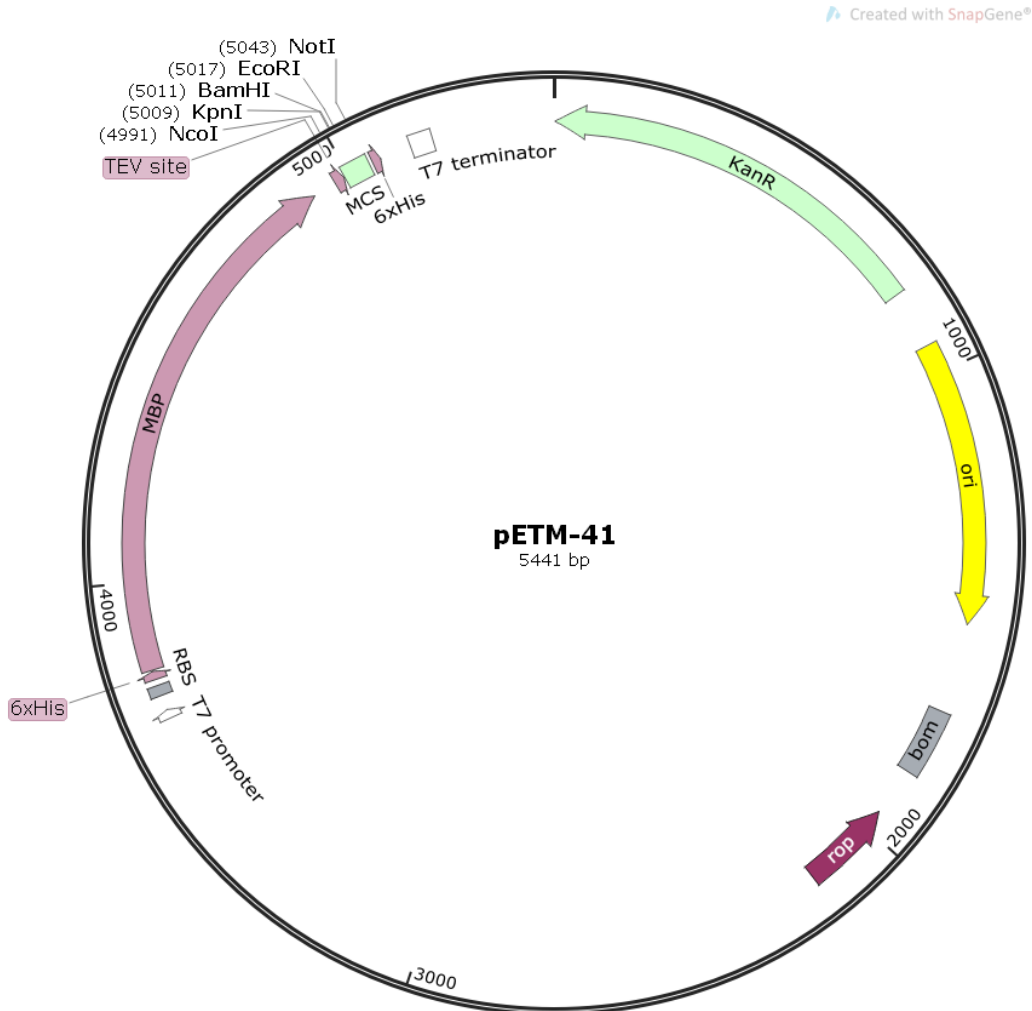
Étapes	Température	Temps
Dénaturation	98°C	3 min
Amplification	98°C	10 secondes
	72°C	30 secondes
Fin d'amplification	72°C	5 min

} 35 cycles



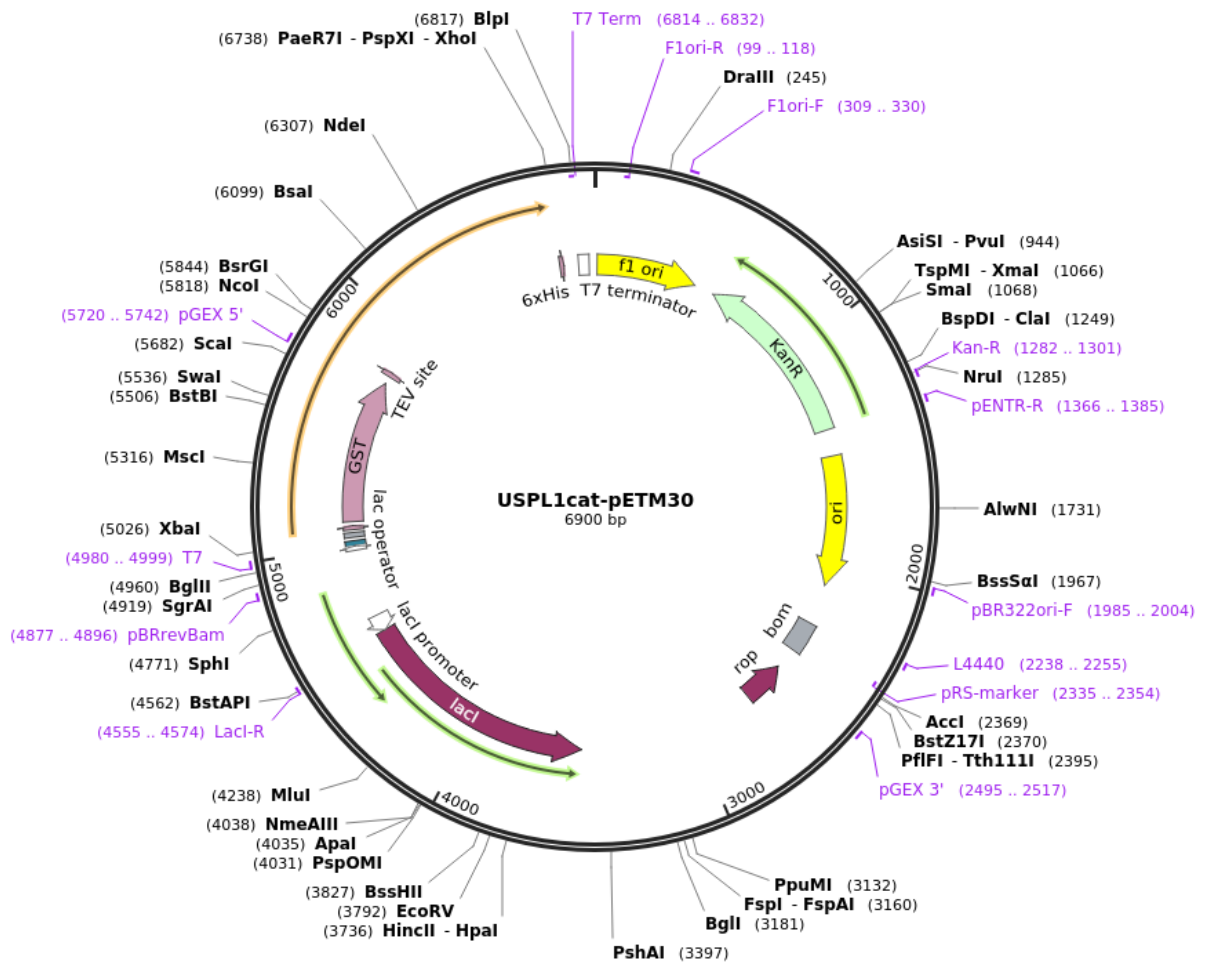
**Figure 15 – Carte du plasmide pET41a.**

Le vecteur pET41a contient un gène de résistance à la kanamycine (KanR) ainsi que la séquence codant l'étiquette poly-Histidine (His-Tag) et la protéine GST. Le partenaire de fusion va nous permettre de purifier la protéine d'intérêt par chromatographie d'affinité au moyen d'une résine à forte affinité pour la GST. Entre la protéine d'intérêt et la GST se trouve un site de clivage reconnu par la TEV protéase (qui coupe après le motif Glu-X-X-Tyr-Gln-Gly/Ser).



**Figure 16 – Carte du plasmide pETM41.**

Le vecteur pETM41 contient un gène de résistance à la kanamycine (KanR) ainsi que la séquence codant l'étiquette poly-Histidine (His-Tag) et la protéine (*Maltose Binding Protein*, MBP). Le partenaire de fusion va nous permettre de purifier la protéine d'intérêt par chromatographie d'affinité au moyen d'une résine à forte affinité pour la MBP. Entre la protéine d'intérêt et la MBP se trouve un site de clivage reconnu par la TEV protéase (qui coupe après le motif Glu-X-X-Tyr-Gln-Gly/Ser).



**Figure 17 – Carte du plasmide pETM30.**

Le vecteur pETM30 contient un gène de résistance à la kanamycine (KanR) ainsi que la séquence codant l'étiquette poly-Histidine (His-Tag) et la protéine de fusion GST. Le partenaire de fusion va nous permettre de purifier la protéine d'intérêt par chromatographie d'affinité au moyen d'une résine à forte affinité pour la GST. Entre la protéine d'intérêt et la GST se trouve un site de clivage reconnu par la TEV protéase (qui coupe après le motif Glu-X-X-Tyr-Gln-Gly/Ser).

**Tableau 13 – Liste des différents paramètres étudiés lors l’optimisation de la production de la h $\beta$ 3GalT6.**

Paramètres		Modifications
Partenaire de fusion (plasmide utilisé)		GST (pET41a ou pETM30) ou MBP (pETM41)
Condition de production	Milieu	LB ou LBNB
	Température	20°C ou 25°C
	Concentration en IPTG	0.5 mM ou 1 mM
Troncation de la protéine		$\beta$ 3GalT6 - $\Delta$ Nter30 (avec la région stem), $\Delta$ Nter34, $\Delta$ Nter43, $\Delta$ Nter50 et $\Delta$ Nter55 (sans la région stem)



glace pendant 2min. 75  $\mu$ L de milieu riche est ajouté et laissé sous agitation à 37°C pendant 1 à 2h. Les bactéries transformées sont étalées sur une boîte de Pétri contenant un milieu gélosé *Luria-Bertani* (LB) – Agar sélectif contenant 75 mg/L de kanamycine et placées à l'étuve à 37°C pendant une nuit.

Trois clones bactériens isolés ont été récupérés pour chaque construction et mis en culture dans 10 mL de milieu LB contenant 75 mg/L de kanamycine. Les cultures bactériennes ont été incubées à 37°C sous agitation pendant une nuit. Les plasmides ont été purifiés en utilisant le kit *QIAprep Spin Miniprep Kit* (Qiagen) avant dosage de la concentration par spectrophotométrie. Trois clones individuels sont envoyés pour un double séquençage de l'ADN (Eurofins, Allemagne).

### II.1.2. Production de la $\beta$ GalT6 recombinante humaine en système bactérien :

#### II.1.2.1. Etape d'optimisation de la production de la h $\beta$ GalT6 pour les différentes troncations de la protéine :

Dans le cadre de l'optimisation de la production de notre protéine, il a été nécessaire de déterminer la souche bactérienne, le vecteur et la troncation de la protéine adaptés pour produire en quantité suffisante notre protéine d'intérêt dans forme soluble. Les différentes conditions réalisées sont résumées dans le tableau 13.

Pour cette étape, les souches bactériennes compétentes : *E. coli* Rosetta2 DE3, BL21 DE3 et C41 DE3 vont être transformées en utilisant les plasmides pETM41, pETM30 ou pET41. Chaque transformation est réalisée suivant la méthode présentée précédemment (voir partie II.1.1).

Des colonies isolées sur chacune des boîtes de Pétri ont été piquées,ensemencées de façon indépendante dans 10 mL de milieu LB contenant 75 mg/L de kanamycine et incubées une nuit à 37°C et sous agitation.

Nous avons également procédé à une optimisation des conditions de production. Brièvement, deux approches ont été testées :

- (i) Une approche de production « standard » en milieu LB, où les paramètres de température et concentration d'inducteur varient : 0.6 mL de la pré-culture bactérienne sontensemencés dans 6 mL de milieu LB contenant 75 mg/L de kanamycine et incubés à 37°C sous agitation. Quand la DO (densité optique) à 600 nm atteint 0.6 unités, 0,1 ou 0,5 mM d'Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) est ajouté et incubé à 20°C ou 25°C une nuit sous agitation.



- (ii) Une approche de production en milieu LBNB (*Luria Bertani* NaCl bétaïne, contenant 500 mM NaCl et 2 mM bétaïne) grâce auquel un choc osmotique est réalisé en plus du choc thermique sur les bactéries : 0.6 mL de la pré-culture bactérienne sont ensemencés dans 6 mL de milieu LB contenant 75 mg/L de kanamycine et incubés à 37°C sous agitation. Quand la DO à 600 nm atteint 0.6 unités, un choc osmotique est réalisé sur les bactéries en culture en ajoutant au milieu LB 0,5 M de NaCl et 2 mM de bétaïne. Le milieu de production LBNB est alors incubé 1h à 47°C sous agitation (choc thermique) puis replacé à 20°C sous agitation pendant 1h. La production est induite par l'ajout de 0,1 ou 0,5 mM d'IPTG. Les cultures bactériennes sont incubées à 20°C une nuit sous agitation.

### II.1.2.2. Production de la $\beta$ GalT6 recombinante humaine : 6His-MBP-h $\beta$ GalT6 $\Delta$ Nter30 par la souche *E. coli* BL21 DE3 :

Des colonies isolées sur chacune des boîtes de Pétri ont été piquées, ensemencées de façon indépendante dans 10 mL de milieu LB contenant 75 mg/L de kanamycine et incubées toute une nuit à 37°C et sous agitation.

La production a été réalisée en milieu LB : 6mL de la pré-culture sont ensemencés dans 600 mL de milieu LB contenant 75 mg/L de kanamycine et incubés à 37°C sous agitation. Quand la DO à 600 nm atteint 0.6 unités, le milieu de culture est incubé 1h sous agitation à 4°C (choc thermique). Après cette incubation, 0.5 mM d'IPTG est ajouté et le milieu est incubé à 20°C une nuit sous agitation.

### II.1.2.3. Extraction de la protéine recombinante humaine :

Dans le cadre de l'étude de la solubilité de la protéine, lorsque les cultures sont arrêtées, les suspensions bactériennes sont centrifugées 10 min à 6.000 x g. Les culots bactériens ont été repris dans 1 mL de tampon de lyse (50 mM Sodium Phosphate, 200 mM NaCl, 0,5% (v/v) Triton X-100, 5% Glycérol). La suspension bactérienne a été lysée par sonication en effectuant 3 cycles de 30 secondes de sonication suivis par 30 secondes sur glace avec la puissance de l'appareil réglée à 80%. Par la suite, la technique de lyse est identique sauf que le culot bactérien est repris dans 5 mL de tampon de lyse et que les cycles de lyse sont de 2 minutes de lyse suivie d'une minute sur glace, ceci répété 3 fois.

Une fois la lyse effectuée, une aliquote du lysat est réservée pour une analyse par SDS-PAGE. Le reste est centrifugé à haute vitesse (12 000 x g) pendant 30 minutes afin d'éliminer les bactéries non lysées, les gros débris et les agrégats protéiques. Les protéines solubles restent



dans le surnageant, une aliquote est conservée pour être également analysée par SDS-PAGE.

### *II.1.3. Purification de la protéine de fusion 6His-MBP-h $\beta$ GalT6 $\Delta$ Nter30 :*

La purification de la protéine de fusion 6His-MBP- $\beta$ 3GalT6 $\Delta$ Nter30 a été réalisée en utilisant une colonne d'affinité MBPTrap™ HP (GE Healthcare). La colonne contient une résine de sépharose sur laquelle sont fixées des molécules de maltose. La protéine de fusion est retenue sur la colonne grâce à l'affinité de la MBP pour le maltose de la résine. La protéine fixée sera éluée grâce à l'ajout d'une solution à forte concentration de maltose dans le tampon d'élution.

Pour réaliser la purification, la colonne a été équilibrée dans le tampon de lyse qui sert de tampon de liaison pour la protéine. Une fois l'équilibration de la colonne réalisée, le surnageant bactérien est injecté dans le système de purification (BioRad®) pour que la protéine de fusion se fixe sur la colonne. Cette étape de fixation se passe sur la nuit. L'élution se fait en utilisant le tampon de lyse (voir II.1.2.3) contenant 500 mM de maltose. La protéine est directement éluée et récupérée en plusieurs fractions.

Une aliquote de chaque fraction d'élution est conservée pour être analysée par SDS-PAGE.

### *II.1.4. Analyse de l'expression et de la solubilité des protéines de fusion 6His-MBP-h $\beta$ GalT6 $\Delta$ Nter30, Nter34, Nter43, Nter50 et Nter55 :*

#### *II.1.4.1. Dosage des protéines totales :*

La méthode utilisée pour doser les protéines totales est celle de Bradford (1975). Il s'agit d'un dosage colorimétrique utilisant le bleu de Coomassie qui se fixe sur les acides aminés basiques. Une gamme étalon est réalisée à partir d'une solution de SAB en parallèle des échantillons et la DO à 595 nm est mesurée 5 minutes après l'ajout du colorant sur un spectrophotomètre UV-visible (Shimadzu UV-Spectrophotometer UV-1800). Les dosages sont réalisés en duplicat. Les concentrations protéiques totales des échantillons à doser sont déduites à partir de la gamme étalon.

#### *II.1.4.2. Analyse de l'expression de la h $\beta$ GalT6 $\Delta$ Nter30, 34, 43, 50 et 55 par SDS-PAGE :*

La technique SDS-PAGE est une technique de séparation des protéines en fonction de leur masse moléculaire apparente. Le gel utilisé est composé de : (1) un gel de concentration (10% en acrylamide/bisacrylamide) qui va permettre de concentrer les protéines et de créer un front



de migration unique et (2) d'un gel de séparation à 12% (p/v) d'acrylamide dans lequel les protéines vont être séparées selon leur masse moléculaire apparente. L'APS (*ammonium persulfate*) et le TEMED (*TEtraMethylEthyleneDiamine*) (Sigma Aldrich) sont utilisés pour permettre la réticulation du réseau d'acrylamide. Un courant électrique de 15 mA est d'abord appliqué pour l'étape de concentration des échantillons, puis un courant de 30 mA pour l'étape de séparation. La migration dure environ 1h30.

Les dépôts des différents échantillons ont été préparés en mélangeant un volume contenant l'équivalent de 10 µg de protéines totales avec le même volume de bleu de charge (25 mM Tris pH 6.8, 1% (p/v) SDS, 10% (v/v) Glycérol, 0.3 M DTT et Bleu de Bromophénol). Les échantillons sont dénaturés 2-3 minutes à 95°C avant d'être déposés.

#### **II.1.4.2.1. Analyse de l'expression des protéines après coloration du gel par le bleu de Coomassie :**

La présence des protéines dans le gel est visualisée grâce au bleu de Coomassie. Après une coloration d'une heure dans un tampon de coloration (45% (v/v) éthanol, 45% (v/v) acide acétique et bleu de Coomassie R250), le gel est lavé avec un tampon contenant 10% (v/v) d'acide acétique et 30% (v/v) d'éthanol sous agitation et jusqu'à décoloration complète du gel. Les bandes correspondant aux protéines peuvent alors être visualisées.

#### **II.1.4.2.2. Analyse de l'expression des protéines par Western Blot :**

Après migration sur gel, les protéines sont transférées sur une membrane PVDF (PolyFluorure de VinyliDène – Sigma Aldrich) en milieu liquide, dans un appareil d'électrotransfert, sous un voltage constant de 50 mV. La membrane doit être préalablement activée par l'éthanol 96% (v/v) puis équilibrée avec le tampon de transfert (25 mM Tris-Base, 192 mM Glycine et 10% (v/v) éthanol). Après le transfert, la membrane est saturée pendant 1 heure, sous agitation, avec une solution de SAB à 5% (m/v) préparée dans le tampon TBST (*Tris Buffer Saline Tween*, 150 mM Tris-HCl (pH8.5), 10 mM NaCl, 0.5% (v/v) Tween 20). La membrane est ensuite hybridée avec un anticorps primaire (de rat) anti-MBP dirigé contre l'étiquette MBP de la protéine de fusion 6His-MBP-β3GalT6ΔNterX (X=30, 34, 43, 50 ou 55) pendant la nuit. A l'issue de cette incubation, la membrane est lavée 3 fois pendant 10 minutes à l'aide du tampon TBST puis hybridée avec un anticorps secondaire anti-IgG de rat, couplé à la HRP (*horse radish peroxidase*) pendant 1 heure sous agitation. Trois lavages de 10 minutes avec du tampon TBST sont de nouveau réalisés pour éliminer l'excédent d'anticorps non fixés. La présence de la protéine est alors révélée par l'ajout du réactif *Amersham ECL*





*WesternBlot detection* (GE Healthcare) qui contient le substrat de la peroxydase ( $H_2O_2$ ) et un chromophore qui sera oxydé. L'enchaînement de réactions enzymatiques permet de déterminer l'expression de l'enzyme par bioluminescence.

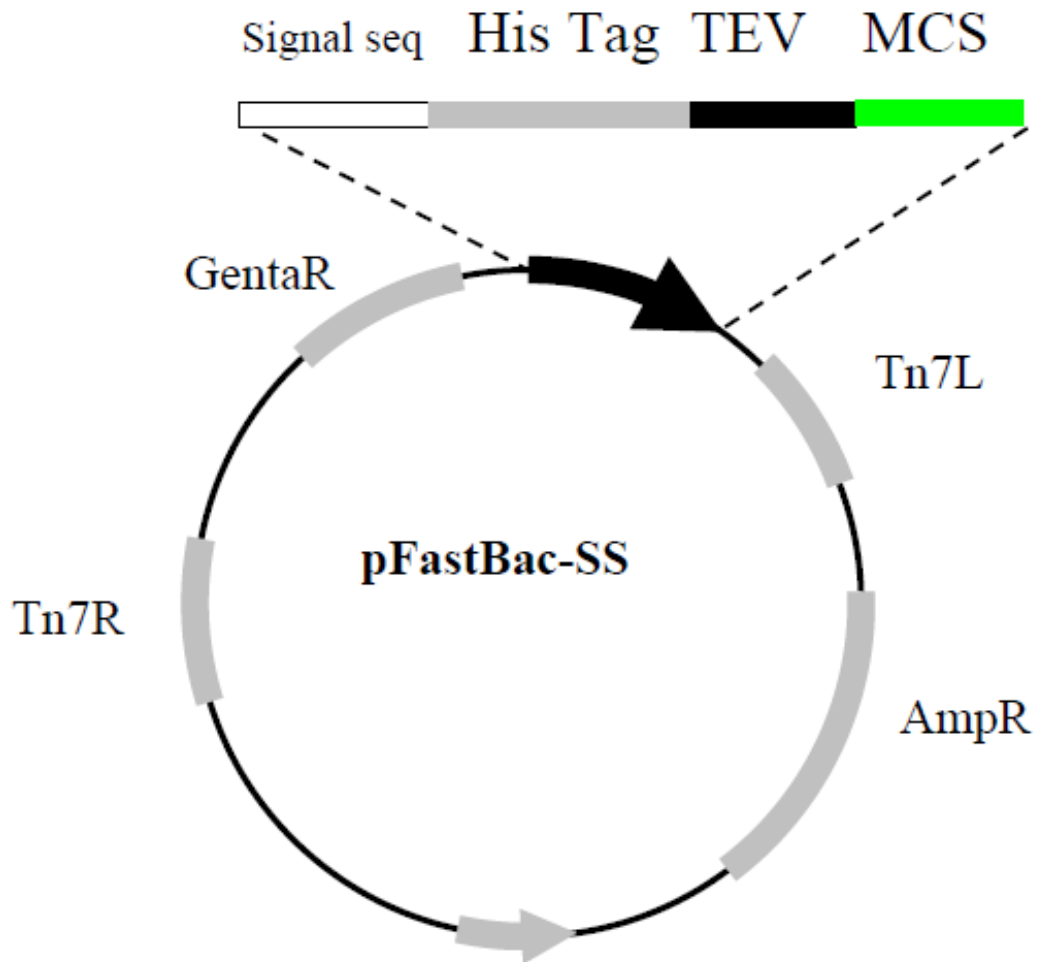
### II.1.5. Evaluation de l'activité galactosyltransférase :

La réaction enzymatique est effectuée en incubant les homogénats cellulaires (environ 50  $\mu$ g de protéines totales) ou la protéine purifiée en présence du substrat donneur (UDP-Gal, 1 mM) et du substrat accepteur de l'enzyme, le disaccharide phosphorylé, Gal-Xyl(2P)-OMN dans un tampon 50 mM Bis Tris (pH 7,2), 5 mM  $MnCl_2$  à 37°C. Pour trouver les meilleures conditions de l'essai enzymatique, plusieurs tests ont été réalisés en faisant varier différents paramètres : le temps de réaction (de 30 min à 2 h), la quantité d'enzyme (2  $\mu$ g) et la quantité de substrat accepteur (10 à 1000  $\mu$ M).

La réaction est arrêtée en plaçant les tubes dans la glace et en ajoutant 5  $\mu$ L d'HCl 2N. Les essais sont centrifugés 5 minutes à 10.000  $\times g$  à 4°C, puis les surnageants contenant les produits de la réaction sont analysés par CLHP en phase inverse sur une colonne C18 (XBridge, 4,6  $\times$  150 mm, 5  $\mu$ m, Waters, Saint-Quentin-en-Yvelines, France) en utilisant une chaîne de CLHP (Alliance Waters e2695) équipée d'un détecteur de fluorescence (Waters 2475 *Multi  $\lambda$  Fluorescence Detector*). La phase mobile permettant la séparation du produit de réaction du substrat non métabolisé est composée d'un mélange 13% (v/v) acétonitrile/ $H_2O$  et 0.02% (v/v) acide trifluoroacétique (TFA). Le produit de réaction est détecté en suivant la fluorescence émise à 356 nm par le groupement naphthyl présent sur le produit, après une excitation à la longueur d'onde de 324 nm. Le produit est quantifié au moyen d'une gamme étalon (0 à 50 nmol) analysée en parallèle des échantillons à partir de dilutions successives de produit de réaction Gal-Gal-Xyl(2P)-OMN. L'activité enzymatique est exprimée en pmoles de produit formées par unité de temps rapportées à la quantité de protéines totales présentes dans l'échantillon.

### II.2. Clonage des séquences codant les différentes tronctions de la $\beta$ GalT6 recombinante humaine dans les vecteurs de transfert, pFastBac-SS :

En parallèle de la production de la h $\beta$ 3GalT6 recombinante en système bactérien, nous avons essayé un système de production de protéines recombinantes connu pour avoir de bons



**Figure 18 – Carte du plasmide pFastBac-SS.**

Le vecteur pFastBac-SS contient un gène de résistance à l'ampiciline (AmpR) ainsi que la séquence codant l'étiquette poly-Histidine (His-Tag). Cette étiquette va nous permettre de purifier la protéine d'intérêt par chromatographie d'affinité au moyen d'une résine ayant une forte affinité pour les histidines. Entre la protéine d'intérêt et la GST se trouve un site de clivage reconnu par la TEV protéase (qui coupe après le motif Glu-X-X-Tyr-Gln-Gly/Ser).

rendements : le système baculovirus/cellules d'insectes. Nous avons donc entrepris la construction d'un vecteur de transfert pour la production de la protéine en cellules d'insectes en utilisant un système d'expression dérivé du baculovirus. Nous nous sommes concentrés sur la réalisation du vecteur de transfert pour transmettre ces constructions à nos collègues du CERMAV à Grenoble (chez lesquels, les premiers essais de production ont été réalisés).

Pour construire les vecteurs, les ADNc codant les différentes troncations de la  $\beta$ 3GalT6 humaine (NM\_080605.3), délétée de ses 30 ou 55 premiers acides aminés N-terminaux, ont été amplifiés par PCR à partir de l'ADNc de la  $\beta$ 3GalT6 humaine en utilisant des amorces (i) sens contenant le site de restriction NcoI et (ii) anti-sens contenant le site de restriction XhoI après le codon stop (Tableau 11).

La réaction de PCR a été effectuée en utilisant la polymérase *Phusion* (ThermoFischer Scientific) durant 40 cycles de PCR en suivant les mêmes étapes que celles présentées dans les étapes de clonage des séquences pour la production de la protéine en bactéries (voir partie II.1.1.).

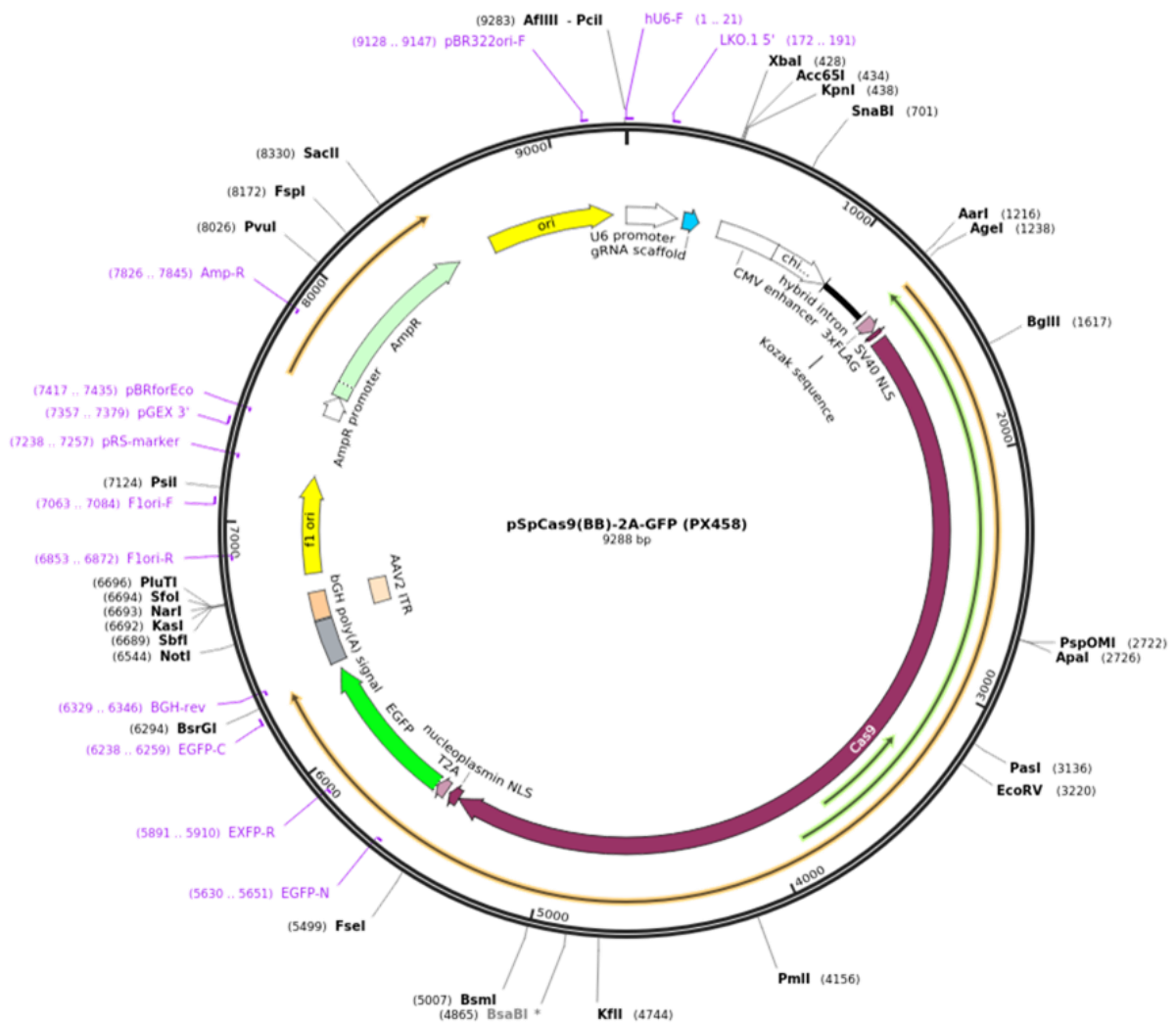
Les fragments de PCR amplifiés et le plasmide pFastBac-SS (Figure 18) ont été digérés individuellement par les enzymes de restriction NcoI et XhoI. La digestion a été réalisée en utilisant 500 ng d'ADN (fragment de PCR ou plasmide) dans le tampon de digestion *Green FastDigest* (ThermoFischer Scientific). Le clonage dans les vecteurs d'expression a été effectué avec le kit *DNA Ligation Kit Mighty Mix* (Takara, Paris, France). Le mélange de ligation a été incubé à 16°C durant la nuit et le produit de ces ligations a été utilisé pour transformer des bactéries compétentes Top10 (Invitrogen) pour la production et la purification des plasmides (suivant le protocole de la partie I.1.1). Les plasmides obtenus ont été envoyés pour un double séquençage de l'ADN (Eurofins, Allemagne).

Les plasmides de transfert ont été envoyés au CERMAV pour la réalisation d'essais de production de nos protéines tronquées en cellules d'insectes.

### **II.3. Génération d'un modèle cellulaire déficient en $\beta$ 3GalT6 par la technique CRISPR/cas9 :**

#### **II.3.1. Entretien et congélation des cellules HeLa :**

Les cellules HeLa et les cellules déficientes en  $\beta$ 3GalT6 (KO) sont cultivées dans un milieu DMEM (4.5 g/L D-Glucose) auquel sont ajoutés de la Glutamine 1% (v/v), un mélange d'antibiotiques pénicilline/streptomycine 1% (v/v) et 10% (v/v) de SVF préalablement décomplémenté. Les cellules sont ensuite placées dans un incubateur à 37°C sous atmosphère



**Figure 19 – Carte du plasmide pX458.**

Le plasmide pX458 contient le gène de la Cas9 et la séquence codant l'ARN guide dirigé contre le gène *B3GALT6*. Ce plasmide présente également la séquence codante de l'EGFP (*Enhanced Green Fluorescent Protein*) qui va permettre la sélection et le tri des cellules recombinantes.

humide contenant du CO<sub>2</sub> 5% (v/v) dans des flasques 75 cm<sup>2</sup> (Corning Incorporated, NY, Costar®). Le milieu est renouvelé tous les 2 jours.

Des stocks de cellules sont réalisés à différents passages de culture dans un milieu contenant du diméthylsulfoxyde (DMSO) 10% (v/v) et du SVF 90% (v/v). Les cryotubes utilisés pour le stockage sont d'abord placés en chambre de congélation à -80°C puis rangés à plus long terme dans un congélateur à -140°C.

### *II.3.2. Transfection des cellules HeLa par le plasmide pX458\_sghB3GALT6-2 :*

Les cellules HeLa sont transfectées par le plasmide pX458\_sghB3GALT6-2 (Figure 19) (TACGENE, CNRS UMR 7196/INSERM U1154 – Sorbonne Universités) en utilisant le polyéthylèneimine (PEI) comme agent de transfection.

A partir d'une flasque de 75 cm<sup>2</sup> contenant des cellules HeLa à environ 80% de confluence, une étape de trypsination suivie d'une centrifugation à 110 *x g* pendant 5 minutes permet de récupérer le culot cellulaire qui est remis en suspension dans 1 mL de milieu DMEM complet. Pour le comptage des cellules, 10 µL de la suspension cellulaire et 10 µL de Bleu Trypan sont mélangés et placés dans une cellule de comptage manuel (cellule de Malassez). Une fois le comptage effectué, les cellules sontensemencées en plaques 6 puits (Corning Incorporated, NY, Costar®) à une densité d'environ 300 000 cellules vivantes par puits et placées durant toute la nuit dans l'incubateur à 37°C avant la réalisation de la transfection le lendemain.

Le mélange de transfection contenant 1 µg de plasmide pX458\_sghB3GALT6-2 et 3 µg de l'agent de transfection (PEI) est incubé 15 minutes à température ambiante avant d'être déposé directement dans le milieu de culture des cellules. Les cellules sont ensuite laissées en culture à 37°C pendant 48 heures en présence de 3 mL de milieu de culture complet par puits. Une manipulation contrôle sans plasmide, mais en présence du PEI est réalisée en parallèle afin de s'assurer de la non-toxicité de cet agent sur les cellules dans nos conditions expérimentales.

### *II.3.3. Sélection des cellules recombinantes et isolement des clones cellulaires déficients*

Les cellules recombinantes sont triées en utilisant le trieur de cellules (Astrios, Beckman Coulter) de la plateforme de cytométrie de l'UMS 2008 IBSLor (Campus Brabois Santé, Vandoeuvre-lès-Nancy). Pour le tri, les cellules recombinantes expriment la GFP dont la



séquence est présente dans le vecteur pX458 et sont donc potentiellement des cellules *B3GALT6* déficientes.

Après l'étape de tri cellulaire, les cellules ont été diluées et isolées par dilution limite, technique qui consiste à effectuer des dilutions en cascade pour obtenir statistiquement une cellule par puits dans les plaques de culture. 12 clones ont pu être isolés et amplifiés selon cette technique. Ces clones ont été caractérisés au niveau (i) génomique (extraction d'ADN génomique et séquençage), (ii) transcriptomique (mesure du niveau d'expression de l'ARNm) et (iii) protéique (mesure de l'activité enzymatique) afin de valider l'absence d'expression de la  $\beta$ GalT6 et l'inactivation de son gène dans ces cellules.

#### II.3.4. Caractérisation génomique des cellules déficiente en $h\beta$ GalT6 :

L'ADN génomique a été extrait en suivant les instructions du kit *DNeasy Blood and Tissue Kit* following manufacturer's instructions (Qiagen, Hilden, Allemagne). La concentration en ADN génomique est mesurée en utilisant le *NanoDrop™ One Microvolume UV-Vis Spectrophotometer* (Thermo Fischer Scientific, Waltham, MA, USA). Les PCR sont réalisées avec la *KOD Xtreme™ Hot Start DNA Polymerase* (Merk Millipore, Burlington, MA, USA) et avec les amorces présentées dans le tableau 11. Les conditions de la première PCR sont les suivantes : 96°C pendant 2 minutes, 20 cycles de 98°C pendant 10 secondes, 56°C pendant 30 secondes, 68°C pendant 30 secondes et une étape d'élongation finale de 68°C pendant 7 minutes. Les conditions de la seconde PCR sont les suivantes : 96°C pendant 2 minutes, 32 cycles de 98°C pendant 10 secondes, 50°C pendant 30 secondes, 68°C pendant 30 secondes et une étape d'élongation finale de 68°C pendant 7 minutes. Les produits de PCR sont ensuite clonés dans le plasmide pCR™-Blunt II-TOPO (Thermo Fisher Scientific, Waltham, MA, USA). Cinq clones individuels ont été envoyés pour un double séquençage de l'ADN (Eurofins, Allemagne).

#### II.3.5. Analyse de l'expression des ARNm de la $h\beta$ GalT6 par RT-PCR quantitative :

Le niveau d'expression des ARNm de la  $h\beta$ GalT6 dans les cellules HeLa et les clones précédemment sélectionnés (potentiellement KO) est quantifié par RT-PCR quantitative.

##### II.3.5.1. Extraction des ARN totaux :

Les ARN totaux sont extraits en suivant les instructions du kit *RNeasy® Mini Kit* (Qiagen).

**Tableau 14 – Durées et températures des différentes étapes du protocole de RT-qPCR utilisées pour la quantification des ARNm.**

Étapes	Température	Temps
1	95°C	15 minutes
2	95°C	15 secondes
	60°C	25 secondes
	72°C	20 secondes
		} 45 cycles
3	95°C	15 secondes
	60°C	1 minute
	95°C	15 secondes



La concentration des ARN totaux est ensuite évaluée par la mesure de l'absorbance à 260nm sur le spectrophotomètre Nanodrop® (*Spectrophotometer ND-1000*, ThermoFisher Scientific, Wilmington, USA). La qualité et la pureté des ARN extraits sont évaluées grâce au rapport  $A_{260nm}/A_{280nm}$  qui doit être compris entre 1,8 et 2.

#### II.3.5.2. La RT-PCR quantitative :

La rétro-transcription (RT) de l'ARNm en ADNc est effectuée en utilisant 1 µg d'ARN totaux à l'aide du kit *Super Script VILO® cDNA Synthesis* (Invitrogen) selon les recommandations du fournisseur. Les conditions expérimentales utilisées sont les suivantes : (1) 10 minutes à 25°C, (2) 60 minutes à 42°C, (3) 5 minutes à 85°C. La réaction de PCRq a été réalisée en plaque 96 puits sur l'appareil *Applied Biosystems® StepOnePlus™ Real-Time PCR System* dans un volume final de 20 µL, contenant 2 µL de la dilution d'ADNc (dilution au 1/10<sup>ème</sup> ou 1/100<sup>ème</sup> dans de l'eau ultra-pure sans RNase), 10 pmol d'amorces *B3GALT6* (Eurofins genomics) ou 10 pmol d'amorces correspondant aux gènes de ménage (*RPS29* (*Ribosomal Protein S29*), *18S* et *HPRT* (*Hypoxanthine-guanine phosphoribosyltransferase*)) utilisées comme contrôles et 10 µL de SYBR Green (BioRad). La réaction de PCRq comprend les étapes décrites dans le Tableau 14.

La méthode  $2^{-\Delta Ct}$  est utilisée pour déterminer les niveaux d'expression relative des ARNm de la hβ3GalT6 dans les différentes cellules HeLa (lignée initiale et clones potentiellement déficients) après normalisation par rapport aux gènes de référence *RPS 29*, *18S* et *HPRT*.

#### II.3.6. Obtention des plasmides contenant les séquences nucléotidiques codant les mutants de la hβ3GalT6 par mutagenèse dirigée :

##### II.3.6.1. La mutagenèse dirigée :

La mutagenèse dirigée permet de substituer ou de déléter un ou plusieurs résidus d'acides aminés au sein de la séquence de la protéine cible. Dans cette étude, nous reproduisons les mutations ponctuelles observées chez les patients atteints du SEDsp (Tableau 2). Nous procédons ainsi afin de tenter de comprendre au niveau moléculaire l'impact de ces mutations sur l'activité de la protéine.

Le vecteur recombinant pcDNA\_hβ3GalT6-HA contenant l'ADNc de la β3GalT6 humaine portant l'étiquette HA (*Human influenza hemagglutinin*) a été utilisé comme matrice pour les réactions de PCR en utilisant des amorces sens et anti-sens portant le codon contenant la mutation d'intérêt. La mutagenèse dirigée a été réalisée en utilisant le kit « *QuickChange II XL*



*site-directed mutagenesis kit* » (Agilent technologies, Les Ulis, France). Pour ce faire, 10 ng d'ADN plasmidique sont mélangés avec 125 ng de chaque amorce nucléotidique, 10 mM de dNTP, 2 unités de l'ADN polymérase *Pfu Turbo* dans un volume final de 50 µL. Le cycle de PCR utilisé est le suivant : (i) 95°C pendant 1 minute, (ii) 95°C pendant 50 secondes, (iii) 64°C pendant 50 secondes, (iv) 68°C pendant 10 minutes et (v) 68°C pendant 7 minutes. Les étapes ii à iv (inclusive) ont été répétées 18 fois. A la fin de la PCR, l'enzyme de restriction DpnI est ajoutée (incubation une nuit à 37°C) afin de cliver de façon spécifique l'ADN plasmidique matrice méthylé, permettant ainsi de sélectionner l'ADN muté (non méthylé). Le produit de la réaction est ensuite inséré dans des bactéries *E. coli* (Top10) compétentes après transformation par choc thermique pour y être amplifié.

### II.3.6.2. Amplification et extraction d'ADN plasmidique :

Les bactéries (*E. coli One Shot® Top 10*) sont transformées en utilisant les vecteurs pcDNA hβ3GalT6-HA par choc thermique. La transformation est réalisée suivant la méthode présentée au paragraphe II.1.1 à la différence que le milieu gélosé LB-Agar contient 50 mg/L d'ampicilline. Trois clones bactériens isolés ont été récupérés pour chaque construction et mis en culture dans 100 mL de milieu LB contenant 75 mg/L d'ampicilline. Les cultures bactériennes ont été incubées à 37°C sous agitation pour la nuit. Les plasmides ont été purifiés en utilisant le kit *QIAGEN® Plasmid Midi Kit* (Qiagen). Trois clones individuels ont été envoyés pour un double séquençage de l'ADN (Eurofins, Allemagne).

### II.3.7. Analyse de la synthèse des GAGs à partir d'un xyloside exogène, le 4-Methylumbelliferyl-β-D-xylopyranoside (4-MOX) :

L'analyse de la biosynthèse des GAGs repose sur la mesure du taux d'incorporation du soufre radiomarqué ( $\text{Na}_2\text{SO}_4$ [<sup>35</sup>S]) dans les chaînes de GAGs néo-synthétisées.

Pour ce faire, les cellules HeLa déficientes en β3GalT6 et les cellules contrôle sont cultivées pendant 48h dans le milieu Fisher's (+) L-Glutamine (Life technologies™) supplémenté par 2% (v/v) de sérum veau foetal, et contenant le sulfate de sodium radiomarqué  $\text{Na}_2\text{SO}_4$  [<sup>35</sup>S] (Perkin Elmer®) (10 µCi/mL), en présence de concentrations croissantes de 4-MOX (Sigma-Aldrich) (0, 2 et 5 µM final). Après 12h, le milieu de culture des cellules est recueilli et déposé sur une colonne de filtration sur gel Sephadex G-50 fine (GE Healthcare). Des fractions de 1,5 mL sont collectées à la sortie de la colonne et la radioactivité présente dans les différentes fractions est mesurée par comptage en scintillation liquide (dans le mélange de scintillation Ultima Gold™ Perkin Elmer).



Les cellules sont quant à elles grattées dans 100 µL de tampon PBS-1% (p/v) SDS. La concentration en ADN des cellules (proportionnelle au nombre de cellules présentes dans les puits) est évaluée par fluorimétrie à l'aide du kit *Qubit™ dsDNA HS Assay Kit* (Invitrogen™) suivant les instructions du fournisseur sur l'appareil *Qubit® 2.0 Fluorometer*.

Le taux de synthèse des GAGs matriciels des fibroblastes à partir du 4-MOX est finalement évalué en rapportant la radioactivité associée aux GAGs radiomarqués à la quantité d'ADN présente dans les différents puits.

### *II.3.8. Analyse de l'activité galactosyltransférase et de la glycosylation de la décorine dans les cellules déficientes :*

#### *II.3.8.1. Transfection des cellules déficientes avec le plasmide pcDNA décorine et/ou pcDNA β3GalT6 :*

Les cellules HeLa déficientes et contrôle sont transfectées avec le plasmide pcDNA3.1 décorine-His et/ou les plasmides pcDNA3.1 hβ3GalT6-HA (contenant la séquence codant la protéine sauvage ou les mutants 1 et 2) (tableau 11) en utilisant le PEI comme agent de transfection.

Le protocole de transfection est le même que celui utilisé pour la transfection avec le plasmide de CRISPR/Cas9 (voir partie II.2.2). Les cellules ont étéensemencées dans trois puits à une densité de 225 000 cellules par puits pour chaque transfection. Après 24h, les cellules sont transfectées en utilisant le mélange suivant : 2 µg de plasmide pour 4 µg de l'agent de transfection (PEI). Dans le cas de la double transfection, les cellules sont transfectées avec un mélange contenant 2 µg de pcDNA3.1\_Décorine, 2 µg de pcDNA3.1\_β3GalT6 et 8 µg de PEI. Les cellules sont ensuite laissées en culture à 37°C pendant 48h avec 2 mL de milieu de culture complet par puit.

#### *II.3.8.2. Purification de la décorine, préparation des homogénats cellulaires et dosages des protéines totales :*

Pour récupérer la décorine, les milieux de culture des cellules transfectées sont prélevés et centrifugés à 500 x g pendant 5 minutes pour éliminer les cellules mortes. Les surnageants sont ensuite mis en contact avec 100 µL de billes de sépharose liées à du nickel (GE Healthcare, Chicago, Illinois, USA), préalablement lavées au PBS (*phosphate-buffered saline*). Le surnageant et les billes sont mis sous agitation pendant 2h avant d'être centrifugés 3 minutes à 2000 x g. Les billes sont récupérées et de nouveau lavées au PBS avant d'être mises en contact



avec 100  $\mu$ L de bleu de charge (voir composition partie II.1.4.2). Les billes sont chauffées à 95°C pendant 10 minutes pour éluer la décorine fixée et la suspension est conservée pour être déposée sur gel SDS-PAGE.

Les cellules HeLa et les clones cellulaires sélectionnés sont dans un premier temps grattés dans du PBS pour être détachés de leur support de culture avant d'être centrifugés 5 minutes à 500  $x$  g. Les culots cellulaires sont repris dans un tampon 50 mM Bis-Tris (pH 6,5), contenant 10 mM  $MnCl_2$  et Triton X-100 0,5% (v/v). Les membranes cellulaires sont lysées grâce à 5 cycles successifs de congélation-décongélation « *freeze-thaw cycles* » (1 minute dans de l'azote liquide puis 5 minutes à 37°C).

Le dosage des protéines totales est réalisé suivant le protocole décrit dans la partie II.1.4.1.

### *II.3.8.3. Analyse de l'expression de la h $\beta$ GalT6 sauvage ou mutée et de la glycosylation de la décorine par Western Blot :*

Le protocole d'analyse de l'expression de la h $\beta$ GalT6 est identique au protocole décrit dans la partie II.1.4.2.2. La différence est l'utilisation de l'anticorps primaire anti-HA (souris). Dans le cas de la décorine, la migration sur gel est réalisée sur des gels pré-coulés « *Criterion precast gels* » 4-12% (Bio-Rad, Hercules, CA, USA) et l'anticorps utilisé est l'anticorps primaire anti-décorine (souris).

### *II.3.8.4. Évaluation de l'activité galactosyltransférase :*

La réaction enzymatique est décrite dans la partie II.1.5. Pour ces réactions, la quantité de protéines totales des homogénats cellulaires est de 50  $\mu$ g en présence de 1 mM d'UDP-Gal et de 2 mM de substrat accepteur dans le tampon Bis-Tris (pH 7,2) contenant 5 mM  $MnCl_2$ .





# RESULTATS ET DISCUSSION



La  $\beta$ 3GalT6 est une GT impliquée dans l'initiation de la biosynthèse des GAGs qui catalyse l'addition du troisième résidu (galactose) de l'amorce tétrasaccharidique fixée au niveau de la protéine *core* des PGs. Des précédents travaux ont montré que des patients atteints du sous-type SEDsp, présentaient des mutations de l'enzyme avec un tableau clinique variable. Au niveau cellulaire, une diminution globale du taux de GAGs synthétisés (CS et HS) a été observée dans les fibroblastes de derme des patients ainsi qu'une réduction de la capacité de migration des cellules déficientes par rapport aux cellules contrôle (Malfait et al., 2013; Van Damme et al., 2018). Sur la base de ces résultats, l'objectif de ce travail est d'étudier l'impact des mutations de la  $\beta$ 3GalT6 sur la fonction et/ou la structure de l'enzyme. Ce travail est d'une grande importance car, à l'heure actuelle, peu d'études ont été consacrées à la caractérisation structurale et fonctionnelle de cette enzyme contrairement à la  $\beta$ 4GalT7, la GT qui catalyse l'étape précédente dans la voie de biosynthèse des GAGs et dont la structure cristallographique a été résolue (Rahuel-Clermont et al., 2010; Saliba et al., 2015; Tsutsui et al., 2013).

C'est dans ce contexte que j'ai consacré la première partie de ce travail de thèse à la production de la h $\beta$ 3GalT6 recombinante sauvage (et mutée le cas échéant) dans un système bactérien puis à sa purification et sa caractérisation moléculaire. Les paramètres cinétiques de la protéine recombinante humaine sauvage ont été étudiés et pourront être ultérieurement comparés à ceux des mutants. La deuxième partie de ce travail a permis la mise en place et la caractérisation d'un modèle cellulaire déficient en  $\beta$ 3GalT6 pour évaluer les conséquences fonctionnelles de la déficience génétique sur la cellule. Les cellules déficientes en  $\beta$ 3GalT6 ont par ailleurs été utilisées pour ré-exprimer individuellement l'enzyme sauvage et les mutants de la  $\beta$ 3GalT6 décrits chez les patients pour mieux comprendre l'impact de la mutation sur la fonction de l'enzyme et la synthèse des GAGs au niveau cellulaire.

## **I. Caractérisation fonctionnelle de la $\beta$ 3GalT6 humaine recombinante :**

### **I.1. Travaux préliminaires (rappels des résultats de Master 2) :**

Les manipulations réalisées au cours de mon stage de Master 2 avaient mis en évidence que la forme tronquée soluble de la  $\beta$ 3GalT6 recombinante humaine (h $\beta$ 3GalT6), délétée de ses 53 premiers acides aminés, fusionnée à la GST et contenant une étiquette composée de six résidus histidine (6His-GST-h $\beta$ 3GalT6 $\Delta$ Nter53), produite en système bactérien, n'était pas soluble dans les conditions de culture utilisées (production en milieu LB à 25°C sur une nuit après une



induction avec 0,5 mM d'IPTG – source : *Master 2 BSIS BMBI-PT RNA Enzyme Science, Benjamin JOLIVET 2016*).

Suite à ces résultats, il avait alors été décidé de :

(1) Tester d'autres conditions de production en travaillant d'abord en conditions « contraintes » de culture en utilisant le milieu LBNB qui, par chocs thermique et osmotique, induit la production de protéines chaperonnes (comme par exemple les protéines HSP (*Heat Shock Protein*)) (Oganesyan et al., 2007), aidant au repliement correct des protéines produites. Nous estimions en effet que cette approche favoriserait la production de notre protéine sous forme soluble. Malgré ces conditions, notre protéine d'intérêt n'était pas présente dans la fraction soluble de nos productions. Nous avons également tenté de modifier la phase d'induction en utilisant 0,1 mM d'IPTG au lieu de 0,5 mM et/ou en appliquant une température de 12°C au lieu de 25°C pour la phase de production, pour réduire la quantité de protéines produites dans nos cultures, ce qui réduirait l'agrégation, au profit de la solubilité ;

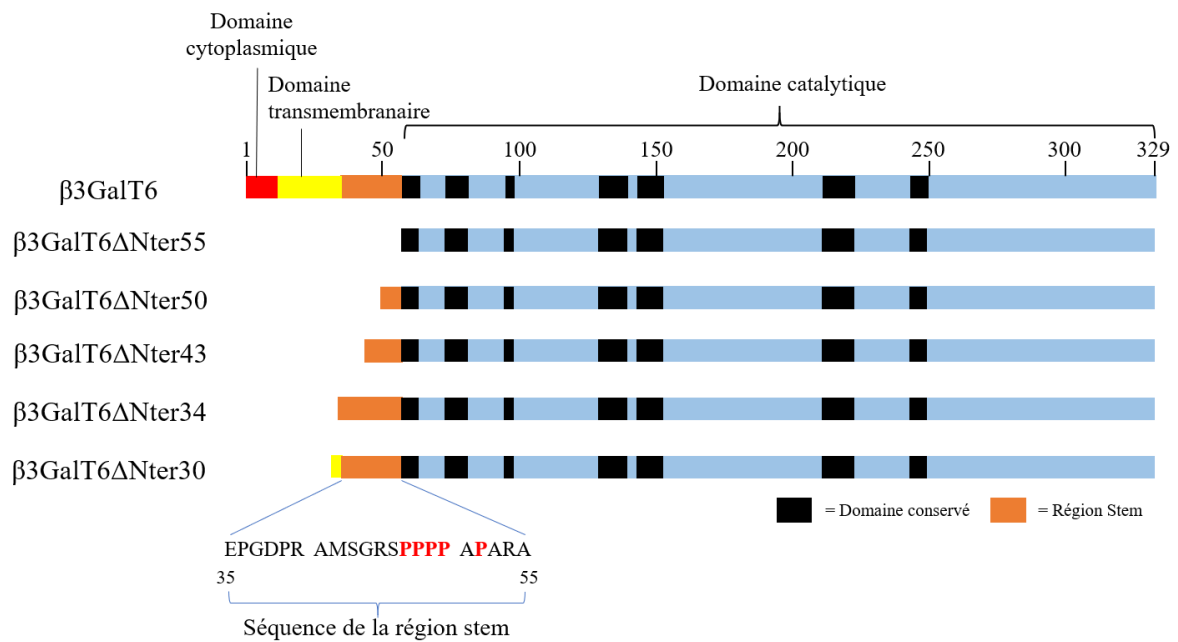
(2) Remplacer la méthode de lyse des bactéries, le système « *French Press* » que nous utilisons habituellement par la sonication pour essayer d'extraire un maximum de protéines ;

(3) Tester différents tampons de lyse en jouant sur la nature du tampon (par exemple, les tampons phosphate de sodium, acétate de sodium, MES (Acide 2-(N-morpholino)éthanosulfonique) et Tris-HCl), le pH de la solution de lyse (pH 5 à 9), la présence ou non de détergent ou d'agent dénaturant plus ou moins doux (Urée à 1-6 M, Triton X-100, NP40 ou Tween-20).

L'analyse de l'expression et de la solubilité de la protéine a été visualisée par SDS-PAGE après coloration des protéines au bleu de Coomassie et/ou par Western Blot en utilisant un anticorps primaire anti-GST dirigé contre le partenaire de fusion de l'enzyme d'intérêt (résultats non montrés). Malheureusement, aucune des conditions testées n'a permis d'obtenir une protéine soluble.

## **I.2. Optimisation des conditions de production pour l'étude de la solubilité de la $\beta$ 3GalT6 recombinante humaine :**

Aussi, dans cette partie de mon travail de thèse, nous avons cherché à vérifier l'influence du domaine N-terminal de la protéine sur sa solubilité. En effet, ce domaine, et notamment la région « *stem* », est connue comme facteur influençant la solubilité et l'activité des GTs. Une étude réalisée sur une GT végétale a montré qu'une importante troncation de la région « *stem* » améliore la solubilité mais diminue l'activité de l'enzyme. Inversement, une courte troncation de cette région affecte peu l'activité mais réduit la solubilité de l'enzyme



**Figure 20 – Représentation schématique des différentes troncations de la hβ3GalT6 utilisées pour optimiser la solubilité et la production de la protéine.**

Une forme tronquée soluble a été produite en enlevant les domaines cytoplasmique (en rouge) et transmembranaire (en jaune). Afin de déterminer l'impact de la région stem (en vert) sur la solubilité de la protéine, différentes troncations ont été réalisées au niveau de cette région afin de déterminer la longueur optimale de troncation pour l'obtention d'une protéine soluble.

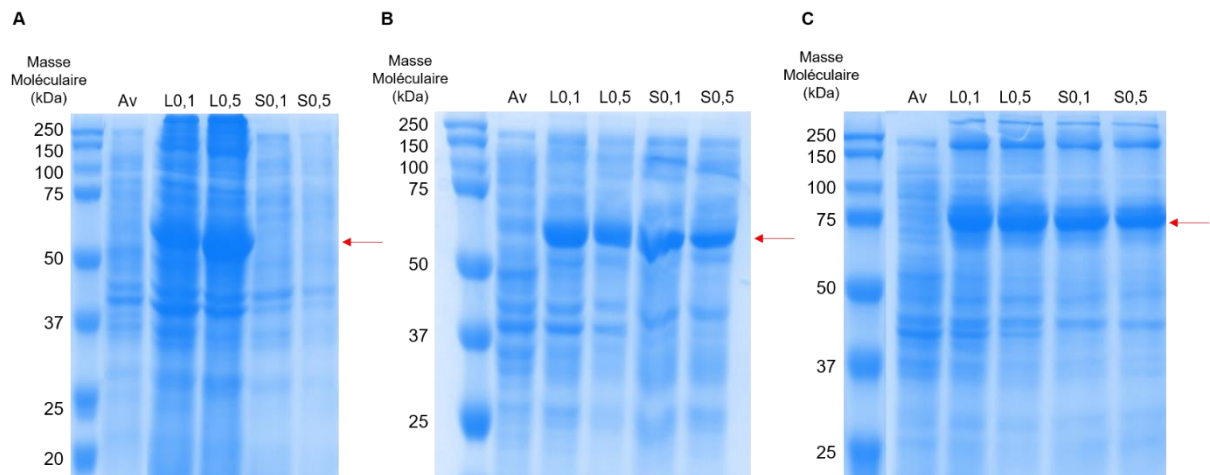
(Culbertson et al., 2016). Pour ce faire, un test de solubilité à moyenne échelle a été mené. Nous avons tenté de produire différentes formes tronquées (solubles) de la protéine (délétion des 30 à 55 premiers acides aminés) (Figure 20) dans différentes conditions résumées dans le tableau 13 (souches bactériennes, milieux de culture, températures de production, concentration en IPTG, partenaires de fusion). Une aliquote du milieu a été systématiquement prélevée avant et après induction et a servi respectivement de contrôle négatif et positif de production.

Suite à la production de la protéine, nous avons vérifié sa solubilité grâce à une étape de centrifugation à 12 000 x *g* qui permet d'éliminer les gros agrégats, les débris cellulaires ainsi que les protéines insolubles (agrégats de protéines ou corps d'inclusion). Le niveau d'expression de la protéine a été visualisé dans un premier temps par SDS-PAGE avec coloration au Bleu de Coomassie.

Nos premiers essais n'ayant pas montré de solubilité de la  $\beta$ 3GalT6 avec le partenaire de fusion GST, nous avons décidé de tester un autre partenaire, la MBP. Ces deux protéines sont couramment utilisées dans les étapes de production de protéines recombinantes en système bactérien parce qu'elles facilitent les étapes de purification par chromatographie d'affinité. Contrairement aux étiquettes peptidiques (par exemple, l'étiquette polyhistidine 6His), ces deux protéines jouent également un rôle dans la solubilisation de la protéine à laquelle elles sont liées. Enfin, la MBP et la GST sont deux protéines pouvant jouer un rôle de protéine « chaperonne » pour assister au bon repliement des protéines produites et ainsi prévenir de l'agrégation de ces dernières comme cela a été précédemment décrit pour d'autres enzymes (Esposito and Chatterjee, 2006; Nallamsetty and Waugh, 2006; Nygren et al., 1994). L'utilisation de ces deux partenaires de fusion nous a conduits à l'utilisation des vecteurs pETM41 et pETM30 contenant chacun le gène codant la MBP et la GST, respectivement. Nous avons également décidé d'utiliser un troisième vecteur, le plasmide pET41a, qui inclut la séquence codant la GST, qui est dans ce cas utilisée comme partenaire de fusion, mais dont le site de coupure protéique entre le partenaire de fusion et la protéine d'intérêt est différent.

### *1.2.1. Etude de la solubilité de la $\beta$ 3GalT6 recombinante humaine en fonction du partenaire de fusion et de la souche bactérienne utilisée pour la production:*

Le premier paramètre étudié est l'impact du partenaire de fusion (GST ou MBP) sur la production de la h $\beta$ 3GalT6 $\Delta$ Nter30. La production de la protéine est réalisée à 25°C pendant 18h après une induction à 0,1 ou 0,5 mM d'IPTG. Les analyses par SDS-PAGE après



**Figure 21 – Analyse de la solubilité de la hβ3GalT6ΔNter30 en fonction du partenaire de fusion.**

Analyse par SDS-PAGE après coloration au bleu de Coomassie. La production a été réalisée à 25°C pendant 18h après une induction de 0,1 ou 0,5 mM d'IPTG en utilisant la souche bactérienne BL21 (DE3). Analyse de la solubilité de la 6His-GST-β3GalT6ΔNter30 (vecteur pETM30) (A), GST-6His-β3GalT6ΔNter30 (vecteur pET41a) (B) et de la 6His-MBP-β3GalT6ΔNter30 (vecteur pETM41) (C). Environ 10 µg de protéines totales ont été déposés dans chaque puit.

Légende : Av : Avant induction ; L0,1 : Lysat avec induction à 0,1 mM d'IPTG ; L0,5 : Lysat avec induction à 0,5 mM d'IPTG ; S0,1 : Surnageant avec induction à 0,1 mM d'IPTG ; S0,5 : Surnageant avec induction à 0,5 mM d'IPTG. La flèche rouge indique la position de la bande correspondant à la protéine de fusion.



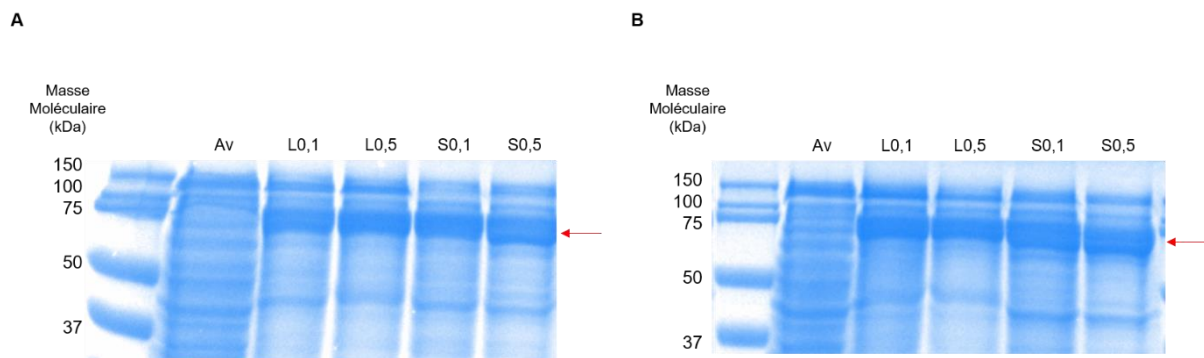
coloration au bleu de Coomassie ont montré la présence d'une bande importante aux environs de 65 kDa correspondant aux protéines de fusion 6His-GST-h $\beta$ 3GalT6 $\Delta$ Nter30 (pETM30 – Figure 21A) et GST-6His-h $\beta$ 3GalT6 $\Delta$ Nter30 (pET41a – Figure 21B) et d'une bande vers 70 kDa correspondant à la protéine de fusion 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 (pETM41 – Figure 21C) dans le lysat cellulaire, attestant de l'expression de la protéine.

Dans le cas de la production au moyen du plasmide pETM30 (utilisé au début de ma thèse), la protéine produite est insoluble (Figure 21A), au contraire des deux autres systèmes (utilisant les vecteurs d'expression pETM41 et pET41a), pour lesquels une grande quantité de protéines est présente dans la fraction soluble.

En parallèle de l'étude de l'impact du partenaire de fusion, nous avons vérifié la production de la protéine en utilisant ces trois vecteurs dans différentes souches bactériennes : les souches *E. coli* BL21(DE3), Rosetta2(DE3) et C41(DE3). Les souches BL21 (DE3) sont communément utilisées pour la production de protéines recombinantes et utilisent la T7 RNA polymérase sous contrôle de l'opérateur lacUV5 (promoteur activé par l'IPTG). Elles présentent également la délétion de deux gènes codant des protéases bactériennes pouvant interférer avec la production de la protéine recombinante et préviennent d'une éventuelle dégradation de la protéine produite. La souche Rosetta2 (DE3) est une souche dérivée des bactéries BL21 spécifiquement utilisée pour la production de protéines eucaryotes. C'est une souche qui exprime des ARNt pouvant reconnaître des codons plus rares sur l'ARNm au moment de la traduction de la protéine. Enfin les souches C41 (DE3), elles aussi dérivées des souches BL21 (DE3), sont plus particulièrement utilisées pour la production de protéines toxiques pour la bactérie. Elles présentent la délétion d'un gène induisant la mort cellulaire en cas de toxicité des protéines produites.

Les analyses par SDS-PAGE suivies d'une coloration au bleu de Coomassie ont montré les mêmes résultats que précédemment. Quelle que soit la souche bactérienne utilisée, les protéines issues des constructions pET41a et pETM41 ont toutes été retrouvées dans la fraction soluble, contrairement à la production utilisant le vecteur pETM30. Par ailleurs, nous n'avons pas remarqué de différence d'expression de nos protéines en fonction de la souche utilisée (résultats non montrés). Nous avons ainsi montré que la souche bactérienne ne semble pas impacter la solubilité de la protéine et qu'aucune ne permet réellement d'augmenter la solubilité de la protéine produite.

Nous retenons donc les vecteurs d'expression pETM41 (pour la production d'une protéine de fusion couplée à la MBP) et pET41a (pour la production d'une protéine de fusion couplée à la GST) que nous utilisons dans les étapes suivantes de l'optimisation de la production de



**Figure 22 – Vérification de la solubilité de la 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 en fonction de la température de production.**

Analyse par SDS-PAGE après coloration au bleu de Coomassie. La production a été réalisée pendant 18h après une induction avec 0,1 ou 0,5 mM d'IPTG en utilisant la souche bactérienne BL21(DE3). Analyse de la solubilité de la 6His-MBP- $\beta$ 3GalT6 $\Delta$ Nter30 produite à 20°C (A) ou à 25°C (B). Environ 10  $\mu$ g de protéines totales ont été déposés dans chaque puit.

Légende : Av : Avant induction ; L0,1 : Lysat avec induction à 0,1 mM d'IPTG ; L0,5 : Lysat avec induction à 0,5 mM d'IPTG ; S0,1 : Surnageant avec induction à 0,1 mM d'IPTG ; S0,5 : Surnageant avec induction à 0,5 mM d'IPTG. La flèche rouge désigne la bande correspondant à la protéine de fusion.

l'enzyme recombinante. Des résultats similaires ayant été obtenus avec les deux vecteurs d'expression, seuls les résultats utilisant le vecteur pETM41 seront présentés dans la suite des résultats.

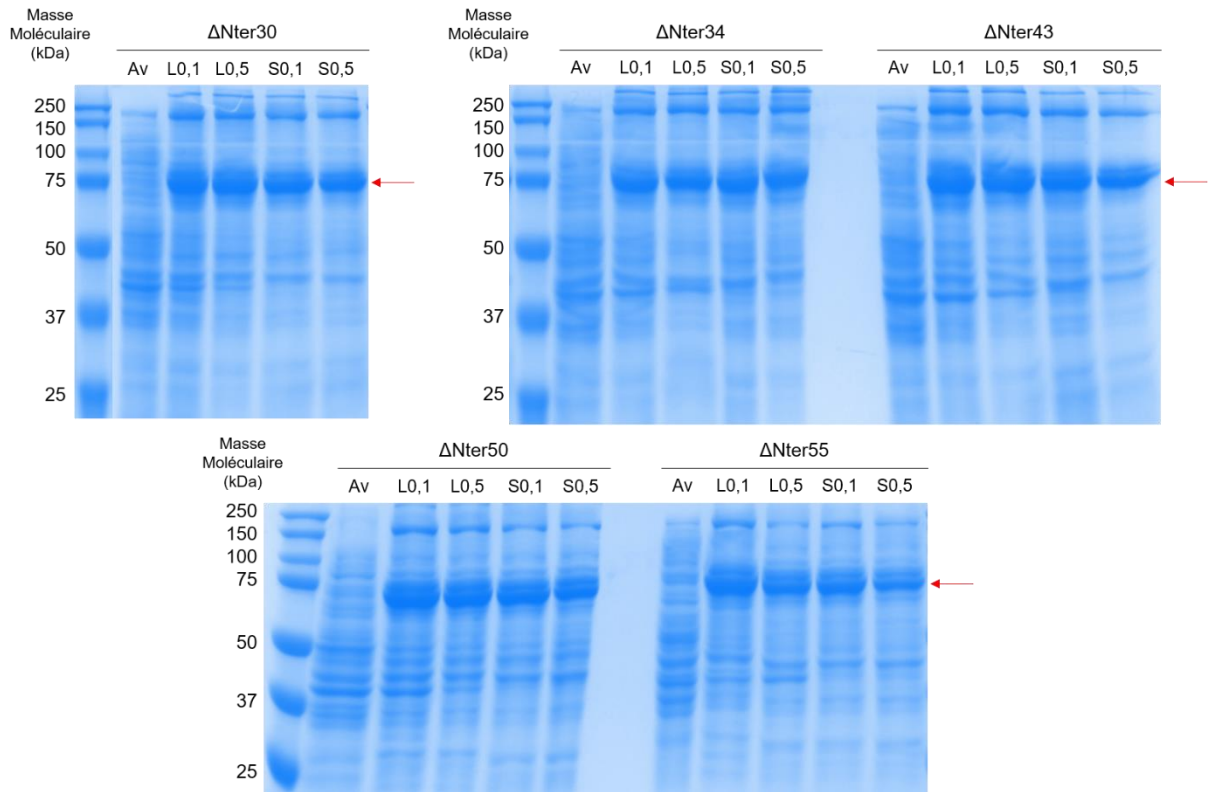
### *1.2.2. Etude de la solubilité de la $\beta$ 3GalT6 recombinante humaine en fonction des conditions de culture :*

Une fois le vecteur d'expression et la souche bactérienne de production choisis, nous avons déterminé l'impact des conditions de culture (température, nature du milieu et concentration en inducteur) sur la production de la  $\beta$ 3GalT6 $\Delta$ Nter30. La figure 22 montre l'impact de la température de production sur la solubilité de la protéine. Les analyses par SDS-PAGE suivies d'une coloration au bleu de Coomassie ont montré la présence de deux bandes vers 70 kDa correspondant à la masse moléculaire apparente de la protéine de fusion 6His-MBP- $\beta$ 3GalT6 $\Delta$ Nter30 dans le lysat bactérien, indiquant que la protéine d'intérêt a bien été produite quelle que soit la température de culture utilisée. Une bande dont la taille correspond à celle de notre protéine de fusion (piste S0,1 ou S0,5 sur la figure 22A et 22B) est également observée dans la fraction soluble. Ainsi, nos deux conditions permettent de récupérer de façon correcte notre protéine de fusion à 20°C et 25°C sans différence apparente en termes de production. L'impact de la concentration en inducteur (0,1 mM d'IPTG : pistes L0,1 et S0,1 ; 0,5 mM d'IPTG : L0,5 et S0,5 – Figure 22) a également été étudié. Nous avons remarqué que la concentration en inducteur n'a d'effet ni sur la quantité de protéines produites, ni sur la solubilité de la protéine dans la mesure où la protéine de fusion est produite suite à une induction avec 0,1 et 0,5 mM d'IPTG. Ainsi, la solubilité de la protéine ne semble pas être modifiée par la température de culture, ni par la concentration en inducteur. Des résultats similaires ont été obtenus avec le vecteur d'expression pET41a (utilisé pour la production de la protéine de fusion GST-6His-h $\beta$ 3GalT6 $\Delta$ Nter30).

Pour la suite, nous avons conservé une température de production de 20°C et la concentration de 0,1 mM d'IPTG pour produire notre protéine d'intérêt.

### *1.2.3. Etude de la solubilité de la $\beta$ 3GalT6 recombinante humaine en fonction des troncations de la protéine :*

En s'appuyant sur les conditions déterminées précédemment, nous avons ensuite vérifié si le niveau de troncation de la h $\beta$ 3GalT6 aidait à la production de protéines solubles. Nous avons ainsi produit l'enzyme dépourvue de ses 30, 34, 43, 50 ou 55 premiers acides aminés N-



**Figure 23 – Analyse de la solubilité des différentes formes tronquées de la 6His-MBP-hβ3GalT6.**

Analyse par SDS-PAGE après coloration au bleu de Coomassie. La production a été réalisée à 20°C pendant 18h après une induction de 0,1 ou 0,5 mM d'IPTG en utilisant la souche bactérienne BL21(DE3). 10 µg de protéines totales ont été déposés dans chaque puit.

Légende : Av : Avant induction ; L0,1 : Lysat avec induction à 0,1 mM d'IPTG ; L0,5 : Lysat avec induction à 0,5 mM d'IPTG ; S0,1 : Surnageant avec induction à 0,1 mM d'IPTG ; S0,5 : Surnageant avec induction à 0,5 mM d'IPTG. La flèche rouge désigne la bande correspondant à la protéine de fusion.

terminaux. Le niveau de production ainsi que la présence de protéines dans la fraction soluble ont été suivis par SDS-PAGE. Les résultats présentés dans la figure 23 correspondent à la production des formes tronquées de la h $\beta$ 3GalT6 dans le vecteur pETM41 (partenaire de fusion : 6His-MBP). Une bande correspondant à une protéine d'environ 70 kDa a été observée pour toutes les formes dans les lysats bactériens (Figure 23).

Toutes les formes tronquées sont produites dans nos conditions et nous retrouvons nos protéines d'intérêt dans la fraction soluble (surnageant) quelle que soit la troncation avec des niveaux de production apparemment similaires. Le niveau de troncation de la protéine ne semble donc pas affecter sa solubilité dans nos conditions de production.

En résumé et après optimisation, nous disposons désormais de deux systèmes d'expression permettant d'obtenir la protéine tronquée soluble.

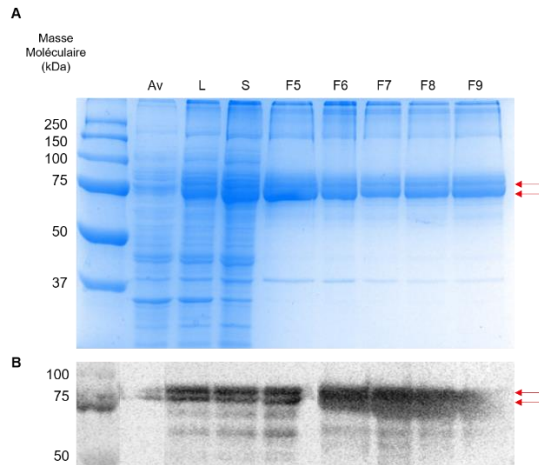
Ces différents tests ont également montré que produire la protéine sous la forme 6His-GST- $\beta$ 3GalT6 $\Delta$ Nter53 en utilisant le vecteur pETM30, n'était pas envisageable car la protéine n'était pas produite sous forme soluble. De plus, la forme soluble de la GST-6His-h $\beta$ 3GalT6 $\Delta$ Nter30 produite en utilisant le vecteur pET41a présentait une hydrolyse partielle de la protéine de fusion lors de la production (résultats non montrés).

En conclusion de cette partie, et pour la suite de ce travail, nous produirons et purifierons notre enzyme d'intérêt sous la forme 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 en employant la souche bactérienne Rosetta2 (DE3), par culture en milieu LB, à 20°C, sur la nuit et avec une induction à 0,1 mM d'IPTG.

### ***1.3. Purification de la 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 :***

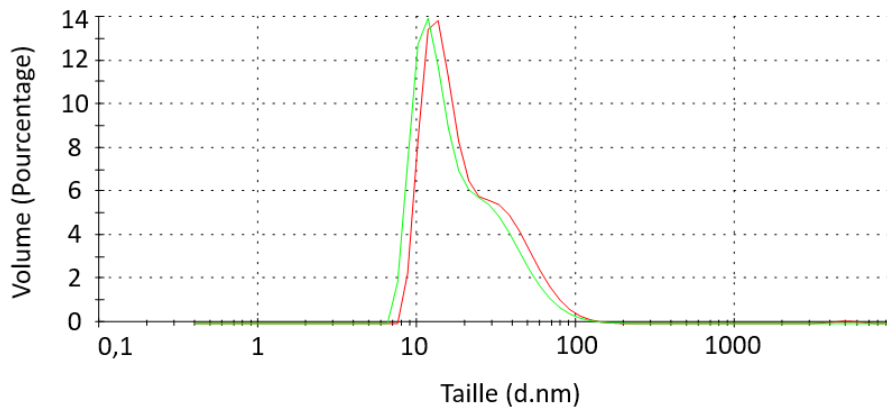
La purification de la protéine de fusion 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 a été réalisée par chromatographie d'affinité en utilisant une colonne MBPTrap™ HP qui présente une affinité pour le partenaire MBP associé à notre protéine recombinante. Pour purifier notre protéine, nous avons utilisé un tampon composé de 50 mM phosphate de sodium pH 7.2 à 4°C, 200 mM NaCl, 1 mM EDTA, 0,5% Triton X-100 et 5% glycérol, dans lequel nous avons ajouté 500 mM de maltose, le ligand naturel de la MBP pour l'élution. Plusieurs fractions d'élution consécutives (pour une même purification) ont été collectées et analysées par SDS-PAGE pour analyser le niveau de pureté de la protéine.

L'analyse par SDS-PAGE suivie d'une coloration au bleu de Coomassie montre que la protéine de fusion (masse attendue 75 kDa) est présente dans les fractions de purification notées F5 à F9 (Figure 24A). Ce résultat a été confirmé par Western blot (Figure 24B).



**Figure 24 – Analyse de la purification de la 6His-MBP- $\beta$ 3GalT6 $\Delta$ Nter30 par chromatographie d'affinité.**

Analyse par SDS-PAGE après coloration au bleu de Coomassie (A) et par Western blot en utilisant un anticorps primaire anti-MBP (B). 10  $\mu$ g de protéines totales ont été déposés dans chaque puit. Av : Avant induction ; L : Lysat bactérien ; S : Surnageant bactérien ; FX : Fraction de purification n°X



**Figure 25 – Analyse par DLS de la distribution de taille des macromolécules dans la fraction d'élution de la 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30.**

Deux fractions d'élution (courbes verte et rouge) provenant de purifications différentes ont été analysées par DLS.

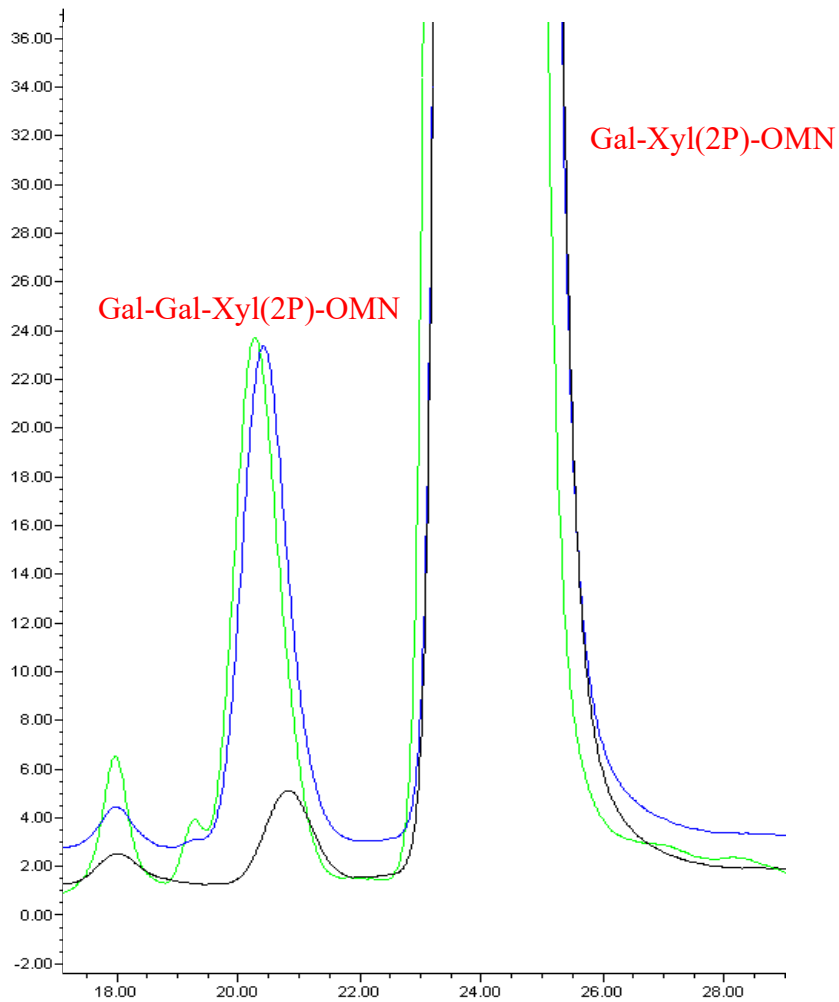
La présence d'une bande correspondant à une protéine d'une masse moléculaire légèrement supérieure à celle de notre protéine de fusion a été également observée. Afin de déterminer la nature de cette protéine, une fraction de purification a été analysée par spectrométrie de masse au sein de la Plateforme de Protéomique (UMS IBSLor – Dr. Jean-Baptiste Vincourt). Les analyses ont montré que les deux bandes correspondaient bien à la protéine de fusion 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 mais n'ont pas permis de déterminer si cette différence de masse était liée à une différence de séquences en acides aminés ou à une(des) modification(s) post-traductionnelle(s) telles que des glycosylations ou des phosphorylations (résultat non montré). L'existence de telles modifications est possible, celles-ci ayant été récemment décrites dans la bactérie (Nothaft and Szymanski, 2010). La différence de masse entre ces deux bandes, assez faible, pourrait également être due à la présence d'un élément tronqué entre ces deux structures. Il peut être suspecté que la séquence poly-His N-terminale soit tronquée, expliquant ainsi l'apparition de deux bandes protéiques de masse moléculaire légèrement différente.

Une bande de faible intensité correspondant à une protéine d'environ 40 kDa, qui correspondrait à la masse molaire de la MBP est toutefois observée dans les fractions de purification. La présence de cette bande montre une faible dégradation de la protéine de fusion.

Des bandes correspondant à des protéines de haut poids moléculaire (entre 150 et 250 kDa) ont également été observées sur la partie supérieure du gel et pourraient correspondre à des formes agrégées de la protéine produite ou à des contaminants protéiques qui ont été élués avec la protéine d'intérêt lors de la purification.

Nous avons également vérifié dans deux fractions de purification, si notre protéine était présente sous forme soluble ou présentait des agrégats protéiques. Une analyse par DLS a été réalisée pour déterminer la taille des molécules dans la solution protéique et évaluer l'homogénéité de la fraction purifiée (Figure 25). La présence d'un pic majoritaire correspondant à une population de protéines d'environ 10 nm de diamètre, pouvant correspondre à notre protéine d'intérêt, a été mise en évidence. Cependant, un épaulement correspondant à des éléments protéiques de taille plus importante, vraisemblablement des agrégats, a également été observé et pourrait expliquer la présence de bandes de masse moléculaire élevée mise en évidence par SDS-PAGE.

Nous avons donc trouvé une première condition de purification de la 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30, mais qui engendre une agrégation partielle de la protéine. Dans ces conditions, nous avons pu récupérer jusqu'à 0,7 mg de protéines dans chaque fraction de purification (soit un total de 2,1 mg de protéines purifiées par purification à partir de 600 mL



**Figure 26 – Chromatogramme HPLC de l’activité de la 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30.**

Profil chromatographique de deux essais enzymatiques de la 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 (courbe bleue et verte) et d’un témoin d’activité (absence de substrat donneur – courbe noire) analysés par CLHP



de production). Une optimisation des conditions de purification et/ou de production devra être menée notamment si des essais de cristallisation sont envisagés par la suite.

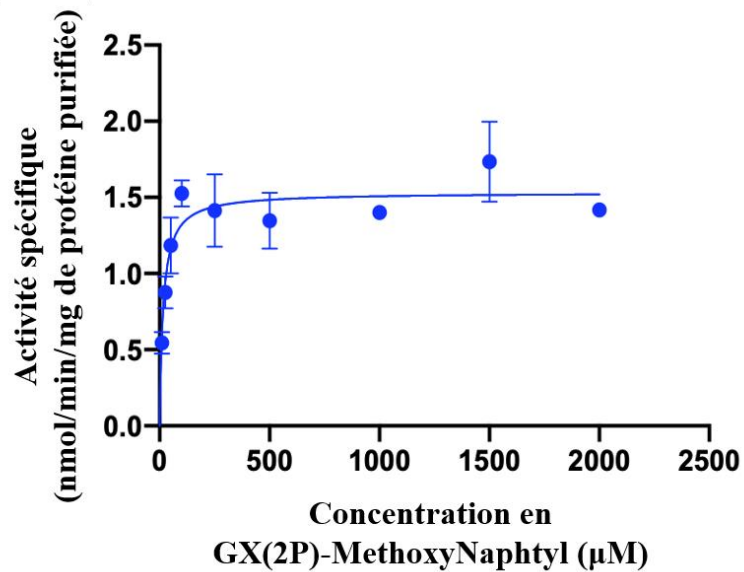
#### ***1.4. Analyse de l'activité enzymatique de la 6His-MBP-β3GalT6 ΔNter30 recombinante humaine :***

L'objectif de cette étape était de vérifier si l'enzyme produite est active et dans l'affirmative, de déterminer les paramètres cinétiques de l'enzyme ( $V_{max}$ ,  $K_m$  et  $k_{cat}$ ) vis-à-vis de ses substrats, en commençant par le substrat accepteur. Une étape d'optimisation a d'abord été réalisée pour mettre au point le test enzymatique en déterminant la concentration d'enzyme à utiliser, la quantité de substrat et le temps de la réaction.

##### ***1.4.1. Mise au point du test enzymatique de la 6His-MBP-β3GalT6ΔNter30 :***

La mesure de l'activité galactosyltransférase de la protéine recombinante humaine a été réalisée vis-à-vis d'un analogue de substrat accepteur, un disaccharidique phosphorylé en position 2 au niveau du xylose couplé à une aglycone hydrophobe de type méthoxy-naphthyle (Gal-Xyl(2P)-OMN) synthétisé par nos collègues chimistes organiciens à l'Institut de Chimie Organique et Analytique (UMR 7311 ICOA) d'Orléans. Comme décrit précédemment, la présence du groupement phosphate sur le xylose améliore la prise en charge du substrat accepteur par l'enzyme (Wen et al., 2014). Nous avons évalué la capacité de l'enzyme à transférer un résidu galactose, à partir de l'UDP-Gal (substrat donneur), sur cet accepteur dans des homogénats de cellules HeLa, exprimant l'enzyme sauvage de façon endogène, test qui nous a servi de contrôle positif. L'activité enzymatique a été évaluée par CLHP en phase inverse couplée à un système de détection par fluorescence comme décrit dans la partie matériels et méthodes. Une fois ce test d'activité mis en place, l'activité enzymatique a été évaluée sur l'enzyme purifiée (Figure 26).

Nous avons mis en place différents tests afin de déterminer les conditions optimales pour observer l'activité enzymatique de la protéine. Nous avons pu déterminer grâce à ces essais (résultats non montrés) que les meilleures conditions pour nos essais étaient obtenues en présence de 2 µg de protéine purifiée dans un tampon Bis-Tris à pH 7,2 avec un temps de réaction d'une heure. Dans ces conditions en utilisant 1 mM de substrat accepteur et 1 mM de substrat donneur, il a été possible d'obtenir une activité enzymatique de 2,7 nmoles de produit formé/min/mg de protéine purifiée.



**Figure 27 – Analyse de l’activité enzymatique de la 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 purifiée en fonction de la quantité de substrat accepteur.**

Les essais enzymatiques ont été réalisés en utilisant des quantités croissantes de substrat accepteur (10  $\mu$ M à 2 mM), 1 mM de substrat donneur (UDP-Gal) avec 2  $\mu$ g de protéine purifiée dans 50 mM de tampon Bis-Tris, pH 7,2. Après un temps de réaction de 60 min, l’activité enzymatique est évaluée par CLHP en phase inverse. Les essais enzymatiques ont été réalisés en triplicat sur deux lots différents d’enzyme purifiée.

#### **I.4.2. Analyse de l'activité galactosyltransférase en fonction de la quantité de substrat accepteur :**

Après avoir optimisé les essais enzymatiques, des tests cinétiques ont été réalisés en faisant varier la concentration de substrat accepteur de 10  $\mu\text{M}$  à 1 mM (la concentration de substrat donneur étant fixée à 1 mM) afin de déterminer les constantes cinétiques de l'enzyme vis-à-vis de ce substrat (Figure 27). L'activité enzymatique a augmenté avec la concentration en substrat accepteur jusqu'à atteindre un plateau à une concentration de substrat de 250  $\mu\text{M}$ . Un traitement non-linéaire des données cinétiques a permis de déterminer un  $K_m$  d'environ 30  $\mu\text{M}$ . Ainsi, comparativement à d'autres enzymes de la même famille telle que la  $\beta 4\text{GalT7}$  recombinante humaine purifiée ( $K_m \sim 290 \mu\text{M}$ ), la  $\beta 3\text{GalT6}$  présente une bonne affinité pour son substrat accepteur. En revanche, avec une  $V_{\max}$  aux alentours de 1,53 nmol/min/mg de protéine, notre enzyme a présenté une vitesse de catalyse moindre comparée à celle de la  $\beta 4\text{GalT7}$  ( $\sim 9$  nmol/min/mg de protéine). Avec ces analyses, nous avons pu faire une estimation de la constante catalytique,  $k_{cat}$  de notre protéine qui est de l'ordre de  $0,05 \text{ min}^{-1}$ . Nous pouvons ainsi voir que notre protéine ne semble pas être une enzyme très active. A titre de comparaison, la  $\beta 4\text{GalT7}$  présente une constante catalytique d'environ  $66 \text{ min}^{-1}$ . La réalisation d'un test avec un autre lot de protéines a permis de retrouver des valeurs de  $V_{\max}$ ,  $K_m$  et  $k_{cat}$  comparables. Néanmoins, ces résultats doivent être tempérés par l'observation qu'une partie de notre protéine se trouve sous forme aggrégée. Il est ainsi probable que la valeur réelle de la  $V_{\max}$  de la  $\beta 3\text{GalT6}$  soit supérieure à la celle que nous avons obtenue.

Des tests de conservation ont été menés en parallèle en utilisant un lot de protéine purifiée dont les différentes aliquotes ont été conservées à différentes températures ( $4^\circ\text{C}$ ,  $-20^\circ\text{C}$  et  $-80^\circ\text{C}$ ) dans le tampon d'éluion (voir partie II.1.2 du Matériels et Méthodes) pendant une semaine. Des tests enzymatiques ont été ensuite réalisés à deux concentrations de protéines (100 et 500 ng). Les résultats ont montré une absence d'activité quelle que soit la température de conservation. Dans ces conditions, il sera important de disposer d'une préparation fraîche d'enzyme pour la réalisation des tests enzymatiques. Des essais ultérieurs devront être menés pour résoudre ce problème de conservation et trouver un tampon qui permette de stocker la protéine purifiée à plus long terme.

#### **I.5. Préparation d'un vecteur d'expression recombinant contenant la séquence codant la $h\beta 3\text{GalT6}$ pour la production en cellules d'insectes :**

En parallèle de la production de la protéine en système bactérien, nous avons envisagé de



mettre en place la production de la protéine en cellules d'insectes en collaboration avec l'équipe du Pr Christelle Breton (équipe Glycobiologie moléculaire et structurale, CERMAV, Grenoble). Le système baculovirus/cellules d'insectes a pour principe de détourner la machinerie de production du baculovirus pour produire des protéines recombinantes lors d'une infection virale des cellules d'insectes. Deux étapes sont nécessaires pour ce faire : (i) la préparation du baculovirus recombinant (réalisée en parallèle de l'optimisation de la production en bactéries) et (ii) l'infection des cellules d'insectes par ce virus recombinant et extraction des protéines produites (réalisée à Grenoble). Pour permettre la production de la protéine d'intérêt, les gènes viraux des protéines P10 et de la polyédrine (protéines virales produites en grande quantité et/ou rapidement après l'infection) sont remplacés par le gène d'intérêt, ce qui permettra de produire en grande quantité la protéine recombinante au moment de la traduction. Les avantages du système sont (i) la production de protéines possédant des modifications post-traductionnelles plus complexes dans les cellules d'insectes par rapport aux bactéries (Chambers et al., 2018). A ce stade, nous ne savons pas à quel point la présence des modifications post-traductionnelles (comme les glycosylations par exemple) peut affecter la solubilité de la  $\beta$ 3GalT6 ou son activité. (ii) Le système de production en cellules d'insectes permet la sécrétion de la protéine produite dans le milieu de culture des cellules, ce qui facilite l'extraction et la purification de la protéine. Un autre avantage du système réside dans le fait que la  $\beta$ 3GalT6 a déjà été produite en cellules d'insectes sous forme active, ce qui prouve que ce système peut être envisagé pour l'expression de cette protéine, ce qui a orienté notre choix pour tenter de reproduire la procédure décrite par Wen et al. (2014). Dans notre stratégie, un vecteur de transfert (pFastBac\_SS) contenant les séquences codant la protéine recombinante humaine délétée de ses 30 ou 55 premiers acides aminés (pFastBac\_SS-h $\beta$ 3GalT6 $\Delta$ Nter30 ou pFastBac\_SS-h $\beta$ 3GalT6 $\Delta$ Nter55) a été construit. Les plasmides obtenus ont été envoyés au CERMAV (Grenoble) pour les premiers essais de production de l'enzyme en cellules d'insectes. Leurs essais de production n'ayant pas abouti (aucune protéine n'a été produite par les cellules), d'autres étapes d'optimisation seraient conduites ultérieurement si ce système devait être envisagé à nouveau pour produire la  $\beta$ 3GalT6 recombinante humaine.

### **1.6. Principaux résultats et Discussion de l'axe 1 :**

L'objectif de mon premier axe de travail était de caractériser d'un point de vue fonctionnel la  $\beta$ 3GalT6 recombinante humaine sauvage. La production de la forme complète de la  $\beta$ 3GalT6 a déjà été réalisée en 2001 par Bai et al. (2001) dans les cellules COS-7 (cellules de rein de singe). Dans cette étude, la protéine a également été produite sous forme tronquée (délétée des

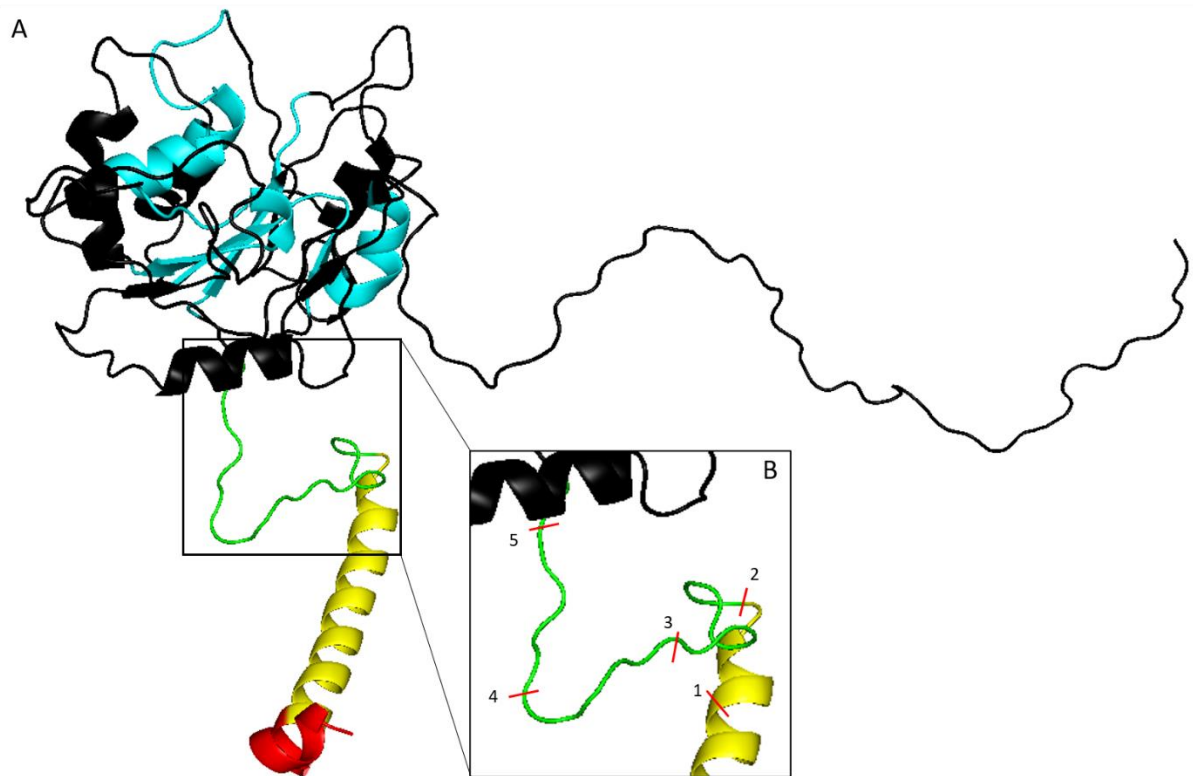


30 premiers acides aminés). Par la suite, elle a également été produite sous forme tronquée (toujours délétée de ses 30 premiers acides aminés) en cellules d'insectes en utilisant le système baculovirus (Wen et al., 2014). Ces études ont permis de produire la protéine sous forme active en système eucaryote, mais pour le moment aucune publication n'a décrit la production de la protéine en système bactérien. C'est dans ce contexte que j'ai mis au point un système d'expression de la  $\beta$ 3GalT6 recombinante humaine en bactéries qui a permis de produire la protéine d'intérêt sous forme tronquée soluble et de la purifier sous forme active et en quantité suffisante (~ 2,1 mg à partir de 600 mL de culture) pour la réalisation de cinétiques enzymatiques.

Les premiers résultats obtenus au cours de mon Master 2 en utilisant le plasmide pETM30 avaient montré un niveau de production important de la protéine dans les bactéries dans toutes les conditions testées. Néanmoins, aucune protéine soluble n'avait pu être obtenue à l'issue des différentes manipulations. Ceci pouvait être dû soit à une lyse incomplète, soit à une insolubilité de la protéine résultant d'une agrégation dans nos conditions expérimentales.

Pour répondre à la possibilité d'une lyse bactérienne incomplète, nous avons essayé de mettre au point un protocole de lyse utilisant la  $\beta$ 4GalT7, une GT dont les conditions de production et de solubilité sont connues et maîtrisées dans l'équipe et qui nous a servi de contrôle. Une fois les conditions de lyse optimisées, nous avons montré que, malgré ce protocole, la  $\beta$ 3GalT6 n'était pas soluble et donc que les problèmes de solubilité ne provenaient pas de la lyse bactérienne, mais plutôt de mauvaises conditions de production et/ou de la composition non adéquate du tampon de lyse utilisé.

Nous avons donc opté pour une modification du tampon de lyse des bactéries en changeant un à un plusieurs paramètres (présence de détergent ou d'agent dénaturant, nature et pH du tampon). En parallèle, d'autres conditions de culture bactérienne ont été testées (température de production réduite, temps de culture allongé, concentration réduite en IPTG, chocs thermique et osmotique) pour essayer de ralentir la production des protéines et ainsi limiter leur agrégation. Malgré ces optimisations, aucune condition n'a permis la solubilisation de la protéine de fusion. A ce stade, nous nous sommes demandés si la troncation réalisée pour l'étude était adaptée. Dans nos premières expériences, nous avons produit la h $\beta$ 3GalT6 délétée de ses 53 premiers acides aminés (h $\beta$ 3GalT6 $\Delta$ Nter53), correspondant à une enzyme dépourvue de la quasi-totalité de sa région stem. Cette troncation a été choisie sur la base des études précédemment réalisées sur la  $\beta$ 4GalT7 qui avaient montré que l'absence de la région stem permettait la solubilité de la protéine (Saliba et al., 2015; Talhaoui et al., 2010). Dans le cas de



**Figure 28 – Modèle de la structure tridimensionnelle de la h $\beta$ 3GalT6.**

(A) Le modèle montre les différents domaines de la protéine (voir Introduction partie II.2.) avec le domaine N-terminal cytosolique (en rouge), transmembranaire (en jaune), la région *stem* (en vert – séquence de la région stem : 35-EPGDPRAMSGRSPPPPAPARA-55) et le domaine catalytique (en noir). Les domaines conservés dans la famille des  $\beta$ 3GalTs sont représentés en bleu. (B) Un agrandissement de la région *stem* non structurée montre la présence d'un coude créé par la succession de résidus proline (vers 4). Les barres rouges indiquent les différentes troncations réalisées pour la production en bactéries (1, délétée des 30 premiers acides aminés ; 2, des 34 premiers acides aminés ; 3, des 43 premiers acides aminés ; 4, des 50 premiers acides aminés et 5, des 55 premiers acides aminés).



notre protéine, il est possible que la troncation de la région *stem* soit trop importante et qu'une séquence nécessaire au bon repliement de la protéine ne soit pas présente si cette région est déléetée. C'est ce qui a pu être observé dans le cas de la *xyloglucan xylosyltransferase 2*, une GT de plante, pour laquelle la présence de la région *stem* augmentait la solubilité de la protéine et l'activité catalytique (Culbertson et al., 2016; Wen et al., 2014). Des observations similaires ont par ailleurs été décrites dans le cas d'une GT humaine, la  $\beta$ 1,4-galactosyltransférase 1 (Boeggeman et al., 2003).

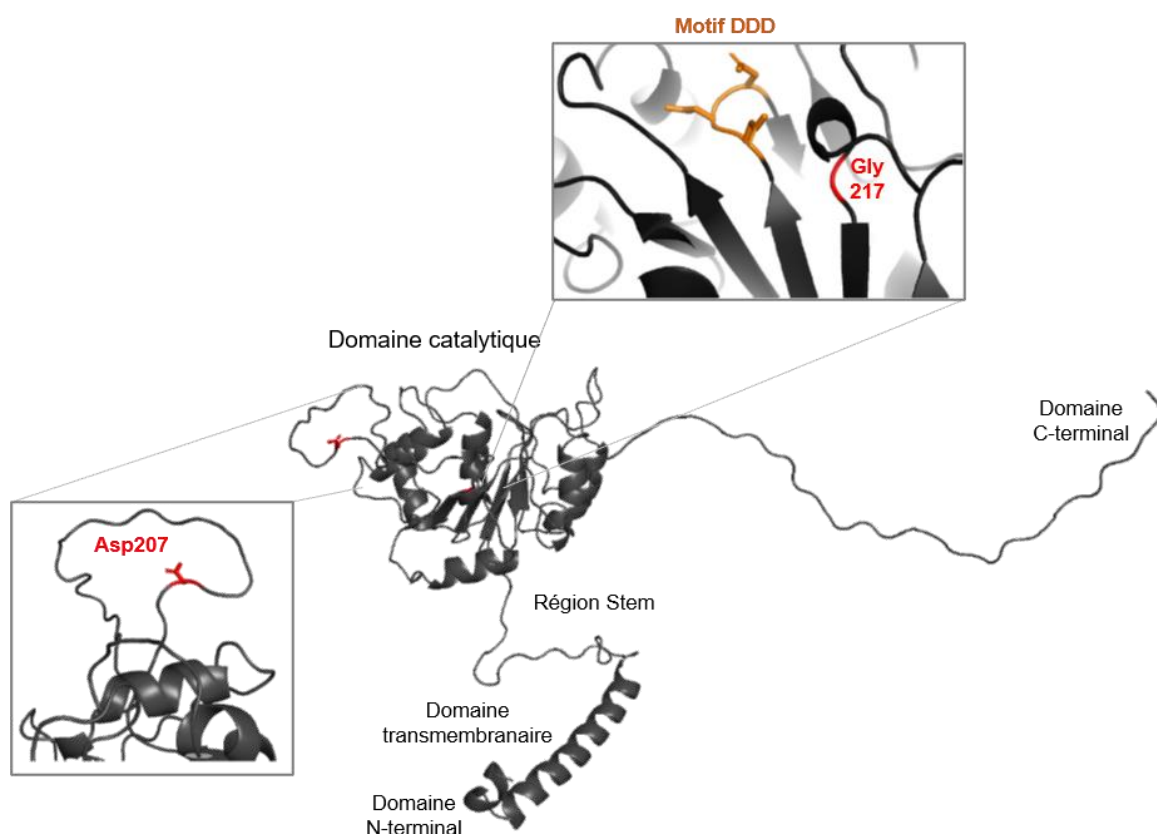
Durant mes travaux de thèse, nous avons bénéficié d'une collaboration avec le Dr François Hoh (Centre de Biochimie Structurale, Montpellier) pour l'établissement d'un modèle moléculaire de la  $\beta$ 3GalT6. Le modèle structural de la protéine a été généré grâce au logiciel Raptor (<http://RaptorX.uchicago.edu/>, (Källberg et al., 2012; Peng and Xu, 2011)) à partir du domaine catalytique d'une  $\beta$ 1,3-N-acétylglucosaminyltransférase murine, *mouse Manic Fringe* (référéncée 2J0A) (Jinek and Conti, 2006). Le modèle de la  $\beta$ 3GalT6 montre une organisation structurale globale de type  $\alpha/\beta/\alpha$  au niveau du domaine catalytique avec la présence d'un feuillet  $\beta$  antiparallèle central (composé de quatre brins  $\beta$ ), entouré de quatre hélices  $\alpha$  d'un côté et trois de l'autre (Figure 28). Cette structure du domaine catalytique est commune aux GTs adoptant un repliement de type GT-A (Voir partie III.1 de l'introduction) (Breton et al., 2012). Dans cette structure, il est possible de retrouver le motif DXD, motif important pour l'enzyme car il intervient dans la fixation du cofacteur métallique ( $Mn^{2+}$ ) et du substrat donneur (ici l'UDP-Gal). Dans la h $\beta$ 3GalT6, ce motif est constitué des résidus aspartates 156, 157 et 158. Dans ce modèle, le domaine C-terminal ne présenterait aucune structure secondaire et ne serait pas structuré. Au niveau de la région *stem*, qui apparait également non structurée dans ce modèle, la présence d'une séquence polyproline pourrait induire la formation d'une structure secondaire (un coude) (Figure 28). La présence de ce coude pourrait être importante pour le bon repliement de la protéine et contribuer à la stabilité du domaine catalytique de la  $\beta$ 3GalT6.

Nous avons décidé de réaliser plusieurs troncations au niveau de la région *stem*, afin de vérifier le rôle de cette région dans le repliement et donc sur la solubilité de la  $\beta$ 3GalT6. Pour déterminer si la troncation a un impact sur la solubilité de la protéine et quelle est la troncation la plus adaptée dans les conditions de production utilisées, nous avons mis en place un système de criblage dans lequel nous avons testé la solubilité de chaque troncation. Pour ce faire, différents plasmides (pETM30, pET41a et pETM41) contenant les séquences codantes correspondant aux protéines tronquées (Figure 17) ont été préparés. Les productions ont été réalisées dans différentes souches bactériennes : Rosetta2 (DE3), BL21(DE3) et C41(DE3) en faisant varier les paramètres de production (température, concentration en IPTG, milieu de



production). Nous avons pu montrer que la troncation de la protéine ne semblait pas avoir d'impact sur la solubilité de la protéine contrairement au vecteur d'expression utilisé. L'utilisation du plasmide pETM30 (employé pour les premières manipulations) n'a pas permis d'obtenir la protéine sous forme soluble contrairement aux deux autres constructions (utilisation des vecteurs d'expression pETM41 et pET41a). Nous avons également remarqué que les autres paramètres de production influençaient peu (ou pas) la solubilité de la protéine.

Par la suite, nous avons pu observer que les productions utilisant le vecteur pET41a conduisaient à la production d'une protéine de fusion plus « sensible » à la coupure du partenaire de fusion (GST) que la seconde (pETM41 avec le partenaire de fusion MBP). Pour récupérer plus facilement la protéine de fusion, nous avons opté pour l'utilisation de la construction pETM41 qui produit une protéine de fusion plus stable. La protéine produite dans ce cas est une forme tronquée soluble de la h $\beta$ 3GalT6 en fusion avec la MBP (6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30). Nous avons donc décidé d'utiliser le plasmide pETM41 pour la suite de nos travaux. Les analyses de DLS de la protéine soluble purifiée 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 ont montré que la protéine de fusion est soluble, mais qu'elle est également partiellement agrégée. Ainsi, les conditions de production et/ou de purification ne sont sans doute pas encore complètement optimales et des améliorations de ces étapes seront nécessaires. Nous avons également montré que la protéine de fusion 6His-MBP-h $\beta$ 3GalT6 $\Delta$ Nter30 ne présentait plus d'activité après une conservation à court et moyen terme (au-delà d'une semaine). Les essais de conservation de la protéine à différentes températures, même en présence de glycérol (utilisé comme agent cryoprotecteur) n'ont pas permis de maintenir l'activité enzymatique de la protéine. Ce problème pourrait venir de l'instabilité de la protéine qui aurait tendance à s'agréger, comme nous l'avons montré avec les tests de DLS. La stabilité des autres formes tronquées de la protéine n'a pas été étudiée, mais le sera ultérieurement afin de voir si la troncation de la protéine a un impact sur ce phénomène de conservation et/ou de stabilité. En plus de la stabilité, l'activité enzymatique des autres troncations sera également étudiée afin de vérifier si la région *stem* a un impact sur la fonctionnalité de la protéine. En parallèle des mises au point de la purification, nous avons optimisé les conditions du test enzymatique et réalisé des premiers tests pour déterminer les paramètres cinétiques de notre enzyme. Ces essais ont permis de mettre en évidence une activité de la protéine purifiée produite vis-à-vis d'un substrat accepteur modifié synthétique, le composé Gal-Xyl(2P)-OMN. Nous avons pu mesurer une activité enzymatique d'environ 1,5 nmoles de produit formé/min/mg de protéine purifiée, ce qui est très proche des données de la littérature qui donnent des valeurs du même ordre pour la



**Figure 29 – Modèle moléculaire de la hβ3GalT6 sauvage et position des mutations décrites dans les syndromes d'Ehlers-Danlos (publication en cours).**

Modèle structural de la hβ3GalT6 généré par le logiciel Raptor (<http://RaptorX.uchicago.edu/>) à partir du domaine catalytique d'une β1,3-N-acétylglucosaminyltransférase murine, *mouse Manic Fringe* (référéncée 2J0A). Le site actif de la β3GalT6 s'organise selon une structure de type α/β/α caractéristique des enzymes adoptant un repliement de type GT-A. La position des résidus mutés et l'orientation des chaînes latérales de ces acides aminés (Asp207 et Gly217) sont montrées dans les inserts et dans la structure générale de la protéine. Le résidu Asp207 se trouve en périphérie au niveau d'une boucle (insert Asp207). Le résidu Gly217 est enfoui dans le site actif et se trouve au niveau d'une petite boucle comprise entre deux structures secondaires (hélice et brin β). Il se trouve à proximité du motif DDD (insert Gly217, en orange).

protéine produite en cellules d'insectes (vis-à-vis d'un substrat osidique portant un résidu galactose terminal couplé à un groupement benzyle et non méthoxy-naphthyle comme notre analogue de substrat accepteur – Figure 12B) (Bai et al., 2001; Wen et al., 2014). Par ailleurs, une première estimation du  $K_M$  de l'enzyme purifiée indique une valeur de 30  $\mu\text{M}$ , attestant de la bonne affinité de l'enzyme pour ce substrat accepteur. A titre de comparaison, Wen et al. (2014) ont déterminé un  $K_M$  de 215  $\mu\text{M}$  vis-à-vis du substrat portant un groupement benzyle avec une enzyme purifiée produite en cellules d'insectes. La variation de l'affinité de la protéine pour son substrat accepteur peut s'expliquer par la nature du groupement aromatique couplé au sucre. En effet, le groupement méthoxy-naphthyle présente un cycle aromatique supplémentaire par rapport au groupement benzyle, ce qui le rend plus hydrophobe, et peut être plus affin pour la protéine.

En parallèle de la production en système bactérien, nous avons mis en place une collaboration avec le CERMAV (Grenoble) pour réaliser des premiers essais de production de la protéine en cellules d'insectes en utilisant le système baculovirus. L'utilisation de ce modèle pour produire la h $\beta$ 3GalT6 est très pertinent car la protéine a déjà été produite dans ce système tout comme d'autres GTs, telles que la  $\beta$ 4GalT1 ou la XylT-I (Kim et al., 2003; Kuhn et al., 2003; Morokuma et al., 2017). Cependant, les premiers essais qui ont été mis en place n'ont pas permis la production de la protéine et d'autres tests devront être menés ultérieurement.

De même que les études réalisées avec l'enzyme sauvage, les formes mutées de la protéine observées chez les patients atteints de SEDsp (Asp207His et Gly217Ser en particulier (Malfait et al., 2013)) ont également été modélisées, en suivant la même démarche que pour le modèle sauvage de la h $\beta$ 3GalT6, afin d'évaluer l'impact des mutations sur la structure de l'enzyme. Les modèles obtenus pour les protéines mutées n'ont pas montré de perturbation importante de la structure globale des protéines modélisées avec des structures mutées qui se superposent avec le modèle de la protéine sauvage. D'après le modèle moléculaire généré, le résidu Asp207 est localisé dans une boucle à la surface externe de la protéine. De telles boucles peuvent potentiellement être impliquées dans des changements conformationnels après la fixation du substrat donneur par exemple, ce qui a déjà été décrit pour la  $\beta$ 4GalT7 (Rahuel-Clermont et al., 2010; Tsutsui et al., 2013). La mutation de l'aspartate en histidine pourrait donc localement affecter la flexibilité de la boucle et potentiellement influencer la fixation des substrats sur l'enzyme. Une étude de modélisation moléculaire dynamique de la protéine pourrait permettre d'apporter des réponses sur l'implication de cette mutation sur la structure et la fonction de la protéine en relation avec les mouvements de la boucle dans laquelle se trouve le résidu muté. Contrairement au résidu Asp207 qui se trouve en périphérie de la protéine, la glycine 217 est



un résidu enfoui dans la structure modélisée et localisé entre deux structures secondaires, à la fin d'une courte hélice  $\alpha$  et avant le début d'un brin  $\beta$  (Figure 29). Le remplacement de la glycine par un résidu sérine plus volumineux (et potentiellement chargé) dans le mutant Gly217Ser, pourrait, par encombrement stérique, perturber l'arrangement local de la boucle contenant le motif DDD (Asp156-Asp157-Asp158 dans la  $\beta$ 3GalT6 sauvage), localisée dans un environnement spatial proche du résidu Gly217. Le motif DXD est connu pour être impliqué dans la coordination du cofacteur métallique, le manganèse, nécessaire à la fixation du substrat donneur, l'UDP-Gal (Zea et al., 2008). Il est raisonnable de supposer que la mutation Gly217Ser pourrait affecter l'interaction entre le substrat donneur, le cofacteur et la protéine. Ainsi, il serait possible d'expliquer la perte de l'activité enzymatique chez le patient présentant cette mutation par un défaut de fixation d'un des substrats ou par son mauvais positionnement suite à la perturbation de son interaction avec la  $\beta$ 3GalT6 (Patient 5 - Malfait et al., 2013).

Des études cinétiques des mutants et des études structurales ultérieures seront réalisées afin de tester ces hypothèses. En effet, il faut rester prudent car ces données sont issues d'un modèle structural, généré sur la base d'une similarité de structures. Ainsi, même si la prédiction peut être relativement précise, il est possible que de fortes variations conformationnelles puissent être observées suite à la fixation de l'un ou l'autre des substrats. Seuls des approches de *docking*, l'établissement d'un modèle dynamique et la détermination de la structure cristallographique de cette protéine pourront aider à mieux comprendre l'implication de ces résidus dans d'éventuelles modifications conformationnelles qu'elles soient localisées ou plus globales.

Dans cet axe de notre travail, nous avons démarré l'étude moléculaire de la h $\beta$ 3GalT6 recombinante humaine afin de comprendre l'impact des mutations observées chez les patients atteints de SEDsp sur la structure et la fonction de la protéine. Cependant, la stratégie d'expression dans la bactérie ne peut fournir d'informations sur l'impact que ces mutations ont sur la synthèse des GAGs ou sur les différentes fonctions cellulaires, comme la migration ou la prolifération. Or, ces informations sont fondamentales pour une meilleure compréhension de l'impact physiopathologique des mutations de la  $\beta$ 3GalT6. Il est donc nécessaire d'obtenir des données à l'échelle de la cellule en plus des données moléculaires fournies par l'étude de l'enzyme recombinante. Dans cet objectif, nous avons décidé de générer et de caractériser un modèle déficient en h $\beta$ 3GalT6.

## II. Génération et caractérisation d'un modèle cellulaire déficient en $\beta$ 3GalT6 :

L'objectif de cette partie de mon travail est de comprendre l'impact de la déficience génétique en  $\beta$ 3GalT6 sur les fonctions métaboliques des cellules invalidées, en particulier leur



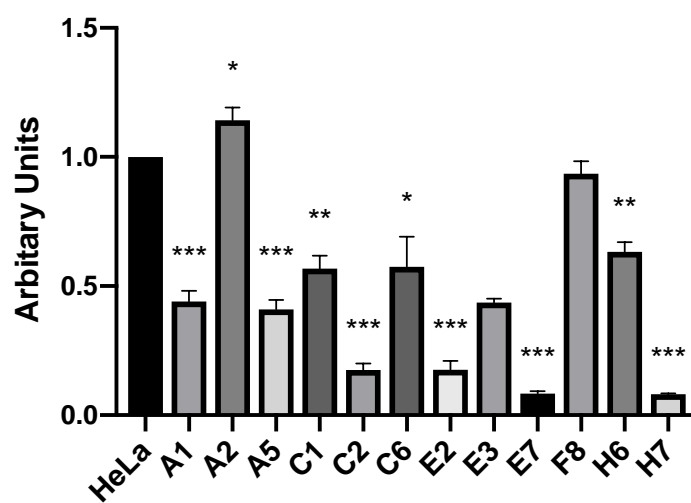


capacité à produire des GAGs. Pour réaliser cette étude, la lignée cellulaire sélectionnée est la lignée HeLa (cellules humaines issues d'un carcinome du col de l'utérus) dans laquelle j'ai souhaité invalider le gène *B3GALT6* par la technique CRISPR/Cas9. Ce choix sera argumenté dans la discussion.

Pour mener à bien cette étape, nous avons fait appel à la plateforme TACGENE (MNHN – CNRS UMR 7196 / INSERM U1154 – Sorbonne Universités) qui a conçu et construit trois plasmides permettant d'obtenir des lignées *knock-out* (KO). Ces plasmides contenaient : (i) une séquence codant un ARN guide (ARNg) ciblant le gène *B3GALT6*, (ii) le gène de la Cas9 (*CRISPR associated protein 9*) et (iii) le gène de l'EGFP (*Enhanced Green Fluorescent Protein*), qui permet la détection et la sélection des cellules ayant intégré le plasmide d'intérêt. TACGENE a également vérifié la capacité des plasmides à induire une coupure du gène *B3GALT6*. Une fois les plasmides reçus, nous avons sélectionné le plasmide qui présentait le taux de coupure d'ADN génomique le plus élevé. Le plasmide choisi a été utilisé pour transfecter les cellules HeLa pour obtenir des clones déficients (selon le protocole décrit dans la partie matériels et méthodes). Les cellules transfectées ont ensuite été triées en utilisant le trieur de cellules de la Plateforme de cytométrie de l'UMS IBSLor en suivant l'expression de la GFP dans les cellules recombinantes. Une fois les clones recombinants triés, la sélection clonale a été réalisée par dilution limite afin d'obtenir des clones isolés.

Douze clones potentiellement KO pour le gène *B3GALT6* ont été obtenus à l'issue de l'étape de sélection clonale. Nous avons ensuite analysé ces clones d'un point de vue moléculaire et fonctionnel afin de déterminer si nous avons bien une lignée cellulaire *B3GALT6*-déficiente. Pour ce faire, plusieurs étapes ont été successivement réalisées : (i) étude de l'impact de la mutation sur le niveau d'expression de l'ARNm de la  $\beta$ GalT6 par RT-qPCR, (ii) séquençage du gène *B3GALT6* afin de déterminer la nature de l'évènement qui a conduit à l'invalidation du gène cible (insertion, délétion ou mutation du gène), (iii) mesure de l'activité de l'enzyme par CLHP et (iv) estimation de la biosynthèse des GAGs (endogène et à partir d'un xyloside exogène) par incorporation d'un précurseur radiomarqué dans les chaînes glycaniques néosynthétisées.

La réalisation de ces différents tests confirmera si les clones obtenus sont invalidés pour le gène *B3GALT6* et permettra de déterminer les conséquences de cette invalidation sur les fonctions cellulaires. En outre, la réexpression de l'enzyme sauvage (ainsi que celle des mutants) dans ces cellules invalidées pourra ensuite être réalisée pour caractériser sur le plan fonctionnel les protéines sauvages et mutées, afin (i) de comprendre l'impact de chaque mutation sur l'activité galactosyltransférase et (ii) d'étudier les déficiences des cellules



**Figure 30 – Analyse du taux d’expression des ARNm de la  $\beta$ 3GalT6 dans les cellules HeLa suite à l’étape d’invalidation.**

L’expérience a été réalisée une fois avec des essais en triplicat, \* $p < 0.05$ , \*\* $p < 0.01$  et \*\*\* $p < 0.0001$ , test  $t$  de Student.

recombinantes en nous concentrant sur la capacité de ces cellules exprimant une enzyme mutée à produire des GAGs. Pour cela, avec l'aide de notre équipe, l'activité galactosyltransférase de ces protéines mutées a été mesurée et évaluée leur capacité à initier la glycosylation d'un protéoglycane modèle, la décorine.

### **II.1. Analyse de l'expression de l'ARNm de la $\beta$ GalT6 des clones recombinants de cellules HeLa :**

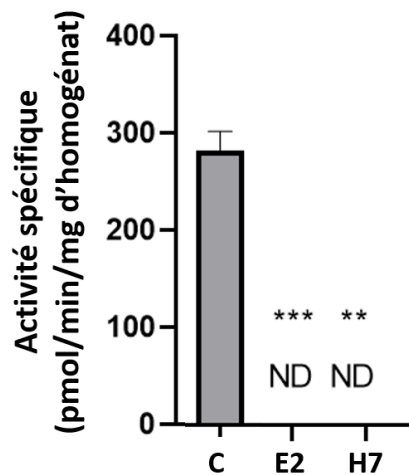
Nous avons évalué le taux d'expression de l'ARNm de la  $\beta$ GalT6 dans des cellules HeLa sauvage et dans les douze clones précédemment sélectionnés (notés A1, A2, A5, C1, C2, C6, E2, E3, E7, F8, H6 et H7 – Figure 30). Le clone F8 présente une expression de l'ARNm inchangée par rapport aux cellules HeLa, alors que le clone A2 montre une augmentation faible mais significative de l'expression de l'ARNm par rapport au contrôle. Il est possible que ces clones ne présentent aucune mutation du gène *B3GALT6* ou que la mutation n'impacte pas l'expression de l'ARNm. Six clones (A1, A5, C1, C6, E3 et H6) présentent une expression de l'ARNm réduite d'environ 50% par rapport à l'expression de l'ARNm dans les cellules HeLa donc doivent posséder une mutation qui impacte le taux d'expression de l'ARNm. Pour finir, quatre clones présentent une diminution de l'expression de l'ARNm à hauteur de 80% (C2 et E2) ou de 90% (E7 et H7). Ces 4 derniers clones doivent présenter une mutation avec un fort impact sur le taux d'expression de l'ARNm. La réalisation du séquençage au niveau du gène *B3GALT6* qui a été réalisée par la suite a permis de répondre à cette question et de corrélérer la nature de la mutation à l'impact sur le niveau d'expression de l'ARNm (débuté et encore en cours).

Pour la réalisation du séquençage génomique, nous nous sommes concentrés sur les clones qui présentaient le plus faible niveau d'expression de l'ARNm (clones E2 et H7). Nous n'avons pas poursuivi les études avec le clone E7 qui n'a pu être remis en culture.

### **II.2. Caractérisation génomique des cellules HeLa *B3GALT6* déficientes :**

Les résultats du séquençage du gène *B3GALT6* du clone E2 ont montré la délétion d'un nucléotide (c.222del) prédit pour induire un décalage du cadre de lecture avec le changement du résidu 74 qui sera immédiatement suivi d'un codon stop (mutant p.Ser74Argfs\*1). Le second clone, H7, présente une délétion de 52 paires de bases dans le gène *B3GALT6* (c.182\_233del), conduisant à un changement d'acide aminé avec décalage du cadre de lecture qui sera suivi d'un codon stop en position 260 (mutant p.Val61Alafs\*199).

Ces résultats montrent que nous avons obtenu deux clones déficients présentant deux défauts



**Figure 31 – Analyse de l’activité enzymatique de la  $\beta$ 3GalT6 dans les homogénats cellulaires issus des clones déficients.**

Les essais enzymatiques ont été réalisés en utilisant 2 mM substrat accepteur, 1 mM de substrat donneur (UDP-Gal) et 50  $\mu$ g de protéines totales d’homogénats issus de cellules contrôle (C, HeLa) ou déficientes (clones E2 et H7). Après un temps de réaction de 60 min, l’activité enzymatique est évaluée par CLHP en phase inverse. Les essais de cette expérience ont été réalisés en triplicat, \* $p < 0.05$ , \*\* $p < 0.01$  et \*\*\* $p < 0.0001$ , test  $t$  de Student.

génétiques différents sur le gène *B3GALT6*, associés à une réduction importante de l'expression de l'ARNm de la  $\beta$ GalT6. Nous allons ensuite pouvoir étudier les conséquences de cette déficience sur l'activité enzymatique de la  $\beta$ GalT6 et la biosynthèse des GAGs par les cellules *B3GALT6* déficientes. Pour ce faire, nous avons regardé la capacité des cellules déficientes (et contrôle) (i) à initier le transfert du galactose à partir de l'UDP-Gal sur un substrat accepteur synthétique, le Gal-Xyl(2P)-OMN, (ii) à initier la biosynthèse des GAGs à partir d'un xyloside exogène (4MOX) et (iii) à réaliser la glycanation de la protéine *core* d'un PG que nous avons choisi comme modèle, la décorine.

### **II.3. Analyse de l'activité enzymatique de la $\beta$ GalT6 dans les homogénats cellulaires des clones déficients :**

Des mesures de l'activité galactosyltransférase ont été réalisées dans les cellules HeLa déficientes (clones C2, E2 et H7) ou non (HeLa contrôle, clone A2) comme décrit dans la partie matériels et méthodes, en utilisant un analogue du substrat accepteur le Gal-Xyl(2P)-OMN. La capacité des cellules à réaliser le transfert d'un résidu galactose sur cet accepteur a été évaluée *in vitro* sur des homogénats cellulaires préparés à partir de ces différents clones (Figure 31).

Les cellules contrôle présentent une activité galactosyltransférase comprise entre 250 et 300 pmol/min/mg de protéines totales. Aucune activité enzymatique n'est détectable dans les clones E2 et H7, (Figure 31). Ainsi, les clones E2 et H7 présentent une perte totale d'activité galactosyltransférase en accord avec les modifications du gène identifiées par le séquençage. Ces résultats indiquent que les clones E2 et H7 sont bien déficients, présentant un faible niveau d'expression de l'ARNm et une absence d'activité enzymatique.

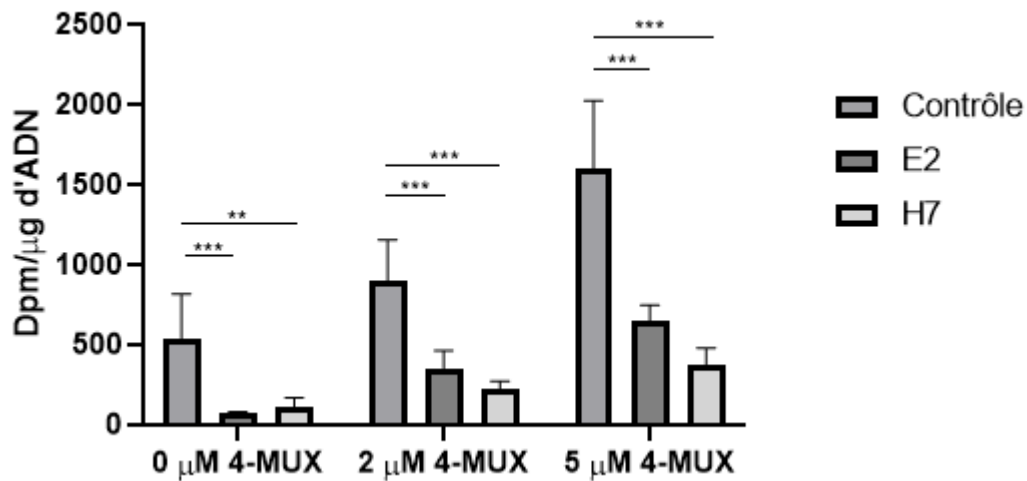
### **II.4. Analyse de la synthèse des GAGs dans les lignées HeLa *B3GALT6* déficientes :**

#### **II.4.1. Analyse de la synthèse des GAGs à partir d'un xyloside exogène, le 4-méthylumbelliferyl- $\beta$ -D-xylopyranoside (4-MOX) :**

Afin d'évaluer les conséquences de la déficience génétique causée par l'inactivation du gène *B3GALT6* sur le métabolisme cellulaire, nous avons d'abord mesuré la capacité de synthèse des GAGs matriciels des clones HeLa déficients (E2 et H7) par rapport à des cellules contrôle (non déficientes).

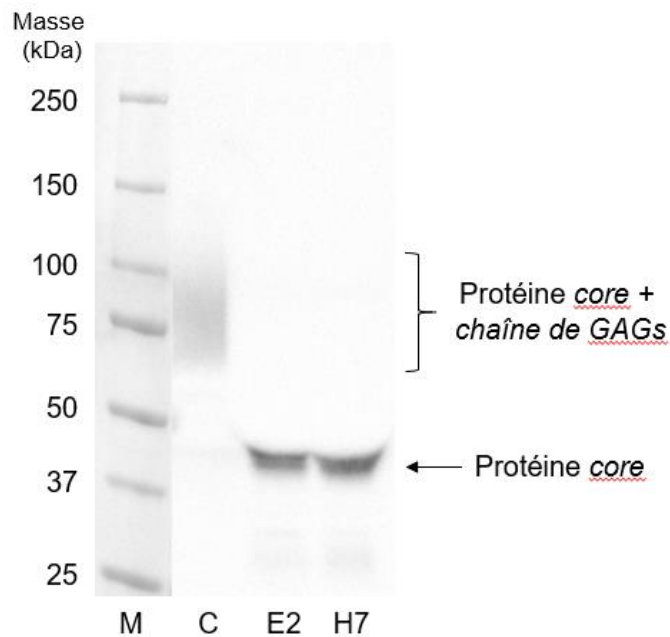
Les résultats de la figure 32 montrent qu'il y a une perte d'environ 85% du niveau de synthèse basal des GAGs dans les clones E2 et H7 par rapport aux cellules contrôle en absence





**Figure 32 – Analyse de la conséquence de la déficience en  $\beta 3\text{GalT6}$  dans les lignées HeLa KO sur l’anabolisme des GAGs.**

Le taux de synthèse des GAGs matriciels est analysé dans les cellules contrôle et déficientes (E2 et H7), en présence (ou absence) de xyloside exogène (4-MOX) et de soufre radiomarqué comme décrit dans la partie Matériels et Méthodes. Trois expériences indépendantes ont été réalisées (n=3), \*p<0.05, \*\*p<0.01 et \*\*\*p<0.0001, *Anova One-Way* (avec correction de Tukey).



**Figure 33 – Analyse de la glycosylation de la protéine *core* de la décorine dans les cellules HeLa contrôle et déficientes, E2 et H7.**

La glycanation de la décorine dans les cellules contrôle (C) et les clones déficients (E2 et H7) a été analysée par Western Blot dans des cellules transfectées avec un plasmide contenant le gène codant la protéine *core* de la décorine. La protéine *core*, non glycosylée, présente une masse moléculaire apparente d'environ 38 kDa alors que le PG (protéine *core* et chaîne de GAGs) a une masse moléculaire entre 65 et 100 kDa.



de xyloside exogène dans le milieu de culture. En présence de xyloside 4-MOX, une augmentation dose-dépendante de la biosynthèse des GAGs est observée entre 0, 2 et 5  $\mu\text{M}$  de 4-MOX. Néanmoins, la synthèse des GAGs à partir de 4-MOX est fortement réduite dans les clones déficients E2 et H7 par rapport aux cellules contrôle. A la plus haute concentration en 4-MOX (5  $\mu\text{M}$ ), la synthèse des GAGs dans les cellules KO est par exemple réduite d'environ 60% pour le clone E2 et de 80% pour le clone H7.

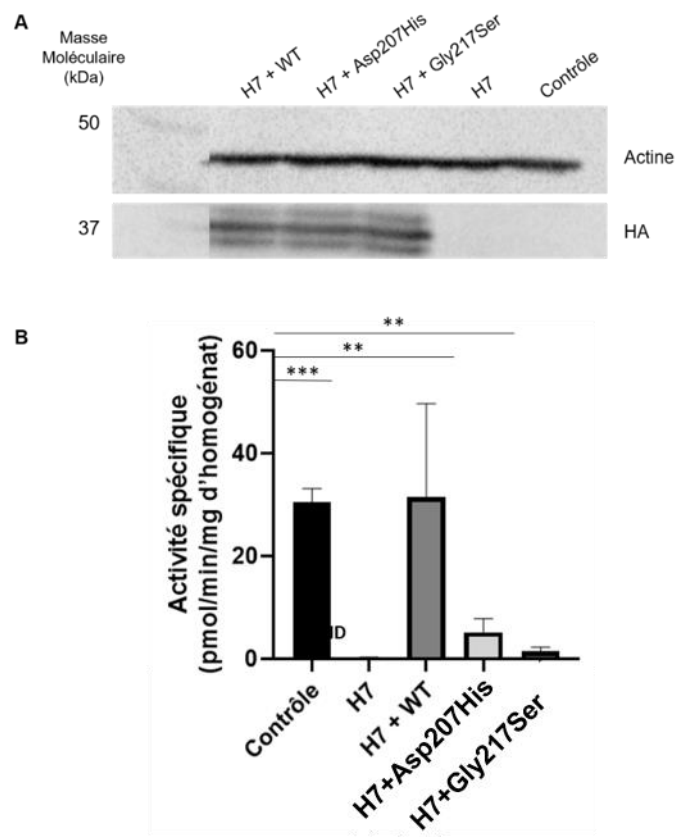
Ces résultats montrent que les deux clones déficients restent capables de produire des chaînes de GAGs malgré la perte totale d'activité enzymatique de la  $\beta\text{3GalT6}$ , mais en plus faible quantité par rapport aux cellules contrôle dans nos conditions expérimentales. Dans un deuxième temps, nous nous sommes intéressés à la capacité des cellules déficientes à initier la biosynthèse des chaînes de GAGs, non plus à partir d'un xyloside exogène, mais sur la protéine *core* d'un PG modèle. Dans ce travail, nous avons étudié la glycanation de la décorine, un petit PG impliqué dans la fibrillogenèse des collagènes et dans l'organisation matricielle (voir Introduction, parties II et III).

#### II.4.2. Analyse de la glycosylation de la protéine core de la décorine :

La capacité des cellules *B3GALT6* déficientes à initier la glycosylation de la protéine *core* de la décorine a été évaluée par immunoblot dans des cellules transfectées par un plasmide contenant la séquence codant la protéine *core* de la décorine couplée à une étiquette polyhistidine. Pour étudier la glycanation de la décorine, la protéine *core* a été extraite du milieu de culture des cellules par chromatographie d'affinité utilisant des billes de nickel possédant une affinité pour l'étiquette histidine avant analyse de la glycosylation de la décorine par Western Blot.

Les résultats de la figure 33 montrent une bande large au niveau des protéines dont la masse médiane est autour de 75 kDa pour les cellules contrôle. Cette bande correspond à la protéine *core* de la décorine ayant été glycosylée dans des cellules dans lesquelles la  $\beta\text{3GalT6}$  est active. Dans le cas des deux clones déficients, nous n'observons pas, dans nos conditions expérimentales, de bande à cette masse moléculaire mais une bande plus fine correspondant à une protéine dont la masse est d'environ 38 kDa, correspondant à la taille attendue pour la protéine *core* de la décorine sans chaîne de GAGs. L'absence de chaîne de GAGs dans les deux clones déficients montre que ces cellules ne sont pas capables d'initier la biosynthèse des GAGs sur la protéine *core* dans des cellules dans lesquelles la  $\beta\text{3GalT6}$  n'est pas active.

En résumé, nous avons réussi à obtenir deux clones déficients présentant : (i) une réduction d'expression d'ARNm du gène *B3GALT6* de plus de 80%, (ii) une absence d'activité



**Figure 34 – Analyse de l’expression et de l’activité enzymatique des mutants de la  $\beta$ 3GalT6 exprimés dans les cellules HeLa *B3GALT6*-déficientes (clone H7).**

(A) Analyse de la surexpression de la  $\beta$ 3GalT6 sauvage (WT) ou des mutants Asp207His et Gly217Ser par Western Blot dans les cellules déficientes en  $\beta$ 3GalT6 (clone H7). L’analyse de l’expression de la  $\beta$ 3GalT6 a été réalisée par détection de l’étiquette HA (étiquette en C-terminal de la protéine) et le contrôle de charge a été réalisé à l’aide d’un anticorps dirigé contre l’actine. (B) Les essais enzymatiques ont été réalisés en utilisant 2 mM de substrat accepteur, 1 mM de substrat donneur (UDP-Gal) avec 50  $\mu$ g de protéines d’homogénats issus de cellules déficientes transfectées avec le plasmide codant la protéine sauvage (H7 + WT) ou les protéines mutées (H7 + Asp207His et H7 + Gly217Ser). L’activité enzymatique est évaluée par CLHP en phase inverse comme décrit dans la partie Matériels et Méthodes. Ces expériences ont été réalisées sur un lot de protéines avec des essais en duplicat, \* $p < 0.05$ , \*\* $p < 0.01$  et \*\*\* $p < 0.0001$ , *Anova One-Way* (avec correction de Tukey).

galactosyltransférase et (iii) un anabolisme réduit en termes de capacité des cellules à initier la biosynthèse des GAGs. Ces clones déficients pourront être utilisés comme modèle cellulaire par la suite pour l'étude de l'impact des mutations de la  $\beta$ 3GalT6 observées chez les patients en déterminant si les protéines mutantes ré-exprimées dans ces cellules dépourvues d'activité endogène (i) présentent (ou non) une activité enzymatique et (ii) sont capables (ou non) d'initier la biosynthèse des GAGs *in cellulo*.

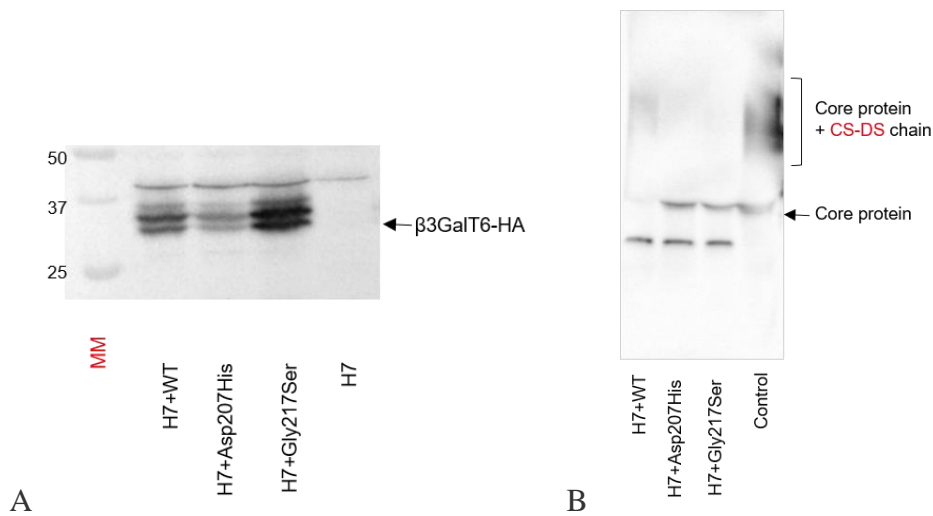
## **II.5. Analyse de la surexpression des mutants de la $\beta$ 3GalT6 dans la lignée HeLa B3GALT6 déficiente :**

Dans cette partie, nous nous sommes concentrés sur les mutations ponctuelles Asp207His et Gly217Ser présentes chez les premiers patients analysés dans notre équipe, décrites dans les travaux de Malfait et al. (2013). Ces mutants ont été obtenus par mutagenèse dirigée, comme décrit dans la partie Matériels et Méthodes. Les plasmides contenant la séquence codant les protéines mutées ont été transfectés dans les cellules de la lignée HeLa déficiente H7.

### **II.5.1. Analyse de l'activité galactosyltransférase des protéines mutantes :**

L'activité enzymatique a été évaluée dans les homogénats des cellules HeLa déficientes (clone H7) transfectées par les plasmides contenant la séquence codant les mutants Asp207His et Gly217Ser, et a été comparée à celle de l'enzyme sauvage (Figure 34). Il a d'abord été vérifié que la transfection des plasmides contenant la séquence codant la protéine sauvage ou celle des mutants permet de surexprimer les protéines correspondantes dans les cellules déficientes (clone H7). La présence des différentes protéines dans les cellules du clone H7 (Figure 34A) indique que les protéines ont bien été (sur)exprimées dans les conditions utilisées.

Lorsque la transfection de la protéine sauvage a été réalisée dans la lignée B3GALT6 déficiente, une restauration de l'activité enzymatique a été observée par rapport aux cellules contrôle (activité spécifique d'environ 30 pmol/min/mg de protéines). Une activité enzymatique de 5 pmol/min/mg de protéines totales a été mesurée pour la protéine mutée Asp207Ser (Figure 34B, H7+Asp207His), montrant que ce mutant présente une perte importante d'activité enzymatique. Pour le second mutant Gly217Ser, l'enzyme présente une activité spécifique plus faible que le mutant Asp207His (1,5 pmol/min/mg de protéines, Figure 34B, H7+Gly217Ser), indiquant un impact plus important de cette mutation sur l'activité galactosyltransférase que la mutation Asp207His.



**Figure 35 – Analyse de la glycosylation de la protéine *core* de la décorine dans les cellules HeLa déficientes transfectées par la  $\beta$ 3GalT6 sauvage ou mutée.**

(A) Analyse de la surexpression de la  $\beta$ 3GalT6 sauvage (WT) ou de ses mutants (Asp207His et Gly217Ser) par Western Blot dans les cellules déficientes en  $\beta$ 3GalT6 (clone H7). L'analyse de l'expression de la  $\beta$ 3GalT6 a été réalisée par détection de l'étiquette HA (étiquette en C-terminal de la protéine) et le contrôle de charge a été réalisé à l'aide d'un anticorps dirigé contre l'actine. (B) Analyse de la glycanation de la protéine *core* de la décorine dans les cellules contrôles (control, cellules HeLa) et les clones déficients transfectés (H7+WT, H7+Asp207His et H7+Gly217Ser) par Western Blot. La protéine *core*, non glycosylée, présente une masse moléculaire apparente d'environ 38 kDa alors que le PG a une masse moléculaire apparente de 65 à 100 kDa.

Ces résultats, qui restent préliminaires, montrent que les deux protéines mutées Asp207His et Gly217Ser présentent une activité galactosyltransférase très faible en comparaison de celle de l'enzyme sauvage. Il est intéressant de noter ici que, contrairement à ce qui est décrit dans les homogénats cellulaires préparés à partir de fibroblastes de derme de patients atteints de SEDsp pour lesquels aucune activité galactosyltransférase n'est détectée, les protéines mutées conservent tout de même une activité enzymatique résiduelle. Pour compléter ces données, nous avons regardé si ces protéines sont capables d'initier la glycanation de la protéine core de la décorine dans les cellules déficientes (clone H7).

### II.5.2. Analyse de la glycosylation de la protéine core de la décorine par les mutants de la $\beta$ 3GalT6 :

La capacité des cellules déficientes, transfectées avec un plasmide contenant la séquence codant la  $\beta$ 3GalT6 sauvage ou mutée, à initier la glycosylation de la protéine *core* d'un PG modèle, la décorine, a été déterminée par Western Blot. Pour cela, les cellules ont été co-transfectées avec un plasmide contenant la séquence codant la protéine *core* de la décorine, comme précédemment décrit.

Il a été d'abord vérifié que la transfection des plasmides contenant la séquence codant la protéine sauvage ou celle des mutants permet de surexprimer les protéines correspondantes dans les cellules déficientes (clone H7). Nous avons ainsi observé la présence des différentes protéines dans les cellules du clone H7 (Figure 35A) indiquant que les protéines ont bien été (sur)exprimées dans nos conditions. L'analyse de la glycanation de la décorine montre une large bande autour de 75 kDa dans les cellules contrôle (Figure 35B), correspondant à la masse moléculaire attendue de la protéine *core* glycosylée. Une glycosylation de la protéine *core* de la décorine est également observée dans le cas de la surexpression de la protéine sauvage dans les cellules déficientes (Figure 34B, H7+WT), même si elle apparait avec une plus faible intensité que celle des cellules contrôle (Figure 35B). Par contre, aucune glycanation de la protéine *core* n'est observée quand les protéines mutées Asp207His et Gly217Ser sont exprimées dans les cellules déficientes (Figure 35B), même si le mutant Asp207His semble moins bien exprimé dans les cellules du clone déficient H7 par rapport au mutant Gly217Ser. Ainsi, les mutants ne semblent pas en mesure d'initier la synthèse des GAGs sur la protéine *core* de la décorine, ce qui pourrait s'expliquer par la faible activité mesurée des protéines mutées ou à la sensibilité de notre méthode d'analyse par immunoblot.



## II.6. Discussion :

L'objectif de cette partie de mon travail a été d'évaluer l'impact cellulaire de la déficience génétique en  $\beta$ 3GalT6 et de fournir une lignée de cellules-hôte dénuée d'activité endogène.

Les précédents travaux réalisés au sein de l'équipe avaient notamment analysé l'impact des mutations du gène *B3GALT6* dans des fibroblastes de derme de patients sur leur capacité à synthétiser des GAGs et sur le phénotype cellulaire (Thèse Xiaomeng Pang, 2016). Pour ce faire, le gène *B3GALT6* avait été inactivé en utilisant une approche de transfection par siRNA. Ces premiers résultats avaient montré que, malgré une réduction d'environ 85-90% du niveau d'ARNm, la protéine était encore exprimée, mais associée à un niveau de synthèse de GAGs plus faible que celui observé dans les cellules transfectées par un siRNA contrôle dans ces cellules. Il n'avait pas été possible de générer un modèle cellulaire déficient en  $\beta$ 3GalT6 par cette méthode et d'obtenir des cellules qui n'exprimaient plus ou très peu l'enzyme. C'est pourquoi nous avons décidé de générer une lignée *B3GALT6* déficiente grâce à la technologie CRISPR/Cas9 qui est une méthode de référence pour inactiver spécifiquement et efficacement l'expression d'un gène (Stolfa et al., 2016; Termini et al., 2017).

La première étape de ce projet a été de choisir la lignée cellulaire à utiliser pour notre étude en privilégiant certains critères de sélection. De nombreuses études ont utilisé des cellules déficientes issues de lignées CHO (*Chinese Hamster Ovary cells*) pour étudier les GAGs et l'absence des enzymes impliquées dans leur biosynthèse (Esko et al., 1985; Zhang et al., 2006). Plus récemment, l'utilisation des cellules CHO a permis de générer une banque de cellules déficientes pour toutes les enzymes impliquées dans la biosynthèse et la maturation des GAGs (Chen et al., 2018). L'utilisation de ces cellules aurait été pertinente pour l'étude des GAGs, cependant, ces cellules ne sont pas d'origine humaine, point qui constitue l'un des premiers critères de sélection fixé pour le choix de l'établissement de notre lignée *B3GALT6* déficiente. L'utilisation d'une lignée humaine a pour intérêt de pouvoir comparer nos résultats avec ceux obtenus dans les fibroblastes de patients. L'utilisation des fibroblastes aurait été pertinente, mais la principale limite de l'utilisation de ces cellules est qu'elles sont difficilement transfectables et que les précédents essais de génération de cellules déficientes par la technologie CRISPR/Cas9 avec ces cellules n'ont pas donné de résultats satisfaisants. De plus, le phénotype des fibroblastes varie au cours des passages. En plus d'avoir une lignée cellulaire humaine, il était nécessaire de disposer d'une lignée qui puisse être transfectée facilement et dont la culture ne constitue pas une étape limitante. Notre choix s'était donc porté au départ sur deux lignées : les cellules HeLa et HEK293. Une étude a pu montrer que la lignée HEK293 (*Human Embryonic Kidney 293*) présentait la capacité à initier la synthèse des GAGs





(Chen et al., 2018). De plus, ces cellules sont connues pour être facilement cultivables et transfectables, ce qui en aurait fait une bonne candidate pour notre étude. A titre de comparaison, les cellules HeLa, des cellules épithéliales cancéreuses du col de l'utérus, sont également des cellules facilement cultivables et transfectables avec une activité galactosyltransférase ( $\beta 3\text{GalT6}$ ) basale facilement détectable (Thèse Xiaomeng Pang, 2016). Ces cellules sont en outre capables de synthétiser des GAGs en quantité suffisante pour voir une modulation de leur expression éventuelle. C'est donc sur ce dernier paramètre que nous avons décidé d'utiliser la lignée HeLa pour notre étude.

Au cours de ce travail, nous avons obtenu deux clones invalidés pour le gène *B3GALT6*. Dans ces cellules, malgré l'absence d'activité galactosyltransférase, les cellules restent encore capables de synthétiser, en faible quantité, des GAGs à partir d'un xyloside exogène. Ce résultat suggère qu'il existerait un moyen d'initier la synthèse sans la présence de l'enzyme. Récemment, une étude a montré la présence d'une amorce tronquée sur la protéine *core* de labikunine, un petit PG circulant, dans les urines de personnes saines (Persson et al., 2019). Ces études ont décrit une amorce trisaccharidique [GlcA-Gal-Xyl] minoritaire, avec la présence d'un seul galactose au lieu de deux dans la zone de liaison classique entre la protéine *core* et les chaînes de GAGs. Ces résultats sont les premiers à décrire une amorce « protéine *core*-GAG » comportant un seul résidu Gal, et indiquent que la GlcAT-1 serait capable d'ajouter un résidu GlcA sur le disaccharide Xyl-Gal. Une autre étude, ayant pour but d'établir une banque de lignées CHO-déficientes pour chacune des enzymes impliquées dans la synthèse et la maturation des GAGs, est en accord avec cette hypothèse. Dans ce travail, des cellules CHO KO pour la  $\beta 3\text{GalT6}$  conservent une capacité de synthèse des GAGs, bien que très fortement réduite, contrairement à l'inactivation du gène *B4GALT7* qui conduit à une absence totale de chaînes de GAGs (Chen et al., 2018). De plus, cette étude montre également que l'inactivation des gènes des autres GTs impliquées dans la biosynthèse de l'amorce, la *XYLT2* et la *B3GAT3*, entraîne également une absence totale de chaînes de GAGs. Il est donc intéressant de voir que l'inactivation du gène *B3GALT6* semble avoir un impact moindre sur la biosynthèse des GAGs que les trois autres enzymes responsables de la biosynthèse de l'amorce tétrasaccharidique. Ainsi, il est possible que nos clones déficients présentent un profil assez proche de celui du clone CHO déficient obtenu par Chen et al. (2018). Ainsi, les GAGs présents dans la lignée CHO KO et notre lignée HeLa déficiente pourraient être issus d'une voie de biosynthèse réalisée à partir d'une amorce tronquée, proposée par Persson et al. (2019). La présence de cette amorce tronquée dans nos cellules déficientes expliquerait pourquoi nous observons des chaînes de GAGs malgré l'absence d'une activité de la  $\beta 3\text{GalT6}$ . Il serait intéressant de pouvoir étudier



les GAGs synthétisés par nos clones déficients afin de vérifier la composition de l'amorce de ces GAGs. En plus de ces analyses, la composition des motifs disaccharidiques des GAGs pourrait également être analysée afin de voir quels types de GAGs sont synthétisés et quelles sont les modifications présentes sur les GAGs produits par ces lignées déficientes (travail en cours en collaboration avec le Professeur Göran Larson et le Dr Fredrik Noborn, Institute of Biomedicine, University of Gothenburg).

Nous nous sommes également demandé si les cellules déficientes étaient capables d'initier la synthèse des GAGs sur la protéine *core* d'un PG modèle, la décorine, comme elles pouvaient le faire dans une moindre mesure à partir du xyloside 4-MOX. Ces résultats ont montré une absence de GAGs sur la protéine *core* de la décorine dans nos conditions expérimentales, résultat déjà mis en évidence chez certains patients atteints de SEDsp : P3 (Malfait et al., 2013), PI:1 et PVII:1 (Van Damme et al., 2018), chez lesquels une diminution importante de la glycanation de la décorine dans les fibroblastes de peau de patients est décrite. Néanmoins, il faut rester prudent sur ces résultats car l'absence de glycanation de la décorine peut être due à la limite de détection de la méthode d'analyse (ici, le Western Blot). L'ensemble de ces résultats a montré que notre modèle cellulaire est robuste et proche de ce que l'on peut observer dans les fibroblastes de patients atteints de SEDsp. En effet, nous avons généré une lignée qui présente une perte d'activité de la  $\beta$ 3GalT6 et une diminution de la synthèse des GAGs à partir d'un substrat exogène.

Notre étude s'est concentrée sur l'impact cellulaire de la déficience génétique en  $\beta$ 3GalT6. Ce modèle *B3GALT6* déficient vise à mieux comprendre l'impact de la déficience sur la synthèse des GAGs et pourrait servir de point de départ pour différentes études « omiques ». En effet, nous pourrions étudier l'implication de la déficience en  $\beta$ 3GalT6 sur l'expression de certains ARNm et voir si elle a une implication sur le métabolisme cellulaire. Par exemple, il a pu être observé dans les fibroblastes de patients atteints de SEMD-JL1, une surexpression des gènes *CSGALNACT1* et *CHPF*, deux gènes codant des protéines impliquées dans l'initiation et dans l'élongation de la synthèse des CS respectivement (Voir introduction partie II.3.2) (Ritelli et al., 2015). La surexpression de ces protéines peut potentiellement favoriser la formation de chaînes de CS à partir de l'amorce tétrasaccharidique. Il est donc possible de comprendre que la cellule puisse orienter la synthèse des GAGs vers celle des CS au dépend des chaînes de HS. Si ce phénomène est également observé chez les patients, la surexpression de ces enzymes permettrait la glycanation de la décorine et une diminution de la synthèse des chaînes de HS en parallèle. En favorisant la synthèse de décorine glycosylée, la fibrillogenèse du collagène pourrait potentiellement être améliorée. Cependant, l'organisation des fibres de collagène n'est



visiblement pas correctement réalisée chez certains patients, comme l'ont montré des analyses en microscopie électronique (voir Introduction partie III.4.3). Ainsi, il doit y avoir un autre mécanisme qui impacte la fibrillogenèse du collagène à ce niveau. Ritelli et al. (2015) ont également montré que les gènes *COL5A1* et *COL15A1* étaient moins exprimés chez des jumelles atteintes de SEMD-JL1. Si les patients, décrits par Malfait et al. (2013), présentent le même profil, il est possible de comprendre la sévérité des symptômes observés. En effet, le gène *COL5A1*, dont les mutations conduisent à une forme classique de SED (Voir Introduction partie I.2.), code un collagène impliqué dans la fibrillogenèse du collagène I. Chez ces patients, la synthèse du collagène V est diminuée et impacte ainsi la fibrillogenèse du collagène de type I (majoritaire dans la peau). Ainsi, si les patients atteints de SEDsp présentent une diminution de l'expression du collagène V, il est compréhensible d'observer une hypermobilité articulaire et une hyperextensibilité de la peau chez ces patients dont le réseau de collagène n'est pas suffisamment structuré. De façon complémentaire, des mutations de ce même gène, qui code le collagène XV, une protéine matricielle exprimée par les ostéoblastes durant la formation des nouveaux tissus osseux ont également été décrites (Lisignoli et al., 2009). Ainsi, il est possible de supposer que la diminution du *COL15A1* pourrait avoir un impact sur le remodelage osseux et donc être en lien avec la dysplasie squelettique observée chez les patients atteints de SED. Ainsi, chez les patients pour lesquels la synthèse des GAGs est fortement impactée, l'expression de certains gènes pourrait être modifiée (comme celle des gènes *COL15A1* et *COL5A1*) et donc la fibrillogenèse des collagènes pourrait être moins importante. Il serait alors possible d'expliquer la sévérité des symptômes observés par rapport à la mutation du gène identifié. D'où l'intérêt de notre modèle pour une étude transcriptomique. Par ailleurs, pour compléter notre étude, il serait intéressant de regarder l'impact de la déficience sur l'expression des protéines matricielles et sur leur organisation dans la matrice. Notre modèle ne nous donne pas d'information sur de potentiels effets sur la matrice extracellulaire et notamment de l'impact du défaut de synthèse des GAGs sur la fibrillogenèse du collagène (Furth et al., 1991). D'où l'intérêt d'avoir en complément de notre modèle, un second modèle cellulaire qui serait plus favorable à ce type d'études, comme la lignée ATDC5. Cette lignée murine est principalement utilisée pour l'étude de la chondrogenèse et présente la capacité de synthétiser des collagènes et des PGs (principalement des SLRPs, famille à laquelle appartient la décorine) (Wilhelm et al., 2020; Yao and Wang, 2013). Ainsi, la déficience en  $\beta$ 3GalT6 dans cette lignée permettrait de comprendre l'implication du défaut de synthèse de GAGs sur la désorganisation matricielle et son impact sur l'organisation des collagènes. Il peut également être envisagé le développement d'un modèle animal afin d'avoir une meilleure compréhension de l'implication



de cette déficience dans le mécanisme de la pathologie des SEDsp. C'est ce qui a déjà pu être réalisé pour quelques GTs. Dans le cas de la  $\beta$ 4GalT7, pour laquelle un modèle animal a pu être établi chez le poisson zèbre, il a pu être observé des anomalies morphologiques proches de ce qui est décrit chez les patients atteints de SEDsp-*B4GALT7* (Delbaere et al., 2019b). Dans le cas de souris déficientes en XylT-I, il a été montré qu'elles présentaient un défaut de développement du squelette qui s'aggravait avec l'âge et un nanisme qui apparaissait dès le stade embryonnaire (Mis et al., 2014). Ce modèle présente les mêmes anomalies que les patients atteints de DBSD2 (Bui et al., 2010; Jamsheer et al., 2016). Par ailleurs, la réalisation d'un modèle murin KO en GlcAT-1 a permis d'observer une importante diminution de la biosynthèse des chaînes de CS et HS. L'absence de ces chaînes de GAGs induit l'échec de la cytokinèse et donc la létalité de l'embryon vers le stade 8 cellules (Izumikawa et al., 2010). Ainsi, ces modèles cellulaires et d'animaux ont permis de mieux comprendre l'implication de ces GTs dans les pathologies qui ont pu être retrouvées chez des patients atteints de SED, mais également de mieux comprendre leurs implications dans le développement embryonnaire. Actuellement, aucun modèle animal KO pour la  $\beta$ 3GalT6 n'a été publié, mais ce travail est démarré dans notre équipe.

Au final, nous disposons actuellement de lignées *B3GALT6* déficientes caractérisées par une absence d'activité galactosyltransférase et par une synthèse de GAGs très réduite. Ce modèle, ne présentant plus d'activité  $\beta$ 3GalT6 endogène, est pertinent pour évaluer l'impact individuel des mutations de la  $\beta$ 3GalT6 observées chez les patients après réexpression transitoire de l'enzyme sauvage ou des mutants. Les résultats préliminaires obtenus semblent indiquer que la réexpression de l'enzyme sauvage dans les cellules déficientes restaure partiellement la glycanation de la protéine *core* de la décorine alors que la réexpression des mutants Asp207His et Gly217Ser ne le permet pas. Nous avons pu discuter précédemment de l'implication de ces mutations sur la structure de la protéine. Il est intéressant de noter que la mutation Gly217Ser semble conduire à une importante réduction de l'activité de la protéine. Ainsi, la présence d'une sérine en position 217 semble être incompatible avec l'activité de l'enzyme *in cellulo*, ce qui explique la non-glycanation de la protéine *core* de la décorine dans nos conditions. Par contre, ces résultats diffèrent de ceux observés chez le patient 5 (patient porteur de la mutation homozygote p.Gly217Ser, Malfait et al, 2013). Chez ce patient, la glycanation de la décorine est en effet encore détectée, mais en quantité moindre à celle trouvée dans les cellules contrôle (Malfait et al., 2013). Pour la seconde mutation (p.Asp207His), observée chez le patient 2 (mutation homozygote) et 3 (mutation hétérozygote avec la mutation p.Ala108Glyfs\*163), la décorine ne présente qu'une très faible glycanation dans les fibroblastes de peau de ces patients.

**Tableau 15 : Résumé des caractéristiques des patients atteints de SEDsp  
B3GALT6 et des études des mutations de la  $\beta$ 3GalT6 dans la lignée  
déficiente (clone H7)**

<b>Patients / lignée</b>	<b>P2</b>	<b>P3</b>	<b>P5</b>
<b>Mutation génomique</b>	C.619G>C C.619G>C	C.619G>C c.323_344del	c.649G>A c.649G>A
<b>Mutation protéique</b>	p.Asp207His p.Asp207His	p.Asp207His p.Ala108Glyfs*163	p.Gly217Ser p.Gly217Ser
<b>Sévérité du SED</b>	<b>Peu affecté</b> +	<b>Très sévèrement affecté</b> +++	<b>Sévèrement affecté</b> ++
<b>Activité résiduelle <i>in vitro</i> de la <math>\beta</math>3GalT6 <sup>1</sup></b>	ND	ND	ND
<b>Taux d'anabolisme résiduel des GAGs <sup>1</sup></b>	60%	16%	33%
<b>Expression des HS <sup>1,2</sup></b>	↘	↘↘↘	↘
<b>Glycanation de la décorine <sup>1</sup></b>	ND	↘↘↘	↘
<b>Expression de la protéine <sup>1,2</sup></b>	↘↘	↘	↘↘
<b>Données obtenues après surexpression des mutants dans la lignée HeLa B3GALT6 déficiente</b>			
	<b>p.Asp207His</b>		<b>p.Gly217Ser</b>
<b>Activité résiduelle <i>in vitro</i> de la <math>\beta</math>3GalT6</b>	~15%		~5%
<b>Glycanation de la décorine</b>	Absente		Absente

<sup>1</sup> Données des patients P2, P3 et P5 issus de Malfait et al. (2013)

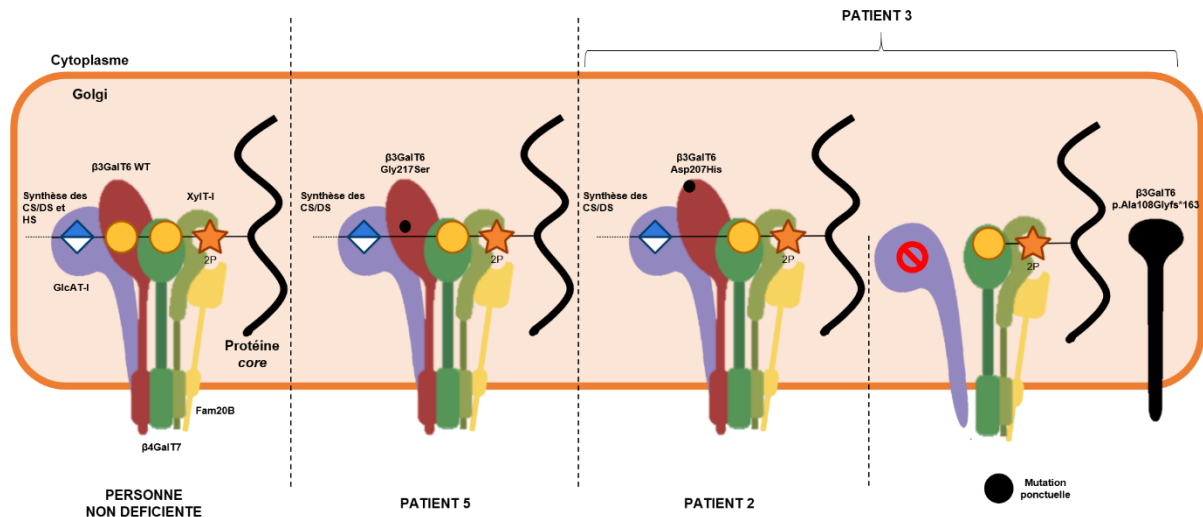
<sup>2</sup> Observée par des analyse d'immunofluorescence de fibroblastes de peau de patients  
ND : non déterminé



Les résultats obtenus pour la mutation ponctuelle donnent des résultats proches, avec une absence de glycanation de la décorine dans les cellules HeLa. Nous pouvons donc en conclure que ce résidu est important pour l'activité galactosyltransférase de l'enzyme.

Bien sûr, ces résultats doivent encore être confirmés, mais tendent à montrer que les protéines mutées ne seraient plus capables d'initier la biosynthèse des GAGs sur la protéine *core* des PGs.

Actuellement, il est encore difficile de comprendre pourquoi et comment certaines mutations altèrent la biosynthèse des GAGs, en particulier en l'absence d'une activité enzymatique de la  $\beta$ GalT6 détectable. Il est possible de trouver une explication en se basant sur quelques hypothèses établies sur la biosynthèse des GAGs. Il est actuellement proposé que les enzymes impliquées dans la biosynthèse des GAGs interagiraient les unes avec les autres afin de former un complexe enzymatique nommé GAGosome (Esko and Selleck, 2002). La capacité de ces enzymes à s'associer les unes aux autres, la stœchiométrie et la présence d'isoformes de certaines enzymes pourraient réguler la fonction du GAGosome et ceci de façon dynamique dans l'espace et dans le temps. Actuellement, certaines études ont permis de montrer des interactions entre des GTs impliquées dans la biosynthèse des GAGs, ce qui renforce cette hypothèse. Par exemple, il a été décrit que les enzymes impliquées dans la biosynthèse des Hep/HS, les GTs EXT1 et EXT2, présentent une activité enzymatique sous forme de complexe, mais ne sont pas actives si elles sont seules (Kobayashi et al., 2000; McCormick et al., 2000; Senay et al., 2000). Leur interaction est indispensable à la biosynthèse des HS. Une interaction entre l'enzyme EXT2 (impliquée dans l'élongation des HS) et la NDST1 (responsable de la première étape de maturation des HS) a également été décrite dans le contexte de la synthèse des HS (Presto et al., 2008). Dans ce cas, il a été montré une augmentation de la N-déacétylation et de la N-sulfatation des chaînes de HS au niveau des résidus N-acétylglucosamine. De la même façon, une interaction entre la C5-épimérase et la 2-OST, des enzymes de maturation des chaînes de HS (Pinhal et al., 2001) a été décrite et conduit à une augmentation de la sulfatation de la position 2 des résidus IdoA après l'action de la C5-épimérase. Toutes ces études vont dans le sens d'une association des enzymes de biosynthèse et/ou de maturation des GAGs pour optimiser l'enchaînement des différentes étapes de biosynthèse et le cas échéant, réguler et/ou orienter cette biosynthèse. Plus récemment, des interactions entre les enzymes impliquées dans les premières étapes de la biosynthèse des GAGs ont pu être observées. Ainsi, l'interaction entre la GlcAT-1 et la Xylose phosphatase (XylP) a été mise en évidence (Koike et al., 2014). Il a été montré que la formation de ce complexe est importante pour une meilleure activité de la phosphatase et une bonne localisation cellulaire des enzymes. Ainsi, la formation de ce



**Figure 36 – Impact potentiel des mutations de la  $\beta 3\text{GalT}6$  présentes chez les patients sur le GAGosome et la synthèse des GAGs.**

La notion de GAGosome propose que les enzymes impliquées dans la biosynthèse des GAGs interagiraient les unes avec les autres afin de former un complexe enzymatique nommé GAGosome (Esko and Selleck, 2002). Si l'existence d'un tel complexe est vérifiée au niveau des personnes non porteuses de la déficience génétique, les enzymes impliquées dans la formation de l'amorce tétrasaccharidique interagiraient entre elles pour permettre la synthèse des GAGs. Dans le cas des patients porteurs de mutation(s) au niveau du gène *B3GALT6*, il serait possible d'expliquer l'effet des mutations sur la synthèse des GAGs par une désorganisation du GAGosome, perturbant ainsi la biosynthèse des GAGs. Sur la base des résultats obtenus avec les fibroblastes de peau de patients atteints de SEDsp (Malfait et al., 2013) et les données issues de notre modèle cellulaire déficient, il est possible d'émettre l'hypothèse que les mutations Asp207His et Gly217Ser pourraient avoir un impact sur l'activité de la protéine, mais ne joueraient pas de rôle sur la stabilité du complexe GAGosome. Ainsi, la synthèse des GAGs serait maintenue même si elle est plus faible (Patient 2 et 5). Par contre, la mutation p.Ala108Glyfs\*163, observée chez le patient 3, ne doit pas permettre de conserver l'intégrité du GAGosome. Dans ce cas-là, il est possible que la synthèse des GAGs soit totalement abolie en l'absence de formation de ce complexe multi protéique.

complexe semble indispensable pour la maturation correcte de l'amorce tétrasaccharidique. Si nous partons du principe de l'existence du GAGosome, nous pouvons penser que la  $\beta$ 3GalT6 pourrait de la même façon interagir avec les enzymes impliquées dans l'initiation (ou l'élongation) de la biosynthèse des GAGs. Ainsi, la présence de mutations dans la protéine pourrait impacter la structure ou l'intégrité de ce complexe et donc avoir des conséquences importantes sur la biosynthèse des GAGs.

Dans le cas de la mutation Gly217Ser, qui ne semble pas induire de changement structural d'après le modèle décrit précédemment (voir partie I.5 et Figure 9), la mutation de la protéine impacte fortement l'activité enzymatique, mais de façon moindre la glycanation de la décorine chez le patient 5 (Tableau 15 - Malfait et al., 2013). Comme la mutation ne semble pas impacter la structure de l'enzyme, il peut être supposé que la protéine, même si elle est plus faiblement exprimée que dans les cellules contrôle comme cela a été montré (Malfait et al., 2013), permette le maintien de l'intégrité d'un éventuel complexe de biosynthèse. Avec le maintien de l'intégrité du GAGosome, le complexe pourrait conduire à la biosynthèse des GAGs à partir de la voie alternative utilisant une amorce trisaccharidique, mais avec une plus faible efficacité (Persson et al., 2019). Cette hypothèse pourrait expliquer la synthèse de GAGs encore observée chez ce patient (Figure 36).

Dans le cas de la mutation Asp207His, présente chez les patients 2 et 3, la protéine mutée ne présente qu'une très faible d'activité enzymatique (environ 15%) et une glycanation de la décorine absente en présence de la protéine mutée (ou non détectable dans nos conditions d'analyse) si nous nous basons sur les résultats obtenus avec notre modèle cellulaire déficient. D'après ces données et le modèle présenté dans la figure 36, il est probable que la mutation ait un impact sur l'activité de l'enzyme mais ne l'abolisse pas totalement. Ainsi, la structure du GAGosome pourrait être maintenue. Si nous comparons les données de notre modèle cellulaire aux résultats obtenus avec les cellules de patient, il est intéressant de voir que le patient 2 (homozygote pour la mutation Asp207His) présente une diminution de la synthèse globale des GAGs modérée et synthétise encore des HS (aucune donnée n'est disponible pour les CS – Tableau 15). Ces résultats montrent une discordance avec ce que nous avons pu observer dans notre modèle cellulaire HeLa *B3GALT6* déficient pour lequel aucune glycanation de la décorine n'est observée après réexpression du mutant Asp207His. Le fait que nous ne retrouvons pas de glycanation de la décorine dans les cellules *B3GALT6* déficientes pourrait venir d'un seuil de détection trop faible de la technique utilisée (Western blot), ou peut-être d'une absence d'inclusion de la protéine réexprimée dans la formation du complexe GAGosome dans ces cellules-hôtes. Avec l'absence de la formation du GAGosome, les protéines recombinantes



produites ne pourraient pas synthétiser correctement des GAGs et ne permettraient pas non plus de mettre en place une voie de biosynthèse normale ou alternative de ces chaînes (Figure 36). En plus du maintien d'une synthèse des GAGs, le patient porteur de la mutation Asp207His semble moins affecté par la mutation de la protéine d'un point de vue clinique. Ainsi, la mutation Asp207His semble permettre de maintenir une synthèse de GAGs résiduelle. Il est possible que la mutation n'affecte pas l'intégrité du GAGosome. Donc, comme pour la mutation Gly217Ser, la biosynthèse des GAGs pourrait se faire à partir de la voie alternative incluant la formation d'une amorce trisaccharidique pour permettre l'initiation de la biosynthèse (Figure 36). Par ailleurs, la présence d'une activité enzymatique résiduelle de la protéine dans les cellules HeLa *B3GALT6* déficientes peut nous laisser supposer que la protéine présente également une activité réduite dans les fibroblastes de patient. Ainsi, il est possible que cette faible activité enzymatique puisse permettre un certain taux de synthèse normale des GAGs. Alors, le maintien de l'intégrité du GAGosome et la présence d'une synthèse de GAGs résiduelle peut expliquer des symptômes moins sévères chez le patient 2.

Par contre, le patient 3 (mutation hétérozygote avec la mutation p.Ala108Glyfs\*163 et p.Asp207His) est très sévèrement atteint et présente des symptômes cliniques beaucoup plus sévères. Il présente par ailleurs une forte diminution de la synthèse des HS et des CS, ainsi qu'une absence de glycanation de la protéine *core* de la décorine (Tableau 15). Comme nous avons pu le voir précédemment, la mutation ponctuelle Asp207His que présente ce patient permettrait le maintien de la structure du GAGosome et pourrait permettre une synthèse partielle des GAGs. Si la protéine mutée Asp207His était exprimée de façon majoritaire chez ce patient, il pourrait présenter des symptômes moins sévères, comme ceux observés chez le patient 2. Il est donc probable que la sévérité de la maladie vienne, soit de la délétion, soit de la présence des deux mutations. Il est intéressant de noter que l'expression de la  $\beta$ 3GalT6 mutée est plus importante chez ce patient que chez les autres patients, même si elle est diminuée par rapport à celle des cellules contrôle (Tableau 15 – Malfait et al., 2013). Ainsi, il est intéressant de noter qu'une forme de la protéine semble plus exprimée au niveau cellulaire. Etant donné que dans le cas du patient 2, la  $\beta$ 3GalT6 mutée Asp207His est faiblement exprimée et que la perte de la synthèse des GAGs est importante, la protéine tronquée (p.Ala108Glyfs\*163) doit être majoritairement exprimée. Il serait intéressant d'étudier cette délétion dans notre modèle cellulaire *B3GALT6* déficient, afin de voir si cette dernière permet réellement d'obtenir une protéine stable dans les cellules. La présence de la protéine tronquée (p.Ala108Glyfs\*163) dans les cellules peut potentiellement induire une altération de la formation du GAGosome qui conduirait à une perte importante de la synthèse des GAGs car cette synthèse ne se ferait plus



ni par la voie classique, ni par la voie alternative. La synthèse de la protéine Asp207His, qui serait minoritaire, permettrait de maintenir une synthèse résiduelle, qu'il est possible d'observer chez le patient.

Dans ces études, notre modèle nous a permis, en le confrontant avec les données obtenues dans les fibroblastes de derme de patients, d'évaluer les conséquences des mutations sur la biosynthèse des GAGs. Il serait intéressant d'essayer de voir si nous retrouvons également des données concordantes en étudiant par immunofluorescence (i) l'expression des GAGs dans nos modèles déficients (suite à la surexpression l'une ou l'autre des protéines mutées) et (ii) la localisation des protéines mutées et (iii) l'expression des chaînes de HS. Si nous retrouvons des résultats concordants, ceci nous confortera dans les hypothèses émises sur l'impact des mutations sur la  $\beta$ 3GalT6, le GAGosome et la biosynthèse des GAGs. Il serait également intéressant de voir ce qu'il se passe dans le cas de l'expression des protéines portant les deux mutations présentes chez le patient 3 dans nos cellules déficientes, afin de comprendre quel est l'impact de ces mutations sur les fonctions de la cellule et mieux prédire la sévérité des symptômes cliniques. Bien sûr toutes les hypothèses développées précédemment nécessitent d'être vérifiées et confirmées par des travaux complémentaires pour mettre en évidence l'existence et la structuration éventuelle du GAGosome par des expériences d'immunoprécipitation par exemple. Dans ce contexte, notre modèle cellulaire pourrait être utilisé pour montrer l'importance des interactions potentielles entre la  $\beta$ 3GalT6 et les autres GTs impliquées dans l'initiation et/ou l'élongation de la biosynthèse des GAGs et les conséquences de ces pertes d'interaction lorsque la protéine mutée est moins ou non exprimée. La mise en évidence de l'existence de ce macro-complexe permettrait de mieux comprendre la synthèse des GAGs, ainsi que l'impact des mutations de la  $\beta$ 3GalT6 sur la synthèse des chaînes de HS et CS.





# CONCLUSIONS ET PERSPECTIVES



La première partie de ce travail de thèse a été consacrée à l'étude fonctionnelle et cinétique de la  $\beta$ 3GalT6 recombinante humaine sauvage produite en système bactérien sous forme tronquée soluble. La majeure partie de cet axe de travail a été dédiée à la mise en place des conditions de production de la protéine recombinante. Nous avons déterminé les différents paramètres pouvant influencer la solubilité de la  $\beta$ 3GalT6, tels que la troncation de la protéine, le vecteur d'expression et le partenaire de fusion (MBP ou GST), la souche bactérienne ou encore les conditions de production de la protéine (température, concentration en inducteur). Ces différents essais ont permis de déterminer des conditions permettant d'obtenir la h $\beta$ 3GalT6 recombinante sous forme soluble. En parallèle, nous avons mis en place une collaboration avec le CERMAV (Grenoble) pour réaliser la production de la protéine en cellules d'insectes en utilisant le système baculovirus. Les vecteurs viraux recombinants ont été préparés et transmis à nos collègues, mais les essais réalisés n'ayant pas permis d'aboutir à la production de notre protéine, cette stratégie de production sera reconduite ultérieurement si ce système devait être envisagé à nouveau pour produire la  $\beta$ 3GalT6 recombinante humaine sous forme tronquée soluble. Finalement, nous avons, avec le système bactérien, pu mettre en évidence une condition de production adaptée pour obtenir notre protéine d'intérêt sous forme soluble. Après avoir analysé l'impact des troncations de la partie N-terminale de la protéine, en particulier de la région *stem* sur la solubilité, nous avons choisi la forme délétée des 30 premiers acides aminés N-terminaux (h $\beta$ 3GalT6 $\Delta$ NTer30) que nous avons produite et purifiée. Les premiers résultats ont montré une purification correcte de la protéine qui semble néanmoins présente sous deux formes qui n'ont pas pu être différenciées par spectrométrie de masse. Ce point devra être investigué par la suite pour savoir s'il s'agit d'une forme glycosylée, d'une autre modification post-traductionnelle de la protéine produite, ou du clivage de l'étiquette histidine. Nous avons ensuite réalisé des études cinétiques, qui ont montré que la protéine présentait une bonne affinité pour un analogue de substrat synthétique, le Gal-Xyl(2P)-O-MN. Par rapport à d'autres GTs, dont la  $\beta$ 4GalT7, la  $\beta$ 3GalT6 présente un  $K_M$  plus faible d'environ 30  $\mu$ M vis-à-vis de son substrat accepteur. Par contre, la  $\beta$ 3GalT6 semble être une enzyme peu véloce avec une constante catalytique  $k_{cat}$  faible de 0,05  $\text{min}^{-1}$ . Ces résultats sont préliminaires et nécessitent une étude plus approfondie car une agrégation et une instabilité des protéines ont pu être observées, ce qui peut être lié à nos conditions de production, de lyse, de purification ou encore de troncation de notre protéine.

Néanmoins, ce travail a permis de produire *in vitro* notre enzyme d'intérêt dans sa version recombinante humaine et de déterminer ses paramètres cinétiques. Afin de compléter ces travaux, l'impact des différentes troncations de la h $\beta$ 3GalT6 sur sa solubilité devra être évaluée



en s'appuyant désormais sur son activité enzymatique. Cette étude permettra ensuite de déterminer quelle forme de la protéine sera utilisée pour étudier les différentes mutations de la  $\beta$ 3GalT6 humaine décrites chez les patients atteints de SEDsp. Il sera alors possible de déterminer l'impact de ces mutations sur la fonctionnalité et les paramètres cinétiques de l'enzyme. En parallèle de l'étude des troncations, l'étude de la solubilité et de la stabilité de la protéine purifiée devra être poursuivie afin d'envisager des études structurales par cristallographie de la protéine. L'obtention de la structure tridimensionnelle de la protéine sauvage en complexe (ou non) avec son (ses) substrat(s), nous permettra dans ce contexte de disposer de plus d'informations sur l'impact des mutations de la protéine sur sa structure.

Nous avons consacré la deuxième partie de ce travail à l'établissement d'un modèle cellulaire déficient en  $\beta$ 3GalT6. Ce modèle a pour but d'étudier les conséquences de l'extinction du gène *B3GALT6* dans des cellules HeLa suite à une invalidation du gène par la technologie CRISPR/Cas9. Pour ce faire, nous avons fait appel au service d'une plateforme (TACGENE) qui a construit des vecteurs permettant la réalisation des lignées déficientes. La transfection des cellules avec le meilleur de ces vecteurs a permis, après enrichissement et sélection clonale, d'obtenir 12 clones. L'analyse de l'expression des ARNm de ces clones a été menée afin de discriminer ceux qui présentaient une expression réduite. Nous avons ainsi pu isoler deux clones qui, après séquençage génomique, ont montré des délétions au niveau du gène *B3GALT6* (c.222del et c.182\_233del pour les clones E2 et H7, respectivement) qui conduisent à un décalage du cadre de lecture et à la formation d'une protéine aberrante. Ces deux clones ne présentent plus d'activité galactosyltransférase, mais restent cependant capables de synthétiser des GAGs à partir d'un xyloside exogène (4-MOX), de façon réduite par rapport à des cellules contrôles non déficientes. Ces résultats sont en accord avec ce qui a pu être obtenu dans d'autres lignées (Chen et al., 2018) et laissent à penser qu'il existe une voie alternative de biosynthèse des GAGs qui a été proposée très récemment avec la description d'une amorce trisaccharidique au niveau de la zone d'ancrage protéine *core*-GAG de certains PGs (Persson et al., 2019).

Une fois la caractérisation de nos lignées déficientes réalisée, avec l'aide de notre équipe, l'étude de l'effet de la restauration de l'expression de la  $\beta$ 3GalT6 sauvage et des mutants de la protéine présents chez les patients atteints de SEDsp (Asp207His et Gly217Ser) a été initiée. Nous avons pu voir que les protéines mutantes ne sont plus capables de catalyser le transfert du galactose que ce soit sur l'analogue du substrat accepteur *in vitro* ou sur la protéine *core* de la décorine, *in cellulo*. Ces résultats sont proches de ceux qui ont pu être observés dans les fibroblastes de certains patients présentant ces mutations.



Pour résumer, nous avons généré un modèle cellulaire que nous pouvons utiliser pour l'étude des mutations de la  $\beta$ 3GalT6 sur la biosynthèse des GAGs *in cellulo*. Nous avons démarré l'étude des premières mutations décrites par Malfait et al. en 2013 (Asp207His et Gly217Ser). Les vecteurs contenant les séquences des autres mutations ont été préparés, mais n'ont pas encore pu être utilisés dans ce contexte.

Afin de vérifier l'hypothèse selon laquelle la biosynthèse des GAGs pourrait être amorcée sur une zone d'ancrage plus courte, nous avons établi une collaboration avec le Dr Frederik Noborn et le Dr Göran Larson (*Institute of Biomedicine, Sahlgrenska Academy, Université de Gothenburg*). Cette étude va permettre d'analyser le type d'amorce présente au niveau des GAGs encore synthétisés par nos deux lignées déficientes. En parallèle, des échantillons préparés à partir de cellules de patients ont été envoyés à nos collègues afin de voir si les GAGs synthétisés par les fibroblastes porteurs des mutations de la  $\beta$ 3GalT6 présentent une amorce tétrasaccharidique normale ou une amorce plus courte. Ces résultats pourraient nous permettre de comprendre pourquoi certains fibroblastes de patients restent capables de réaliser une synthèse de GAGs sur les PGs malgré une activité de la  $\beta$ 3GalT6 absente ou très fortement réduite.

Chez certains patients, la diminution de la synthèse des GAGs est accompagnée d'une altération de l'organisation du réseau de collagène (en termes de composition et d'organisation matricielle). Ces défauts sont à l'origine de la perte de fonctionnalité globale des tissus conjonctifs observée chez les patients atteints de SEDsp. Pour envisager de telles études, il nous faudra avoir recours à une autre lignée cellulaire plus adaptée pour l'étude des collagènes et de la matrice, la lignée ATDC5 par exemple, lignée produisant une matrice abondante et dont la synthèse et la maturation des collagènes peuvent être analysées dans notre institut.

Pour conclure, l'ensemble des résultats obtenus nous a permis de mieux comprendre l'implication de la  $\beta$ 3GalT6 dans le processus physiopathologique du SEDsp. Nous aimerions par ailleurs comprendre la relation existant entre (i) la nature des mutations du gène *B3GALT6*, (ii) l'altération de la structure et de la fonction de la MEC (en lien avec les modifications du métabolisme des GAGs (et des PGs qui les portent) et des collagènes) et (iii) la sévérité des symptômes cliniques observés chez les patients. A terme, ces travaux contribueront à la prédiction de l'évolution de la maladie et pourront conduire à une meilleure prise en charge clinique des patients.





# BIBLIOGRAPHIE



- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). *Molecular Biology of the Cell* (Garland Science).
- Albesa-Jové, D., Giganti, D., Jackson, M., Alzari, P.M., and Guerin, M.E. (2014). Structure–function relationships of membrane-associated GT-B glycosyltransferases. *Glycobiology* *24*, 108–124.
- Almeida, R., Lavery, S.B., Mandel, U., Kresse, H., Schwientek, T., Bennett, E.P., and Clausen, H. (1999). Cloning and expression of a proteoglycan UDP-galactose:beta-xylose beta1,4-galactosyltransferase I. A seventh member of the human beta4-galactosyltransferase gene family. *J. Biol. Chem.* *274*, 26165–26171.
- Amado, M., Almeida, R., Carneiro, F., Lavery, S.B., Holmes, E.H., Nomoto, M., Hollingsworth, M.A., Hassan, H., Schwientek, T., Nielsen, P.A., et al. (1998). A family of human beta3-galactosyltransferases. Characterization of four members of a UDP-galactose:beta-N-acetyl-glucosamine/beta-nacetyl-galactosamine beta-1,3-galactosyltransferase family. *J. Biol. Chem.* *273*, 12770–12778.
- Amado, M., Almeida, R., Schwientek, T., and Clausen, H. (1999). Identification and characterization of large galactosyltransferase gene families: galactosyltransferases for all functions. *Biochim. Biophys. Acta* *1473*, 35–53.
- Aspberg, A. (2012). The Different Roles of Aggrecan Interaction Domains. *J. Histochem. Cytochem.* *60*, 987–996.
- Baasanjav, S., Al-Gazali, L., Hashiguchi, T., Mizumoto, S., Fischer, B., Horn, D., Seelow, D., Ali, B.R., Aziz, S.A.A., Langer, R., et al. (2011). Faulty Initiation of Proteoglycan Synthesis Causes Cardiac and Joint Defects. *Am. J. Hum. Genet.* *89*, 15–27.
- Bachy, S., Letourneur, F., and Rousselle, P. (2008). Syndecan-1 interaction with the LG4/5 domain in laminin-332 is essential for keratinocyte migration. *J. Cell. Physiol.* *214*, 238–249.
- Bai, X., Zhou, D., Brown, J.R., Crawford, B.E., Hennes, T., and Esko, J.D. (2001). Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth member of the beta 1,3-galactosyltransferase family (beta 3GalT6). *J. Biol. Chem.* *276*, 48189–48195.
- Bame, K.J., Zhang, L., David, G., and Esko, J.D. (1994). Sulphated and undersulphated heparan sulphate proteoglycans in a Chinese hamster ovary cell mutant defective in N-sulphotransferase. *Biochem. J.* *303* ( Pt 1), 81–87.
- Beighton, P., Gericke, G., Kozlowski, K., and Grobler, L. (1984). The manifestations and natural history of spondylo-epi-metaphyseal dysplasia with joint laxity. *Clin. Genet.* *26*, 308–317.
- Beighton, P., De Paepe, A., Steinmann, B., Tsipouras, P., and Wenstrup, R.J. (1998). Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). *Am. J. Med. Genet.* *77*, 31–37.
- Ben-Mahmoud, A., Ben-Salem, S., Al-Sorkhy, M., John, A., Ali, B.R., and Al-Gazali, L. (2018). A B3GALT6 variant in patient originally described as Al-Gazali syndrome and

implicating the endoplasmic reticulum quality control in the mechanism of some  $\beta$ 3GalT6-pathway mutations. *Clin. Genet.* *93*, 1148–1158.

Bocian, C., Urbanowitz, A.-K., Owens, R.T., Iozzo, R.V., Götte, M., and Seidler, D.G. (2013). Decorin potentiates interferon- $\gamma$  activity in a model of allergic inflammation. *J. Biol. Chem.* *288*, 12699–12711.

Boeggeman, E.E., Ramakrishnan, B., and Qasba, P.K. (2003). The N-terminal stem region of bovine and human  $\beta$ 1,4-galactosyltransferase I increases the in vitro folding efficiency of their catalytic domain from inclusion bodies. *Protein Expr. Purif.* *30*, 219–229.

Brady, A.F., Demirdas, S., Fournel-Gigleux, S., Ghali, N., Giunta, C., Kapferer-Seebacher, I., Kosho, T., Mendoza-Londono, R., Pope, M.F., Rohrbach, M., et al. (2017). The Ehlers–Danlos syndromes, rare types. *Am. J. Med. Genet. C Semin. Med. Genet.* *175*, 70–115.

Breton, C., Fournel-Gigleux, S., and Palcic, M.M. (2012). Recent structures, evolution and mechanisms of glycosyltransferases. *Curr. Opin. Struct. Biol.* *22*, 540–549.

Brown, E.L., Wooten, R.M., Johnson, B.J., Iozzo, R.V., Smith, A., Dolan, M.C., Guo, B.P., Weis, J.J., and Höök, M. (2001). Resistance to Lyme disease in decorin-deficient mice. *J. Clin. Invest.* *107*, 845–852.

Budde, B.S., Mizumoto, S., Kogawa, R., Becker, C., Altmüller, J., Thiele, H., Rüschenhoff, F., Toliat, M.R., Kaleschke, G., Hämmerle, J.M., et al. (2015). Skeletal dysplasia in a consanguineous clan from the island of Nias/Indonesia is caused by a novel mutation in B3GAT3. *Hum. Genet.* *134*, 691–704.

Bui, C., Talhaoui, I., Chabel, M., Mulliert, G., Coughtrie, M.W.H., Ouzzine, M., and Fournel-Gigleux, S. (2010). Molecular characterization of  $\beta$ 1,4-galactosyltransferase 7 genetic mutations linked to the progeroid form of Ehlers-Danlos syndrome (EDS). *FEBS Lett.* *584*, 3962–3968.

Bui, C., Huber, C., Tuysuz, B., Alanay, Y., Bole-Feysot, C., Leroy, J.G., Mortier, G., Nitschke, P., Munnich, A., and Cormier-Daire, V. (2014). XYLT1 Mutations in Desbuquois Dysplasia Type 2. *Am. J. Hum. Genet.* *94*, 405–414.

Burch, G.H., Gong, Y., Liu, W., Dettman, R.W., Curry, C.J., Smith, L., Miller, W.L., and Bristow, J. (1997). Tenascin-X deficiency is associated with Ehlers–Danlos syndrome. *Nat. Genet.* *17*, 104–108.

Burkitt Wright, E.M.M., Spencer, H.L., Daly, S.B., Manson, F.D.C., Zeef, L.A.H., Urquhart, J., Zoppi, N., Bonshek, R., Tosounidis, I., Mohan, M., et al. (2011). Mutations in PRDM5 in Brittle Cornea Syndrome Identify a Pathway Regulating Extracellular Matrix Development and Maintenance. *Am. J. Hum. Genet.* *88*, 767–777.

Busse, M., Feta, A., Presto, J., Wilén, M., Grønning, M., Kjellén, L., and Kusche-Gullberg, M. (2007). Contribution of EXT1, EXT2, and EXTL3 to Heparan Sulfate Chain Elongation. *J. Biol. Chem.* *282*, 32802–32810.

Byers, P.H. (1993). Vascular Ehlers-Danlos Syndrome. In *GeneReviews®*, M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J. Bean, K. Stephens, and A. Amemiya, eds. (Seattle (WA): University of Washington, Seattle), p.

- Byers, P.H., and Murray, M.L. (2014). Ehlers–Danlos syndrome: A showcase of conditions that lead to understanding matrix biology. *Matrix Biol.* *33*, 10–15.
- Carey, D.J., Evans, D.M., Stahl, R.C., Asundi, V.K., Conner, K.J., Garbes, P., and Cizmeci-Smith, G. (1992). Molecular cloning and characterization of N-syndecan, a novel transmembrane heparan sulfate proteoglycan. *J. Cell Biol.* *117*, 191–201.
- Cartault, F., Munier, P., Jacquemont, M.-L., Vellayoudom, J., Doray, B., Payet, C., Randrianaivo, H., Laville, J.-M., Munnich, A., and Cormier-Daire, V. (2015). Expanding the clinical spectrum of B4GALT7 deficiency: homozygous p.R270C mutation with founder effect causes Larsen of Reunion Island syndrome. *Eur. J. Hum. Genet.* *23*, 49–53.
- Castori, M. (2012). Ehlers-Danlos Syndrome, Hypermobility Type: An Underdiagnosed Hereditary Connective Tissue Disorder with Mucocutaneous, Articular, and Systemic Manifestations. *ISRN Dermatol.* *2012*.
- Chambers, A.C., Aksular, M., Graves, L.P., Irons, S.L., Possee, R.D., and King, L.A. (2018). Overview of the Baculovirus Expression System. *Curr. Protoc. Protein Sci.* *91*, 5.4.1-5.4.6.
- Charnock, S.J., and Davies, G.J. (1999). Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. *Biochemistry* *38*, 6380–6385.
- Chen, E., Hermanson, S., and Ekker, S.C. (2004a). Syndecan-2 is essential for angiogenic sprouting during zebrafish development. *Blood* *103*, 1710–1719.
- Chen, L., Klass, C., and Woods, A. (2004b). Syndecan-2 Regulates Transforming Growth Factor- $\beta$  Signaling. *J. Biol. Chem.* *279*, 15715–15718.
- Chen, Y.-H., Narimatsu, Y., Clausen, T.M., Gomes, C., Karlsson, R., Steentoft, C., Spliid, C.B., Gustavsson, T., Salanti, A., Persson, A., et al. (2018). The GAGome: a cell-based library of displayed glycosaminoglycans. *Nat. Methods* *15*, 881–888.
- Cheng, B., Montmasson, M., Terradot, L., and Rousselle, P. (2016). Syndecans as Cell Surface Receptors in Cancer Biology. A Focus on their Interaction with PDZ Domain Proteins. *Front. Pharmacol.* *7*, 10.
- Chiquet, M., Birk, D.E., Bönnemann, C.G., and Koch, M. (2014). Molecules in Focus: Collagen XII: Protecting bone and muscle integrity by organizing collagen fibrils. *Int. J. Biochem. Cell Biol.* *53*, 51–54.
- Choi, Y., Chung, H., Jung, H., Couchman, J.R., and Oh, E.-S. (2011). Syndecans as cell surface receptors: Unique structure equates with functional diversity. *Matrix Biol. J. Int. Soc. Matrix Biol.* *30*, 93–99.
- Clasper, S., Vekemans, S., Fiore, M., Plebanski, M., Wordsworth, P., David, G., and Jackson, D.G. (1999). Inducible expression of the cell surface heparan sulfate proteoglycan syndecan-2 (fibroglycan) on human activated macrophages can regulate fibroblast growth factor action. *J. Biol. Chem.* *274*, 24113–24123.
- Colige, A., Sieron, A.L., Li, S.W., Schwarze, U., Petty, E., Wertelecki, W., Wilcox, W., Krakow, D., Cohn, D.H., Reardon, W., et al. (1999). Human Ehlers-Danlos syndrome type VII

C and bovine dermatosparaxis are caused by mutations in the procollagen I N-proteinase gene. *Am. J. Hum. Genet.* *65*, 308–317.

Colige, A., Nuytinck, L., Hausser, I., Essen, A.J. van, Thiry, M., Herens, C., Adès, L.C., Malfait, F., Paepe, A.D., Franck, P., et al. (2004). Novel Types of Mutation Responsible for the Dermatosparactic Type of Ehlers–Danlos Syndrome (Type VIIC) and Common Polymorphisms in the ADAMTS2 Gene. *J. Invest. Dermatol.* *123*, 656–663.

Couchman, J.R. (2010). Transmembrane signaling proteoglycans. *Annu. Rev. Cell Dev. Biol.* *26*, 89–114.

Culbertson, A.T., Smith, A.L., Cook, M.D., and Zabolina, O.A. (2016). Truncations of xyloglucan xylosyltransferase 2 provide insights into the roles of the N- and C-terminus. *Phytochemistry* *128*, 12–19.

Danielson, K.G., Baribault, H., Holmes, D.F., Graham, H., Kadler, K.E., and Iozzo, R.V. (1997). Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. *J. Cell Biol.* *136*, 729–743.

David, G., Schueren, B. van der, Marynen, P., Cassiman, J.J., and Berghe, H. van den (1992). Molecular cloning of amphiglycan, a novel integral membrane heparan sulfate proteoglycan expressed by epithelial and fibroblastic cells. *J. Cell Biol.* *118*, 961–969.

Davidson, A.E., Borasio, E., Liskova, P., Khan, A.O., Hassan, H., Cheetham, M.E., Plagnol, V., Alkuraya, F.S., Tuft, S.J., and Hardcastle, A.J. (2015). Brittle Cornea Syndrome ZNF469 Mutation Carrier Phenotype and Segregation Analysis of Rare ZNF469 Variants in Familial Keratoconus. *Invest. Ophthalmol. Vis. Sci.* *56*, 578–586.

De Paepe, A., Nuytinck, L., Hausser, I., Anton-Lamprecht, I., and Naeyaert, J.M. (1997). Mutations in the COL5A1 gene are causal in the Ehlers-Danlos syndromes I and II. *Am. J. Hum. Genet.* *60*, 547–554.

Delbaere, S., Dhooge, T., Syx, D., Petit, F., Goemans, N., Destrée, A., Vanakker, O., Rycke, R.D., Symoens, S., and Malfait, F. (2019a). Novel defects in collagen XII and VI expand the mixed myopathy/Ehlers–Danlos syndrome spectrum and lead to variant-specific alterations in the extracellular matrix. *Genet. Med.* 1–12.

Delbaere, S., Van Damme, T., Syx, D., Symoens, S., Coucke, P., Willaert, A., and Malfait, F. (2019b). Hypomorphic zebrafish models mimic the musculoskeletal phenotype of  $\beta$ 4GalT7-deficient Ehlers-Danlos syndrome. *Matrix Biol.*

Eagleton, M.J. (2016). Arterial complications of vascular Ehlers-Danlos syndrome. *J. Vasc. Surg.* *64*, 1869–1880.

Esco, J.D., and Selleck, S.B. (2002). Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu. Rev. Biochem.* *71*, 435–471.

Esco, J.D., and Zhang, L. (1996). Influence of core protein sequence on glycosaminoglycan assembly. *Curr. Opin. Struct. Biol.* *6*, 663–670.

Esco, J.D., Stewart, T.E., and Taylor, W.H. (1985). Animal cell mutants defective in glycosaminoglycan biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* *82*, 3197–3201.

- Esposito, D., and Chatterjee, D.K. (2006). Enhancement of soluble protein expression through the use of fusion tags. *Curr. Opin. Biotechnol.* 17, 353–358.
- Faiyaz-Ul-Haque, M., Zaidi, S.H.E., Al-Ali, M., Al-Mureikhi, M.S., Kennedy, S., Al-Thani, G., Tsui, L.-C., and Teebi, A.S. (2004). A novel missense mutation in the galactosyltransferase-I (B4GALT7) gene in a family exhibiting facioskeletal anomalies and Ehlers–Danlos syndrome resembling the progeroid type. *Am. J. Med. Genet. A.* 128A, 39–45.
- Ferdous, Z., Lazaro, L.D., Iozzo, R.V., Höök, M., and Grande-Allen, K.J. (2008). Influence of cyclic strain and decorin deficiency on 3D cellularized collagen matrices. *Biomaterials* 29, 2740–2748.
- Ferdous, Z., Peterson, S.B., Tseng, H., Anderson, D.K., Iozzo, R.V., and Grande-Allen, K.J. (2010). A role for decorin in controlling proliferation, adhesion, and migration of murine embryonic fibroblasts. *J. Biomed. Mater. Res. A* 93, 419–428.
- Fukuta, M., Kobayashi, Y., Uchimura, K., Kimata, K., and Habuchi, O. (1998). Molecular cloning and expression of human chondroitin 6-sulfotransferase. *Biochim. Biophys. Acta* 1399, 57–61.
- Funderburgh, J.L. (2002). Keratan Sulfate Biosynthesis. *IUBMB Life Int. Union Biochem. Mol. Biol. Life* 54, 187–194.
- Furth, J.J., Wroth, T.H., and Ackerman, S. (1991). Genes for collagen types I, IV, and V are transcribed in HeLa cells but a postinitiation block prevents the accumulation of type I mRNA. *Exp. Cell Res.* 192, 118–121.
- Gandhi, N.S., and Mancera, R.L. (2008). The structure of glycosaminoglycans and their interactions with proteins. *Chem. Biol. Drug Des.* 72, 455–482.
- Germain, D.P., and Herrera-Guzman, Y. (2004). Vascular Ehlers–Danlos syndrome. *Ann. Génétique* 47, 1–9.
- Giunta, C., Chambaz, C., Pedemonte, M., Scapolan, S., and Steinmann, B. (2008a). The arthrochalasia type of Ehlers–Danlos syndrome (EDS VIIA and VIIB): The diagnostic value of collagen fibril ultrastructure. *Am. J. Med. Genet. A.* 146A, 1341–1346.
- Giunta, C., Elçioglu, N.H., Albrecht, B., Eich, G., Chambaz, C., Janecke, A.R., Yeowell, H., Weis, M., Eyre, D.R., Kraenzlin, M., et al. (2008b). Spondylocheiro Dysplastic Form of the Ehlers-Danlos Syndrome—An Autosomal-Recessive Entity Caused by Mutations in the Zinc Transporter Gene SLC39A13. *Am. J. Hum. Genet.* 82, 1290–1305.
- Gloster, T.M. (2014). Advances in understanding glycosyltransferases from a structural perspective. *Curr. Opin. Struct. Biol.* 28, 131–141.
- Gómez-Garre, P., Seijo, M., Gutiérrez-Delgado, E., Castro del Río, M., de la Torre, C., Gómez-Abad, C., Morales-Corraliza, J., Puig, M., and Serratos, J.M. (2006). Ehlers-Danlos syndrome and periventricular nodular heterotopia in a Spanish family with a single FLNA mutation. *J. Med. Genet.* 43, 232–237.

Götte, M., and Kresse, H. (2005). Defective glycosaminoglycan substitution of decorin in a patient with progeroid syndrome is a direct consequence of two point mutations in the galactosyltransferase I (beta4GalT-7) gene. *Biochem. Genet.* *43*, 65–77.

Götte, M., Spillmann, D., Yip, G.W., Versteeg, E., Echtermeyer, F.G., van Kuppevelt, T.H., and Kiesel, L. (2008). Changes in heparan sulfate are associated with delayed wound repair, altered cell migration, adhesion and contractility in the galactosyltransferase I (beta4GalT-7) deficient form of Ehlers-Danlos syndrome. *Hum. Mol. Genet.* *17*, 996–1009.

Götting, C., Kuhn, J., Zahn, R., Brinkmann, T., and Kleesiek, K. (2000). Molecular cloning and expression of human UDP-d-Xylose:proteoglycan core protein beta-d-xylosyltransferase and its first isoform XT-II. *J. Mol. Biol.* *304*, 517–528.

Gould, S.E., Upholt, W.B., and Kosher, R.A. (1992). Syndecan 3: a member of the syndecan family of membrane-intercalated proteoglycans that is expressed in high amounts at the onset of chicken limb cartilage differentiation. *Proc. Natl. Acad. Sci. U. S. A.* *89*, 3271.

Grootjans, J.J., Zimmermann, P., Reekmans, G., Smets, A., Degeest, G., Dürr, J., and David, G. (1997). Syntenin, a PDZ protein that binds syndecan cytoplasmic domains. *Proc. Natl. Acad. Sci. U. S. A.* *94*, 13683.

Guarnieri, V., Morlino, S., Stolfo, G.D., Mastroianno, S., Mazza, T., and Castori, M. (2019). Cardiac valvular Ehlers-Danlos syndrome is a well-defined condition due to recessive null variants in COL1A2. *Am. J. Med. Genet. A.* *179*, 846–851.

Gubbiotti, M.A., Neill, T., and Iozzo, R.V. (2017). A current view of perlecan in physiology and pathology: A mosaic of functions. *Matrix Biol. J. Int. Soc. Matrix Biol.* *57–58*, 285–298.

Guerrini, R., and Carrozzo, R. (2001). Epileptogenic brain malformations: clinical presentation, malformative patterns and indications for genetic testing. *Seizure* *10*, 532–543; quiz 544–547.

Guerrini, R., and Dobyns, W.B. (1998). Bilateral periventricular nodular heterotopia with mental retardation and frontonasal malformation. *Neurology* *51*, 499–503.

Gulberti, S., Lattard, V., Fondeur, M., Jacquinet, J.-C., Mulliert, G., Netter, P., Magdalou, J., Ouzzine, M., and Fournel-Gigleux, S. (2005). Phosphorylation and sulfation of oligosaccharide substrates critically influence the activity of human beta1,4-galactosyltransferase 7 (GalT-I) and beta1,3-glucuronosyltransferase I (GlcAT-I) involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J. Biol. Chem.* *280*, 1417–1425.

Guo, M.H., Stoler, J., Lui, J., Nilsson, O., Bianchi, D.W., Hirschhorn, J.N., and Dauber, A. (2013). Redefining the Progeroid Form of Ehlers-Danlos Syndrome: Report of the Fourth Patient with B4GALT7 Deficiency and Review of the Literature. *Am. J. Med. Genet. A.* *161*, 2519–2527.

Han, C., Belenkaya, T.Y., Khodoun, M., Tauchi, M., Lin, X., and Lin, X. (2004). Distinct and collaborative roles of Drosophila EXT family proteins in morphogen signalling and gradient formation. *Dev. Camb. Engl.* *131*, 1563–1575.

Hanset, R., and Lapiere, C.M. (1974). Inheritance of Dermatosparaxis in the CalfA genetic defect of connective tissues. *J. Hered.* *65*, 356–358.



- Hata, R.-I., Kurata, S., and Shinkai, H. (1988). Existence of malfunctioning pro $\alpha$ 2(I) collagen genes in a patient with a pro $\alpha$ 2(I)-chain-defective variant of Ehlers-Danlos syndrome. *Eur. J. Biochem.* *174*, 231–237.
- Heinegård, D. (2009). Fell-Muir Lecture: Proteoglycans and more--from molecules to biology. *Int. J. Exp. Pathol.* *90*, 575–586.
- Hennet, T. (2002). The galactosyltransferase family. *Cell. Mol. Life Sci. CMLS* *59*, 1081–1095.
- Hennet, T., Dinter, A., Kuhnert, P., Mattu, T.S., Rudd, P.M., and Berger, E.G. (1998). Genomic cloning and expression of three murine UDP-galactose: beta-N-acetylglucosamine beta1,3-galactosyltransferase genes. *J. Biol. Chem.* *273*, 58–65.
- Hicks, D., Farsani, G.T., Laval, S., Collins, J., Sarkozy, A., Martoni, E., Shah, A., Zou, Y., Koch, M., Bönnemann, C.G., et al. (2014). Mutations in the collagen XII gene define a new form of extracellular matrix-related myopathy. *Hum. Mol. Genet.* *23*, 2353–2363.
- Hiraoka, N., Nakagawa, H., Ong, E., Akama, T.O., Fukuda, M.N., and Fukuda, M. (2000). Molecular cloning and expression of two distinct human chondroitin 4-O-sulfotransferases that belong to the HNK-1 sulfotransferase gene family. *J. Biol. Chem.* *275*, 20188–20196.
- Holmborn, K., Habicher, J., Kasza, Z., Eriksson, A.S., Filipek-Gorniok, B., Gopal, S., Couchman, J.R., Ahlberg, P.E., Wiweger, M., Spillmann, D., et al. (2012). On the roles and regulation of chondroitin sulfate and heparan sulfate in zebrafish pharyngeal cartilage morphogenesis. *J. Biol. Chem.* *287*, 33905–33916.
- Iozzo, R.V. (2005). Basement membrane proteoglycans: from cellar to ceiling. *Nat. Rev. Mol. Cell Biol.* *6*, 646–656.
- Iozzo, R.V., and Schaefer, L. (2015). Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. *Matrix Biol. J. Int. Soc. Matrix Biol.* *42*, 11–55.
- Isshiki, S., Togayachi, A., Kudo, T., Nishihara, S., Watanabe, M., Kubota, T., Kitajima, M., Shiraishi, N., Sasaki, K., Andoh, T., et al. (1999). Cloning, expression, and characterization of a novel UDP-galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase (beta3Gal-T5) responsible for synthesis of type 1 chain in colorectal and pancreatic epithelia and tumor cells derived therefrom. *J. Biol. Chem.* *274*, 12499–12507.
- Izumikawa, T., Kanagawa, N., Watamoto, Y., Okada, M., Saeki, M., Sakano, M., Sugahara, K., Sugihara, K., Asano, M., and Kitagawa, H. (2010). Impairment of Embryonic Cell Division and Glycosaminoglycan Biosynthesis in Glucuronyltransferase-I-deficient Mice. *J. Biol. Chem.* *285*, 12190–12196.
- Izumikawa, T., Okuura, Y., Koike, T., Sakoda, N., and Kitagawa, H. (2011). Chondroitin 4-O-sulfotransferase-1 regulates the chain length of chondroitin sulfate in co-operation with chondroitin N-acetylgalactosaminyltransferase-2. *Biochem. J.* *434*, 321–331.
- Izumikawa, T., Sato, B., Mikami, T., Tamura, J., Igarashi, M., and Kitagawa, H. (2015). GlcUA $\beta$ 1-3Gal $\beta$ 1-3Gal $\beta$ 1-4Xyl(2-O-phosphate) is the preferred substrate for chondroitin N-acetylgalactosaminyltransferase-1. *J. Biol. Chem.* *290*, 5438–5448.

- Jamsheer, A., Olech, E.M., Kozłowski, K., Niedziela, M., Sowińska-Seidler, A., Obara-Moszyńska, M., Latos-Bieleńska, A., Karczewski, M., and Zemojtel, T. (2016). Exome sequencing reveals two novel compound heterozygous XYLT1 mutations in a Polish patient with Desbuquois dysplasia type 2 and growth hormone deficiency. *J. Hum. Genet.* *61*, 577–583.
- Janecke, A.R., Li, B., Boehm, M., Krabichler, B., Rohrbach, M., Müller, T., Fuchs, I., Golas, G., Katagiri, Y., Ziegler, S.G., et al. (2016). The Phenotype of the Musculocontractural Type of Ehlers-Danlos Syndrome due to CHST14 Mutations. *Am. J. Med. Genet. A.* *170A*, 103–115.
- Järveläinen, H., Puolakkainen, P., Pakkanen, S., Brown, E.L., Höök, M., Iozzo, R.V., Sage, E.H., and Wight, T.N. (2006). A role for decorin in cutaneous wound healing and angiogenesis. *Wound Repair Regen. Off. Publ. Wound Heal. Soc. Eur. Tissue Repair Soc.* *14*, 443–452.
- Järveläinen, H., Sainio, A., and Wight, T.N. (2015). Pivotal role for decorin in angiogenesis. *Matrix Biol. J. Int. Soc. Matrix Biol.* *43*, 15–26.
- Jepsen, K.J., Wu, F., Peragallo, J.H., Paul, J., Roberts, L., Ezura, Y., Oldberg, A., Birk, D.E., and Chakravarti, S. (2002). A Syndrome of Joint Laxity and Impaired Tendon Integrity in Lumican- and Fibromodulin-deficient Mice. *J. Biol. Chem.* *277*, 35532–35540.
- Jiang, D., Liang, J., and Noble, P.W. (2011). Hyaluronan as an Immune Regulator in Human Diseases. *Physiol. Rev.* *91*, 221–264.
- Jinek, M., and Conti, E. (2006). Eukaryotic expression, purification, crystallization and preliminary X-ray analysis of murine Manic Fringe. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* *62*, 774–777.
- Jones, K.L., Schwarze, U., Adam, M.P., Byers, P.H., and Mefford, H.C. (2015). A homozygous B3GAT3 mutation causes a severe syndrome with multiple fractures, expanding the phenotype of linkeropathy syndromes. *Am. J. Med. Genet. A.* *167A*, 2691–2696.
- Jong, L. de, and Kemp, A. (1984). Stoichiometry and kinetics of the prolyl 4-hydroxylase partial reaction. *Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol.* *787*, 105–111.
- Jørgensen, A., Fagerheim, T., Rand-Hendriksen, S., Lunde, P.I., Vorren, T.O., Pepin, M.G., Leistritz, D.F., and Byers, P.H. (2015). Vascular Ehlers–Danlos Syndrome in siblings with biallelic COL3A1 sequence variants and marked clinical variability in the extended family. *Eur. J. Hum. Genet.* *23*, 796–802.
- Ju, T., Brewer, K., D’Souza, A., Cummings, R.D., and Canfield, W.M. (2002). Cloning and expression of human core 1 beta1,3-galactosyltransferase. *J. Biol. Chem.* *277*, 178–186.
- Kalamajski, S., Liu, C., Tillgren, V., Rubin, K., Oldberg, Å., Rai, J., Weis, M., and Eyre, D.R. (2014). Increased C-telopeptide Cross-linking of Tendon Type I Collagen in Fibromodulin-deficient Mice. *J. Biol. Chem.* *289*, 18873–18879.
- Kalamajski, S., Bihan, D., Bonna, A., Rubin, K., and Farndale, R.W. (2016). Fibromodulin Interacts with Collagen Cross-linking Sites and Activates Lysyl Oxidase. *J. Biol. Chem.* *291*, 7951–7960.

- Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., and Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* 7, 1511–1522.
- Kang, H.-G., Evers, M.R., Xia, G., Baenziger, J.U., and Schachner, M. (2002). Molecular cloning and characterization of chondroitin-4-O-sulfotransferase-3. A novel member of the HNK-1 family of sulfotransferases. *J. Biol. Chem.* 277, 34766–34772.
- Kapferer-Seebacher, I., Pepin, M., Werner, R., Aitman, T.J., Nordgren, A., Stoiber, H., Thielens, N., Gaboriaud, C., Amberger, A., Schossig, A., et al. (2016). Periodontal Ehlers-Danlos Syndrome Is Caused by Mutations in C1R and C1S, which Encode Subcomponents C1r and C1s of Complement. *Am. J. Hum. Genet.* 99, 1005–1014.
- Keene, D.R., San Antonio, J.D., Mayne, R., McQuillan, D.J., Sarris, G., Santoro, S.A., and Iozzo, R.V. (2000). Decorin binds near the C terminus of type I collagen. *J. Biol. Chem.* 275, 21801–21804.
- Kiani, C., Chen, L., Wu, Y.J., Yee, A.J., and Yang, B.B. (2002). Structure and function of aggrecan. *Cell Res.* 12, 19–32.
- Kim, B.T., Kitagawa, H., Tamura, J., Saito, T., Kusche-Gullberg, M., Lindahl, U., and Sugahara, K. (2001). Human tumor suppressor EXT gene family members EXTL1 and EXTL3 encode alpha 1,4- N-acetylglucosaminyltransferases that likely are involved in heparan sulfate/heparin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 98, 7176–7181.
- Kim, H.G., Yang, S.M., Lee, Y.C., Do, S.I., Chung, I.S., and Yang, J.M. (2003). High-level expression of human glycosyltransferases in insect cells as biochemically active form. *Biochem. Biophys. Res. Commun.* 305, 488–493.
- Kitagawa, H., Tone, Y., Tamura, J., Neumann, K.W., Ogawa, T., Oka, S., Kawasaki, T., and Sugahara, K. (1998). Molecular cloning and expression of glucuronyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J. Biol. Chem.* 273, 6615–6618.
- Kitagawa, H., Shimakawa, H., and Sugahara, K. (1999). The tumor suppressor EXT-like gene EXTL2 encodes an alpha1, 4-N-acetylhexosaminyltransferase that transfers N-acetylgalactosamine and N-acetylglucosamine to the common glycosaminoglycan-protein linkage region. The key enzyme for the chain initiation of heparan sulfate. *J. Biol. Chem.* 274, 13933–13937.
- Kitagawa, H., Uyama, T., and Sugahara, K. (2001). Molecular Cloning and Expression of a Human Chondroitin Synthase. *J. Biol. Chem.* 276, 38721–38726.
- Kitagawa, H., Izumikawa, T., Uyama, T., and Sugahara, K. (2003). Molecular Cloning of a Chondroitin Polymerizing Factor That Cooperates with Chondroitin Synthase for Chondroitin Polymerization. *J. Biol. Chem.* 278, 23666–23671.
- Klaassens, M., Reinstein, E., Hilhorst-Hofstee, Y., Schrandt, J.J.P., Malfait, F., Staal, H., ten Have, L.C., Blaauw, J., Roggeveen, H.C.J., Krakow, D., et al. (2012). Ehlers-Danlos Arthrochalasia type (VIIA-B) – expanding the phenotype: from prenatal life through adulthood. *Clin. Genet.* 82, 121–130.

- Klass, C.M., Couchman, J.R., and Woods, A. (2000). Control of extracellular matrix assembly by syndecan-2 proteoglycan. *J. Cell Sci.* *113* ( Pt 3), 493–506.
- Kobayashi, M., Sugumaran, G., Liu, J., Shworak, N.W., Silbert, J.E., and Rosenberg, R.D. (1999). Molecular cloning and characterization of a human uronyl 2-sulfotransferase that sulfates iduronyl and glucuronyl residues in dermatan/chondroitin sulfate. *J. Biol. Chem.* *274*, 10474–10480.
- Kobayashi, S., Morimoto, K., Shimizu, T., Takahashi, M., Kurosawa, H., and Shirasawa, T. (2000). Association of EXT1 and EXT2, hereditary multiple exostoses gene products, in Golgi apparatus. *Biochem. Biophys. Res. Commun.* *268*, 860–867.
- Koike, T., Izumikawa, T., Tamura, J.-I., and Kitagawa, H. (2009). FAM20B is a kinase that phosphorylates xylose in the glycosaminoglycan-protein linkage region. *Biochem. J.* *421*, 157–162.
- Koike, T., Izumikawa, T., Sato, B., and Kitagawa, H. (2014). Identification of phosphatase that dephosphorylates xylose in the glycosaminoglycan-protein linkage region of proteoglycans. *J. Biol. Chem.* *289*, 6695–6708.
- Kojima, T., Shworak, N.W., and Rosenberg, R.D. (1992). Molecular cloning and expression of two distinct cDNA-encoding heparan sulfate proteoglycan core proteins from a rat endothelial cell line. *J. Biol. Chem.* *267*, 4870–4877.
- Kolbinger, F., Streiff, M.B., and Katopodis, A.G. (1998). Cloning of a human UDP-galactose:2-acetamido-2-deoxy-D-glucose 3beta-galactosyltransferase catalyzing the formation of type 1 chains. *J. Biol. Chem.* *273*, 433–440.
- Koningsbruggen, S. van, Knoester, H., Bakx, R., Mook, O., Knegt, L., and Cobben, J.M. (2016). Complete and partial XYLT1 deletion in a patient with neonatal short limb skeletal dysplasia. *Am. J. Med. Genet. A.* *170*, 510–514.
- Korb-Pap, A., Stratis, A., Mühlenberg, K., Niederreiter, B., Hayer, S., Echtermeyer, F., Stange, R., Zwerina, J., Pap, T., Pavenstädt, H., et al. (2012). Early structural changes in cartilage and bone are required for the attachment and invasion of inflamed synovial tissue during destructive inflammatory arthritis. *Ann. Rheum. Dis.* *71*, 1004–1011.
- Kresse, H., Rosthøj, S., Quentin, E., Hollmann, J., Glössl, J., Okada, S., and Tønnesen, T. (1987). Glycosaminoglycan-free small proteoglycan core protein is secreted by fibroblasts from a patient with a syndrome resembling progeroid. *Am. J. Hum. Genet.* *41*, 436–453.
- Kreuger, J., and Kjellén, L. (2012). Heparan sulfate biosynthesis: regulation and variability. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* *60*, 898–907.
- Kuhn, J., Müller, S., Schnölzer, M., Kempf, T., Schön, S., Brinkmann, T., Schöttler, M., Götting, C., and Kleesiek, K. (2003). High-level expression and purification of human xylosyltransferase I in High Five insect cells as biochemically active form. *Biochem. Biophys. Res. Commun.* *312*, 537–544.
- Lairson, L.L., Henrissat, B., Davies, G.J., and Withers, S.G. (2008). Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* *77*, 521–555.

- Lauder, R.M. (2009). Chondroitin sulphate: A complex molecule with potential impacts on a wide range of biological systems. *Complement. Ther. Med.* 17, 56–62.
- Lazarus, M.B., Nam, Y., Jiang, J., Sliz, P., and Walker, S. (2011). Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature* 469, 564–567.
- Levy, H.P. (1993). Hypermobility Ehlers-Danlos Syndrome. In *GeneReviews®*, M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J. Bean, K. Stephens, and A. Amemiya, eds. (Seattle (WA): University of Washington, Seattle), p.
- Li, J.-P., and Kusche-Gullberg, M. (2016). Chapter Six - Heparan Sulfate: Biosynthesis, Structure, and Function. In *International Review of Cell and Molecular Biology*, K.W. Jeon, ed. (Academic Press), pp. 215–273.
- Li, L., Ly, M., and Linhardt, R.J. (2012). Proteoglycan sequence. *Mol. Biosyst.* 8, 1613–1625.
- Lichtman, M.K., Otero-Vinas, M., and Falanga, V. (2016). Transforming growth factor beta (TGF- $\beta$ ) isoforms in wound healing and fibrosis. *Wound Repair Regen.* 24, 215–222.
- Lira-Navarrete, E., Valero-González, J., Villanueva, R., Martínez-Júlvez, M., Tejero, T., Merino, P., Panjkar, S., and Hurtado-Guerrero, R. (2011). Structural insights into the mechanism of protein O-fucosylation. *PLoS One* 6, e25365.
- Lisignoli, G., Codeluppi, K., Todoerti, K., Manferdini, C., Piacentini, A., Zini, N., Grassi, F., Cattini, L., Piva, R., Rizzoli, V., et al. (2009). Gene array profile identifies collagen type XV as a novel human osteoblast-secreted matrix protein. *J. Cell. Physiol.* 220, 401–409.
- Liu, J., and Mushegian, A. (2003). Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci. Publ. Protein Soc.* 12, 1418–1431.
- Liu, C., Sheng, J., Krahn, J.M., Perera, L., Xu, Y., Hsieh, P.-H., Dou, W., Liu, J., and Pedersen, L.C. (2014). Molecular mechanism of substrate specificity for heparan sulfate 2-O-sulfotransferase. *J. Biol. Chem.* 289, 13407–13418.
- Lizak, C., Gerber, S., Numao, S., Aebi, M., and Locher, K.P. (2011). X-ray structure of a bacterial oligosaccharyltransferase. *Nature* 474, 350–355.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495.
- Ma, C., Peng, P., Zhou, Y., Liu, T., Wang, L., and Lu, C. (2020). MicroRNA-93 promotes angiogenesis and attenuates remodeling via inactivation of the Hippo/Yap pathway by targeting Lats2 after myocardial infarction. *Mol. Med. Rep.* 22, 483–493.
- Malfait, F., Wenstrup, R., and De Paepe, A. (1993). Classic Ehlers-Danlos Syndrome. In *GeneReviews®*, M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J. Bean, K. Stephens, and A. Amemiya, eds. (Seattle (WA): University of Washington, Seattle), p.
- Malfait, F., Coster, P.D., Hausser, I., Essen, A.J. van, Franck, P., Colige, A., Nusgens, B., Martens, L., and Paepe, A.D. (2004). The natural history, including orofacial features of three patients with Ehlers-Danlos syndrome, dermatosparaxis type (EDS type VIIC). *Am. J. Med. Genet. A.* 131A, 18–28.

- Malfait, F., Coucke, P., Symoens, S., Loeys, B., Nuytinck, L., and Paepe, A.D. (2005). The molecular basis of classic Ehlers-Danlos syndrome: A comprehensive study of biochemical and molecular findings in 48 unrelated patients. *Hum. Mutat.* *25*, 28–37.
- Malfait, F., Symoens, S., Coucke, P., Nunes, L., De Almeida, S., and De Paepe, A. (2006). Total absence of the  $\alpha 2(I)$  chain of collagen type I causes a rare form of Ehlers-Danlos syndrome with hypermobility and propensity to cardiac valvular problems. *J. Med. Genet.* *43*, e36.
- Malfait, F., Symoens, S., Backer, J.D., Hermanns-Lê, T., Sakalihasan, N., Lapière, C.M., Coucke, P., and Paepe, A.D. (2007). Three arginine to cysteine substitutions in the pro-alpha (I)-collagen chain cause Ehlers-Danlos syndrome with a propensity to arterial rupture in early adulthood. *Hum. Mutat.* *28*, 387–395.
- Malfait, F., Syx, D., Vlummens, P., Symoens, S., Nampoothiri, S., Hermanns-Lê, T., Laer, L.V., and Paepe, A.D. (2010). Musculocontractural Ehlers-Danlos Syndrome (former EDS type VIB) and adducted thumb clubfoot syndrome (ATCS) represent a single clinical entity caused by mutations in the dermatan-4-sulfotransferase 1 encoding CHST14 gene. *Hum. Mutat.* *31*, 1233–1239.
- Malfait, F., Kariminejad, A., Van Damme, T., Gauche, C., Syx, D., Merhi-Soussi, F., Gulberti, S., Symoens, S., Vanhauwaert, S., Willaert, A., et al. (2013). Defective initiation of glycosaminoglycan synthesis due to B3GALT6 mutations causes a pleiotropic Ehlers-Danlos-syndrome-like connective tissue disorder. *Am. J. Hum. Genet.* *92*, 935–945.
- Malfait, F., Francomano, C., Byers, P., Belmont, J., Berglund, B., Black, J., Bloom, L., Bowen, J.M., Brady, A.F., Burrows, N.P., et al. (2017). The 2017 international classification of the Ehlers-Danlos syndromes. *Am. J. Med. Genet. C Semin. Med. Genet.* *175*, 8–26.
- Malmström, A., Bartolini, B., Thelin, M.A., Pacheco, B., and Maccarana, M. (2012). Iduronic acid in chondroitin/dermatan sulfate: biosynthesis and biological function. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* *60*, 916–925.
- Mantle, D., Wilkins, R.M., and Preedy, V. (2005). A novel therapeutic strategy for Ehlers–Danlos syndrome based on nutritional supplements. *Med. Hypotheses* *64*, 279–283.
- Mao, X., Gauche, C., Coughtrie, M.W.H., Bui, C., Gulberti, S., Merhi-Soussi, F., Ramalanjaona, N., Bertin-Jung, I., Diot, A., Dumas, D., et al. (2016). The heparan sulfate sulfotransferase 3-OST3A (HS3ST3A) is a novel tumor regulator and a prognostic marker in breast cancer. *Oncogene* *35*, 5043–5055.
- Marchica, C.L., Pinelli, V., Borges, M., Zimmer, J., Narayanan, V., Iozzo, R.V., and Ludwig, M.S. (2011). A role for decorin in a murine model of allergen-induced asthma. *Am. J. Physiol. Lung Cell. Mol. Physiol.* *300*, L863-873.
- Marynen, P., Zhang, J., Cassiman, J.J., Berghe, H.V. den, and David, G. (1989). Partial primary structure of the 48- and 90-kilodalton core proteins of cell surface-associated heparan sulfate proteoglycans of lung fibroblasts. Prediction of an integral membrane domain and evidence for multiple distinct core proteins at the cell surface of human lung fibroblasts. *J. Biol. Chem.* *264*, 7017–7024.

- McCormick, C., Duncan, G., Goutsos, K.T., and Tufaro, F. (2000). The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate. *Proc. Natl. Acad. Sci. U. S. A.* 97, 668–673.
- Meneghetti, M.C.Z., Hughes, A.J., Rudd, T.R., Nader, H.B., Powell, A.K., Yates, E.A., and Lima, M.A. (2015). Heparan sulfate and heparin interactions with proteins. *J. R. Soc. Interface* 12.
- Merline, R., Moreth, K., Beckmann, J., Nastase, M.V., Zeng-Brouwers, J., Tralhão, J.G., Lemarchand, P., Pfeilschifter, J., Schaefer, R.M., Iozzo, R.V., et al. (2011). Signaling by the matrix proteoglycan decorin controls inflammation and cancer through PDCD4 and MicroRNA-21. *Sci. Signal.* 4, ra75.
- Micheal, S., Siddiqui, S.N., Zafar, S.N., Gabriëla Niewold, I.T., Khan, M.I., and Bergen, A.A.B. (2019). Identification of a Novel ZNF469 Mutation in a Pakistani Family With Brittle Cornea Syndrome. *Cornea* 38, 718–722.
- Mikami, T., and Kitagawa, H. (2013). Biosynthesis and function of chondroitin sulfate. *Biochim. Biophys. Acta* 1830, 4719–4733.
- Mikami, T., Mizumoto, S., Kago, N., Kitagawa, H., and Sugahara, K. (2003). Specificities of three distinct human chondroitin/dermatan N-acetylgalactosamine 4-O-sulfotransferases demonstrated using partially desulfated dermatan sulfate as an acceptor: implication of differential roles in dermatan sulfate biosynthesis. *J. Biol. Chem.* 278, 36115–36127.
- Mis, E.K., Liem, K.F., Kong, Y., Schwartz, N.B., Domowicz, M., and Weatherbee, S.D. (2014). Forward Genetics Defines Xylt1 as a Key, Conserved Regulator of Early Chondrocyte Maturation and Skeletal Length. *Dev. Biol.* 385, 67–82.
- Miyake, N., Kosho, T., and Matsumoto, N. (2014). Ehlers-Danlos syndrome associated with glycosaminoglycan abnormalities. *Adv. Exp. Med. Biol.* 802, 145–159.
- Miyazaki, H., Fukumoto, S., Okada, M., Hasegawa, T., and Furukawa, K. (1997). Expression cloning of rat cDNA encoding UDP-galactose:GD2 beta1,3-galactosyltransferase that determines the expression of GD1b/GM1/GA1. *J. Biol. Chem.* 272, 24794–24799.
- Mongiati, M., Taylor, K., Otto, J., Aho, S., Uitto, J., Whitelock, J.M., and Iozzo, R.V. (2000). The Protein Core of the Proteoglycan Perlecan Binds Specifically to Fibroblast Growth Factor-7. *J. Biol. Chem.* 275, 7095–7100.
- Mongiati, M., Otto, J., Oldershaw, R., Ferrer, F., Sato, J.D., and Iozzo, R.V. (2001). Fibroblast Growth Factor-binding Protein Is a Novel Partner for Perlecan Protein Core. *J. Biol. Chem.* 276, 10263–10271.
- Morokuma, D., Xu, J., Hino, M., Mon, H., Merzaban, J.S., Takahashi, M., Kusakabe, T., and Lee, J.M. (2017). Expression and Characterization of Human  $\beta$ -1, 4-Galactosyltransferase 1 ( $\beta$ 4GalT1) Using Silkworm–Baculovirus Expression System. *Mol. Biotechnol.* 59, 151–158.
- Mukhopadhyay, A., Wong, M.Y., Chan, S.Y., Do, D.V., Khoo, A., Ong, C.T., Cheong, H.H., Lim, I.J., and Phan, T.T. (2010). Syndecan-2 and decorin: proteoglycans with a difference--implications in keloid pathogenesis. *J. Trauma* 68, 999–1008.

- Nakajima, M., Mizumoto, S., Miyake, N., Kogawa, R., Iida, A., Ito, H., Kitoh, H., Hirayama, A., Mitsubuchi, H., Miyazaki, O., et al. (2013). Mutations in B3GALT6, which encodes a glycosaminoglycan linker region enzyme, cause a spectrum of skeletal and connective tissue disorders. *Am. J. Hum. Genet.* *92*, 927–934.
- Nallamsetty, S., and Waugh, D.S. (2006). Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expr. Purif.* *45*, 175–182.
- Neill, T., Painter, H., Buraschi, S., Owens, R.T., Lisanti, M.P., Schaefer, L., and Iozzo, R.V. (2012). Decorin antagonizes the angiogenic network: concurrent inhibition of Met, hypoxia inducible factor 1 $\alpha$ , vascular endothelial growth factor A, and induction of thrombospondin-1 and TIMP3. *J. Biol. Chem.* *287*, 5492–5506.
- Neill, T., Jones, H.R., Crane-Smith, Z., Owens, R.T., Schaefer, L., and Iozzo, R.V. (2013). Decorin induces rapid secretion of thrombospondin-1 in basal breast carcinoma cells via inhibition of Ras homolog gene family, member A/Rho-associated coiled-coil containing protein kinase 1. *FEBS J.* *280*, 2353–2368.
- Nothaft, H., and Szymanski, C.M. (2010). Protein glycosylation in bacteria: sweeter than ever. *Nat. Rev. Microbiol.* *8*, 765–778.
- Nugent, M.A., and Iozzo, R.V. (2000). Fibroblast growth factor-2. *Int. J. Biochem. Cell Biol.* *32*, 115–120.
- Nuytinck, L., Freund, M., Lagae, L., Pierard, G.E., Hermanns-Le, T., and De Paepe, A. (2000). Classical Ehlers-Danlos syndrome caused by a mutation in type I collagen. *Am. J. Hum. Genet.* *66*, 1398–1402.
- Nygren, P.-Å., Stefan, S., and Uhlén, M. (1994). Engineering proteins to facilitate bioprocessing. *Trends Biotechnol.* *12*, 184–188.
- O’Connell, M., Burrows, N.P., Vlijmen-Willems, M.J.J.V., Clark, S.M., and Schalkwijk, J. (2010). Tenascin-X deficiency and Ehlers–Danlos syndrome: a case report and review of the literature. *Br. J. Dermatol.* *163*, 1340–1345.
- von Oettingen, J.E., Tan, W.-H., and Dauber, A. (2014). Skeletal Dysplasia, Global Developmental Delay, and Multiple Congenital Anomalies in a 5 year-old boy– Report of the Second Family with B3GAT3 mutation and Expansion of the Phenotype. *Am. J. Med. Genet. A.* *0*, 1580–1586.
- Oganesyan, N., Ankoudinova, I., Kim, S.-H., and Kim, R. (2007). Effect of Osmotic Stress and Heat Shock in Recombinant Protein Overexpression and Crystallization. *Protein Expr. Purif.* *52*, 280–285.
- Ogawa, T., Tsubota, Y., Hashimoto, J., Kariya, Y., and Miyazaki, K. (2007). The short arm of laminin gamma2 chain of laminin-5 (laminin-332) binds syndecan-1 and regulates cellular adhesion and migration by suppressing phosphorylation of integrin beta4 chain. *Mol. Biol. Cell* *18*, 1621–1633.
- Okajima, T., Fukumoto, S., Furukawa, K., and Urano, T. (1999). Molecular basis for the progeroid variant of Ehlers-Danlos syndrome. Identification and characterization of two mutations in galactosyltransferase I gene. *J. Biol. Chem.* *274*, 28841–28844.



- Paepe, A.D., and Malfait, F. (2004). Bleeding and bruising in patients with Ehlers–Danlos syndrome and other collagen vascular disorders. *Br. J. Haematol.* *127*, 491–500.
- Paepe, A.D., and Malfait, F. (2012). The Ehlers–Danlos syndrome, a disorder with many faces. *Clin. Genet.* *82*, 1–11.
- Patel, R.Y., and Balaji, P.V. (2007). Fold-recognition and comparative modeling of human  $\beta$ 3GalT I, II, IV, V and VI and  $\beta$ 3GalNAcT I: Prediction of residues conferring acceptor substrate specificity. *J. Mol. Graph. Model.* *26*, 255–268.
- Peng, J., and Xu, J. (2011). RaptorX: exploiting structure information for protein alignment by statistical inference. *Proteins* *79*, 161–171.
- Pepin, M., Schwarze, U., Superti-Furga, A., and Byers, P.H. (2009). Clinical and Genetic Features of Ehlers–Danlos Syndrome Type IV, the Vascular Type.
- Persson, A., Nilsson, J., Vorontsov, E., Noborn, F., and Larson, G. (2019). Identification of a non-canonical chondroitin sulfate linkage region trisaccharide. *Glycobiology* *29*, 366–371.
- Pierce, A., Lyon, M., Hampson, I.N., Cowling, G.J., and Gallagher, J.T. (1992). Molecular cloning of the major cell surface heparan sulfate proteoglycan from rat liver. *J. Biol. Chem.* *267*, 3894–3900.
- Pinhal, M.A., Smith, B., Olson, S., Aikawa, J., Kimata, K., and Esko, J.D. (2001). Enzyme interactions in heparan sulfate biosynthesis: uronosyl 5-epimerase and 2-O-sulfotransferase interact in vivo. *Proc. Natl. Acad. Sci. U. S. A.* *98*, 12984–12989.
- Poulain, F.E., and Yost, H.J. (2015). Heparan sulfate proteoglycans: a sugar code for vertebrate development? *Dev. Camb. Engl.* *142*, 3456–3467.
- Presto, J., Thuveson, M., Carlsson, P., Busse, M., Wilén, M., Eriksson, I., Kusche-Gullberg, M., and Kjellén, L. (2008). Heparan sulfate biosynthesis enzymes EXT1 and EXT2 affect NDST1 expression and heparan sulfate sulfation. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 4751–4756.
- Punetha, J., Kesari, A., Hoffman, E.P., Gos, M., Kamińska, A., Kostera-Pruszczyk, A., Hausmanowa-Petrusewicz, I., Hu, Y., Zou, Y., Bönnemann, C.G., et al. (2017). A novel COL12A1 variant expands the clinical picture of congenital myopathies with extracellular matrix defects. *Muscle Nerve* *55*, 277–281.
- Purushothaman, A., Uyama, T., Kobayashi, F., Yamada, S., Sugahara, K., Rapraeger, A.C., and Sanderson, R.D. (2010). Heparanase-enhanced shedding of syndecan-1 by myeloma cells promotes endothelial invasion and angiogenesis. *Blood* *115*, 2449–2457.
- Pyeritz, R.E. (2000). Ehlers-Danlos syndrome. *N. Engl. J. Med.* *342*, 730–732.
- Qian, R., Chen, C., and Colley, K.J. (2001). Location and Mechanism of  $\alpha$ 2,6-Sialyltransferase Dimer Formation : Role of cysteine residues in enzyme dimerization, localization, activity, and processing. *J. Biol. Chem.* *276*, 28641–28649.

- Qin, Y., Ke, J., Gu, X., Fang, J., Wang, W., Cong, Q., Li, J., Tan, J., Brunzelle, J.S., Zhang, C., et al. (2015). Structural and functional study of D-glucuronyl C5-epimerase. *J. Biol. Chem.* *290*, 4620–4630.
- Rahuel-Clermont, S., Daligault, F., Piet, M.-H., Gulberti, S., Netter, P., Branlant, G., Magdalou, J., and Lattard, V. (2010). Biochemical and thermodynamic characterization of mutated  $\beta$ 1,4-galactosyltransferase 7 involved in the progeroid form of the Ehlers-Danlos syndrome. *Biochem. J.* *432*, 303–311.
- Raspanti, M., Viola, M., Forlino, A., Tenni, R., Gruppi, C., and Tira, M.E. (2008). Glycosaminoglycans show a specific periodic interaction with type I collagen fibrils. *J. Struct. Biol.* *164*, 134–139.
- Ringsdorf, W.M., and Cheraskin, E. (1982). Vitamin C and human wound healing. *Oral Surg. Oral Med. Oral Pathol.* *53*, 231–236.
- Ritelli, M., Dordoni, C., Venturini, M., Chiarelli, N., Quinzani, S., Traversa, M., Zoppi, N., Vascellaro, A., Wischmeijer, A., Manfredini, E., et al. (2013). Clinical and molecular characterization of 40 patients with classic Ehlers–Danlos syndrome: identification of 18 COL5A1 and 2 COL5A2 novel mutations. *Orphanet J. Rare Dis.* *8*, 58.
- Ritelli, M., Chiarelli, N., Zoppi, N., Dordoni, C., Quinzani, S., Traversa, M., Venturini, M., Calzavara-Pinton, P., and Colombi, M. (2015). Insights in the etiopathology of galactosyltransferase II (GalT-II) deficiency from transcriptome-wide expression profiling of skin fibroblasts of two sisters with compound heterozygosity for two novel B3GALT6 mutations. *Mol. Genet. Metab. Rep.* *2*, 1–15.
- Ritelli, M., Dordoni, C., Cinquina, V., Venturini, M., Calzavara-Pinton, P., and Colombi, M. (2017). Expanding the clinical and mutational spectrum of B4GALT7-spondylodysplastic Ehlers-Danlos syndrome. *Orphanet J. Rare Dis.* *12*.
- Ritelli, M., Cinquina, V., Giacomuzzi, E., Venturini, M., Chiarelli, N., and Colombi, M. (2019). Further Defining the Phenotypic Spectrum of B3GAT3 Mutations and Literature Review on Linkeropathy Syndromes. *Genes* *10*.
- Robinson, J.A., and Robinson, H.C. (1985). Initiation of chondroitin sulphate synthesis by beta-D-galactosides. Substrates for galactosyltransferase II. *Biochem. J.* *227*, 805–814.
- Rodríguez-Herva, J.J., González-Melendi, P., Cuartas-Lanza, R., Antúnez-Lamas, M., Río-Alvarez, I., Li, Z., López-Torrejón, G., Díaz, I., Del Pozo, J.C., Chakravarthy, S., et al. (2012). A bacterial cysteine protease effector protein interferes with photosynthesis to suppress plant innate immune responses. *Cell. Microbiol.* *14*, 669–681.
- Rohrbach, M., Vandersteen, A., Yiş, U., Serdaroglu, G., Ataman, E., Chopra, M., Garcia, S., Jones, K., Kariminejad, A., Kraenzlin, M., et al. (2011). Phenotypic variability of the kyphoscoliotic type of Ehlers-Danlos syndrome (EDS VIA): clinical, molecular and biochemical delineation. *Orphanet J. Rare Dis.* *6*, 46.
- Rossi, G.D., Evans, A.R., Kay, E., Woodfin, A., McKay, T.R., Nourshargh, S., and Whiteford, J.R. (2014). Shed syndecan-2 inhibits angiogenesis. *J. Cell Sci.* *127*, 4788.

- Rudd, N., Holbrook, K., Nimrod, C., and Byers, P. (1983). Pregnancy complications in type IV Ehlers-Danlos Syndrome. *The Lancet* *321*, 50–53.
- Rühland, C., Schönherr, E., Robenek, H., Hansen, U., Iozzo, R.V., Bruckner, P., and Seidler, D.G. (2007). The glycosaminoglycan chain of decorin plays an important role in collagen fibril formation at the early stages of fibrillogenesis. *FEBS J.* *274*, 4246–4255.
- Saliba, M., Ramalanjaona, N., Gulberti, S., Bertin-Jung, I., Thomas, A., Dahbi, S., Lopin-Bon, C., Jacquinet, J.-C., Breton, C., Ouzzine, M., et al. (2015). Probing the Acceptor Active Site Organization of the Human Recombinant  $\beta$ 1,4-Galactosyltransferase 7 and Design of Xyloside-based Inhibitors. *J. Biol. Chem.* *290*, 7658–7670.
- Salter, C.G., Davies, J.H., Moon, R.J., Fairhurst, J., Bunyan, D., DDD Study, and Foulds, N. (2016). Further defining the phenotypic spectrum of B4GALT7 mutations. *Am. J. Med. Genet. A.* *170*, 1556–1563.
- Saunders, S., Jalkanen, M., O’Farrell, S., and Bernfield, M. (1989). Molecular cloning of syndecan, an integral membrane proteoglycan. *J. Cell Biol.* *108*, 1547–1556.
- Schalkwijk, J., Zweers, M.C., Steijlen, P.M., Dean, W.B., Taylor, G., van Vlijmen, I.M., van Haren, B., Miller, W.L., and Bristow, J. (2009). A Recessive Form of the Ehlers–Danlos Syndrome Caused by Tenascin-X Deficiency.
- Schmid, J., Heider, D., Wendel, N.J., Sperl, N., and Sieber, V. (2016). Bacterial Glycosyltransferases: Challenges and Opportunities of a Highly Diverse Enzyme Class Toward Tailoring Natural Products. *Front. Microbiol.* *7*.
- Schwarze, U., Hata, R.-I., McKusick, V.A., Shinkai, H., Hoyme, H.E., Pyeritz, R.E., and Byers, P.H. (2004). Rare Autosomal Recessive Cardiac Valvular Form of Ehlers-Danlos Syndrome Results from Mutations in the COL1A2 Gene That Activate the Nonsense-Mediated RNA Decay Pathway. *Am. J. Hum. Genet.* *74*, 917–930.
- Scott, J.E. (1988). Proteoglycan-fibrillar collagen interactions. *Biochem. J.* *252*, 313–323.
- Sebestyén, A., Gallai, M., Knittel, T., Ambrust, T., Ramadori, G., and Kovalszky, I. (2000). Cytokine regulation of syndecan expression in cells of liver origin. *Cytokine* *12*, 1557–1560.
- Seidler, D.G., Schaefer, L., Robenek, H., Iozzo, R.V., Kresse, H., and Schönherr, E. (2005). A physiologic three-dimensional cell culture system to investigate the role of decorin in matrix organisation and cell survival. *Biochem. Biophys. Res. Commun.* *332*, 1162–1170.
- Seidler, D.G., Faiyaz-Ul-Haque, M., Hansen, U., Yip, G.W., Zaidi, S.H.E., Teebi, A.S., Kiesel, L., and Götte, M. (2006). Defective glycosylation of decorin and biglycan, altered collagen structure, and abnormal phenotype of the skin fibroblasts of an Ehlers–Danlos syndrome patient carrying the novel Arg270Cys substitution in galactosyltransferase I ( $\beta$ 4GalT-7). *J. Mol. Med.* *84*, 583–594.
- Seidler, D.G., Mohamed, N.A., Bocian, C., Stadtmann, A., Hermann, S., Schäfers, K., Schäfers, M., Iozzo, R.V., Zarbock, A., and Götte, M. (2011). The role for decorin in delayed-type hypersensitivity. *J. Immunol. Baltim. Md 1950* *187*, 6108–6119.

Senay, C., Lind, T., Muguruma, K., Tone, Y., Kitagawa, H., Sugahara, K., Lidholt, K., Lindahl, U., and Kusche-Gullberg, M. (2000). The EXT1/EXT2 tumor suppressors: catalytic activities and role in heparan sulfate biosynthesis. *EMBO Rep.* 1, 282–286.

Shao, S., Qin, T., Qian, W., Li, X., Li, W., Han, L., Zhang, D., Wang, Z., Ma, Q., Wu, Z., et al. (2020). Cav-1 Ablation in Pancreatic Stellate Cells Promotes Pancreatic Cancer Growth through Nrf2-Induced shh Signaling. *Oxid. Med. Cell. Longev.* 2020, 1868764.

Sharma, U., Carrique, L., Vadon-Le Goff, S., Mariano, N., Georges, R.-N., Delolme, F., Koivunen, P., Myllyharju, J., Moali, C., Aghajari, N., et al. (2017). Structural basis of homo- and heterotrimerization of collagen I. *Nat. Commun.* 8.

Shashikiran, U., Rastogi, A., Gupta, R.P., and Sabir, M. (1999). Ehler-Danlos syndrome type VI variant presenting with recurrent respiratory infections and responding to high dose vitamin C. *J. Assoc. Physicians India* 47, 554–555.

Shimo, T., Gentili, C., Iwamoto, M., Wu, C., Koyama, E., and Pacifici, M. (2004). Indian hedgehog and syndecans-3 coregulate chondrocyte proliferation and function during chick limb skeletogenesis. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* 229, 607–617.

Skehel, J.J., and Wiley, D.C. (2000). Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu. Rev. Biochem.* 69, 531–569.

Soya, N., Fang, Y., Palcic, M.M., and Klassen, J.S. (2011). Trapping and characterization of covalent intermediates of mutant retaining glycosyltransferases. *Glycobiology* 21, 547–552.

Stolfa, G., Mondal, N., Zhu, Y., Yu, X., Buffone, A., and Neelamegham, S. (2016). Using CRISPR-Cas9 to quantify the contributions of O-glycans, N-glycans and Glycosphingolipids to human leukocyte-endothelium adhesion. *Sci. Rep.* 6.

Sugahara, K., Yamashina, I., Waard, P.D., Halbeek, H.V., and Vliegthart, J.F. (1988). Structural studies on sulfated glycopeptides from the carbohydrate-protein linkage region of chondroitin 4-sulfate proteoglycans of swarm rat chondrosarcoma. Demonstration of the structure Gal(4-O-sulfate)beta 1-3Gal beta 1-4XYL beta 1-O-Ser. *J. Biol. Chem.* 263, 10168–10174.

Sugahara, K., Ohi, Y., Harada, T., Waard, P. de, and Vliegthart, J.F. (1992). Structural studies on sulfated oligosaccharides derived from the carbohydrate-protein linkage region of chondroitin 6-sulfate proteoglycans of shark cartilage. I. Six compounds containing 0 or 1 sulfate and/or phosphate residues. *J. Biol. Chem.* 267, 6027–6035.

Sugawara, K., Tsuruta, D., Ishii, M., Jones, J.C.R., and Kobayashi, H. (2008). Laminin-332 and -511 in skin. *Exp. Dermatol.* 17, 473–480.

Svensson, L., Aszódi, A., Reinholt, F.P., Fässler, R., Heinegård, D., and Oldberg, Å. (1999). Fibromodulin-null Mice Have Abnormal Collagen Fibrils, Tissue Organization, and Altered Lumican Deposition in Tendon. *J. Biol. Chem.* 274, 9636–9647.

Symoens, S., Syx, D., Malfait, F., Callewaert, B., Backer, J.D., Vanakker, O., Coucke, P., and Paepe, A.D. (2012). Comprehensive molecular analysis demonstrates type V collagen mutations in over 90% of patients with classic EDS and allows to refine diagnostic criteria. *Hum. Mutat.* 33, 1485–1493.

- Syx, D., Damme, T.V., Symoens, S., Maiburg, M.C., Laar, I. van de, Morton, J., Suri, M., Del Campo, M., Hausser, I., Hermanns-Lê, T., et al. (2015). Genetic Heterogeneity and Clinical Variability in Musculocontractural Ehlers–Danlos Syndrome Caused by Impaired Dermatan Sulfate Biosynthesis. *Hum. Mutat.* *36*, 535–547.
- Syx, D., De Wandele, I., Rombaut, L., and Malfait, F. (2017). Hypermobility, The Ehlers–Danlos syndromes and chronic pain. *Clin. Exp. Rheumatol.* *35*, S116–S122.
- Syx, D., De Wandele, I., Symoens, S., De Rycke, R., Hougrand, O., Voermans, N., De Paepe, A., and Malfait, F. (2019). Bi-allelic AEBP1 mutations in two patients with Ehlers–Danlos syndrome. *Hum. Mol. Genet.* *28*, 1853–1864.
- Talhaoui, I., Bui, C., Oriol, R., Mulliert, G., Gulberti, S., Netter, P., Coughtrie, M.W.H., Ouzzine, M., and Fournel-Gigleux, S. (2010). Identification of Key Functional Residues in the Active Site of Human  $\beta$ 1,4-Galactosyltransferase 7. *J. Biol. Chem.* *285*, 37342–37358.
- Termini, J.M., Silver, Z.A., Connor, B., Antonopoulos, A., Haslam, S.M., Dell, A., and Desrosiers, R.C. (2017). HEK293T cell lines defective for O-linked glycosylation. *PLoS ONE* *12*.
- Theocharis, A.D., Skandalis, S.S., Gialeli, C., and Karamanos, N.K. (2016). Extracellular matrix structure. *Adv. Drug Deliv. Rev.* *97*, 4–27.
- Tillgren, V., Mörgelin, M., Önnarfjord, P., Kalamajski, S., and Aspberg, A. (2016). The Tyrosine Sulfate Domain of Fibromodulin Binds Collagen and Enhances Fibril Formation. *J. Biol. Chem.* *291*, 23744–23755.
- Tinkle, B., Castori, M., Berglund, B., Cohen, H., Grahame, R., Kazkaz, H., and Levy, H. (2017). Hypermobility Ehlers–Danlos syndrome (a.k.a. Ehlers–Danlos syndrome Type III and Ehlers–Danlos syndrome hypermobility type): Clinical description and natural history. *Am. J. Med. Genet. C Semin. Med. Genet.* *175*, 48–69.
- Tkachenko Eugene, Elfenbein Arye, Tirziu Daniela, and Simons Michael (2006). Syndecan-4 Clustering Induces Cell Migration in a PDZ-Dependent Manner. *Circ. Res.* *98*, 1398–1404.
- Tone, Y., Pedersen, L.C., Yamamoto, T., Izumikawa, T., Kitagawa, H., Nishihara, J., Tamura, J.-I., Negishi, M., and Sugahara, K. (2008). 2-o-phosphorylation of xylose and 6-o-sulfation of galactose in the protein linkage region of glycosaminoglycans influence the glucuronyltransferase-I activity involved in the linkage region synthesis. *J. Biol. Chem.* *283*, 16801–16807.
- Toriello, H.V., Glover, T.W., Takahara, K., Byers, P.H., Miller, D.E., Higgins, J.V., and Greenspan, D.S. (1996). A translocation interrupts the COL5A1 gene in a patient with Ehlers–Danlos syndrome and hypomelanosis of Ito. *Nat. Genet.* *13*, 361–365.
- Tsutsui, Y., Ramakrishnan, B., and Qasba, P.K. (2013). Crystal structures of  $\beta$ -1,4-galactosyltransferase 7 enzyme reveal conformational changes and substrate binding. *J. Biol. Chem.* *288*, 31963–31970.
- Uchimura, K. (2015). Keratan Sulfate: Biosynthesis, Structures, and Biological Functions. In *Glycosaminoglycans*, K. Balagurunathan, H. Nakato, and U.R. Desai, eds. (New York, NY: Springer New York), pp. 389–400.

- Uchimura, K., Kadomatsu, K., Nishimura, H., Muramatsu, H., Nakamura, E., Kurosawa, N., Habuchi, O., El-Fasakhany, F.M., Yoshikai, Y., and Muramatsu, T. (2002). Functional analysis of the chondroitin 6-sulfotransferase gene in relation to lymphocyte subpopulations, brain development, and oversulfated chondroitin sulfates. *J. Biol. Chem.* 277, 1443–1450.
- Utani, A., Nomizu, M., Matsuura, H., Kato, K., Kobayashi, T., Takeda, U., Aota, S., Nielsen, P.K., and Shinkai, H. (2001). A Unique Sequence of the Laminin  $\alpha 3$  G Domain Binds to Heparin and Promotes Cell Adhesion through Syndecan-2 and -4. *J. Biol. Chem.* 276, 28779–28788.
- Uyama, T., Kitagawa, H., Tamura, J., and Sugahara, K. (2002). Molecular Cloning and Expression of Human Chondroitin N-Acetylgalactosaminyltransferase: The key enzyme for chain initiation and elongation of chondroitin/dermatan sulfate on protein linkage region tetrasaccharide shared by heparin/heparan sulfate. *J. Biol. Chem.* 277, 8841–8846.
- Uyama, T., Kitagawa, H., Tanaka, J., Tamura, J., Ogawa, T., and Sugahara, K. (2003). Molecular Cloning and Expression of a Second Chondroitin N-Acetylgalactosaminyltransferase Involved in the Initiation and Elongation of Chondroitin/Dermatan Sulfate. *J. Biol. Chem.* 278, 3072–3078.
- Van Damme, T., Colige, A., Syx, D., Giunta, C., Lindert, U., Rohrbach, M., Aryani, O., Alanay, Y., Simsek-Kiper, P.Ö., Kroes, H.Y., et al. (2016). Expanding the clinical and mutational spectrum of the Ehlers–Danlos syndrome, dermatosparaxis type. *Genet. Med.* 18, 882–891.
- Van Damme, T., Pang, X., Guillemin, B., Gulberti, S., Syx, D., De Rycke, R., Kaye, O., de Die-Smulders, C.E.M., Pfundt, R., Kariminejad, A., et al. (2018). Biallelic B3GALT6 mutations cause spondylodysplastic Ehlers–Danlos syndrome. *Hum. Mol. Genet.* 27, 3475–3487.
- Vanpouille, C., Deligny, A., Delehedde, M., Denys, A., Melchior, A., Liénard, X., Lyon, M., Mazurier, J., Fernig, D.G., and Allain, F. (2007). The heparin/heparan sulfate sequence that interacts with cyclophilin B contains a 3-O-sulfated N-unsubstituted glucosamine residue. *J. Biol. Chem.* 282, 24416–24429.
- Vasta, J.D., and Raines, R.T. (2016). Human Collagen Prolyl 4-Hydroxylase is Activated by Ligands for its Iron Center. *Biochemistry* 55, 3224–3233.
- Vorster, A.A., Beighton, P., and Ramesar, R.S. (2015). Spondyloepimetaphyseal dysplasia with joint laxity (Beighton type); mutation analysis in eight affected South African families. *Clin. Genet.* 87, 492–495.
- Vrieling, A., Rüger, W., Driessen, H.P., and Freemont, P.S. (1994). Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J.* 13, 3413–3422.
- Wan, Q., Tang, J., Han, Y., Xiao, Q., and Deng, Y. (2018). Brittle cornea syndrome: a case report and review of the literature. *BMC Ophthalmol.* 18.
- Wang, J., Markova, D., Anderson, D.G., Zheng, Z., Shapiro, I.M., and Risbud, M.V. (2011). TNF- $\alpha$  and IL-1 $\beta$  promote a disintegrin-like and metalloprotease with thrombospondin type I motif-5-mediated aggrecan degradation through syndecan-4 in intervertebral disc. *J. Biol. Chem.* 286, 39738–39749.

- Wang, S., Sugahara, K., and Li, F. (2016). Chondroitin sulfate/dermatan sulfate sulfatases from mammals and bacteria. *Glycoconj. J.* *33*, 841–851.
- Wang, X., Li, W., Zhu, D., Zhao, H., Chen, P., and Chen, X. (2020). Characterization of human peripheral blood  $\gamma\delta$  T cells in patients with sepsis. *Exp. Ther. Med.* *19*, 3698–3706.
- Weis, S.M., Zimmerman, S.D., Shah, M., Covell, J.W., Omens, J.H., Ross, J., Dalton, N., Jones, Y., Reed, C.C., Iozzo, R.V., et al. (2005). A role for decorin in the remodeling of myocardial infarction. *Matrix Biol. J. Int. Soc. Matrix Biol.* *24*, 313–324.
- Wen, J., Xiao, J., Rahdar, M., Choudhury, B.P., Cui, J., Taylor, G.S., Esko, J.D., and Dixon, J.E. (2014). Xylose phosphorylation functions as a molecular switch to regulate proteoglycan biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 15723–15728.
- Wenstrup, R.J., Langland, G.T., Willing, M.C., D'Souza, V.N., and Cole, W.G. (1996). A splice-junction mutation in the region of COL5A1 that codes for the carboxyl propeptide of pro alpha 1(V) chains results in the gravis form of the Ehlers-Danlos syndrome (type I). *Hum. Mol. Genet.* *5*, 1733–1736.
- Whitelock, J.M., Murdoch, A.D., Iozzo, R.V., and Underwood, P.A. (1996). The degradation of human endothelial cell-derived perlecan and release of bound basic fibroblast growth factor by stromelysin, collagenase, plasmin, and heparanases. *J. Biol. Chem.* *271*, 10079–10086.
- Wight, T.N., Toole, B.P., and Hascall, V.C. (2011). Hyaluronan and the Aggregating Proteoglycans. In *The Extracellular Matrix: An Overview*, R.P. Mecham, ed. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 147–195.
- Wilhelm, D., Kempf, H., Bianchi, A., and Vincourt, J.-B. (2020). ATDC5 cells as a model of cartilage extracellular matrix neosynthesis, maturation and assembly. *J. Proteomics* 103718.
- Wilkins-Port, C.E., Sanderson, R.D., Tominna-Sebald, E., and McKeown-Longo, P.J. (2003). Vitronectin's basic domain is a syndecan ligand which functions in trans to regulate vitronectin turnover. *Cell Commun. Adhes.* *10*, 85–103.
- Worapamorn, W., Haase, H.R., Li, H., and Bartold, P.M. (2001). Growth factors and cytokines modulate gene expression of cell-surface proteoglycans in human periodontal ligament cells. *J. Cell. Physiol.* *186*, 448–456.
- Wu, J., Yang, J., Zhao, J., Wu, J., Zhang, X., Leung, W.K., and Sun, W. (2018). A Chinese family with periodontal Ehlers–Danlos syndrome associated with missense mutation in the C1R gene. *J. Clin. Periodontol.* *45*, 1311–1318.
- Xian, X., Gopal, S., and Couchman, J.R. (2010). Syndecans as receptors and organizers of the extracellular matrix. *Cell Tissue Res.* *339*, 31–46.
- Xu, X., Wang, Z., and Zan, T. (2019). A case of Ehlers–Danlos syndrome presenting with widened atrophic scars of forehead, elbow, knee, and pretibial area. *Medicine (Baltimore)* *98*.
- Yada, T., Gotoh, M., Sato, T., Shionyu, M., Go, M., Kaseyama, H., Iwasaki, H., Kikuchi, N., Kwon, Y.-D., Togayachi, A., et al. (2003a). Chondroitin sulfate synthase-2. Molecular cloning and characterization of a novel human glycosyltransferase homologous to chondroitin sulfate glucuronyltransferase, which has dual enzymatic activities. *J. Biol. Chem.* *278*, 30235–30247.

- Yada, T., Sato, T., Kaseyama, H., Gotoh, M., Iwasaki, H., Kikuchi, N., Kwon, Y.-D., Togayachi, A., Kudo, T., Watanabe, H., et al. (2003b). Chondroitin sulfate synthase-3. Molecular cloning and characterization. *J. Biol. Chem.* *278*, 39711–39725.
- Yang, Y., Yaccoby, S., Liu, W., Langford, J.K., Pumphrey, C.Y., Theus, A., Epstein, J., and Sanderson, R.D. (2002). Soluble syndecan-1 promotes growth of myeloma tumors in vivo. *Blood* *100*, 610–617.
- Yao, Y., and Wang, Y. (2013). ATDC5: An excellent in vitro model cell line for skeletal development. *J. Cell. Biochem.* *114*, 1223–1229.
- Yeowell, H.N., and Steinmann, B. (1993). PLOD1-Related Kyphoscoliotic Ehlers-Danlos Syndrome. In *GeneReviews®*, M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J. Bean, K. Stephens, and A. Amemiya, eds. (Seattle (WA): University of Washington, Seattle), p.
- Zea, C.J., Camci-Unal, G., and Pohl, N.L. (2008). Thermodynamics of binding of divalent magnesium and manganese to uridine phosphates: implications for diabetes-related hypomagnesaemia and carbohydrate biocatalysis. *Chem. Cent. J.* *2*, 15.
- Zhang, L., and Esko, J.D. (1994). Amino acid determinants that drive heparan sulfate assembly in a proteoglycan. *J. Biol. Chem.* *269*, 19295–19299.
- Zhang, L., David, G., and Esko, J.D. (1995). Repetitive Ser-Gly Sequences Enhance Heparan Sulfate Assembly in Proteoglycans. *J. Biol. Chem.* *270*, 27127–27135.
- Zhang, L., Lawrence, R., Frazier, B.A., and Esko, J.D. (2006). CHO Glycosylation Mutants: Proteoglycans. In *Methods in Enzymology*, (Academic Press), pp. 205–221.
- Zhou, D., Berger, E.G., and Hennet, T. (1999). Molecular cloning of a human UDP-galactose:GlcNAc $\beta$ 1,3GalNAc  $\beta$ 1, 3 galactosyltransferase gene encoding an O-linked core3-elongation enzyme. *Eur. J. Biochem.* *263*, 571–576.
- Zou, Y., Zwolanek, D., Izu, Y., Gandhi, S., Schreiber, G., Brockmann, K., Devoto, M., Tian, Z., Hu, Y., Veit, G., et al. (2014). Recessive and dominant mutations in COL12A1 cause a novel EDS/myopathy overlap syndrome in humans and mice. *Hum. Mol. Genet.* *23*, 2339–2352.





## **Résumé :**

Les protéoglycanes (PGs) sont des composants majeurs des membranes plasmiques des cellules et de la matrice extracellulaire. Ces macromolécules jouent un rôle important dans l'organisation architecturale des tissus conjonctifs et dans la signalisation cellulaire, notamment au cours du développement embryonnaire et post-natal. Les PGs sont composés de chaînes de glycosaminoglycanes (GAGs) attachées de façon covalente à une protéine *core* par une amorce tétrasaccharidique [acide glucuronique- $\beta$ 1,3-galactose- $\beta$ 1,3-galactose- $\beta$ 1,4-xylose- $\beta$ 1-O-]. L'addition du troisième résidu (galactose) est catalysée par la  $\beta$ 1,3-galactosyltransférase 6 ( $\beta$ 3GalT6), une enzyme clé de l'initiation de la synthèse des GAGs. Plusieurs études ont mis en évidence la présence de mutations de la  $\beta$ 3GalT6 associées à la forme spondylodysplastique du Syndrome d'Ehlers-Danlos (SEDsp), une maladie génétique sévère des tissus conjonctifs et caractérisée par une fragilité des tissus, une hyperlaxité articulaire, une hyperextensibilité et des défauts de cicatrisation de la peau ainsi que des altérations du système musculo-squelettique.

L'objectif de ce travail de thèse est de comprendre les conséquences fonctionnelles et structurales des mutations de la  $\beta$ 3GalT6 et leur rôle dans la pathogénie du SEDsp, (i) en réalisant une caractérisation moléculaire et fonctionnelle de la protéine recombinante humaine et (ii) en développant un modèle cellulaire déficient en  $\beta$ 3GalT6 pour étudier l'impact des défauts génétiques sur les fonctions métaboliques des cellules, en particulier sur la synthèse des GAGs.

La première partie du projet vise à la détermination de l'impact des mutations sur la fonction de la  $\beta$ 3GalT6. Pour ce faire, nous avons produit dans la bactérie et purifié différentes formes tronquées solubles de la  $\beta$ 3GalT6 fusionnées avec la *Maltose Binding Protein* (MBP). Les essais enzymatiques ont permis de déterminer les constantes cinétiques  $K_M$  et  $k_{cat}$  de la protéine sauvage. Les mutants de la  $\beta$ 3GalT6 seront étudiés ultérieurement selon la même approche.

La seconde partie de ce travail de thèse a été de réaliser un modèle cellulaire déficient en  $\beta$ 3GalT6 en utilisant la technologie CRISPR/Cas9. Les clones déficients obtenus présentent (i) un très faible niveau d'expression de l'ARNm, (ii) une absence d'activité galactosyltransférase et (iii) un défaut de synthèse des GAGs endogènes ou à partir d'un substrat exogène xylosidique. Nous avons initié l'analyse de la capacité de la  $\beta$ 3GalT6 sauvage et de deux mutants (Asp207His et Gly217Ser) à restaurer la synthèse des GAGs dans ces lignées invalidées.

A partir de ce travail, nous avons acquis une meilleure compréhension de l'implication de la  $\beta$ 3GalT6 dans la pathogénie du SEDsp et mieux compris le lien entre la perte de fonction de la  $\beta$ 3GalT6 et les conséquences métaboliques et cellulaires de cette déficience génétique. Ces résultats, mis en lien avec la sévérité des symptômes cliniques observés chez les patients, devraient aider les cliniciens dans la prise en charge et le suivi clinique des patients atteints de SEDsp.

Mot clés : Syndrome d'Ehlers-Danlos, Protéoglycanes, Glycosaminoglycanes, Glycosyltransférases, B3GALT6

## **Abstract :**

Proteoglycans (PGs) are major components of cell plasma membranes and extracellular matrix. These macromolecules play an important role in matrix organization of connective tissues and in cell signaling or embryonic and post-natal development. PGs are composed of glycosaminoglycan (GAG) chains covalently attached to a core protein through a tetrasaccharide linkage [Glucuronic acid- $\beta$ 1,3-Galactose- $\beta$ 1,3-Galactose- $\beta$ 1,4-Xylose- $\beta$ 1-O-]. The addition of the third residue (galactose) is catalyzed by the  $\beta$ 1,3-Galactosyltransferase 6 ( $\beta$ 3GalT6), a key glycosyltransferase in GAG initiation. Our group and others discovered that mutations of  $\beta$ 3GalT6 are associated to a spondylodysplastic form of Ehlers-Danlos Syndrome (spEDS), a severe connective tissue disorder characterized by skin and bone fragility, musculoskeletal malformations, delayed wound healing, joint hyperlaxity and intellectual disabilities.

The objectives of this project is to understand the functional and structural consequences of  $\beta$ 3GalT6 mutations in the development of spEDS, (i) achieving the molecular and functional characterization of the recombinant human  $\beta$ 3GalT6 and (ii) to develop cellular models (as  $\beta$ 3GalT6 KO cells) to study the impact of genetic deficiency on cells metabolism, precisely on GAGs synthesis.

The first part of the project is dedicated to the determination of mutation impact on the  $\beta$ 3GalT6 function. For this, we produce and purify several truncated soluble forms of h $\beta$ 3GalT6 in fusion to Maltose Binding Protein. The enzymatic activity tests have determined a  $K_M$  of 30  $\mu$ M and a  $k_{cat}$  of 0,05  $\text{min}^{-1}$  on wild-type enzyme.  $\beta$ 3GalT6 mutants will be further analyzed using the same approach.

The second part of the project is achieving to develop a  $\beta$ 3GalT6 deficient cell model using the CRISPR/Cas 9 technology. Deficient clones obtained present (i) a low level of RNA expression, (ii) an absence of galactosyltransferase activity and (iii) a defect on endogenous GAG synthesis or with exogenous substrate. We also analyze the capacity for WT  $\beta$ 3GalT6 and two mutants (Asp207His and Gly217Ser) to restore GAGs synthesis in deficient cells.

From this work, we better understand the implication of  $\beta$ 3GalT6 in the pathology of spEDS and relationships between  $\beta$ 3GalT6 loss of function, cellular consequences of genetic defect. Those results linked with the severity of spEDS clinical symptoms observed in patients, would help clinicians with management and clinical monitoring of spEDS patients.

Key words : Ehlers-Danlos Syndrome, Proteoglycans, Glycosaminoglycans, Glycosyltransferases, B3GALT6