



HAL
open science

Impact de la déficience en β 1,3-galactosyltransférase 6 dans une maladie génétique rare des tissus conjonctifs

Roméo Diana

► **To cite this version:**

Roméo Diana. Impact de la déficience en β 1,3-galactosyltransférase 6 dans une maladie génétique rare des tissus conjonctifs. Biochimie, Biologie Moléculaire. Université de Lorraine, 2024. Français. NNT : 2024LORR0060 . tel-04760670

HAL Id: tel-04760670

<https://hal.univ-lorraine.fr/tel-04760670v1>

Submitted on 30 Oct 2024

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.



**UNIVERSITÉ
DE LORRAINE**

**BIBLIOTHÈQUES
UNIVERSITAIRES**

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 bibliothèque : ddoc-theses-contact@univ-lorraine.fr
(Cette adresse ne permet pas de contacter les auteurs)

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

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

http://www.cfcopies.com/V2/leg/leg_droi.php

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

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 **Roméo DIANA**

Impact de la déficience en β 1,3-galactosyltransférase 6 dans une maladie génétique rare des tissus conjonctifs

Le 14 mai 2024

Membres du jury :

Rapporteurs :

M. Romain VIVÈS Directeur de recherche, CNRS, Grenoble

M^{me} Mirjam CZJZEK Directrice de recherche, CNRS, Roscoff

Examineurs :

M. David COELHO Chargé de recherche, INSERM, Vandœuvre-lès-Nancy

M. Ulrich VALCOURT Professeur, Université Claude Bernard Lyon 1, Lyon, président du jury

M^{me} Sylvie FOURNEL-GIGLEUX Directrice de recherche, INSERM, Vandœuvre-lès-Nancy, directrice de thèse

M^{me} Catherine BUI Maître de conférences, Université de Lorraine, Vandœuvre-lès-Nancy, co-directrice de thèse

UMR 7365 CNRS-Université de Lorraine
Ingénierie Moléculaire, Cellulaire et Physiopathologie (IMoPA)
Campus Biologie-Santé, Faculté de Médecine, Bâtiment Biopôle
9 avenue de la forêt de Haye, 54505 Vandœuvre-lès-Nancy

Remerciements

Je remercie le Dr. Sylvie Fournel-Gigleux, responsable de l'axe 1 de l'équipe Glycobiologie moléculaire, cellulaire et translationnelle (GlycoBio, anciennement MolCelTEG), de m'avoir accueilli lors de mes stages de Master et de m'avoir confié ce projet. Je la remercie de son implication, de ses conseils et de son écoute en tant que directrice de thèse.

Je remercie le Dr. Catherine Bui, ma co-directrice de thèse, de m'avoir fait confiance depuis que nous avons commencé à travailler ensemble en 2019. Je la remercie de m'avoir transmis les valeurs et la méthode de la recherche, ainsi que de son implication active dans le projet, ce qui a permis de compléter nos résultats. Je souhaite également la remercier (ainsi que le Dr. Sylvie Fournel-Gigleux) pour l'opportunité qui m'a été offerte de présenter mes résultats en congrès international. Après avoir été votre premier stagiaire puis votre premier étudiant en thèse, je peux affirmer que travailler avec vous a été un véritable plaisir et que j'en garderai un excellent souvenir.

Je souhaite remercier l'ensemble des membres du jury d'avoir accepté de juger ce travail de thèse. Merci aux rapporteurs les Drs. Romain Vivès (UMR 5075 IBS, Université Grenoble Alpes – CNRS – CEA, Grenoble, France) et Mirjam Czjzek (UMR 8227 LBI2M, Sorbonne Université – CNRS, Roscoff, France) ainsi qu'aux examinateurs le Dr. David Coelho (U1256 NGERE, Université de Lorraine – INSERM, Vandoeuvre-lès-Nancy, France) et le Pr. Ulrich Valcourt (UMR 5305 LBTI, Université Claude-Bernard-Lyon-1 – CNRS, Lyon, France).

Je remercie le Pr. Jean-Yves Jouzeau, ancien directeur de l'UMR, et le Pr. Bruno Charpentier, actuel directeur de l'UMR, de m'avoir accueilli au sein du laboratoire.

Je remercie le Dr. Sophie Rahuel-Clermont, directrice de l'École Doctorale BioSE, pour son implication et ses innovations lorsque j'étais représentant des doctorants.

Je remercie le Dr. Mohamed Ouzzine, responsable de l'axe 2 de l'équipe GlycoBio. I would also like to thank Dr. Guillermo Barreto, head of axis 3, for his advice and assistance with the drafting of my first scientific paper. De façon globale, je remercie les membres de l'équipe GlycoBio.

J'adresse un remerciement particulier au Dr. Jean-Baptiste Vincourt pour ses conseils et critiques, ses interventions pertinentes, et surtout son implication pratique dans mes expériences. Protéoglycanes et collagènes n'ont pas fini de se causer.

Un chaleureux remerciement à Anne Robert qui m'a transmis ses techniques à la paillasse lors de mon arrivée au laboratoire et aux côtés de qui j'ai pu développer mon esprit critique et prendre du recul.

Merci au Dr. Alexandre Kriznik de m'avoir formé au dichroïsme circulaire et au Dr. Jean-Malo Massicard d'avoir suggéré une nouvelle approche et de m'avoir prêté du matériel pour débloquer une situation extrêmement difficile. Vous avez été tous les deux à l'origine d'excellents moments et de discussions passionnantes.

Je remercie nos multiples collaborateurs pour leurs apports qui ont amené à la publication de nos travaux : les Drs. Franziska Malfait et Delfien Syx (*Center of Medical Genetics*, Université de Gand), le Dr. Chrystel Lopin-Bon et Sophie Front (ICOA, Université d'Orléans), le Dr. Sébastien Hergalant (NGERE, Université de Lorraine), le Dr. Grégory Francius (LCPME, Université de Lorraine), les Drs. Yasaman Karami et Hamed Khakzad (LORIA, Université de Lorraine), le Dr. Rebekka Wild et Marie Bourgeois (IBS, Université Grenoble Alpes) et Rachel Onifarasoaniaina (Institut Cochin, Université de Paris).

Merci aux enseignants de la Faculté des Sciences et Technologies de Nancy qui m'ont donné envie de chercher et trouver : Dr. Stéphane Labialle, Prs. Stéphanie Grandemange et Kira Weissman. Une pensée particulière à l'équipe de l'UFR Sciences Fondamentales et Appliquées de Metz : Drs. Hélène Dubois-Pot-Schneider, Delphine Aran, Sandrine Pain-Devin, Prs. Thierry Oster, Simon Devin et Adrien Neguesque.

Merci à tous les collègues (directs ou indirects) que j'ai pu côtoyer et qui ont été essentiels à cette période de ma vie, chacun à leur manière : Kevin, Anne, Ghita, Arnaud, Romain, Serge, Océane, Diana, Guru, Magali, Camille, Valentine, Quentin, Andrea Cal., Jade, Julie, Florent, Jean-Malo, Céline, Lauralie, Lucas, Laura, Félicia, Tiffany, Précillia. Merci à l'ensemble des personnes avec qui j'ai interagi au laboratoire, que ce soit pour travailler ou souffler un coup. Une pensée particulière à Lydia, Déborah, Karine et Christelle pour les nombreuses discussions devant le labo. J'ajoute également toute ma sympathie au personnel administratif du laboratoire, particulièrement Nadia et Ghislaine.

Merci à mes parents de m'avoir formé, de m'avoir soutenu inconditionnellement et de m'avoir offert cette liberté dont je profite chaque jour. Tout aurait été si différent sans vous. Vorrei cogliere l'occasione per ringraziare la mia famiglia e rivolgere un pensiero a Nonna.

Danke an Saskia, entfernte Juristenkollegin und liebste Freundin seit über zehn Jahren. Für deine Unterstützung und deinen Rat, aber vor allem für all die leichten Momente fernab des Büros, bei einem Drink, in Berlin, vor einer Bühne und für all die Momente, die noch kommen werden.

Merci à Gabrielle que j'ai connue en plein milieu de ma thèse et qui m'a directement accompagné de façon plus ou moins proche. Pour de nombreuses découvertes et une autre façon de voir le monde. Bientôt, le marathon, bientôt.

Merci à Adrien H., compagnon musical qui, selon ses dires, me jure qu'il ne comprend rien à ce que je fais mais a hâte de m'appeler Dr. Diana.

Merci à mes proches de façon plus générale, souvent dispersés loin de notre Grand Est, qui auraient certainement profité de voir mes cernes se creuser au fil du temps s'ils avaient pu. Merci à Antoine, Agathe, Léonor, Eloi, Pierre, Doriane, Charles, Romuald, Hugo, Adrien G., Léa, Andréa Cay., Lucie. Je commence à manquer de place, alors un dernier merci à tous mes amis.

Bonne lecture !

Liste des abréviations

| | |
|----------------|---|
| β 3GalT | β 1,3-galactosyltransférase |
| β 3GalT1 | β 1,3-galactosyltransférase 1 |
| β 3GalT6 | β 1,3-galactosyltransférase 6 |
| β 4GalT7 | β 1,4-galactosyltransférase 7 |
| 4-MOX | 4-méthylombélliferyl- β -D-xylopyranoside |
| 4Hyp | 4-hydroxyproline |
| 5Hyl | 5-hydroxylysine |
| ADAMTS-2 | <i>A disintegrin and metalloproteinase with thrombospondin motifs 2</i> |
| ADN | Acide désosyribonucléique |
| AFM | <i>Atomic force microscopy</i> |
| AH | Acide hyaluronique |
| ARN | Acide ribonucléique |
| ARNsg | Acide ribonucléique <i>single guide</i> |
| ASB | Albumine de sérum bovin |
| ATP | Adénosine triphosphate |
| B3GNT2 | β 1,3-N-acétylglucosaminyltransférase 2 |
| BCS | <i>Brittle Cornea Syndrome</i> |
| BMP-1 | <i>Bone morphogenic protein 1</i> |
| C1GalT1 | Glycoprotéine-N-acétylgalactosamine 3- β -galactosyltransférase 1 |
| CAZy | <i>Carbohydrate-Active enZYme</i> |
| CLHP | Chromatographie liquide à haute performance |
| CHO | <i>Chinese Hamster Ovary</i> |
| ChPF | <i>Chondroitin-polymerizing factor</i> |
| CHST14 | <i>Carbohydrate sulfotransferase 14</i> |
| ChSy | <i>Chondroitin sulfate synthase</i> |
| COMP | <i>Cartilage oligomeric matrix protein</i> |
| CS | Chondroïtine sulfate |
| CSGalNAcT | <i>Chondroitin sulfate N-acetylgalactosaminyltransferase</i> |
| DC | Dichroïsme circulaire |
| DLS | <i>Dynamic light scattering</i> |
| DMEM | <i>Dulbecco's Modified Eagle Medium</i> |
| DMEM / F-12 | <i>Dulbecco's Modified Eagle Medium Nutriment Mixture F-12</i> |
| dNTP | Désoxyribonucléotide triphosphate |
| DO | Densité optique |
| DS | Dermatane sulfate |

| | |
|---------------------|---|
| DSE | Dermatane sulfate épimérase |
| DTT | Dithiothréitol |
| DxD | Asp-x-Asp |
| <i>E. coli</i> | <i>Escherichia coli</i> |
| EDTA | Acide éthylènediaminetétraacétique |
| EXT1 | Exotosine 1 |
| EXT2 | Exotosine 2 |
| EXTL | Exotosine-like |
| FACIT | <i>Fibril-associated collagens with interrupted triple helices</i> |
| FAM20B | <i>FAMily with sequence similarity 20 member B</i> |
| GAG | Glycosaminoglycane |
| Gal | Galactose |
| Gal-Gal-Xyl-OMN | Galactose- β 1,3-galactose- β 1,4-xylose- β 1- <i>O</i> -méthoxynaphtyle |
| Gal-Gal-Xyl(2P)-OMN | Galactose- β 1,3-galactose- β 1,4-xylose-2- <i>O</i> -phosphate- β 1- <i>O</i> -méthoxynaphtyle |
| Gal-Xyl-OMN | Galactose- β 1,4-xylose- β 1- <i>O</i> -méthoxynaphtyle |
| Gal-Xyl(2P)-OMN | Galactose- β 1,4-xylose-2- <i>O</i> -phosphate- β 1- <i>O</i> -méthoxynaphtyle |
| GalNAc | <i>N</i> -acétylgalactosamine |
| GCNT1 | β 1,3-galactosyl- <i>O</i> -glycosyl-glycoprotéine β 1,6- <i>N</i> -acétylglucosaminyltransférase |
| GFP | <i>Green fluorescent protein</i> |
| GlcA | Acide glucuronique |
| GlcAT-I | β 1,3-glucuronosyltransférase 1 |
| GlcNAc | <i>N</i> -acétylglucosamine |
| GST | Glutathion <i>S</i> -transférase |
| GT | Glycosyltransférase |
| HA | Hémagglutinine |
| Hep | Héparine |
| HS | Héparane sulfate |
| HxH | His-x-His |
| IdoA | Acide iduronique |
| IPTG | Isopropyl β - <i>D</i> -1-thiogalactopyranoside |
| KS | Kératane sulfate |
| LB | <i>Lysogeny broth</i> |
| LH1 | Lysyl hydroxylase 1 |
| LH2 | Lysyl hydroxylase 2 |
| LH3 | Lysyl hydroxylase 3 |
| LOX | Lysyl oxidase |
| LOXL1 | Lysyl oxidase-like 1 |
| LOXL2 | Lysyl oxidase-like 2 |

| | |
|----------|--|
| LOXL3 | Lysyl oxidase- <i>like</i> 3 |
| LOXL4 | Lysyl oxidase- <i>like</i> 4 |
| LRR | <i>Leucine-rich repeat</i> |
| MBP | <i>Maltose binding protein</i> |
| MEC | Matrice extracellulaire |
| mTLD | <i>Tolloid-like proteinases</i> |
| P4H | Prolyl 4-hydroxylase |
| PBS | <i>Phosphate buffered saline</i> |
| PCR | <i>Polymerase chain reaction</i> |
| PG | Protéoglycane |
| PRELP | <i>Proline arginine-rich-end leucine-rich repeat protein</i> |
| PVDF | Difluorure de polyvinylidène |
| PXYLP1 | 2-phosphoxylose phosphatase 1 |
| RE | Réticulum endoplasmique |
| RMSF | <i>Root-mean square fluctuations</i> |
| SDS-PAGE | <i>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</i> |
| SED | Syndrome d'Ehlers-Danlos |
| SEDa | Syndrome d'Ehlers-Danlos arthrochaliasique |
| SEDC | Syndrome d'Ehlers-Danlos classique |
| SEDcl | Syndrome d'Ehlers-Danlos classique- <i>like</i> |
| SEDcl2 | Syndrome d'Ehlers-Danlos classique- <i>like</i> 2 |
| SEDCv | Syndrome d'Ehlers-Danlos cardiovasculaire |
| SEDD | Syndrome d'Ehlers-Danlos dermatosparaxique |
| SEDH | Syndrome d'Ehlers-Danlos hypermobile |
| SEDK | Syndrome d'Ehlers-Danlos cyphoscoliotique |
| SEDM | Syndrome d'Ehlers-Danlos myopathique |
| SEDMc | Syndrome d'Ehlers-Danlos musculocontractural |
| SEDP | Syndrome d'Ehlers-Danlos parodontal |
| SEDsp | Syndrome d'Ehlers-Danlos spondylodysplasique |
| SEDV | Syndrome d'Ehlers-Danlos vasculaire |
| SEMD-JL1 | <i>Spondyloepimetaphyseal dysplasia with joint laxity type 1</i> |
| SLRP | <i>Small leucine-rich proteoglycan</i> |
| SpsA | <i>Spore coat polysaccharide biosynthesis protein</i> |
| ST | Sulfotransférase |
| SVF | Sérum de veau fœtal |
| TBST | <i>Tris-buffered saline with Tween</i> |
| TC | Tissu conjonctif |
| TEV | <i>Tobacco etch virus</i> |

| | |
|----------|--|
| TGFβ | <i>Transforming growth factor beta</i> |
| UDP | Uridine diphosphate |
| UDP-Gal | Uridine diphosphate galactose |
| UDP-GlcA | Uridine diphosphate acide glucuronique |
| UDP-Xyl | Uridine diphosphate xylose |
| xED | x-Glu-Asp |
| Xyl | Xylose |
| XyIT-I | Xylosyltransférase 1 |
| XyIT-II | Xylosyltransférase 2 |

Les collagènes sont systématiquement abrégés sous la forme « Col xx », où « xx » désigne le type de collagène.

Les bases nucléotidiques sont désignées sous le système à une lettre.

Les acides aminés sont désignés sous le système à une ou trois lettres, selon le contexte.

Les mutations sont décrites avec la nomenclature clinique (c. ; p.) lorsqu'elles font référence à des patients ou à du matériel biologique de patients.

Liste des figures

| | |
|---|----|
| Figure 1. Classification des protéoglycanes. | 12 |
| Figure 2. Représentation schématique de la localisation de plusieurs protéoglycanes. | 13 |
| Figure 3. Schéma simplifié de la synthèse des chaînes de chondroïtine sulfate, dermatane sulfate et héparane sulfate à partir de l'amorce tétrasaccharidique. | 14 |
| Figure 4. Biosynthèse de l'amorce tétrasaccharidique des chaînes de glycosaminoglycanes. | 15 |
| Figure 5. Modèle ruban de la structure de la décorine bovine cristallisée représentant les répétitions riches en leucine. | 19 |
| Figure 6. Organisation structurale des chaînes polypeptidiques α formant les collagènes fibrillaires et des collagènes associés aux fibrilles avec des triples hélices interrompues. | 21 |
| Figure 7. Schéma de la structure protéique du collagène de type XII. | 24 |
| Figure 8. Schéma des grandes étapes de la fibrillogenèse. | 25 |
| Figure 9. Clivage des propeptides et initiation de l'assemblage des fibrilles de collagène. | 26 |
| Figure 10. Modèle ruban de la structure de la décorine bovine cristallisée représentant la séquence de fixation au collagène de type I. | 28 |
| Figure 11. Image du réseau collagénique de la cornée de bovin obtenue par microscopie à transmission électronique. | 29 |
| Figure 12. Mécanisme proposé de la régulation de la fibrillogenèse par le lumican et la fibromoduline. | 30 |
| Figure 13. Modèle proposé du mécanisme de l'interaction entre la protéine-core de l'ostéoadhérine et le collagène de type I lors de la fibrillogenèse. | 31 |
| Figure 14. Schéma des interactions du collagène de type XII dans la matrice extracellulaire. | 34 |
| Figure 15. Structures adoptées par les différentes familles de glycosyltransférases. | 39 |
| Figure 16. Structure du site actif de la β 1,4-galactosyltransférase 7 humaine en conformation ouverte. | 40 |
| Figure 17. Les structures des deux conformations de la β 1,4-galactosyltransférase humaine. | 44 |
| Figure 18. Prédiction de structure de la β 1,3-galactosyltransférase 6 humaine par le logiciel de prédiction AlphaFold. | 45 |
| Figure 19. Alignement des séquences des β 1,3-galactosyltransférases humaines. | 46 |
| Figure 20. Réaction enzymatique catalysée par la β 1,3-galactosyltransférase 6. | 47 |
| Figure 21. Séquences des motifs conservés de la famille des β 1,3-galactosyltransférases. | 49 |
| Figure 22. Schéma du site actif de la β 1,3-N-acétylglucosaminyltransférase 2 humaine. | 50 |

| | |
|---|----|
| Figure 23. Visualisation des domaines catalytiques de la structure résolue de la β 1,3-N-acétylglucosaminyltransférase 2 et d'un modèle de la β 1,3-galactosyltransférase 6. | 51 |
| Figure 24. Carte du vecteur d'expression pETM41 dans lequel a été clonée la séquence nucléotidique de la β 1,3-galactosyltransférase 6 humaine dont les vingt-neuf premiers acides aminés ont été supprimés. | 63 |
| Figure 25. Expression de la β 1,3-galactosyltransférase 6 sauvage ou mutée dans les cellules HeLa invalidées pour le gène <i>B3GALT6</i> . | 78 |
| Figure 26. Radiomarquage des chaînes de glycosaminoglycane synthétisées dans les cellules HeLa invalidées pour le gène <i>B3GALT6</i> à partir d'un substrat exogène après surexpression de la β 1,3-galactosyltransférase 6 sauvage ou mutée. | 79 |
| Figure 27. Analyse du taux de biosynthèse des chaînes de glycosaminoglycane de la décorine exprimée dans les cellules HeLa invalidées pour le gène <i>B3GALT6</i> après expression de la β 1,3-galactosyltransférase 6 sauvage ou mutée. | 81 |
| Figure 28. Profils chromatographiques des composés de synthèse utilisés pour la mesure de l'activité enzymatique de la β 1,3-galactosyltransférase 6. | 82 |
| Figure 29. Exemple de chromatogramme de l'activité de la β 1,3-galactosyltransférase 6 endogène. | 83 |
| Figure 30. Taux de déphosphorylation des composés de synthèse phosphorylés en fonction du pH lors de leur incubation avec les lysats cellulaires H7. | 84 |
| Figure 31. Effet de l'ajout d'adénosine triphosphate sur la déphosphorylation du substrat de synthèse de la réaction de la β 1,3-galactosyltransférase 6. | 85 |
| Figure 32. Représentation schématique de l'organisation de la β 1,3-galactosyltransférase 6 humaine et des protéines recombinantes produites dans le système bactérien. | 86 |
| Figure 33. Analyse de l'expression de la 6His-MBP- β 3GalT6 Δ N _{ter} 29-WT en fonction de la souche bactérienne utilisée. | 87 |
| Figure 34. Analyse de l'expression, de la solubilité et de la pureté des protéines recombinantes 6His-MBP- β 3GalT6 Δ N _{ter} 29-WT et 6His-MBP- β 3GalT6 Δ N _{ter} 50-WT par SDS-PAGE après coloration du gel au bleu de Coomassie. | 88 |
| Figure 35. Analyse de l'expression, de la solubilité et de la pureté des protéines recombinantes 6His-MBP- β 3GalT6 Δ N _{ter} 29-WT, 6His-MBP- β 3GalT6 Δ N _{ter} 29-Y182C, 6His-MBP- β 3GalT6 Δ N _{ter} 29-D207H et 6His-MBP- β 3GalT6 Δ N _{ter} 29-G217S par SDS-PAGE après coloration du gel au bleu de Coomassie. | 90 |
| Figure 36. Activité spécifique de la β 1,3-glucuronosyltransférase I recombinante purifiée. | 92 |

| | |
|--|------------|
| Figure 37. Profils d'intensité de diffusion dynamique de la lumière des β1,3-galactosyltransférases 6 humaines purifiées sauvage ou mutées. | 93 |
| Figure 38. Spectre de dichroïsme circulaire dans l'UV lointain des β1,3-galactosyltransférases 6 recombinantes sauvage et mutées. | 94 |
| Figure 39. Diagramme de la topologie du domaine catalytique de la β1,3-galactosyltransférase 6 humaine. | 95 |
| Figure 40. Séquence primaire et structures secondaires de la β1,3-galactosyltransférase 6. | 96 |
| Figure 41. Forces des liaisons hydrogènes et des ponts salins enregistrées lors des simulations de dynamique moléculaire de la β1,3-galactosyltransférase 6 sauvage et mutée. | 97 |
| Figure 42. Fluctuations moyennes des positions des résidus du domaine catalytique de la β1,3-galactosyltransférase 6 dans les modèles de simulations de dynamique moléculaire. | 98 |
| Figure 43. Immunoblot de la décorine sécrétée par les ATDC5 contrôles ou invalidées pour le gène <i>B3galt6</i>. | 100 |
| Figure 44. Analyse de la fraction guanidine extraite à partir des pseudo-tissus de cellules ATDC5 contrôles ou invalidées pour le gène <i>B3galt6</i>. | 101 |
| Figure 45. Analyses histologiques de coupes transversales de pseudo-tissus obtenus à partir des cellules ATDC5 contrôles ou invalidées pour le gène <i>B3galt6</i>. | 102 |
| Figure 46. Mesure de l'extensibilité des pseudo-tissus de cellules ATDC5 contrôles ou invalidées pour le gène <i>B3galt6</i>. | 102 |
| Figure 47. Analyses topographiques de pseudo-tissus ATDC5 contrôles ou invalidées pour le gène <i>B3galt6</i>. | 104 |
| Figure 48. Courbe conventionnelle de déformation d'un matériau selon l'intensité de la contrainte appliquée. | 105 |
| Figure 49. Analyses de la hauteur et de la rigidité des pseudo-tissus obtenus ATDC5. | 106 |

Liste des tableaux

| | |
|--|-----------|
| Tableau I. La classification internationale de 2017 des syndromes d'Ehlers-Danlos, dite de New-York. | 3 |
| Tableau II. Liste des mutations du gène B3GALT6 répertoriées chez l'Homme. | 7 |
| Tableau III. Liste des patients identifiés porteurs de mutations de B3GALT6 dans l'ordre chronologique de leur recensement. | 10 |
| Tableau IV. Structure et composition des chaînes de glycosaminoglycanes en fonction de la nature de la chaîne considérée. | 18 |
| Tableau V. Comparaison des résidus de la β1,3-N-acétylglucosaminyltransférase 2 impliqués dans la reconnaissance des substrats avec les résidus de la β1,3-galactosyltransférase 6. | 51 |
| Tableau VI. Liste des amorces nucléotidiques utilisées pour la mutagenèse dirigée de la β1,3-galactosyltransférase 6. | 57 |
| Tableau VII. Amorces nucléotidiques utilisées pour générer les tronctions de la β1,3-galactosyltransférase 6 humaine. | 63 |
| Tableau VIII. Bilan de purification de la β1,3-galactosyltransférase 6 humaine tronquée et fusionnée à la MBP. | 88 |
| Tableau IX. Paramètres cinétiques des protéines recombinantes 6His-MBP-β3GalT6ΔN_{ter}29-WT et 6His-MBP-β3GalT6ΔN_{ter}50-WT purifiées. | 89 |

Table des matières

REMERCIEMENTS

LISTE DES ABREVIATIONS

LISTE DES FIGURES

LISTE DES TABLEAUX

TABLE DES MATIERES

INTRODUCTION BIBLIOGRAPHIQUE.....1

| | | |
|------------|---|-----------|
| I. | LES SYNDROMES D'ÉHLERS-DANLOS, MALADIES GENETIQUES RARES DES TISSUS CONJONCTIFS | 1 |
| 1. | HISTORIQUE DES SYNDROMES D'ÉHLERS-DANLOS | 1 |
| 2. | CLASSIFICATION DES SYNDROMES D'ÉHLERS-DANLOS | 1 |
| 3. | LE SYNDROME D'ÉHLERS-DANLOS SPONDYLODYSPLASIQUE..... | 3 |
| a. | Plusieurs causes génétiques | 3 |
| b. | Une symptomatologie complexe | 4 |
| c. | Les difficultés diagnostiques | 5 |
| d. | Un nombre conséquent de mutations de <i>B3GALT6</i> répertoriées | 6 |
| II. | LES MATRICES EXTRACELLULAIRES..... | 11 |
| 1. | LES PROTEOGLYCANES | 11 |
| a. | Structure et classification..... | 12 |
| i. | La protéine- <i>core</i> , partie protéique des protéoglycanes | 12 |
| ii. | La différentes chaînes de glycosaminoglycanes, modifications post-traductionnelles glucidiques des protéoglycanes | 13 |
| b. | La structure des petits protéoglycanes riches en leucine..... | 18 |
| 2. | LES COLLAGENES | 20 |
| a. | Classification des collagènes | 20 |
| b. | Synthèse et maturation des collagènes fibrillaires | 21 |
| c. | La structure du collagène de type XII, collagène et protéoglycane à la fois | 22 |

| | | | |
|-------------|------|--|-----------|
| | i. | Structure générale | 22 |
| | ii. | Épissage alternatif du collagène de type XII | 22 |
| | iii. | Les chaînes polypeptidiques $\alpha 1(\text{XIIA})$ sont des protéoglycanes | 24 |
| 3. | | LA RELATION FONCTIONNELLE ENTRE LES PROTEOGLYCANES ET LES COLLAGENES..... | 25 |
| | a. | La fibrillogenèse | 25 |
| | i. | Les nucléateurs | 26 |
| | ii. | Les organisateurs | 27 |
| | iii. | Les régulateurs..... | 27 |
| | iv. | Conclusion..... | 34 |
| | b. | La matrice extracellulaire en situation pathologique | 35 |
| | i. | Défauts de la synthèse des collagènes fibrillaires | 35 |
| | ii. | Défauts de l'assemblage des fibres de collagène dus à des mutations du gène <i>COL12A1</i> | 36 |
| | iii. | Défauts de l'assemblage des fibres de collagène dus à des défauts de la synthèse des chaînes de glycosaminoglycanes..... | 36 |
| III. | | LA SUPERFAMILLE DES GLYCOSYLTRANSFERASES..... | 38 |
| 1. | | GENERALITES | 38 |
| | a. | Activité enzymatique | 38 |
| | b. | Structures et repliements | 38 |
| | i. | Les GT-A | 39 |
| | ii. | Les GT-B | 41 |
| | iii. | Les GT-C..... | 41 |
| | c. | Interactions entre des glycosyltransférases | 41 |
| | d. | L'organisation structurale des glycosyltransférases de la synthèse de l'amorce tétrasaccharidique | 42 |
| | i. | La xylosyltransférase I..... | 42 |
| | ii. | La $\beta 1,4$ -galactosyltransférase 7 | 43 |
| | iii. | La glucuronosyltransférase I | 44 |
| 2. | | LA $\beta 1,3$ -GALACTOSYLTRANSFERASE 6, ENZYME DE LA FAMILLE GT31 | 45 |
| | a. | Une enzyme conservée à travers les espèces..... | 46 |
| | b. | Activité enzymatique | 47 |
| | c. | Structure et repliement | 48 |
| | d. | Les $\beta 1,3$ -galactosyltransférases | 48 |
| | e. | Deux structures de GT31 résolues | 50 |
| | i. | La $\beta 1,3$ -N-acétylglucosaminyltransférase 2 humaine..... | 50 |

| | |
|--|-----------|
| ii. La glycoprotéine- <i>N</i> -acétylgalactosamine 3- β -galactosyltransférase 1 de <i>D. melanogaster</i> | 52 |
| 3. CONCLUSION | 52 |
| PROBLEMATIQUE | 53 |
| MATERIELS ET METHODES | 56 |
| I. ÉTUDE DE L'IMPACT DE MUTATIONS DE LA B1,3-GALACTOSYLTRANSFERASE 6 DANS UN MODELE CELLULAIRE INVALIDE POUR LE GENE <i>B3GALT6</i> | 56 |
| 1. LE MODELE CELLULAIRE HELA | 56 |
| a. Génération d'un modèle de cellules HeLa invalidé pour le gène <i>B3GALT6</i> | 56 |
| b. Culture cellulaire | 56 |
| 2. OBTENTION DES VECTEURS D'EXPRESSION DES MUTANTS DE LA B1,3-GALACTOSYLTRANSFERASE 6 HUMAINE..... | 57 |
| a. Mutagenèse dirigée | 57 |
| b. Préparation des vecteurs plasmidiques | 58 |
| 3. ÉTUDE DE LA SYNTHÈSE DES CHAINES DE GLYCOSAMINOGLYCANES | 58 |
| a. Synthèse à partir d'un substrat exogène | 58 |
| b. Synthèse à partir de la décorine recombinante..... | 59 |
| 4. DETERMINATION DE L'ACTIVITE ENZYMATIQUE DE LA B1,3-GALACTOSYLTRANSFERASE 6 A PARTIR D'HOMOGENATS CELLULAIRES | 60 |
| a. Préparation des homogénats cellulaires et dosage des protéines totales des échantillons | 60 |
| b. Analyse de l'expression de la β 1,3-galactosyltransférase 6 recombinante par immunoblot | 61 |
| c. Évaluation de l'activité galactosyltransférase par chromatographie liquide à haute performance | 61 |
| II. ÉTUDES DE LA FONCTION DES GLYCOSYLTRANSFERASES RECOMBINANTES IMPLIQUEES DANS LA SYNTHÈSE DE L'AMORCE TETRASACCHARIDIQUE | 62 |
| 1. LA B1,3-GALACTOSYLTRANSFERASE 6 RECOMBINANTE..... | 62 |
| a. Obtention des plasmides codant la 6His-MBP- β 3GalT6 Δ N _{ter} 29 et la 6His-MBP- β 3GalT6 Δ N _{ter} 50 | 62 |

| | | |
|-------------|--|-----------|
| b. | Obtention des plasmides codant les mutants de la 6His-MBP- β3GalT6ΔN _{ter} 29 | 64 |
| i. | Mutagenèse dirigée par PCR | 64 |
| ii. | Préparation de bactéries <i>Escherichia coli</i> électrocompétentes . | 64 |
| iii. | Amplification, extraction et purification des plasmides | 64 |
| c. | Production et purification des enzymes sauvage et mutées recombinantes | 65 |
| i. | Tests de surexpression des enzymes sauvages recombinantes .. | 65 |
| ii. | Production et purification des enzymes recombinantes sauvages et mutées | 66 |
| d. | Caractérisation des enzymes recombinantes purifiées | 67 |
| i. | Dosage protéique par la méthode de Bradford..... | 67 |
| ii. | Analyse des étapes de production et purification sur gel de polyacrylamide en conditions dénaturantes | 68 |
| iii. | Étude de la distribution de taille par diffusion dynamique de la lumière | 68 |
| iv. | Étude de la structure secondaire par dichroïsme circulaire | 68 |
| v. | Détermination des paramètres cinétiques des enzymes purifiées par chromatographie liquide à haute performance | 69 |
| 2. | LA B1,3-GLUCURONOSYLTRANSFERASE 1 RECOMBINANTE | 69 |
| a. | Production et purification de l'enzyme sauvage tronquée | 69 |
| b. | Détermination de l'activité enzymatique par bioluminescence | 69 |
| III. | SIMULATIONS DE DYNAMIQUE MOLECULAIRE DE LA B1,3-GALACTOSYLTRANSFERASE 6 SAUVAGE ET MUTEE | 70 |
| IV. | CARACTERISATION D'UN MODELE DE SYNTHESE MATRICIELLE MURIN INVALIDE POUR LE GENE B3GALT6 | 71 |
| 1. | CULTURE CELLULAIRE DES CELLULES ATDC5..... | 71 |
| a. | Matériel..... | 71 |
| b. | Stratégie d'inactivation du gène <i>B3galt6</i> par transduction virale | 72 |
| c. | Infection adénovirale | 72 |
| d. | Induction de la synthèse matricielle et formation des pseudo-tissus | 72 |
| 2. | ANALYSES BIOCHIMIQUES DE PSEUDO-TISSUS | 73 |
| a. | Récupération des milieux et précipitation du sécrétome..... | 73 |
| b. | Fractionnement protéique..... | 73 |
| c. | Analyse de l'expression protéique par immunoblot..... | 74 |

| | | |
|----|---|----|
| 3. | ANALYSES BIOMECANIQUES ET MORPHOLOGIQUES DES PSEUDO-TISSUS | 75 |
| a. | Coloration histologique..... | 75 |
| b. | Évaluation de l'extensibilité | 75 |
| c. | Caractérisation biomécanique et morphologique des pseudo-tissus par microscopie à force atomique | 75 |

RESULTATS EXPERIMENTAUX **77**

| | | |
|------------|---|-----------|
| I. | ÉTUDE DES CONSEQUENCES FONCTIONNELLES DE MUTATIONS PATHOGENES DE LA B1,3- GALACTOSYLTRANSFERASE 6 HUMAINE | 77 |
| 1. | UTILISATION D'UN MODELE CELLULAIRE EUCARYOTE INVALIDE POUR LE GENE <i>B3GALT6</i> . | 77 |
| a. | Analyse de l'expression de la β 1,3-galactosyltransférase 6 recombinante sauvage ou mutée | 77 |
| b. | Mesure du taux de biosynthèse des chaînes de glycosaminoglycanes après transfection dans les cellules invalidées pour le gène <i>B3GALT6</i> | 78 |
| i. | Synthèse à partir d'un xyloside exogène | 78 |
| ii. | Synthèse à partir de la protéine- <i>core</i> de la décorine | 80 |
| iii. | Conclusion..... | 81 |
| c. | Mesure de l'activité enzymatique de la β 1,3-galactosyltransférase 6 recombinante humaine | 82 |
| 2. | UTILISATION D'UN MODELE BACTERIEN POUR LA PRODUCTION ET LA PURIFICATION D'ENZYMES RECOMBINANTES | 86 |
| a. | Optimisation des conditions de production en système bactérien ... | 87 |
| b. | Détermination des effets de la troncation sur les propriétés enzymatiques..... | 88 |
| c. | Évaluation de l'effet de mutations sur l'activité enzymatique | 90 |
| 3. | EXPLORATION DU ROLE DE LA B1,3-GLUCURONOSYLTRANSFERASE 1 EN CAS DE DEFICIENCE EN B1,3-GALACTOSYLTRANSFERASE 6 | 91 |
| II. | ÉTUDE DES CONSEQUENCES STRUCTURALES DE MUTATIONS PATHOGENES DE LA B1,3- GALACTOSYLTRANSFERASE 6 | 92 |
| 1. | ÉTUDE DE LA DISTRIBUTION DE TAILLE DES ENZYMES RECOMBINANTES PAR DIFFUSION DYNAMIQUE DE LA LUMIERE | 92 |
| 2. | ÉTUDE DE LA STRUCTURE SECONDAIRE DES ENZYMES RECOMBINANTES SAUVAGE ET MUTEES PAR DICHROÏSME CIRCULAIRE..... | 93 |

| | | |
|-------------|---|-------------------|
| 3. | ÉTUDE DE LA DYNAMIQUE MOLECULAIRE DE LA B1,3-GALACTOSYLTRANSFERASE 6 SAUVAGE OU MUTEE..... | 94 |
| III. | CONSEQUENCES DE LA DEFICIENCE EN B1,3-GALACTOSYLTRANSFERASE 6 SUR LA MATRICE EXTRACELLULAIRE DANS UN MODELE CELLULAIRE DE SYNTHESE MATRICIELLE | 99 |
| 1. | ÉVALUATION DE LA SYNTHESE DES CHAINES DE GLYCOSAMINOGLYCANES PAR LES PSEUDO- TISSUS INVALIDES POUR LE GENE <i>B3GALT6</i> | 99 |
| 2. | ANALYSES DES PSEUDO-TISSUS INVALIDES POUR LE GENE <i>B3GALT6</i> | 101 |
| a. | Analyses histologiques | 101 |
| b. | Tests d'extensibilité | 102 |
| c. | Analyses morphologiques et biomécaniques par microscopie à force atomique..... | 103 |
| | <u>DISCUSSION</u> | <u>107</u> |
| | <u>CONCLUSIONS GENERALES ET PERSPECTIVES DE TRAVAIL</u> | <u>122</u> |
| | <u>REFERENCES BIBLIOGRAPHIQUES</u> | |
| | <u>ANNEXES</u> | |

Introduction bibliographique

I. Les syndromes d'Ehlers-Danlos, maladies génétiques rares des tissus conjonctifs

1. Historique des syndromes d'Ehlers-Danlos

Les syndromes d'Ehlers-Danlos (SED) représentent un groupe hétérogène de maladies génétiques rares des tissus conjonctifs (TC). La première mention de personnes présentant des symptômes de SED remonte à 400 av. J.-C. avec le *Traité des Airs, des Eaux et des Lieux* rédigé par **Hippocrate** qui observa une hyperlaxité articulaire chez des nomades Scythes.

Au fil du temps, divers médecins rapportèrent des cas de personnes aux caractéristiques cliniques similaires. Ainsi, en 1657, lors d'un congrès de l'Université de Leyde aux Pays-Bas, le chirurgien néerlandais **Job van Meekren** présenta les cas d'un marin espagnol à la peau hyperextensible et d'une contorsionniste aux articulations lâches (1). En 1892, le dermatologue russe **Alexandre Chernogubov** fournit une description détaillée de maladies des TC chez un adolescent de 17 ans présentant des dislocations articulaires à répétition, des nodules cutanés, une peau hyperextensible et fragile et de nombreuses cicatrices (2). En Russie, les SED sont encore connus sous le nom de syndrome de Chernogubov (3).

Ce n'est qu'au début du XX^{ème} siècle que deux autres dermatologues, le Danois **Edvard Ehlers** et le Français **Henri-Alexandre Danlos**, apportèrent des contributions significatives à la compréhension de ces maladies. Ehlers décrit le cas d'un jeune homme de 21 ans souffrant notamment de *cutis laxa* (du latin « peau lâche », peau hyperextensible et ridée), de luxations répétées et d'hématomes fréquents (4). De son côté, Danlos étudia un individu très mince présentant également une *cutis laxa* (5). Dans sa thèse de médecine en 1933, intitulée *Les syndromes d'Ehlers-Danlos*, **Achille Miget** fut le premier à utiliser le terme SED pour désigner ces maladies (6).

2. Classification des syndromes d'Ehlers-Danlos

Les SED sont dus à des mutations de gènes codant pour des protéines composant les matrices extracellulaires (MEC), en premier lieu les collagènes puis des enzymes participant à la maturation de ces protéines. Sur la base de descriptions cliniques, plusieurs classifications des SED ont été établies pour accompagner le clinicien dans son diagnostic différentiel des patients, avant confirmation par l'identification du gène muté. Deux premières classifications

de SED se sont succédées : la nosologie de Berlin (7) et la nosologie de Villefranche (8). Depuis 2017, la nosologie internationale de New-York (9) est la classification en vigueur.

Cette classification divise les SED en treize sous-types qui diffèrent selon les gènes affectés et les manifestations cliniques. Il est important de noter que la classification des SED continue d'évoluer et qu'un quatorzième sous-type a été proposé : le SED classique-like 2 (SEDcl2) (10) (Tableau I).

| Sous-type clinique de syndrome d'Ehlers-Danlos (SED) | Gène(s) | Protéine(s) | Critères cliniques majeurs | Mode de transmission |
|--|---|-----------------------------|--|------------------------------------|
| SED classique (SEDc) | <i>COL5A1</i> <i>COL5A2</i> <i>COL1A1</i> | Col V Col I | Hyperextensibilité de la peau Cicatrices atrophiques Hypermobilité générale des articulations | Autosomique dominante |
| SED vasculaire (SEDv) | <i>COL3A1</i> <i>COL1A1</i> | Col III Col I | Rupture précoce des artères, de la paroi du côlon ou de l'utérus | Autosomique dominante |
| SED arthrochalasique (SEDa) | <i>COL1A1</i> <i>COL1A2</i> | Col I | Dislocation congénitale et bilatérale du bassin Hypermobilité sévère des articulations avec dislocations multiples Hyperextensibilité de la peau | Autosomique dominante |
| SED dermatosparaxique (SEEd) | <i>ADAMTS2</i> | ADAMTS-2 | Fragilité extrême de la peau Croissance postnatale retardée | Autosomique récessive |
| SED cardiovasculaire (SEDcv) | <i>COL1A2</i> | Col I | Insuffisance cardiaque progressive et sévère Hypermobilité des articulations | Autosomique récessive |
| SED cyphoscoliotique (SEDk) | <i>PLOD1</i> <i>FKBP14</i> | LH1 FKBP22 | Hypotonie des muscles Cyphoscoliose précoce Hypermobilité des articulations avec (sub)luxations | Autosomique récessive |
| SED classique-like (SEDcl) | <i>TNXB</i> | Ténascine X | Hyperextensibilité de la peau Peau duveteuse sans cicatrices atrophiques Hypermobilité généralisée des articulations Ecchymoses | Autosomique récessive |
| SED classique-like 2 (SEDcl2)* | <i>AEBP1</i> | ACLp | Hyperextensibilité de la peau Cicatrices atrophiques Hypermobilité généralisée des articulations Ostéopénie précoce | Autosomique récessive |
| SED myopathique (SEEdm) | <i>COL12A1</i> | Col XII | Hypotonie et/ou atrophie des muscles Hypermobilité des articulations distales | Autosomique dominante ou récessive |
| SED spondylodysplasique (SEDsp)[§] | <i>B4GALT7</i> <i>B3GALT6</i> <i>SLC39A13</i> | β4GalT7 β3GalT6 ZIP13 | Petite taille Hypotonie des muscles Déformation des membres | Autosomique récessive |

| | | | | |
|--|-------------------------------|-------------------|---|-----------------------|
| SED musculocontractural (SEDmc) | <i>CHST14</i> <i>DSE</i> | CHST14 DSE | Multiples contractures congénitales Dysmorphie faciale | Autosomique récessive |
| SED parodontal (SEDp) | <i>C1R</i> <i>C1S</i> | C1r C1s | Rétractation des gencives Parodontite | Autosomique dominante |
| Syndrome de la cornée fragile (<i>Brittle Cornea Syndrome</i>, BCS) | <i>ZNF469</i> <i>PRDM5</i> | ZNF469 PRDM5 | Cornée fine et fragile Kératocône précoce Sclérotique bleue | Autosomique récessive |
| SED hypermobile (SEDh) | Non identifié(s) | Non identifiée(s) | Hypermobilité généralisée des articulations | Autosomique dominante |

Tableau I. La classification internationale de 2017 des syndromes d'Ehlers-Danlos, dite de New-York. * Sous-type décrit après la publication de la classification actuelle (10). [§] Sous-type faisant l'objet de ce travail de thèse. ADAMTS-2 : *a disintegrin and metalloproteinase with thrombospondin motifs 2*. LH1 : lysyl hydroxylase 1. β 4GalT7 : β 1,4-galactosyltransférase 7. β 3GalT6 : β 1,3-galactosyltransférase 6. CHST14 : *carbohydrate sulfotransferase 14*. DSE : dermatane sulfate épimérase. Adapté de (9).

3. Le syndrome d'Ehlers-Danlos spondylodysplasique

Le syndrome d'Ehlers-Danlos spondylodysplasique (SEDsp) fut initialement identifié comme une forme de SED associée à des caractéristiques progéroïdes, c'est-à-dire un aspect de vieillissement prématuré de la peau. Entre 1979 et 1986, cinq individus présentant des aspects inhabituels tels que des rides faciales, des sourcils et cheveux fins ainsi qu'un déficit intellectuel furent étudiés. Leurs symptômes différant des manifestations typiques du SED, les médecins les regroupèrent dans la nosologie de Villefranche (8) sous le nom de SED progéroïde (11–13).

a. Plusieurs causes génétiques

Les protéoglycanes (PG) sont des glycoprotéines particulières et composants majeurs des MEC. Ils sont constitués d'une partie protéique appelée protéine-*core* et d'une longue chaîne glucidique linéaire appelée glycosaminoglycane (GAG) dont la structure est détaillée dans la section II.1.a de cette introduction bibliographique.

En 1987, une équipe de chercheurs étudia le cas d'un patient de petite taille présentant des anomalies cutanées et squelettiques. Par immunoblot, ils observèrent que les fibroblastes du patient présentaient une déficience de la biosynthèse de la chaîne de GAG d'un PG (14). Des recherches ultérieures sur les cellules de ce même patient révélèrent une activité déficiente d'une des enzymes de la biosynthèse des chaînes de GAG, la β 1,4-galactosyltransférase 7 (β 4GalT7) (EC 2.4.1.133) (15). Cette diminution significative d'activité

enzymatique était due aux mutations hétérozygotes composites p.[(Ala186Asp) ; (Leu206Pro)] du gène *B4GALT7* codant la β 4GalT7 (16).

Par la suite, des mutations du gène *B3GALT6* codant pour l'enzyme catalysant la réaction suivante dans la voie de biosynthèse des chaînes de GAG, la β 1,3-galactosyltransférase 6 (β 3GalT6) (EC 2.4.1.134), ont été identifiées (17, 18). Ces mutations, ainsi que celles du gène *B4GALT7*, sont désormais regroupées dans un même sous-type de SED : le **SEDsp dû à des défauts de la biosynthèse des chaînes de GAG** (9). Dans notre travail, nous nous concentrerons sur l'exploration des anomalies fonctionnelles de la β 3GalT6.

Enfin, le SEDsp regroupe également des mutations du gène *SLC39A13*, qui code pour un transporteur du zinc de l'appareil de Golgi appelé ZIP13 (19). Jusqu'à présent, seulement trois mutations de ce gène ont été identifiées. Ce **SEDsp dû à des défauts de processus intracellulaires** est associé à une sous-hydroxylation des résidus Lys et Pro des collagènes, ce qui entraîne un mauvais assemblage de ces derniers (20, 21).

Ces trois formes sont désignées sous les noms de SEDsp-*B4GALT7*, SEDsp-*B3GALT6* et SEDsp-*SLC39A13* (9).

b. Une symptomatologie complexe

En plus des symptômes majeurs communs à tous les types de SED (peau hyperextensible, articulations hypermobiles, fragilité généralisée des tissus) (9, 22), les patients atteints de SEDsp présentent également une **petite taille** apparaissant progressivement pendant l'enfance, une **hypotonie musculaire** congénitale sévère ou apparaissant tardivement de façon plus légère et une **déformation des membres**.

Des symptômes mineurs de SEDsp peuvent être présents à différents degrés de sévérité, tels qu'une peau douce, fine et duveteuse, des pieds plats, un retard du développement moteur, une ostéopénie, des os moins denses et un retard du développement cognitif.

Les symptômes mineurs mais spécifiques du SEDsp-*B3GALT6* comprennent une cyphoscoliose congénitale ou apparaissant tôt et progressive, une hypermobilité des articulations généralisée ou restreinte aux articulations distales avec une dislocation des articulations, une contracture congénitale ou progressive des articulations surtout au niveau des mains, des doigts fins coniques et arachnodactyles (allongement métacarpo-phalangien avec hyperlaxité articulaire) avec de larges phalanges distales, un *talipes equinovarus* (ou pied-bot), des dysmorphies faciales, une décoloration ou une dysplasie dentaire et des anomalies

radiologiques caractéristiques (platyspondylie, ilium court, dysplasie acétabulaire, dysplasie métaphysaire des têtes fémorales).

Enfin, la douleur chronique (23, 24) est un symptôme présent à différents niveaux chez les patients atteints de SEDsp. Malfait *et al.* (25) précisait en 2021 qu'aucune étude expérimentale de douleur sur des individus souffrant d'un SED confirmé n'avait été entreprise. Bien que la prise d'analgésiques et d'antalgiques soit recommandée, elle ne procure pas de résultats satisfaisants en termes de qualité de vie pour les patients (26, 27).

c. Les difficultés diagnostiques

Pour diagnostiquer un SEDsp, la classification internationale de New-York recommande de constater la présence d'au moins deux critères majeurs ainsi qu'au moins quatre critères mineurs (dont les anomalies caractéristiques radiologiques) (9). Le séquençage des gènes *B4GALT7*, *B3GALT6* et *SLC39A13* est ensuite nécessaire pour confirmer le diagnostic clinique. Cependant, d'autres sous-types de SED présentent des symptômes similaires. Par conséquent, le consortium international des SED recommande d'inclure d'autres gènes dans le test génétique. Il s'agit des gènes *PLOD1* et *FKBP14* qui sont associés au SED cyphoscoliotique (SEdk), des gènes *ZNF469* et *PRDM5* qui sont liés au syndrome de la cornée fragile (*Brittle Cornea Syndrome*, BCS) et des gènes *CHST14* et *DSE* qui sont impliqués dans le SED musculocontractural (SEDmc) (9).

Il faut noter que les mutations du gène *B3GALT6* peuvent produire un large spectre clinique. L'identification des premiers variants alléliques pathogènes de *B3GALT6* par deux équipes de recherche distinctes a permis le diagnostic du **SEDsp-B3GALT6** chez huit patients issus de cinq familles différentes (17, 18) et de la **dysplasie spondyloépimétaphysaire avec hyperlaxité ligamentaire de type 1** (*spondyloepimetaphyseal dysplasia with joint laxity type 1*, SEMD-JL1) chez neuf patients issus de huit autres familles (18, 28). A ce jour, on dénombre vingt-quatre patients atteints de SEDsp-B3GALT6 et trente-et-un patients atteints de SEMD-JL1. De plus, une troisième forme rare a été identifiée chez trois patients supplémentaires : le **syndrome d'Al-Gazali** (29) (Tableau III).

La SEMD-JL1 a été décrite pour la première fois par Beighton et Kozlowski (30). Il s'agit d'une pathologie aux symptômes extrêmement similaires au SEDsp-B3GALT6, parmi lesquels on trouve parfois une hyperlaxité articulaire surtout au niveau des mains, une cyphoscoliose

sévère et progressive, une fente palatine, des yeux proéminents avec sclérotique bleue ou encore un pied-bot (31). Nakajima *et al.* (18) ont mené une étude comparative de trois patients diagnostiqués avec le SEDsp-*B3GALT6* et neuf patients diagnostiqués avec la SEMD-JL1. Bien que le tableau clinique se chevauche entre les deux maladies, certains critères spécifiques ont pu être mis en évidence. Par exemple, les patients atteints de SEDsp-*B3GALT6* ont généralement des atteintes squelettiques et cutanées plus prononcées.

Le syndrome d'Al-Gazali a été identifié chez trois patients, dont deux sont décédés peu après la naissance (32). Les symptômes présentent des similitudes avec le SEDsp-*B3GALT6*, tels que les dislocations articulaires, le pied-bot, les atteintes des mains, la fragilité osseuse et les symptômes oculaires (29).

Dans ce contexte, l'hétérogénéité au niveau de l'expression clinique peut entraîner des incertitudes et/ou des erreurs diagnostiques de la part des cliniciens (33–38), ce qui nuit à la prise en charge des patients et à leur état psychologique relatif à l'errance diagnostique qu'ils subissent (39–41).

d. Un nombre conséquent de mutations de *B3GALT6* répertoriées

Depuis la première description des mutations de *B3GALT6* en 2013, quarante-sept mutations ponctuelles ont été recensées chez soixante-huit patients issus de cinquante-deux familles différentes (17, 18, 28, 29, 33–38, 42–52) (Tableaux II et III).

| Mutation sur la séquence nucléotidique | Mutation sur la séquence protéique | Nombre de patients porteurs de la mutation (homozygote ou hétérozygote composite) | Première description par | Mutation également étudiée par |
|--|------------------------------------|---|--------------------------|--------------------------------|
| c.1A>G | p.(Met1?) | 5 | (18) | / |
| c.3G>A | p.? | 1 | (50) | / |
| c.16C>T | p.(Arg6Trp) | 3 | (18) | (48, 51) |
| c.77T>C | p.(Leu26Pro) | 2 | (50) | (33) |
| c.181G>C | p.(Val61Leu) | 1 | (50) | / |
| c.193A>G | p.(Ser65Gly) | 1 | (18) | / |
| c.197_253del | p.(Arg66_Arg84del) | 3 | (50) | / |
| c.200C>T | p.(Pro67Leu) | 7 | (18) | (45, 51) |
| c.227delT | p.(Ile76Thrfs*202) | 2 | (35) | / |
| c.235A>G | p.(Thr79Ala) | 9 | (51) | (33, 45) |
| c.308C>T | p.(Ala103Val) | 1 | (28) | / |
| c.323_344del | p.(Ala108Glyfs*163) | 2 | (17) | / |

| | | | | |
|-------------------|----------------------|---|------|--------------|
| c.353delA | p.(Asp118Alafs*160) | 2 | (18) | (28) |
| c.415_423del | p.(Met139_Ala141del) | 1 | (18) | / |
| c.430G>A | p.(Asp144Asn) | 1 | (50) | / |
| c.466G>A | p.(Asp156Asn) | 1 | (18) | / |
| c.? | p.(Asp159Tyr) | 1 | (34) | / |
| c.477del | p.(Phe160Serfs*118) | 1 | (50) | / |
| c.510_517del | p.(Leu170fs*268) | 5 | (37) | / |
| c.511C>T | p.(Arg171Cys) | 4 | (49) | (47) |
| c.513_520del | p.(Glu174Alafs*266) | 4 | (50) | / |
| c.536_541dup | p.(Arg179_Arg180dup) | 1 | (42) | (46) |
| c.539_540insCCT | p.? | 1 | (52) | / |
| c.545A>G | p.(Tyr182Cys) | 1 | (50) | / |
| c.? | p.(Phe185Cys) | 1 | (43) | / |
| c.556T>C | p.(Phe186Leu) | 4 | (42) | (46) |
| c.589C>T | p.(Arg197Cys) | 1 | (36) | / |
| c.588dup | p.(Arg197Alafs*246) | 1 | (48) | / |
| c.588del | p.(Arg197Alafs*81) | 2 | (18) | / |
| c.618C>G | p.(Cys206Trp) | 3 | (29) | / |
| c.619G>C | p.(Asp207His) | 4 | (17) | / |
| c.631C>T | p.(Pro211Ser) | 2 | (50) | (33) |
| c.649G>A | p.(Gly217Ser) | 1 | (17) | / |
| c.694C>T | p.(Arg232Cys) | 6 | (18) | (43, 52) |
| c.766C>T | p.(Arg256Trp) | 2 | (35) | / |
| c.782G>A | p.(Arg261His) | 1 | (50) | / |
| c.795A>C | p.(Glu265Asp) | 3 | (34) | (33, 38, 50) |
| c.? | p.(Gly270Ser) | 2 | (50) | (44) |
| c.845_846delinsTA | p.(Ser282Ile) | 1 | (47) | / |
| c.883C>T | p.(Arg295Cys) | 3 | (37) | / |
| c.899G>C | p.(Cys300Ser) | 1 | (18) | / |
| c.901_902dup | p.(Lys301_307dup) | 4 | (49) | (47) |
| c.925T>A | p.(Ser309Thr) | 3 | (18) | (28, 50) |
| c.929A>G | p.(Tyr310Cys) | 1 | (38) | (50) |
| c.942G>C | p.(Trp314Cys) | 1 | (44) | / |
| c.953C>T | p.(Pro318Leu) | 3 | (50) | / |
| c.987_989delCTG | p.(*330Alaext*73) | 1 | (28) | / |

Tableau II. Liste des mutations du gène *B3GALT6* répertoriées chez l'Homme. Lorsque les informations sont disponibles, les mutations sont décrites aux niveaux nucléotidique et protéique. S'il est connu, le nombre de patients porteurs de la mutation est précisé.

Parmi les quarante-sept mutations présentées dans le tableau, dix mutations sont retrouvées à l'état homozygote chez dix-neuf patients. A l'exception de la mutation p.(Arg179_180dup), qui consiste en une insertion de deux nucléotides (42), on retrouve très souvent des mutations par substitution : p.(Thr79Ala) (33, 51), p.(Tyr182Cys) (50), p.(Phe186Leu) (42), p.(Arg197Cys) (36), p.(Cys206Trp) (29), p.(Asp207His), p.(Gly217Ser) (17), p.(Glu265Asp) (33) et p.(Ser282Ile) (47). La majorité des patients atteints sont donc porteurs

de mutations hétérozygotes composites avec une grande hétérogénéité au niveau des variants (Tableau III).

| Désignation du patient | Mutation sur la séquence nucléotidique | Mutation sur la séquence protéique | Diagnostic clinique proposé | Référence |
|------------------------|--|--------------------------------------|---|-----------|
| Family 1 P1 | c.619G>C c.619G>C | p.(Asp207His) p.(Asp207His) | SEDsp-B3GALT6 | (17) |
| Family 1 P2 | c.619G>C c.619G>C | p.(Asp207His) p.(Asp207His) | SEDsp-B3GALT6 | |
| Family 2 P3 | c.323_344del c.619G>C | p.(Ala108Glyfs*163) p.(Asp207His) | SEDsp-B3GALT6 | |
| Family 2 P4 | c.323_344del c.619G>C | p.(Ala108Glyfs*163) p.(Asp207His) | SEDsp-B3GALT6 | |
| Family 3 P5 | c.649G>A c.649G>A | p.(Gly217Ser) p.(Gly217Ser) | SEDsp-B3GALT6 | |
| F1 P1 | c.1A>G c.694C>T | p.(Met1?) p.(Arg232Cys) | SEMD-JL1 | (18, 28) |
| F1 P2 | c.1A>G c.694C>T | p.(Met1?) p.(Arg232Cys) | SEMD-JL1 | |
| F2 P3 | c.1A>G c.466G>A | p.(Met1?) p.(Asp156Asn) | SEMD-JL1 | |
| F3 P4 | c.1A>G WT | p.(Met1?) WT | SEMD-JL1 | |
| F4 P5 | c.1A>G c.694C>T | p.(Met1?) p.(Arg232Cys) | SEMD-JL1 | |
| F5 P6 | c.694C>T c.899G>C | p.(Arg232Cys) p.(Cys300Ser) | SEMD-JL1 | |
| F6 P7 | c.1A>G c.193A>G | p.(Met1?) p.(Ser65Gly) | SEMD-JL1 | |
| F7 P8 | c.200C>T c.694C>T | p.(Pro67Leu) p.(Arg232Cys) | SEMD-JL1 | |
| F8 P9 | c.353delA c.925T>A | p.(Asp188Alafs*160) p.(Ser309Thr) | SEMD-JL1 | |
| F9 P10 | c.588delG c.925T>A | p.(Arg197Alafs*81) p.(Ser309Thr) | SEDsp-B3GALT6 | |
| F9 P11 | c.588delG c.925T>A | p.(Arg197Alafs*81) p.(Ser309Thr) | SEDsp-B3GALT6 | |
| F10 P12 | c.16C>T c.415_423del | p.(Arg6Trp) p.(Met139Ala141del) | SEDsp-B3GALT6 | |
| Patient | c.? c.795A>C | p.(Ser159Tyr) p.(Glu265Asp) | SEDsp-B3GALT6 SEMD-JL1 Syndrome d'Al-Gazali | (34) |
| F1 | c.16C>T c.200C>T | p.(Arg6Trp) p.(Pro67Leu) | SEMD-JL1 | (51) |
| F3 | c.235A>G c.235A>G | p.(Thr79Ala) p.(Thr79Ala) | SEMD-JL1 | |
| F4 | c.200C>T c.235A>G | p.(Pro67Leu) p.(Thr79Ala) | SEMD-JL1 | |
| F5 | c.200C>T c.235A>G | p.(Pro67Leu) p.(Thr79Ala) | SEMD-JL1 | |
| F8 | c.235A>G c.235A>G | p.(Thr79Ala) p.(Thr79Ala) | SEMD-JL1 | |

| | | | | |
|-------------|--|--|---------------------------|----------|
| F10 | c.235A>G c.235A>G | p.(Thr79Ala) p.(Thr79Ala) | SEMD-JL1 | |
| F1 P1 | c.227del c.766C>T | p.(Ile76Thrfs*202) p.(Arg256Trp) | SEDsp-B3GALT6 SEMD-JL1 | (35) |
| F1 P2 | c.227del c.766C>T | p.(Ile76Thrfs*202) p.(Arg256Trp) | SEDsp-B3GALT6 SEMD-JL1 | |
| Family 6 P1 | c.556T>C c.556T>C | p.(Phe186Leu) p.(Phe186Leu) | SEMD-JL1 | (42, 46) |
| Family 7 P1 | c.556T>C c.556T>C | p.(Phe186Leu) p.(Phe186Leu) | SEMD-JL1 | |
| Family 7 P2 | c.556T>C c.556T>C | p.(Phe186Leu) p.(Phe186Leu) | SEMD-JL1 | |
| Family 8 P1 | c.556T>C c.556T>C | p.(Phe186Leu) p.(Phe186Leu) | SEDsp-B3GALT6 | |
| Family 9 P1 | c.536_541dupGCCGCC c.536_541dupGCCGCC | p.(Ag179_Arg180dup) p.(Ag179_Arg180dup) | SEMD-JL1 | |
| F1 P1 | c.200C>T c.235A>G | p.(Pro67Leu) p.(Thr79Ala) | SEMD-JL1 | (45) |
| F1 P2 | c.200C>T c.235A>G | p.(Pro67Leu) p.(Thr79Ala) | SEMD-JL1 | |
| F2 P1 | c.200C>T c.235A>G | p.(Pro67Leu) p.(Thr79Ala) | SEMD-JL1 | |
| P7 | c.845_846delinsTA c.845_846delinsTA | p.(Ser282Ile) p.(Ser282Ile) | SEMD-JL1 | (47) |
| P8 | c.901_921dup c.511C>T | p.(Lys301_Arg307dup) p.(Arg171Cys) | / | |
| F1 IV-2 | c.511C>T c.901_921dup | p.(Arg171Cys) p.(Lys301_307dup) | SEMD-JL1 | (49) |
| F1 IV-3 | c.511C>T c.901_921dup | p.(Arg171Cys) p.(Lys301_307dup) | SEMD-JL1 | |
| F1 IV-5 | c.511C>T c.901_921dup | p.(Arg171Cys) p.(Lys301_307dup) | SEMD-JL1 | |
| Patient 1 | c.618G>C c.618G>C | p.(Cys206Trp) p.(Cys206Trp) | Syndrome d'Al-Gazali | (29) |
| Patient 2 | c.618G>C c.618G>C | p.(Cys206Trp) p.(Cys206Trp) | Syndrome d'Al-Gazali | |
| Patient 3 | c.618G>C c.618G>C | p.(Cys206Trp) p.(Cys206Trp) | Syndrome d'Al-Gazali | |
| PI:1 | c.795A>C c.929A>G | p.(Glu265Asp) p.(Tyr310Cys) | SEDsp-B3GALT6 | (38, 50) |
| PII:1 | c.513_520del c.925T>A | p.(Glu174Alafs*266) p.(Ser309Thr) | SEDsp-B3GALT6 | |
| PIII:1 | c.477del c.782G>A | p.(Phe160Serfs*118) p.(Arg261His) | SEDsp-B3GALT6 | |
| PIV:1 | c.197_253del c.953C>T | p.(Ala66_Arg84del) p.(Pro318Leu) | SEDsp-B3GALT6 | |
| PIV:2 | c.197_253del c.953C>T | p.(Ala66_Arg84del) p.(Pro318Leu) | SEDsp-B3GALT6 | |
| PIV:3 | c.197_253del c.953C>T | p.(Ala66_Arg84del) p.(Pro318Leu) | SEDsp-B3GALT6 | |
| PV:1 | c.3G>A c.631C>T | p.? p.(Pro211Ser) | SEDsp-B3GALT6 | |
| PVI:1 | c.513_520del c.430G>A | p.(Glu174Alafs*266) p.(Asp144Asn) | SEDsp-B3GALT6 | |

| | | | | |
|------------|---|---|---|------|
| PVII:1 | c.545A>G c.545A>G | p.(Tyr182Cys) p.(Tyr182Cys) | SEMD-JL1 | |
| PVIII:1 | c.77T>C c.808G>A | p.(Leu26Pro) p.(Gly270Ser) | SEDsp-B3GALT6 | |
| PIX:1 | c.513_520del c.181G>C | p.(Glu174Alafs*266) p.(Val61Leu) | SEDsp-B3GALT6 | |
| PIX:2 | c.513_520del c.181G>C | p.(Glu174Alafs*266) p.(Val61Leu) | SEDsp-B3GALT6 | |
| Patient 3 | c.[308C>T ; 353delA] c.987_989delCTG | p.[(Ala103Val) ; (Asp118Alafs*160)] p.(*330Alaext*73) | SEDsp-B3GALT6 | (28) |
| Patient | c.694C>T c.539_540insCCT | / | SEMD-JL1 | (52) |
| Patient | / | p.(Arg232Cys) p.(Phe185Cys) | SEDsp-B3GALT6 | (43) |
| Patient 67 | c.16C>T c.588dup | p.(Arg6Trp) p.(Arg197Alafs*246) | SEMD-JL1 | (48) |
| Patient | c.589C>T c.589C>T | p.(Arg197Cys) p.(Arg197Cys) | SEDsp-B3GALT6 SEMD-JL1 | (36) |
| Patient | c.808G>A c.942G>C | p.(Gly270Ser) p.(Trp314Cys) | SEMD-JL1 | (44) |
| I:1 | WT c.510_517del | WT p.(Leu170fs*268) | / | (37) |
| II:1 | WT c.510_517del | WT p.(Leu170fs*268) | / | |
| II:2 | WT c.510_517del | WT p.(Leu170fs*268) | / | |
| III:1 | c.883C>T c.510_517del | p.(Arg295Cys) p.(Leu170fs*268) | SEDsp-B3GALT6 | |
| III:2 | c.883C>T c.510_517del | p.(Arg295Cys) p.(Leu170fs*268) | SEDsp-B3GALT6 | |
| Patient 1 | c.235A>G c.235A>G | p.(Thr79Ala) p.(Thr79Ala) | SEMD-JL1 | (33) |
| Patient 2 | c.77T>C c.631C>T | p.(Leu26Pro) p.(Pro211Ser) | / | |
| Patient 3 | c.795A>C c.795A>C | p.(Glu265Asp) p.(Glu265Asp) | SEDsp-B3GALT6 SEMD-JL1 Syndrome d'Al-Gazali | |

Tableau III. Liste des patients identifiés porteurs de mutations de B3GALT6 dans l'ordre chronologique de leur recensement. Lorsque les informations sont disponibles, les mutations homozygotes ou hétérozygotes composites sont décrites aux niveaux nucléotidique et protéique. Si un diagnostic a été proposé, il est également mentionné. Le SEDsp-B3GALT6 a été attribué à vingt-quatre patients (en vert), le SEMD-JL1 à trente-et-un patients (en bleu) et le syndrome d'Al-Gazali à trois patients (en rouge). Pour les dix derniers patients, le diagnostic est incertain (en noir).

II. Les matrices extracellulaires

Les SED sont des maladies qui causent une désorganisation des MEC et un dérèglement de leur homéostasie. Il convient donc de s'intéresser de plus près à ces réseaux tridimensionnels ordonnés et complexes qui entourent les cellules et qui sont principalement constitués de trois types de macromolécules : les collagènes, les fibres élastiques et les glycoprotéines. Dans les TC, ces dernières sont surtout représentées par les PG. Les MEC sont le siège de nombreux processus physiologiques essentiels et complexes tels que la prolifération, la différenciation et la migration cellulaire (53).

Les connaissances sur la composition diverse et la complexité structurale des MEC ont été compilées par Karamanos *et al.* (54). Pour les besoins de notre étude, cette section se concentrera sur les voies de biosynthèse et les structures des PG et des collagènes, les deux macromolécules les plus abondantes des MEC. Leurs rôles dans l'organisation et l'architecture de la MEC seront explorés. Par la suite, nous évoquerons la relation fonctionnelle entre ces deux familles afin de mieux comprendre les mécanismes moléculaires impliqués au cours des maladies génétiques rares des TC telles que les SED (55).

1. Les protéoglycanes

Les PG sont des macromolécules ubiquitaires principalement retrouvées dans les MEC ou intégrées aux membranes plasmiques des cellules. Ils se composent d'une partie protéique liée de manière covalente à une ou plusieurs chaînes polysaccharidiques linéaires appelées GAG (56).

La synthèse des chaînes de GAG est initiée dans les compartiments tardifs du réticulum endoplasmique (RE) puis se poursuit dans l'appareil de Golgi. Les GAG résultent d'un processus de maturation post-traductionnel de la protéine *core* des PG réalisé par différentes enzymes de biosynthèse, en premier lieu des glycosyltransférases (GT). Cette glycosylation est complétée par l'action de sulfotransférases (ST), d'une kinase et d'une phosphatase.

Il existe une classification des PG établie selon leur localisation, leur degré d'homologie génique ou protéique, la présence potentielle de modules protéiques spécifiques et la nature de leurs chaînes de GAG (57). La diversité des PG se reflète dans l'éventail de fonctions qu'ils assurent au sein des cellules ou tissus, ce qui complexifie leur étude (Figure 1).

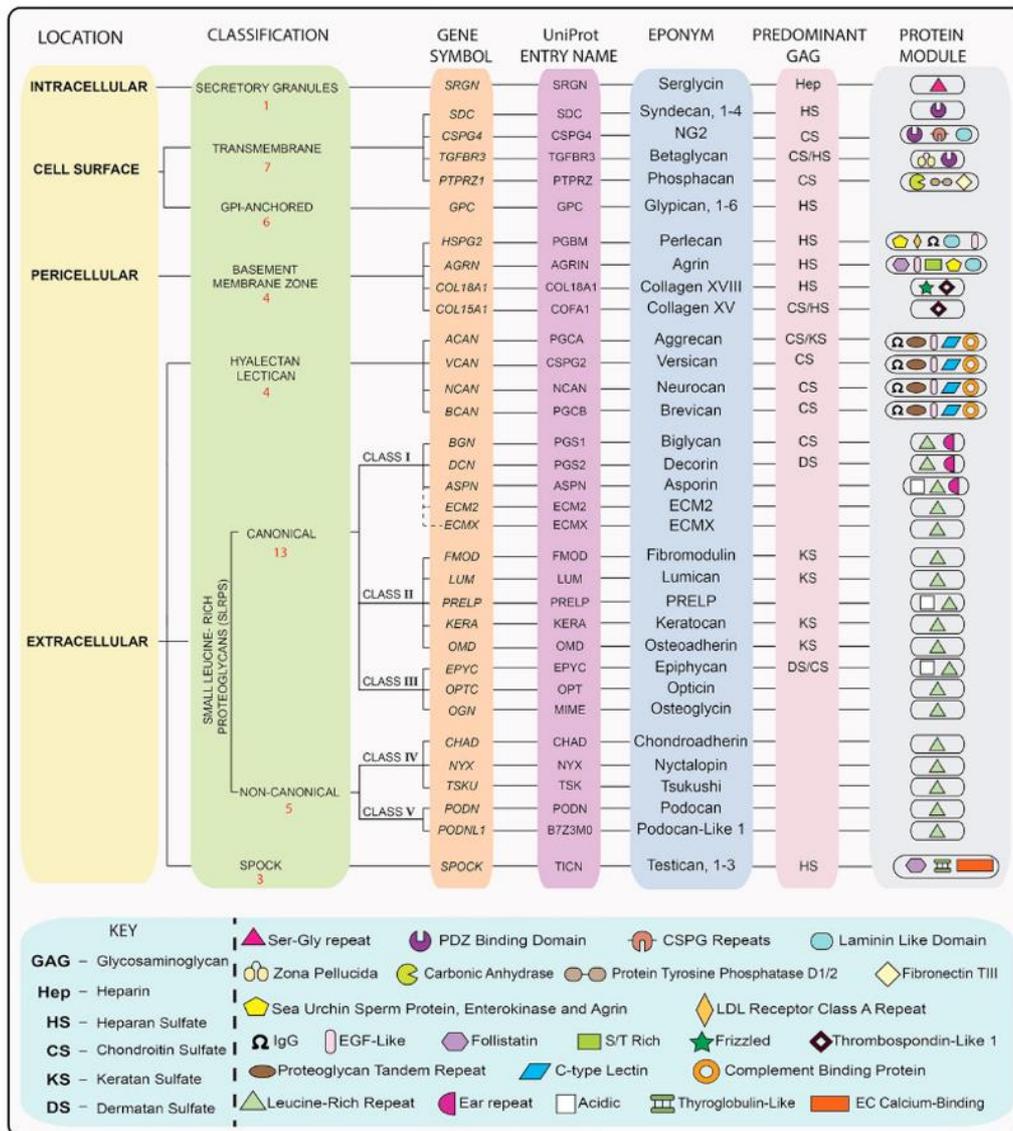


Figure 1. Classification des protéoglycans. Les familles sont déterminées par la localisation, l'homologie génique et protéique et les différents modules retrouvés le long de la séquence protéique. Par exemple, un protéoglycane (PG) transmembranaire est le syndécan 1, codé par le gène *SDC*. Il porte des chaînes de glycosaminoglycane (GAG) de type héparane sulfate (HS). Un PG extracellulaire est la décorine, faisant partie de la famille des *small leucine-rich proteoglycans* (SLRP). Il est codé par le gène *DCN* et arbore principalement des chaînes de GAG de type dermatane sulfate (DS). Adapté de (57).

a. Structure et classification

i. La protéine-core, partie protéique des protéoglycans

La partie protéique est nommée protéine-core, en raison de son rôle central identitaire et fonctionnel pour le PG. Elle porte des séquences d'attachement des chaînes de GAG et détermine la localisation du PG mature : intracellulaire, membranaire ou extracellulaire (57).

En fonction de sa protéine-core, le PG peut être attaché à la membrane plasmique *via* une ancre glycosylphosphatidylinositol comme dans le cas des glypicans (58), être transmembranaire et potentiellement clivé puis secrété dans la MEC comme les syndécans (59) ou encore être directement secrété dans la MEC comme la décorine (60) (Figure 2).

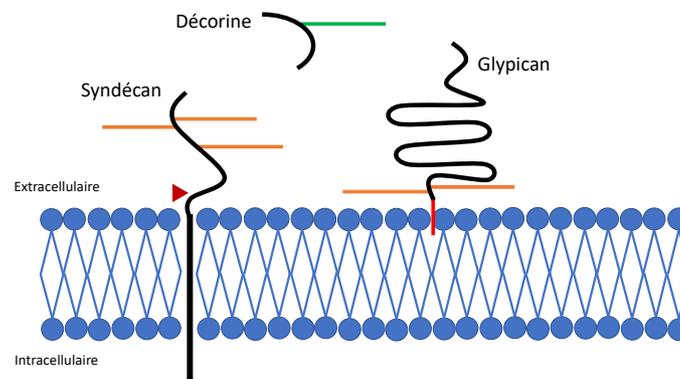


Figure 2. Représentation schématique de la localisation de plusieurs protéoglycanes. Les parties protéiques sont représentées en noir et la bicouche lipidique de la membrane plasmique est représentée en bleu. Deux protéoglycanes (PG) transmembranaires sont représentés : le syndécan et le glypican. Le syndécan possède un segment transmembranaire et plusieurs chaînes de glycosaminoglycanes (GAG) extracellulaires, majoritairement de type héparane sulfate (HS) en orange. Le domaine extracellulaire possède un site de clivage (pointe de flèche rouge) qui permet au PG d'être secrété dans la matrice extracellulaire (MEC). Le glypican est attaché du côté extracellulaire de la membrane par une ancre glycosylphosphatidylinositol en rouge et porte également des chaînes de GAG de type HS. La décorine est un PG extracellulaire arborant une chaîne de GAG de type chondroïtine (CS) / dermatane sulfate (DS) en vert.

ii. Les différentes chaînes de glycosaminoglycanes, modifications post-traductionnelles glucidiques des protéoglycanes

On distingue plusieurs catégories de chaînes de GAG en fonction de leur composition glucidique : les chaînes de GAG de type **chondroïtine sulfate (CS) / dermatane sulfate (DS)**, de type **héparane sulfate (HS) / héparine (Hep)**, de type **kératane sulfate (KS)** et l'**acide hyaluronique (AH)**.

Les chaînes de GAG de type CS, DS, HS et Hep partagent une structure commune en tout début de chaîne : l'amorce tétrasaccharidique (61, 62). Leur synthèse se poursuit à partir de cette amorce par l'ajout successif de motifs disaccharidiques sulfatés (Figure 3). Les chaînes de GAG de type KS ne possèdent pas d'amorce tétrasaccharidique (63). L'AH est, quant à lui, une chaîne de GAG non sulfatée et non reliée à une protéine-core (64).

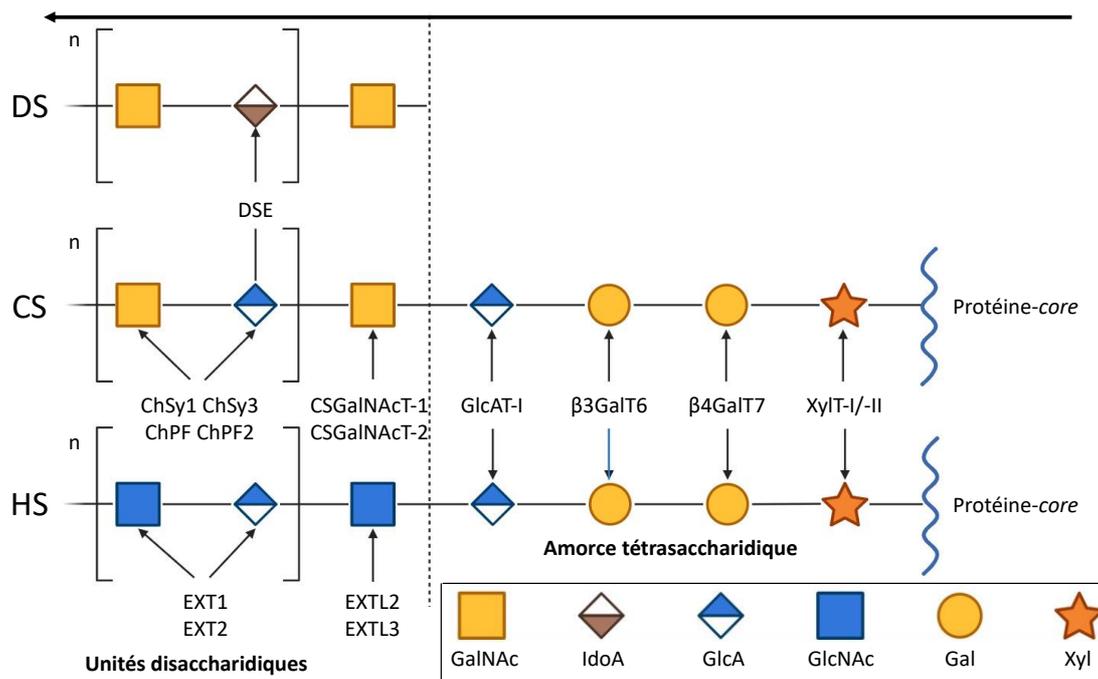


Figure 3. Schéma simplifié de la synthèse des chaînes de chondroïtine sulfate, dermatane sulfate et héparane sulfate à partir de l'amorce tétrasaccharidique. L'ajout des différents résidus saccharidiques est réalisé de façon successive par des glycosyltransférases (GT) spécifiques de chacune des voies de biosynthèse. Des modifications non représentées telles que des sulfatations et une phosphorylation au niveau de l'amorce et des chaînes de glycosaminoglycane (GAG) en cours de synthèse sont également observées. Les chaînes de GAG de type héparane sulfate (HS) sont composées d'une répétition d'un motif disaccharidique *N*-acétylglucosamine (GlcNAc) – acide glucuronique (GlcA) alors que les chaînes de type chondroïtine sulfate (CS) sont composées d'un motif disaccharidique *N*-acétylgalactosamine (GalNAc) – GlcA. Au sein des chaînes de dermatane sulfate (DS), le GlcA est épimérisé en acide iduronique (IdoA). La flèche noire en gras indique le sens d'élongation de la chaîne polysaccharidique. Xyl : xylose. Gal : galactose. XylT-I/-II : xylosyltransferase 1 et xylosyltransferase 2. β4GalT7 : β1,4-galactosyltransferase 7. β3GalT6 : β1,3-galactosyltransferase 6. GlcAT-I : β1,3-glucuronosyltransferase 1. DSE : dermatane sulfate épimérase. CSGalNAcT : *chondroïtine sulfate N-acétylgalactosaminyltransferase*. ChSy : chondroïtine sulfate synthase. ChPF : *chondroïtine-polymerizing factor*. EXTL : exotosine-like. EXT : exotosine.

L'amorce tétrasaccharidique :

L'amorce tétrasaccharidique est une succession de quatre oses spécifiques faisant office de zone de liaison entre la protéine-core et la suite de la chaîne de GAG de type CS, DS, HS ou Hep (Figure 3). Elle débute par l'ajout d'un xylose (Xyl) sur un résidu Ser d'un motif Ser-Gly de la protéine-core par deux isoformes de xylosyltransferase, **la xylosyltransferase 1 (XylT-I)** ou **la xylosyltransferase 2 (XylT-II)** (EC 2.4.2.26) (65, 66). L'ajout du second résidu saccharidique, un galactose (Gal), est catalysé par **la β4GalT7** (16). Un deuxième Gal est ensuite ajouté par **la**

β 3GalT6 (67), l'enzyme faisant l'objet de mon travail. Le dernier monosaccharide de l'amorce est l'acide glucuronique (GlcA), dont l'ajout est catalysé par la **β 1,3-glucuronosyltransférase 1** (GlcAT-I) (EC 2.4.1.135) (68).

De façon intéressante, le Xyl peut être phosphorylé par la **xylosylkinase *FAMILY with sequence similarity 20 member B*** (FAM20B) (EC. 2.7.1) (69) vraisemblablement après l'ajout du premier Gal, étape qui favorise l'activité de la β 3GalT6 (70) et de la GlcAT-I (71). La déphosphorylation du Xyl, assurée par la **2-phosphoxylose phosphatase 1** (PXLYP1) (EC. 3.1.3), aurait lieu après la réaction catalysée par la GlcAT-I afin de permettre l'élongation de la chaîne (72, 73). La phosphorylation transitoire du Xyl joue donc un rôle majeur dans la régulation des étapes précoces de la synthèse des chaînes de GAG. Néanmoins, la chronologie précise des réactions de phospho-/déphosphorylation n'est pas complètement établie (Figure 4).

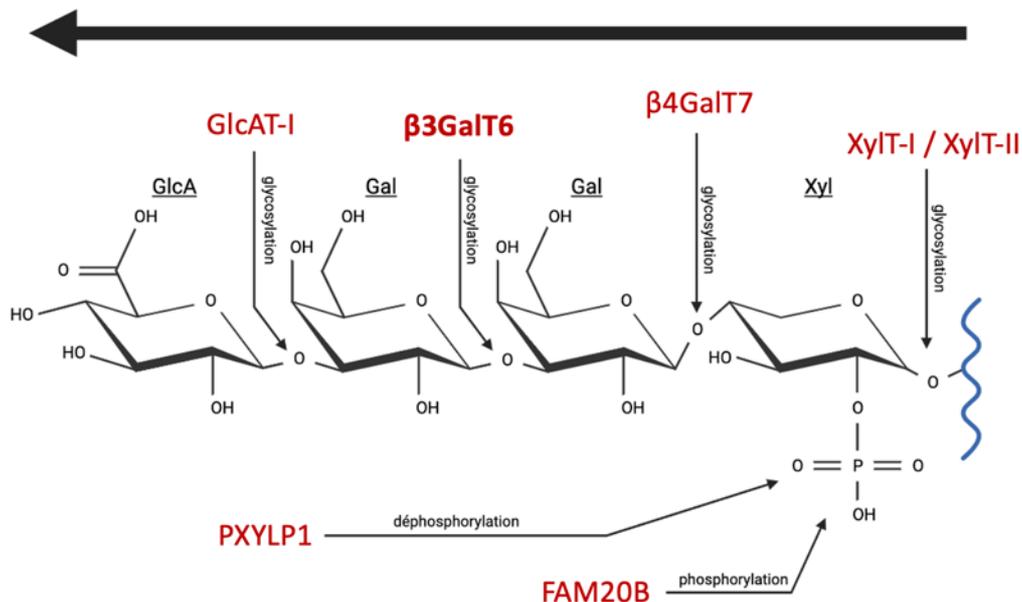


Figure 4. Biosynthèse de l'amorce tétrasaccharidique des chaînes de glycosaminoglycane. L'initiation de la biosynthèse a lieu dans les compartiments tardifs du réticulum endoplasmique (RE) au niveau d'un résidu Ser de la protéine-*core*. Quatre réactions enzymatiques ont lieu grâce à la xylosyltransférase 1 (XylT-I) ou la xylosyltransférase 2 (XylT-II), la β 1,4-galactosyltransférase 7 (β 4GalT7), la β 1,3-galactosyltransférase 6 (β 3GalT6) et la β 1,3-glucuronosyltransférase 1 (GlcAT-I) catalysant respectivement les ajouts successifs de résidus xylose (Xyl), galactose (Gal), Gal et acide glucuronique (GlcA). La *FAMILY with sequence similarity 20 member B* (FAM20B) et la 2-phosphoxylose phosphatase 1 (PXLYP1) sont respectivement la kinase et la phosphatase qui permettent la phosphorylation et déphosphorylation du Xyl. La flèche noire en gras indique le sens de la synthèse de l'amorce tétrasaccharidique.

L'élongation des chaînes de GAG :

Les chaînes de GAG sont des répétitions de motifs disaccharidiques constitués d'un acide uronique et d'une hexosamine qui sont différents selon la nature de la chaîne et se présentent de la manière suivante (64) :

- (i) Les **CS** sont composés d'une succession du disaccharide *N*-acétylgalactosamine (GalNAc) – acide glucuronique (GlcA).
- (ii) Les **DS** sont composés d'une succession du disaccharide GalNAc – acide iduronique (IdoA).
- (iii) Les **HS**, ainsi que l'**Hep**, sont composés d'une succession du disaccharide *N*-acétylglucosamine (GlcNAc) – GlcA. Ce dernier peut être épimérisé en IdoA.
- (iv) Les **KS** sont composées d'une succession du disaccharide Gal – GlcNAc.
- (v) L'**AH** est composé du disaccharide GlcA – GlcNAc.

Pour les CS et DS, le premier saccharide ajouté par une *chondroïtine sulfatase N-acétylgalactosaminyltransferase* (CSGalNAcT) est un GalNAc tandis que pour les HS et l'Hep, le premier saccharide ajouté par l'*exotosine-like* (EXTL) 3 est un GlcNAc. Une question récurrente a été de savoir quels sont les mécanismes moléculaires qui déterminent si une chaîne de GAG deviendra un CS ou un HS, *i.e.* si un GalNAc ou un GlcNAc est ajouté sur l'amorce tétrasaccharidique. Il semble que la synthèse se dirige vers la formation d'une chaîne de type HS en fonction de la composition de la séquence protéique autour du motif Ser-Gly. En effet, l'EXTL3 interagit de façon spécifique avec des résidus basiques de la protéine. De l'autre côté, les CSGalNAcT ne semblent reconnaître que l'amorce tétrasaccharidique (74).

Les modifications des chaînes de GAG :

A l'exception de l'AH, les chaînes de GAG portent des groupements sulfates qui contribuent à leur charge globale négative (75) (Tableau IV). Les sulfatations représentent ainsi une étape importante de maturation qui est réalisée par plusieurs familles de ST à différents endroits de la chaîne.

Les CS peuvent être 4-*O*-sulfatés ou 6-*O*-sulfatés sur leurs GalNAc, ainsi que 2-*O*-sulfatés sur leurs GlcA. Les possibilités de sulfatation des DS sont les mêmes : le GlcA étant épimérisé en IdoA, la 2-*O*-sulfatation est alors réalisée sur l'IdoA (76).

Le processus de sulfatation des HS mobilise un grand nombre d'enzymes. Tout d'abord, les GlcNAc peuvent être *N*-déacétylés puis *N*-sulfatés par la famille des *N*-déacétylases-*N*-

sulfotransférases (NDST), et ce de façon non-exhaustive, pour former des glucosamines *N*-sulfatées ou non. Les HS possèdent ainsi des régions fonctionnelles distinctes, comprenant des domaines sulfatés (domaines S) entrecoupés de région *N*-acétylées faiblement sulfatées (domaines N). Les glucosamines peuvent ensuite être 6-*O*-sulfatées et 3-*O*-sulfatées. De plus, les GlcA et les IdoA peuvent être 2-*O*-sulfatés. De nombreuses isoformes de ST sont responsables de ces modifications (76). Enfin, après passage dans le milieu extracellulaire, le PG de type HS peut être modifié par une famille de désulfatases, les Sulf, capables de réaliser une 6-*O*-désulfatation (77).

Ces modifications confèrent de multiples capacités d'interaction aux chaînes de GAG. Par exemple, la 6-*O*-sulfatation des HS augmente leur affinité envers le FGF, participant ainsi à la migration et la division cellulaire (78). La désulfatation opérée par Sulf-1 atténue alors la prolifération des cellules, et l'on observe une diminution du taux d'expression de Sulf-1 dans des cas de cancers (79).

Un autre exemple est l'effet anticoagulant et anti-inflammatoire des 3-*O*-sulfatations et des 6-*O*-sulfatations des chaînes de GAG de type HS (80). De plus, bien que rares (81) les 3-*O*-sulfatations ont des rôles pro- ou anti-oncogènes, selon le contexte cellulaire et tumoral considéré (82, 83). L'expression des gènes codant pour les 3-*O*-sulfotransférases est souvent épigénétiquement modifiée lors d'un cancer, faisant de ces gènes de potentiels biomarqueurs et procurant des informations sur les mécanismes de progression tumorale (82, 84).

Enfin, c'est en grande partie grâce à la charge négative des sulfatations que les chaînes de GAG peuvent interagir avec les molécules d'eau pour conférer aux TC une résistance aux forces de compression (85). Les sulfatations sont donc des modifications extrêmement importantes pour les fonctions biologiques des chaînes de GAG.

| Glycosaminoglycane (GAG) | Acide uronique | Galactose | Hexosamine | Composition disaccharidique | Amorce tétrasaccharidique | Protéine-core |
|--|--|-----------------|--------------------------------|---|---------------------------|---------------|
| Héparane sulfate (HS) / héparine (Hep) | Acide glucuronique (GlcA) ou Acide iduronique (IdoA) | x | N-acétylglucosamine (GlcNAc) | GlcA $\beta(1\rightarrow4)$ GlcNAc $\alpha(1\rightarrow4)$ IdoA $\alpha(1\rightarrow4)$ GlcNAc $\alpha(1\rightarrow4)$ | ✓ | ✓ |
| Kératane sulfate (KS) | x | Galactose (Gal) | N-acétylglucosamine (GlcNAc) | Gal $\beta(1\rightarrow4)$ GlcNAc $\beta(1\rightarrow3)$ | x | ✓ |
| Chondroïtine sulfate (CS) | Acide glucuronique (GlcA) | x | N-acétylgalactosamine (GalNAc) | GlcA $\beta(1\rightarrow3)$ GalNAc $\beta(1\rightarrow4)$ | ✓ | ✓ |
| Dermatane sulfate (DS) | Acide iduronique (IdoA) | x | N-acétylgalactosamine (GalNAc) | IdoA $\beta(1\rightarrow3)$ GalNAc $\beta(1\rightarrow4)$ | ✓ | ✓ |
| Acide hyaluronique (AH) | Acide glucuronique (GlcA) | x | N-acétylglucosamine (GlcNAc) | GlcA $\beta(1\rightarrow3)$ GlcNAc $\beta(1\rightarrow4)$ | x | x |

Tableau IV. Structure et composition des chaînes de glycosaminoglycane en fonction de la nature de la chaîne considérée. Le motif disaccharidique est composé d'un acide uronique et d'une hexosamine dans le cas des CS/DS et HS ou d'un Gal et d'une hexosamine dans le cas des KS. La colonne « composition disaccharidique » indique les sites de sulfatation (représentés par des encadrés en pointillés rouges) et la nature des liaisons de part et d'autre de chaque saccharide. x : élément absent. ✓ : élément présent. Adapté de (62).

b. La structure des petits protéoglycane riches en leucine

Les petits PG riches en leucine (*small leucine-rich proteoglycans*, SLRP) sont la plus grande famille de PG. Nous pouvons citer la décorine (86), le biglycan (87) ou encore la fibromoduline (88) parmi les SLRP les plus étudiés. Exclusivement retrouvés dans la MEC, les SLRP ont pour particularité d'avoir une protéine-core plus petite que celles d'autres PG matriciels comme l'aggrécane. Leur masse moléculaire excède rarement les 40 kDa (89).

Il existe dix-huit gènes qui codent pour des protéines-core de SLRP exprimées de façon ubiquitaire dans nos tissus (90, 91). La classification actuelle les répartit en cinq classes notamment en fonction de la conservation des gènes au cours de l'évolution, des homologies géniques et protéiques et de l'organisation du gène au niveau du chromosome. Les classes I, II et III comprennent les gènes « canoniques » codant pour dix SLRP, dont la décorine et le

biglycan qui portent respectivement une ou deux chaînes de GAG de type CS/DS. Les classes IV et V, quant à elles, regroupent les huit autres SLRP qui ne possèdent pas de chaînes de GAG. Cependant, des homologues de structure au niveau de la protéine-core ainsi que des fonctions proches des SLRP canoniques leur permettent d'en faire partie (92) (Figure 1).

Les SLRP possèdent des séquences protéiques caractéristiques. Les plus représentatives sont des régions de répétitions riches en leucines (*leucine-rich repeats*, LRR), très conservées et constituées d'un motif LxxLxLxxNxL, où « L » est une Leu, une Ile, une Val ou une Phe, « N » est une Asn, une Thr, une Ser ou une Cys et « x » est un résidu quelconque (93). Les LRR contiennent vingt à trente résidus et sont présentes en moyenne dix à douze fois par séquence protéique de SLRP (94). Un second motif conservé est nommé *ear repeat* à l'extrémité C_{ter}, correspondant au passage de l'avant-dernière à la dernière LRR. Remarquées pour la première fois lors de la cristallisation de la décorine, elles ressortent « comme des oreilles » de la forme solénoïde régulière qu'adoptent les SLRP. A titre d'exemple, l'avant-dernière et la dernière LRR de la décorine contiennent chacune une Cys entre lesquelles un pont disulfure est formé. Ces deux LRR sont plus longues que les autres, particulièrement dans la séquence des SLRP kératocan et *Proline arginine-Rich-end Leucine-rich repeat Protein* (PRELP). De ce fait, elles ressortent clairement de la structure (95) (Figure 5). On peut également noter la présence d'une région riche en Cys à l'extrémité N_{ter} (96).

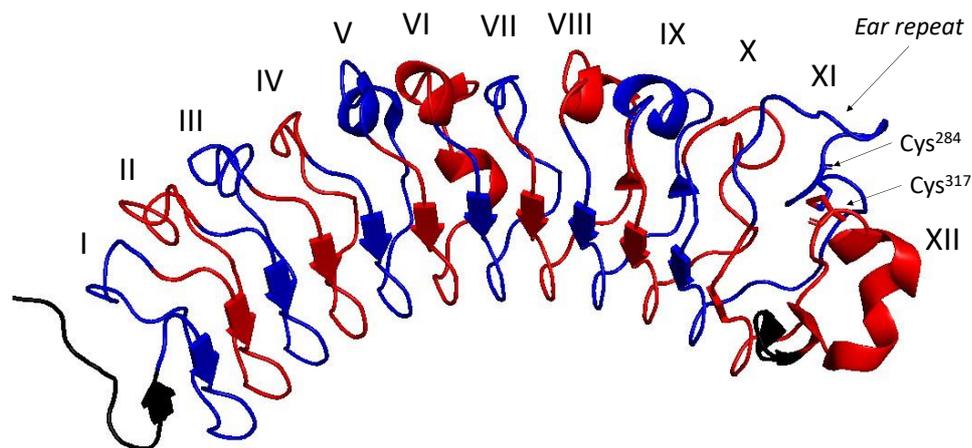


Figure 5. Modèle ruban de la structure de la décorine bovine cristallisée représentant les répétitions riches en leucine. Les flèches et les hélices indiquent les brins β et les hélices α , respectivement. Les douze répétitions riches en leucine (*leucine-rich repeats*, LRR) sont numérotées en chiffres romains et sont discernées par une alternance de leur couleur (rouge et bleu). La structure *ear repeat*, qui ressort de la structure solénoïde et fait le lien entre les LRR XI et XII, est annotée. Les deux cystéines Cys²⁸⁴ et Cys³¹⁷ formant un pont disulfure pour

permettre d'adopter cette structure particulière sont annotées et représentées en bâtonnets. Modèle préparé sur PyMOL (PDB : 1XKU) (95).

2. Les collagènes

Les collagènes sont la superfamille de protéines les plus abondantes des MEC, représentant environ 30 % de la masse protéique totale des mammifères (97). Ils sont très conservés à travers les espèces et organismes et leurs rôles ont été décrits extensivement. Au cours du temps, la découverte de nouveaux collagènes et l'exploration de leurs structures et fonctions a permis de mieux répondre à la question posée en 1983 par Steffen Gay et Edward J. Miller : « *What is collagen, what is not ?* » (98).

a. Classification des collagènes

Il existe vingt-huit collagènes distincts chez les vertébrés, chacun étant identifié par un chiffre romain. Ce sont des composants matriciels qui ont pour caractéristique de s'associer en une triple hélice à partir de trois chaînes polypeptidiques α .

En fonction du collagène considéré, l'association des chaînes α se fait de façon spécifique. Par exemple, il existe un seul type de chaîne α pour le Col II, la chaîne $\alpha 1(\text{II})$ codée par le gène *COL2A1*. Trois de ces chaînes s'associent pour former le Col II sous la forme homotrimérique $\alpha 1(\text{II})_3$ (99). D'autres collagènes sont composés de plusieurs types de chaînes α . C'est le cas du Col I qui est composé de chaînes $\alpha 1(\text{I})$ et $\alpha 2(\text{I})$, respectivement codées par les gènes *COL1A1* et *COL1A2*. On le retrouve alors sous forme hétérotrimérique $\alpha 1(\text{I})_2, \alpha 2(\text{I})$ (100). Certains collagènes sont le résultat de l'association de trois chaînes polypeptidiques de collagènes différents, comme le Col V/XI. La composition de ce collagène est $\alpha 1(\text{XI})\alpha 1(\text{V})\alpha 3(\text{XI})$ (101, 102).

Les collagènes sont regroupés en famille selon leur structure, leurs fonctions et leur localisation (103). On retrouve notamment (Figure 6) :

- Les collagènes formant des fibrilles, ou **collagènes fibrillaires** comme les Col I, Col II, Col III et Col V. Ils possèdent un long domaine principal en triple hélice.
- Les **collagènes associés aux fibrilles avec des triples hélices interrompues** (*fibril-associated collagens with interrupted triple helices*, FACIT) comme les Col IX, Col XII et Col XIV. Contrairement aux collagènes fibrillaires, ils possèdent plusieurs domaines en triples hélices séparés par des séquences n'en formant pas.

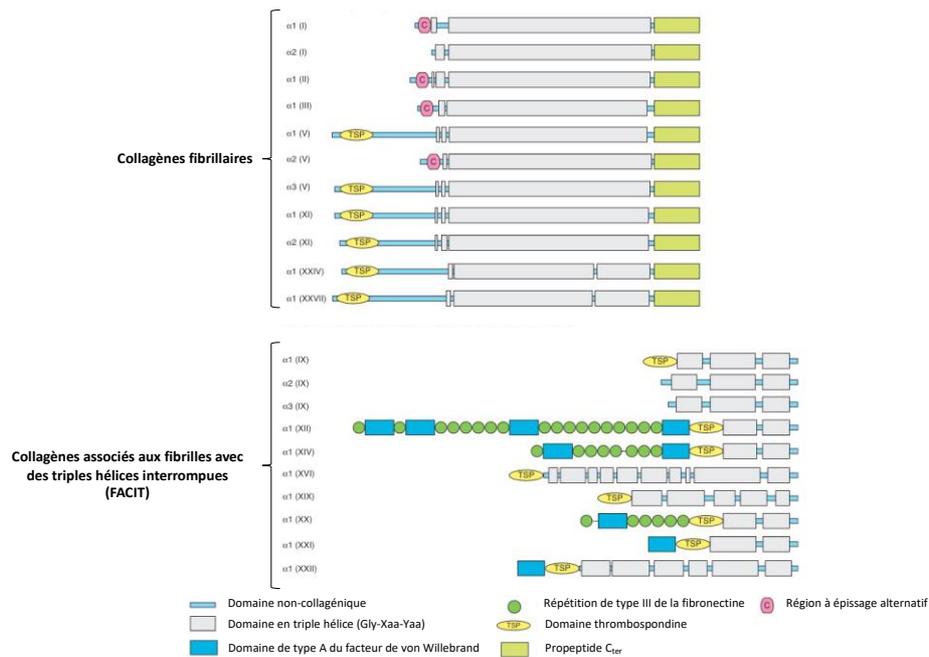


Figure 6. Organisation structurale des chaînes polypeptidiques α formant les collagènes fibrillaires et des collagènes associés aux fibrilles avec des triples hélices interrompues. Les structures des chaînes polypeptidiques α sont représentées des extrémités N_{ter} vers C_{ter}. L'association des chaînes en triples hélices se fait grâce aux domaines composés de répétitions de triplet Gly-Xaa-Yaa (gris), séparés par des domaines non-collagéniques (bleu clair). Les chaînes α formant les collagènes fibrillaires possèdent un propeptide (vert clair) clivé lors de leur maturation et certains d'entre eux peuvent subir un épissage alternatif (rose). Les FACIT possèdent souvent des répétitions de type III de la fibronectine (vert) et/ou des domaines de type A du facteur de von Willebrand (bleu). La plupart de ces chaînes possède un domaine de fixation à la thrombospondine (jaune). Adapté de (103).

b. Synthèse et maturation des collagènes fibrillaires

De manière générale, la synthèse des collagènes fibrillaires débute par la transcription et la traduction des chaînes α , dont une caractéristique est de posséder un nombre important de triplets Gly-Xaa-Yaa formant les domaines collagéniques nécessaires à la trimérisation. Le résidu Xaa est souvent une Pro et le résidu Yaa est souvent une Pro hydroxylée dans le RE par la prolyl 4-hydroxylase (P4H), formant une 4-hydroxyproline (4Hyp) (104). Ces séquences polyprolines séparées par des Gly, un acide aminé peu encombrant, permettent à la chaîne α d'adopter une structure hélicoïdale extrêmement stable (105).

Des résidus Lys retrouvés dans la séquence ou au niveau des télépeptides des chaînes α peuvent être hydroxylés en 5-hydroxylysines (5Hyl) par la lysyl hydroxylase 1 (LH1), 2 (LH2) ou 3 (LH3) (106). La LH3 peut ensuite ajouter un Gal puis potentiellement un glucose à des 5Hyl

spécifiques (107). Si les 5Hyl participent activement à l'association des chaînes α , les glycosylations permettent de réguler ce processus (108).

Aidés notamment par la stabilisation apportée par les nombreuses 4Hyp en position Yaa (109), par l'hydroxylation des Lys et par la formation de ponts disulfures entre les extrémités C_{ter} des chaînes α , les collagènes fibrillaires vont s'associer en hélices trimériques nommées procollagènes par le biais de liaisons croisées (110) préalablement à leur sécrétion dans la MEC *via* un transport par des vésicules sécrétoires (111). Elles y subissent une étape de clivage des extrémités N_{ter} par une métalloprotéinase (*A Disintegrin And Metalloproteinase with Thrombospondin motifs 2* (ADAMTS-2)) (112) et C_{ter} par deux métalloprotéinases (*Bone morphogenic protein 1 / Tolloid-like proteinases* (BMP-1/mTLD)) (113). L'hélice trimérique est alors appelée tropocollagène.

Suivant ces clivages, les Lys et 5Hyl des chaînes α peuvent subir une nouvelle modification. La lysyl oxidase (LOX) et ses isoformes lysyl oxidase-like 1 (LOXL1), 2 (LOXL2), 3 (LOXL3) et 4 (LOXL4) sont des enzymes extracellulaires capables d'oxyder les Lys et 5Hyl en allysines et 5-hydroxyallysines, respectivement (114). Cette modification est indispensable à l'assemblage du réseau collagénique car elle participe à la création de liaisons croisées entre les triples hélices de collagènes fibrillaires, donc aux propriétés mécaniques de la MEC (115). Ce processus est la fibrillogénèse qui sera détaillée dans la section II.3.a (116).

c. La structure du collagène de type XII, collagène et protéoglycane à la fois

i. Structure générale

Le Col XII est un FACIT, un collagène non fibrillaire mais qui s'associe aux collagènes fibrillaires, dont la structure est détaillée dans la Figure 7. L'unique chaîne α du Col XII est la chaîne polypeptidique $\alpha 1(\text{XII})$ codée par le gène *COL12A1*. Elle est constituée de deux petits domaines collagéniques COL1 et COL2 interrompus par deux domaines non-collagéniques NC1 et NC2 et est terminée par un long domaine non-collagénique NC3 (117). Ce dernier est constitué de dix-huit répétitions de type III de la fibronectine interrompues par quatre domaines de type A du facteur de von Willebrand.

ii. Épissage alternatif du collagène de type XII

Il existe quatre isoformes de la chaîne polypeptidique $\alpha 1(\text{XII})$ dû à deux évènements d'épissages alternatifs distincts (118) :

- Deux isoformes différant au niveau du domaine NC3 consistent en une forme longue nommé **$\alpha 1(\text{XIIA})$** et une forme courte **$\alpha 1(\text{XIIB})$** (Figure 7A). La partie commune se trouve du côté 3' de la séquence nucléotidique, codant pour les domaines COL1 et COL2 (119). Le domaine NC3 de la chaîne $\alpha 1(\text{XIIA})$ est deux fois plus long que celui de la chaîne $\alpha 1(\text{XIIB})$. Ces deux isoformes peuvent s'associer entre elles. En effet, ce sont les domaines collagéniques qui permettent l'association des chaînes pour former une molécule de collagène trimérique à l'aide de ponts disulfures à l'intersection entre le domaine NC3 et le domaine COL2 et de liaisons croisées entre les domaines collagéniques de chaque monomère (120). On trouve ainsi des homotrimères de Col XII $\alpha 1(\text{XIIA})_3$ ou $\alpha 1(\text{XIIB})_3$ ainsi que des hétérotrimères, par exemple $[\alpha 1(\text{XIIA})]_2, \alpha 1(\text{XIIB})$ (Figure 7B) (117).
- Deux isoformes différant au niveau du domaine NC1 avec un long domaine NC1 ou un court domaine NC1, sans impact majeur sur la structure globale de la molécule (121). Les informations structurales et fonctionnelles de ces isoformes étant limitées, elles ne seront pas traitées dans ce manuscrit.

Le Col XII est exprimé dans les tissus osseux, ligamenteux, tendineux, cartilagineux, musculaires et cutanés lors du développement embryonnaire (122, 123). De façon intéressante, son expression post-natale est fortement régulée. Il devient alors spécifique de certains tissus comme les ligaments péri-dentaires (124) et la cornée (125). Chacune des deux isoformes $\alpha 1(\text{XIIA})$ et $\alpha 1(\text{XIIB})$ ont elles-mêmes des taux d'expression différents au cours du développement du poussin (117). Cette différence d'expression peut être due à l'action de facteurs de croissance comme le *transforming growth factor beta* (TFG β) (126).

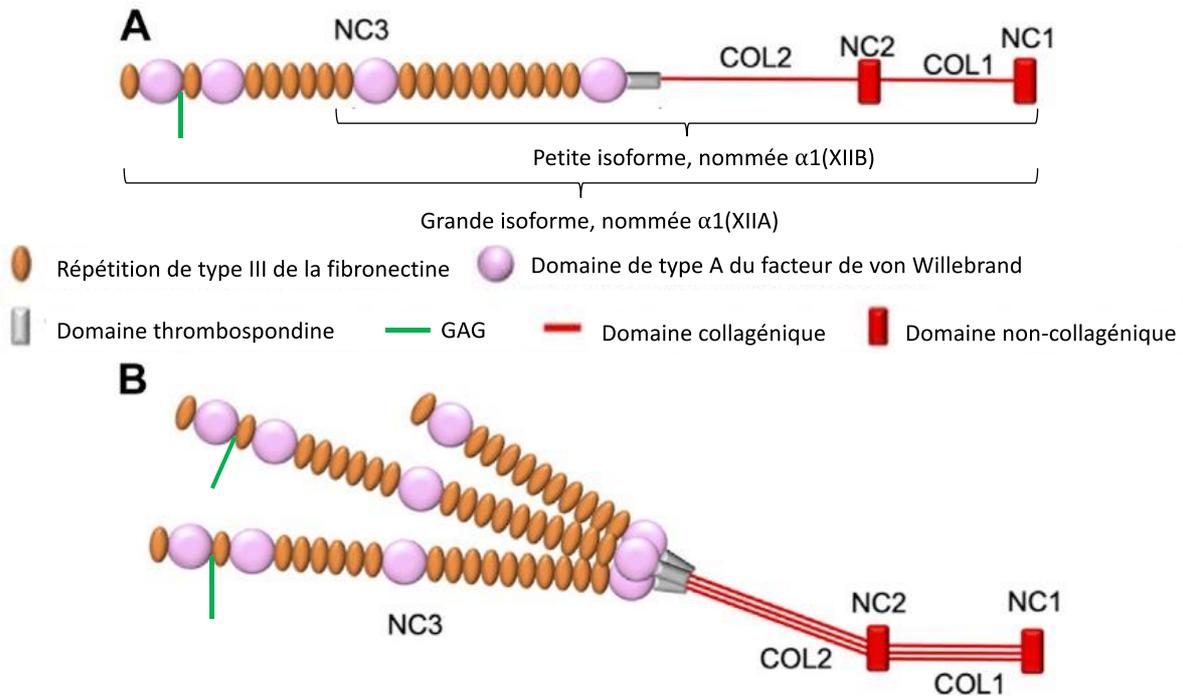


Figure 7. Schéma de la structure protéique du collagène de type XII. (A) Schéma distinguant les isoformes majeures des chaînes α du Col XII. Le Col XII consiste en deux domaines collagéniques COL1 et COL2 séparés par deux petits domaines non-collagéniques NC1 et NC2. Un troisième domaine non-collagénique NC3 constitué d'un domaine thrombospondine, de répétitions de type III de la fibronectine et de domaines de type A du facteur de von Willebrand est situé en N_{ter}. Le gène *COL12A1* subi un épissage alternatif donnant naissance à deux isoformes : un $\alpha 1(XIIB)$ de petite taille ou un $\alpha 1(XIIA)$ de grande taille. La chaîne $\alpha 1(XIIA)$ est substituée par au moins une chaîne de glycosaminoglycane (GAG) de type chondroïtine sulfate (CS) située au niveau du domaine NC3, entre un domaine de type A du facteur de von Willebrand et une répétition de type III de la fibronectine. Il existe un autre épissage alternatif aux conséquences mineures (non représenté). (B) Schéma de la structure hétérotrimerique du Col XII, ici $[\alpha 1(XIIA)]_2, \alpha 1(XIIB)$. Les chaînes α s'associent grâce à leurs domaines COL1 et COL2. Adapté de (127).

iii. Les chaînes polypeptidiques $\alpha 1(XIIA)$ sont des protéoglycanes

L'isoforme longue $\alpha 1(XIIA)$ possède une chaîne de GAG de type CS sur le domaine NC3 (Figure 7) et trois autres sites potentiels d'attachement de GAG ont été identifiés sur ce même domaine NC3 (128, 129). Les Col XII ayant au moins une chaîne longue $\alpha 1(XIIA)$ sont donc des PG dont la glycosylation débute par la synthèse de l'amorce tétrasaccharidique décrite dans la section II.1.a. Il n'y a pas de données expérimentales quant à la fonction de la chaîne de GAG identifiée.

3. La relation fonctionnelle entre les protéoglycanes et les collagènes

L'homéostasie d'un tissu est assurée grâce aux multiples acteurs qui le composent. Parmi ceux-ci, les collagènes et les PG occupent une place majeure. On dénombre vingt-huit collagènes (103) et plus de quarante PG (57). Ces molécules sont exprimées de façon différente selon l'âge de l'individu et le tissu considéré. Certains PG interagissent avec les collagènes pour créer et consolider des fonctions spécifiques, notamment lors de la fibrillogenèse et de l'ancrage du réseau collagénique dans les MEC. Lorsque cette relation est compromise, des atteintes multiples au niveau de la MEC peuvent se déclarer.

a. La fibrillogenèse

La fibrillogenèse est le processus complexe et extrêmement contrôlé d'assemblage des collagènes fibrillaires en un réseau régulier dans l'espace extracellulaire, adapté au tissu (130). Il s'agit de l'association de molécules de tropocollagènes fibrillaires en fibrilles, puis en fibres. Des expériences *in vitro* ont pointé vers un assemblage spontané des fibres de collagène (131, 132), mais d'autres acteurs interviennent également dans la fibrillogenèse *in vivo* : des collagènes fibrillaires et non fibrillaires (133), des PG (134) et autres glycoprotéines, ainsi que d'autres fibres.

Cette section traite principalement de la fibrillogenèse du Col I, largement retrouvé dans les tissus vasculaires, cutanés, osseux et tendineux (135). Un schéma récapitulatif est représenté en Figure 8.

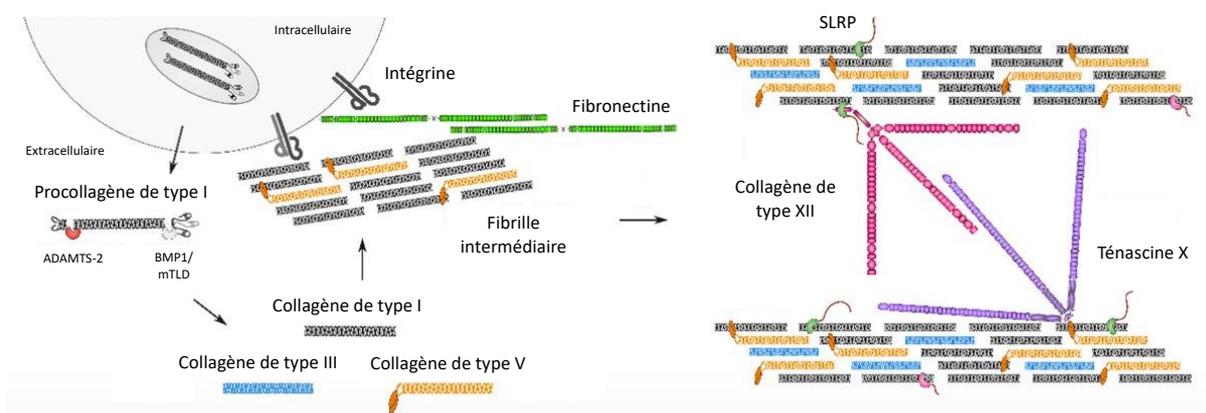


Figure 8. Schéma des grandes étapes de la fibrillogenèse. Les molécules de Col I synthétisées sont transportées vers la matrice extracellulaire (MEC) dans des vésicules. Leurs N-propeptides et C-propeptides sont clivés par deux métalloprotéinases (*A Disintegrin And Metalloproteinase with Thrombospondin motifs 2* (ADAMTS-2) et *Bone morphogenic protein-1 / Tolloid-like proteinases* (BMP1/mTLD)). Les fibres de Col I, de Col III et de Col V

s'associent lors de la nucléation. L'organisation est assurée par la fibronectine et les intégrines. Des paramètres comme le diamètre, la taille, la régularité et la vitesse de formation du réseau sont ensuite régulés par différents acteurs. Les *small leucine-rich proteoglycans* (SLRP) comme la décorine s'associent aux fibres de Col I grâce à leur partie protéique et au Col XII grâce à leurs parties glucidiques. La décorine s'associe également avec la ténascine X (non montré). Le Col XII s'associe avec les fibres de Col I grâce à ses domaines collagéniques COL1 et COL2. L'isoforme longue Col XIA s'associe également à la ténascine X grâce à son domaine non collagénique NC3. Adapté de (55).

i. Les nucléateurs

Après le clivage des *N*-propeptides et des *C*-propeptides et sa sécrétion sous forme de tropocollagène, le Col I va s'associer notamment avec le Col V, de façon tissu-dépendante à l'aide de plusieurs protéines (136). Le Col V est un nucléateur lors de cet assemblage, c'est-à-dire qu'il est impliqué activement dans l'initiation, mais également de la régulation de la taille, de la largeur et de la régularité des fibrilles de collagène (Figure 8) (137, 138). Il en résulte la formation de fibrilles hétérotypiques I/V composées de collagène fibrillaire « majoritaire » (ici le Col I) et de collagène fibrillaire « minoritaire » (ici le Col V). Le clivage des *N*-propeptides du Col V n'est pas total (139). Il conserve ainsi un domaine non-collagénique qui s'insère entre deux Col I et qui va être exposé à la surface des fibrilles (Figure 9). Le Col XI est également impliqué dans ces étapes préliminaires de l'assemblage fibrillaire notamment avec le Col II pour former les fibrilles hétérotypiques II/XI (140).

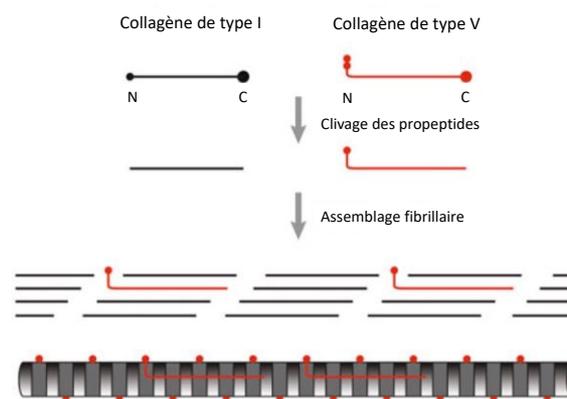


Figure 9. Clivage des propeptides et initiation de l'assemblage des fibrilles de collagène. Après leur sécrétion dans la matrice extracellulaire (MEC), les propeptides des Col I sont clivés. Ceux des Col V le sont totalement à l'extrémité C_{ter} et partiellement à l'extrémité N_{ter} . Les fibrilles s'assemblent avec une majorité de Col I et une minorité de Col V dont le *N*-propeptide est exposé à la surface. Il en résulte une fibrille hétérotypique I/V. Adapté de (130).

ii. Les organisateurs

Les organisateurs sont la fibronectine et les intégrines. Elles ont pour rôle d'assister à la nucléation et d'organiser le réseau naissant dans l'espace péricellulaire (141) (Figure 8).

La fibronectine est une glycoprotéine fibrillaire matricielle et dimérique, capable de se lier à la chaîne $\alpha 1(I)$ du Col I (142). Cette interaction permet d'accélérer la nucléation entamée au début de la fibrillogenèse et participe à l'organisation globale du réseau (143, 144). En outre, cette interaction est nécessaire pour limiter la dégradation des collagènes (145, 146), ce qui protège les tissus d'un vieillissement prématuré. En effet, le taux de synthèse des collagènes baisse graduellement avec l'âge (147). Il est donc nécessaire de conserver au maximum les collagènes synthétisés car leur production par l'organisme est limitée dans le temps.

Certaines intégrines sont capables de se lier au Col I pour participer à la fibrillogenèse (148). D'autres jouent indirectement un rôle en se liant à la fibronectine, l'absence de cette interaction intégrine/fibronectine entraînant la perte du réseau collagénique (149).

iii. Les régulateurs

Les FACIT et les SLRP sont les principaux régulateurs de la composition et de l'organisation de la MEC grâce à des mécanismes de contrôle du taux d'assemblage, de la taille, du diamètre et de la structure des fibrilles de collagène (Figure 8) (132, 150, 151).

Les SLRP :

Lors de la fibrillogenèse, les parties protéiques des SLRP peuvent interagir avec les collagènes fibrillaires (152). La décorine établit une interaction avec le Col I grâce à son motif conservé ²⁰³SYIRIADTNIT²¹³. Cette séquence se situe à cheval sur les LRR n°5 et n°6 et est retrouvée de façon homologue chez d'autres SLRP tels que le biglycan et l'asporine (153) (Figure 10).

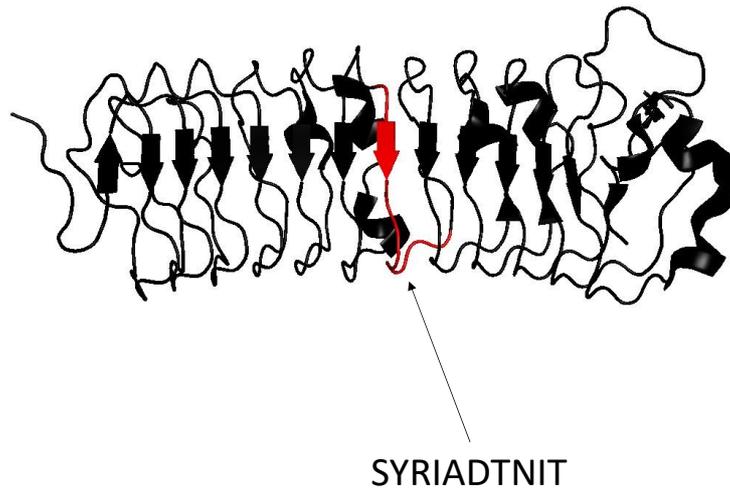


Figure 10. Modèle ruban de la structure de la décorine bovine cristallisée représentant la séquence de fixation au collagène de type I. Les flèches et les hélices indiquent les brins β et les hélices α , respectivement. La séquence SYRIADTNIT de fixation au Col I est représentée en rouge. Modèle préparé sur PyMOL (PDB : 1XKU) (95).

Une modélisation du complexe entre la protéine-*core* de la décorine et les fibrilles de Col I suggère que l'interaction se fait entre la partie protéique de la décorine et plusieurs fibrilles de collagène à la manière d'une pince qui dicterait le diamètre des fibres de collagène matures (154). La capacité d'interaction des SLRP avec plusieurs fibrilles de Col I (155, 156) de façon périodique le long des fibrilles (157) consolide le réseau naissant. Des ponts interfibrillaires sont créés par leurs chaînes de GAG (158). Le tout participe à l'organisation des collagènes fibrillaires et à l'ancrage de la MEC (159) (Figure 11). La protéine-*core* du biglycan adopte une forme solénoïde lorsque ses chaînes de GAG sont correctement synthétisées : la stabilité conformationnelle de la partie protéique est dépendante de la synthèse de la partie glucidique, ce qui serait critique pour sa fonction biologique (160). L'absence ou la mauvaise synthèse des chaînes de GAG peut donc perturber l'interaction de la protéine-*core* aux fibres de Col I en plus d'avoir un impact direct sur la formation des ponts interfibrillaires.

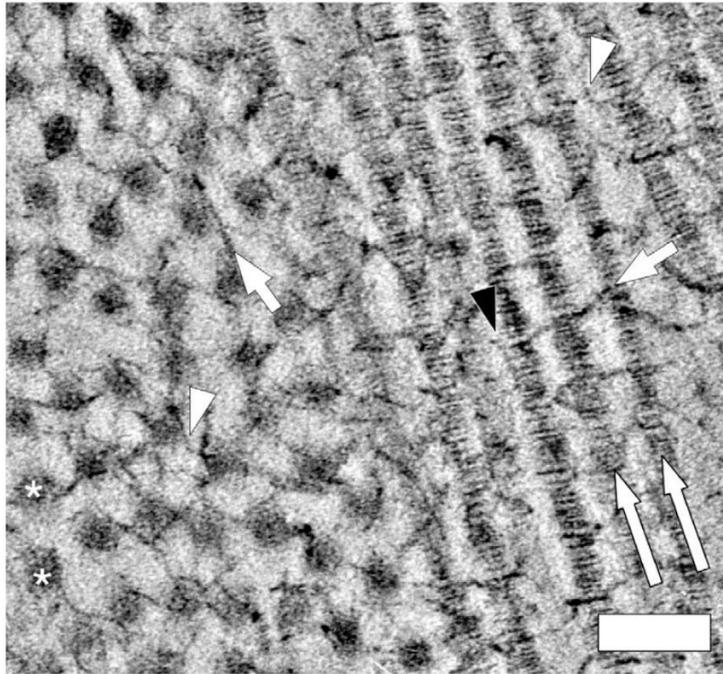


Figure 11. Image du réseau collagénique de la cornée de bovin obtenue par microscopie à transmission électronique. Une coupe transversale est visible à gauche de l'image, tandis qu'une coupe longitudinale est visible à droite de l'image. Astérisques : fibre de collagène. Longues flèches blanches : fibre de collagène. Petites flèches blanches : protéoglycanes (PG) reliant deux fibres de collagènes ou plus. Pointes de flèche blanche : PG reliant deux fibres de collagène. Pointe de flèche noire : PG se fixant aux fibres de collagène de manière longitudinale. L'échelle blanche représente une distance de 100 nm. Adapté de (156).

De manière générale, la liaison des SLRP avec les fibrilles de collagène se fait principalement par le biais d'interactions électrostatiques établies par des résidus chargés de la protéine-*core* (161) : le Glu¹⁸⁰ de la décorine humaine crée une forte interaction avec le Col VI (162) et l'Arg²⁰⁷ et l'Asp²¹⁰ du motif ²⁰⁴SYIRIADTNIT²¹⁴ de la décorine bovine sont essentiels pour l'interaction avec le Col I (153).

La fibromoduline et le lumican sont des SLRP de classe II (57). La fibromoduline contient les résidus Glu³⁵³ et Lys³⁵⁵, indispensables pour son interaction avec le Col I (163). Le Glu²⁵¹ de la fibromoduline est un autre site de fixation au Col I qui entre en compétition avec l'Asp²¹³ du lumican. Le modèle proposé est que le Glu³⁵³ et la Lys³⁵⁵ de la fibromoduline et l'Asp²¹³ du lumican leur permettent de se fixer à des monomères de Col I différents pour initier la régulation des liaisons croisées. La fibromoduline a une meilleure affinité avec le Col I avec ses trois résidus cités qui forment deux sites de fixation distincts. Le Glu²⁵¹ du second site de fixation de la fibromoduline déloge alors le lumican pour continuer la formation de liaisons

croisées entre plus de deux monomères de Col I (164) (Figure 12). Ce mécanisme est d'autant plus probable qu'il a été prouvé que la fibromoduline interagit spécifiquement avec les sites potentiels de liaisons croisées du Col I ainsi qu'avec la LOX, l'enzyme responsable de l'oxydation des Lys et 5Hyl. Après avoir délogé le lumican, la fibromoduline recrute donc la LOX au niveau des résidus du collagène à oxyder et permet d'augmenter son activité enzymatique. La fibromoduline joue donc un rôle très important dans la création du réseau collagénique en rapprochant les fibrilles de collagène dans l'espace et en favorisant l'oxydation des résidus, ce qui permet la formation de liaisons croisées inter-fibrillaires (165).

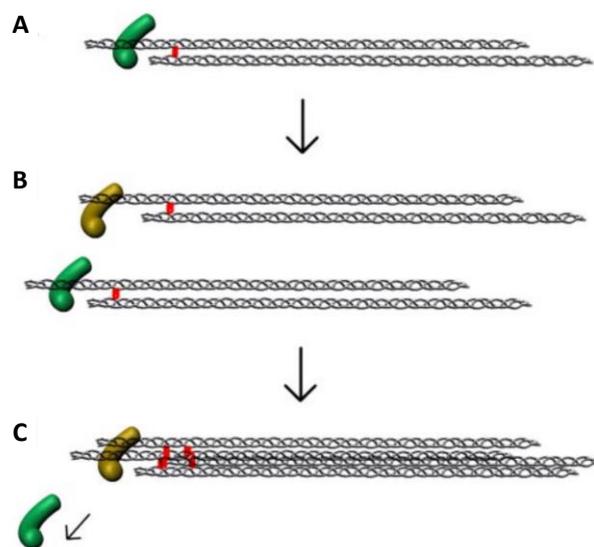


Figure 12. Mécanisme proposé de la régulation de la fibrillogenèse par le lumican et la fibromoduline. (A) Le lumican (vert) interagit avec un monomère de Col I grâce à son Asp²¹³ et initie la formation de liaisons croisées (rouge). **(B)** La fibromoduline (jaune) s'associe à un monomère de Col I avec ses Glu³⁵³ et Lys³⁵⁵ et initie la formation de liaisons croisées. **(C)** La fibromoduline déloge le lumican grâce à son Glu²⁵¹ qui possède une meilleure affinité pour le Col I. Les deux unités de Col I peuvent alors s'associer par le biais de nouvelles liaisons croisées. Adapté de (164).

L'ostéoadhérine, un SLRP de classe II (57), a un mécanisme de liaison au Col I qui a été étudié *in vitro*. Les interactions électrostatiques apportent un équilibre entre le taux de fixation et la dissociation rapide des complexes ostéoadhérine / Col I. Lors de l'interaction, une déshydratation de la surface des fibres de collagène favorable à leur association est observée. En présence d'ostéoadhérine, l'association des fibres est ralentie mais aboutit à des

réseaux uniformes. Étant donné les similitudes entre les structures des protéines-core des SLRP, l'action de l'ostéoadhérine *in vitro* en tant que régulateur de la fibrillogenèse des fibres de collagène a été proposé comme modèle (166) (Figure 13).

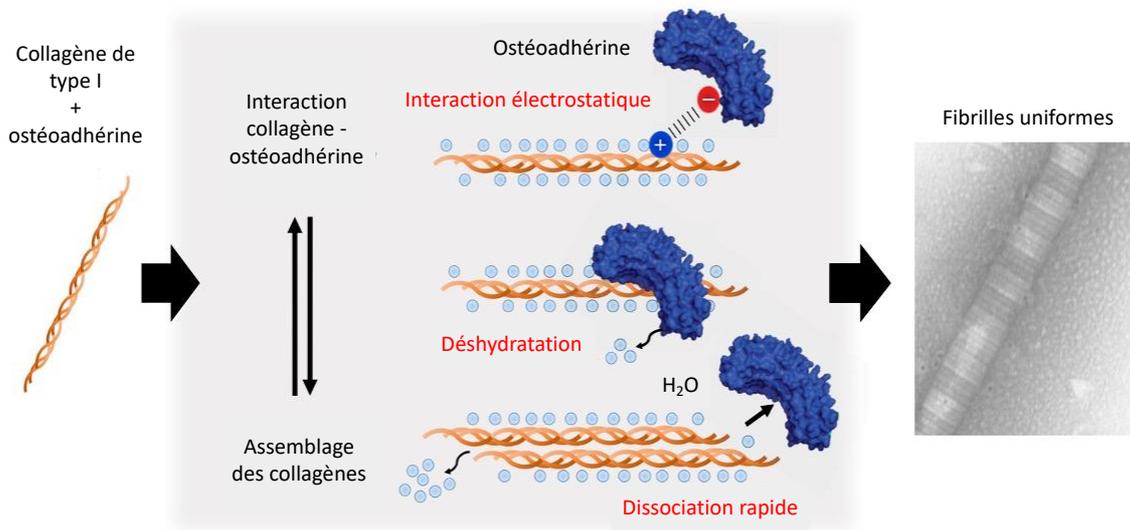


Figure 13. Modèle proposé du mécanisme de l'interaction entre la protéine-core de l'ostéoadhérine et le collagène de type I lors de la fibrillogenèse. Le protéoglycane (PG) s'associe avec des fibres de Col I. L'interaction est électrostatique, ce qui chasse les molécules d'eau qui entourent les fibres hors du réseau. Le PG se dissocie et les fibres de Col I s'associent spontanément. Adapté de (166).

Selon une étude computationnelle, le pontage entre les GAG de la décorine et les fibres de collagène du tendon serait nécessaire au transfert de force entre ces dernières (167). Le rôle critique des GAG dans ce contexte a été confirmé par la suite. Les GAG des SLRP sont donc aussi des connecteurs mécaniques pour le transfert de la force et de la charge entre les fibres adjacentes et permettent également le glissement des fibres grâce à leur caractère hydrophile. Le tendon peut ainsi remplir son rôle de transmission de la force entre les muscles et les os (168). L'énergie élastique se trouve stockée et l'excès est dissipé grâce à la collaboration entre les PG et les collagènes fibrillaires, ce qui préserve l'intégrité des TC et prévient les ruptures (169). Ces connecteurs restent fixés au réseau afin de le maintenir en place dans le temps (170).

Dans le cartilage, l'association entre le SLRP épiphycane, un SLRP de classe III (57), et les collagènes entraîne une agrégation des très nombreuses chaînes de GAG du PG majoritaire du cartilage, l'aggrécane. Cette agrégation, régulée par la décorine (171), confère au cartilage

sa capacité de résistance aux chocs et aux forces de compression (85). Les résidus Lys du collagène sont impliqués dans l'interaction avec les chaînes de GAG de l'épiphycane ce qui signifie que leur modification influence la formation du réseau matriciel (172).

Il est important de mentionner que d'autres PG matriciels que les SLRP exercent une influence sur la fibrillogenèse. C'est le cas de l'aggrécane et du versican, deux PG extracellulaires qui appartiennent à la famille des lecticans (57). Le versican permettrait d'augmenter le taux de fibrillogenèse, tandis que l'aggrécane ralentirait le processus. L'hypothèse avancée serait que le très grand nombre de chaînes de GAG sulfatées de l'aggrécane entraînerait une répulsion des collagènes, tandis que le versican aurait l'effet inverse car il en possède moins, malgré leurs structures protéiques similaires (173).

En résumé, les PG matriciels agissent de concert pour participer à l'organisation du réseau collagénique. Les parties protéiques et glucidiques des SLRP interagissent avec des fibrilles de collagène, notamment par le biais d'interactions électrostatiques, pour les rapprocher et favoriser des modifications post-traductionnelles nécessaires à la mise en place de liaisons croisées inter-fibrillaires pour agrandir le réseau. La croissance se fait de façon ordonnée et régulière, les SLRP influençant le diamètre des fibres de collagène, leur longueur et leur régularité. En leur absence, le réseau s'assemble plus rapidement mais de façon erratique, aboutissant à des fibres désorganisées, de taille et de composition variable.

Les FACIT :

Il existe huit FACIT (Figure 6). Parmi eux, le Col IX, la forme longue du Col XII (nommée Col XIIA), le Col XIV, le Col XV et le Col XVIII possèdent des chaînes de GAG, et doivent donc être considérés comme des PG (103).

Les FACIT possèdent deux motifs retrouvés dans de nombreuses protéines extracellulaires : les répétitions de type III de la fibronectine et les domaines de type A du facteur de von Willebrand. Les répétitions de type III de la fibronectine sont impliquées dans l'interaction avec différents acteurs physiologiques tels que l'intégrine $\beta 1$ (174, 175) lors de la différenciation cellulaire (176). Les domaines de type A du facteur de von Willebrand sont responsables de l'interaction de plusieurs molécules extracellulaires avec un large spectre de ligands impliqués dans différents processus physiologiques comme l'adhésion cellulaire, la

migration cellulaire et la transduction du signal (177). Ils sont notamment capables de s'associer avec les thrombocytes (178), le Col VI (179) ou encore le Col III (180). Cependant, ces séquences sont peu étudiées dans le contexte du rôle des FACIT dans la fibrillogenèse.

Dans les MEC, le Col XII est lié par ses domaines collagéniques COL1 et COL2 à la protéine matricielle oligomérique du cartilage (*cartilage oligomeric matrix protein*, COMP). C'est également le cas des Col IX et Col XIV (181, 182). La COMP est une glycoprotéine pentamérique matricielle repérée à l'origine dans le cartilage mais présente dans d'autres tissus comme la peau ou les tendons (183). Elle est capable de se lier aux fibrilles de Col I et de Col II qui ne font pas encore partie du réseau en cours de formation (184). Cette association aux collagènes fibrillaires et aux FACIT favorise la fibrillogenèse : la COMP les accompagne vers le réseau naissant. *In vitro*, la COMP n'est ensuite plus associée au réseau fibrillaire, impliquant que son rôle premier est d'augmenter le taux d'association des fibres et non de maintenir le réseau formé (185). La COMP est donc un intermédiaire entre les FACIT et le Col I en cours de fibrillogenèse (186).

Le Col XIV et les deux isoformes du Col XII restent associés aux collagènes fibrillaires et établissent des interactions avec des SLRP qui se trouvent à la surface des fibres de collagène, comme décrit dans la section précédente. Le Col XIV est capable d'interagir avec la chaîne de GAG de la décorine (187) grâce à la dernière répétition de type III de la fibronectine, à l'extrémité N_{ter} de sa séquence protéique (188). Il serait ainsi impliqué dans les premières étapes de la régulation de la fibrillogenèse (189). De son côté, le Col XII interagit grâce à ses domaines collagéniques avec la chaîne de GAG de la décorine, ainsi qu'avec la protéine-core de la fibromoduline (118, 190) (Figure 14).

Grâce à son long domaine NC3, l'isoforme longue Col XIIA établit des liaisons avec la ténascine X (191). La ténascine X est une large glycoprotéine matricielle qui crée une distance entre les fibres de collagène en maintenant des interactions avec les FACIT, les collagènes fibrillaires (192) et la décorine (193) (Figure 14).

L'isoforme longue $\alpha 1$ (XIIA) possède au moins une chaîne de GAG de type CS. A ce jour, le rôle de cette chaîne de GAG n'est pas connu. Il est établi que le stress mécanique influence l'expression du Col XII au niveau des ligaments péri-dentaires (124) ou des muscles squelettiques (194), suggérant une fonction de ce collagène dans la résistance des tissus aux

chocs et aux pressions. La charge négative de la chaîne de GAG pourrait contribuer à cette résistance (117).

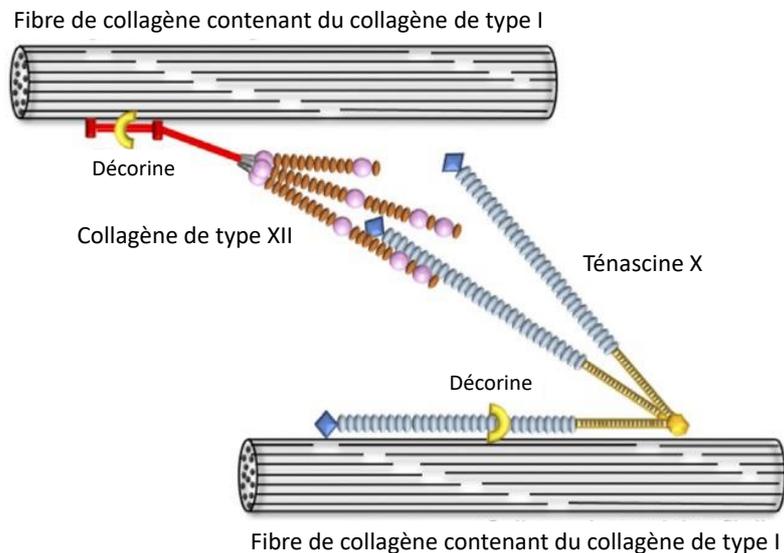


Figure 14. Schéma des interactions du collagène de type XII dans la matrice extracellulaire. Le Col XII interagit via ses domaines collagéniques avec les fibres de collagènes contenant du Col I. La décorine interagit avec les fibres de Col I grâce à sa partie protéique et avec le Col XII grâce à sa chaîne de glycosaminoglycane (GAG). Le domaine non-collagénique NC3 du Col XII interagit avec la ténaïcine X, qui interagit elle-même avec les fibres de collagène contenant du Col I grâce à la décorine. Adapté de (127).

iv. Conclusion

De nombreuses molécules matricielles interviennent au cours de la régulation de la formation du réseau matriciel. Certaines jouent un rôle transitoire, à l'instar de la COMP, et d'autres restent attachées aux fibres, comme les SLRP. Ces PG permettent de contrôler la taille et le diamètre des fibres, mais aussi d'établir des liaisons stables avec des FACIT comme le Col XII et des glycoprotéines comme la ténaïcine X. Ces dernières utilisent leurs grandes tailles et leurs capacités interactives afin de maintenir une distance adéquate entre les fibres pour obtenir un réseau régulier et fonctionnel (Figure 8).

Chaque molécule est importante pour l'architecture de la MEC. Au fil du temps, des équipes de recherche se sont intéressées aux mécanismes à l'origine de la désorganisation de ce réseau, notamment en utilisant des modèles murins invalidés pour des gènes codant certaines protéines matricielles.

Dans des échantillons de peau et de tendon de souris invalidées pour le gène *Dcn* codant pour la décorine, les fibres de collagènes présentent une morphologie anormale, avec des fibres plus ou moins longues aux contours irréguliers. Ceci conforte le rôle important de la décorine dans la régulation de la structure des fibres de collagène (195, 196).

En parallèle, d'autres SLRP ont été invalidés dans des modèles animaux. Les modèles murins invalidés pour le gène *Lum* codant pour le lumican présentent des symptômes similaires à certains sous-types de SED comme l'hyperélasticité et la fragilité de la peau. Des anomalies oculaires sont également observées, telles que l'opacification de la cornée. Ces symptômes sont causés par un assemblage aberrant de fibres de collagènes anormalement épaisses dans ces TC (197). Dans un modèle murin, l'invalidation de la fibromoduline provoque la désorganisation et l'affinement de fibres de collagène. De plus, le lumican est quatre fois plus exprimé dans ces modèles que chez la souris sauvage, ce qui indique une action coordonnée de plusieurs SLRP dans la fibrillogenèse et l'organisation du collagène dans différents TC (198).

Les souris invalidées pour le gène *Col12a1* codant pour le Col XII ont une force plus faible que des souris sauvages (199) et leurs os longs sont plus courts et fins. Les MEC osseuse, articulaire et tendineuse sont désorganisées par rapport à celles des souris contrôles (200–202).

Plusieurs maladies génétiques sont dues à des mutations qui causent des défauts de la synthèse et de la maturation des composants de la MEC, telles que l'ostéogenèse imparfaite (203), des chondrodysplasies (204) et les SED (205). La MEC subit alors une perte de production et/ou des défauts de la structure de ses composants. L'étude des mutations et de leurs conséquences est essentielle pour comprendre les rôles des protéines et des glycoprotéines matricielles (206). Les connaissances actuelles sur le protéome et l'interactome matriciels permettent de comprendre certains aspects de ces maladies qui étaient inconnus auparavant (54, 207).

b. La matrice extracellulaire en situation pathologique

i. Défauts de la synthèse des collagènes fibrillaires

Des mutations des gènes *COL5A1* ou *COL5A2* causent le **SED classique** (SEDC) (9, 208). Les patients sont hétérozygotes mais haploinsuffisants, c'est-à-dire que la copie sauvage du gène

ne permet pas de conserver le phénotype sain de l'individu (209). Une absence totale de Col V chez la souris entraîne une létalité embryonnaire, en corrélation avec l'absence de patients homozygotes ou hétérozygotes composites (137).

Plus rarement, le SEDc peut être la conséquence d'une mutation par substitution du gène *COL1A1* (210), tout comme le **SED vasculaire** (SEdv) (211), le **SED cardiovasculaire** (SEdcv) (212) et le **SED arthrochaliasique** (SEDa) (213). Ces maladies sont classées depuis peu sous l'étiquette de « maladies chevauchantes liées au Col I » (214).

La diminution du clivage des *N*-propeptides du procollagène de type I sécrété dans la MEC est associée à des mutations bialléliques du gène *ADAMTS2*, avec pour conséquences un mauvais assemblage des fibres de collagène dans la MEC et un effet négatif sur la formation des liaisons croisées entre les fibrilles (215). De plus, ADAMTS-2 possède une activité *N*-protéinase envers les Col II, III et V (216, 217). Différents collagènes fibrillaires sont donc touchés par les mutations de *ADAMTS2* à l'origine du **SED dermatoparaxique** (SEdd) (9, 218).

ii. Défauts de l'assemblage des fibres de collagène dus à des mutations du gène *COL12A1*

Le **SED myopathique** (SEdm) est dû à des mutations du gène *COL12A1* (219). De façon intéressante, il s'agit à ce jour du seul sous-type de SED dont des cas de transmissions autosomiques dominantes et récessives ont été découverts (9). Les patients porteurs de mutations de *COL12A1* présentent des atteintes musculaires : une faiblesse musculaire durant l'enfance qui s'améliore avec l'âge et parfois une hypotonie musculaire néonatale. En outre, ils présentent une hyperlaxité articulaire caractéristique des SED (220).

iii. Défauts de l'assemblage des fibres de collagène dus à des défauts de la synthèse des chaînes de glycosaminoglycanes

De nombreuses maladies génétiques rares des TC sont dues à des défauts de la synthèse des GAG qui s'accompagnent de défauts de l'assemblage des fibres de collagène. L'amorce tétrasaccharidique des chaînes de GAG de type CS, DS et HS est un élément crucial dont les défauts de biosynthèse sont spécifiquement impliqués dans des maladies génétiques rares des TC que l'on nomme *linkeropathies*. Parmi les GT impliquées, on retrouve des mutations de la XylT-I qui sont associées au **syndrome de Desbuquois de type 2**, caractérisé par une petite taille, une dysmorphie faciale, une laxité articulaire et des dislocations (221, 222).

Comme mentionné précédemment, des mutations des galactosyltransférases β 4GalT7 et β 3GalT6 sont responsables des **SEDsp-B4GALT7** ou **SEDsp-B3GALT6**, respectivement. Chez ces patients, la glycosylation réduite de la décorine est observée en parallèle de fibres de collagène aux contours anormaux et irréguliers, avec des diamètres erratiques (16, 17, 223). Deux autres expressions cliniques proches sont liées à des mutations de *B3GALT6* : la **SEMD-JL1** (18) et le **syndrome d'Al-Gazali** (29). Enfin, les mutations de la GlcAT-I sont liées au **syndrome de Larsen-like de type B3GAT3** qui est une dysplasie osseuse avec laxité de la peau, dislocations et douleurs articulaires (224).

La maturation des chaînes de GAG est affectée dans le **SED musculocontractural** (SEDmc). Des mutations d'une 4OST des GAG de type DS, la *carbohydrate sulfotransferase 14* (CHST14) (225), ainsi que de la dermatane sulfate épimérase (DSE) (226) sont associées à ce sous-type.

Étant donné que certains FACIT sont aussi des PG, il est raisonnable de suggérer que des mutations d'enzymes impliquées dans la synthèse de l'amorce tétrasaccharidique, dans la sulfatation des chaînes de GAG ou dans l'épimérisation de sucres impactent également la composition et la fonction de ces collagènes non fibrillaires. En effet, la désorganisation du réseau collagénique peut être liée à une altération de la distance séparant normalement les fibres, due notamment aux interactions entre l'isoforme long α 1(XIIA) du Col XII et la ténascine X (191). Il n'existe pas d'étude à notre connaissance sur ce sujet.

Dans ce contexte, nous proposons de détailler les mécanismes fonctionnels et structuraux des GT, puis de s'intéresser aux connaissances actuelles concernant les GT de la biosynthèse de l'amorce tétrasaccharidique des chaînes de GAG. Cette section se concentrera enfin sur la β 3GalT6 pour comprendre son implication dans le **SEDsp-B3GALT6**.

III. La superfamille des glycosyltransférases

1. Généralités

Les GT (EC 2.4.X.y) sont un groupe d'enzymes très répandues, présentes dans les trois domaines du vivant : les archées, les bactéries et les eucaryotes (227). Elles sont classées dans la banque de données *Carbohydrate-Active enZYme* (CAZy)¹ qui répertorie les différentes enzymes impliquées dans le métabolisme, la synthèse ou encore le transport de glucides. Les GT y sont regroupées en familles selon leurs activités enzymatiques et leur structure (228).

a. Activité enzymatique

Les GT sont responsables de la biosynthèse et/ou la modification des oligosaccharides, des polysaccharides et des composés glycoconjugués. Ce processus post-traductionnel est connu sous le nom de glycosylation et implique le transfert d'un glucide d'un substrat donneur vers un substrat accepteur. La diversité fonctionnelle des GT résulte notamment de la multiplicité de leurs substrats accepteurs, tels que des saccharides, des protéines, des lipides et les acides nucléiques (acides ribo- et désoxyribonucléiques (ARN, ADN)) (56, 229).

Les substrats donneurs les plus courants sont des nucléotides diphosphate sucre, par exemple l'uridine diphosphate galactose (UDP-Gal). Lorsqu'un glucide est lié à un nucléotide, il est « activé ». La découverte de ces nucléotides-oses a valu à Luis Federico Leloir et à ses collaborateurs le prix Nobel de chimie en 1970². D'autres substrats donneurs ont depuis été décrits comme les nucléotides monophosphate sucre.

Le transfert du glucide conduit à sa liaison au substrat accepteur par une liaison osidique. Le glucide transféré possède une configuration anomérique qui est préservée ou inversée à la fin de la réaction. Ceci confère à la GT un mécanisme de stéréoisomérisation (230).

b. Structures et repliements

La base de données CAZy, qui est continuellement mise à jour depuis 1998, fournit aujourd'hui des informations sur 135 familles de GT répertoriées. Seulement trois repliements de GT sont actuellement décrits (Figure 15).

¹ <http://www.cazy.org/>

² <https://www.nobelprize.org/prizes/chemistry/1970/summary/>

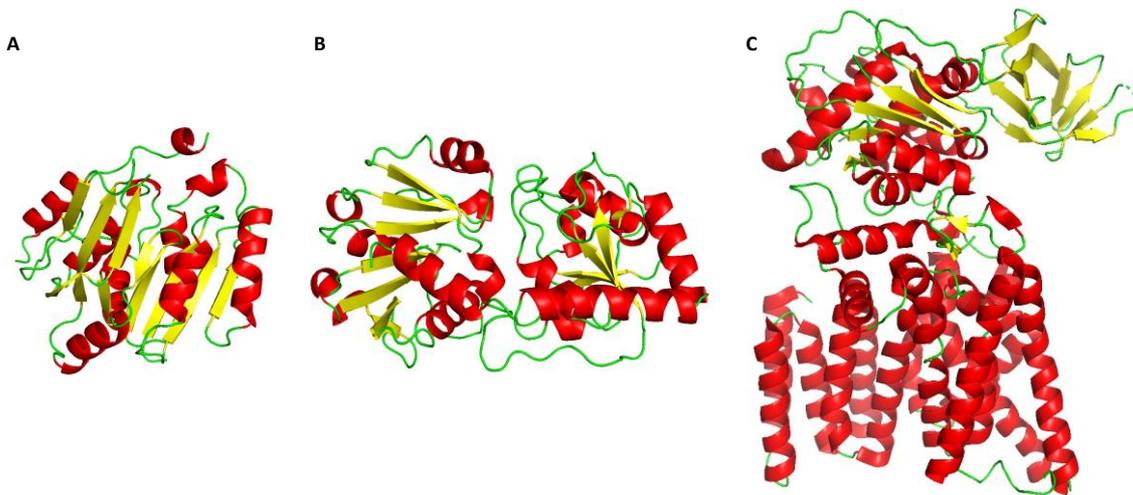


Figure 15. Structures adoptées par les différentes familles de glycosyltransférases. Les hélices α et les feuillets β sont représentés en rouge et en jaune, respectivement. Modèles préparés sur PyMOL. **(A)** Repliement de type GT-A adopté par la *spore coat polysaccharide biosynthesis protein* (SpsA) de *Bacillus subtilis* (PDB : 1QGS) (231). **(B)** Repliement de type GT-B adopté par la β -glucosyltransférase du bactériophage T4 (PDB : 2BGU) (232). **(C)** Repliement de type GT-C adopté par l'oligosaccharyltransférase de *Campylobacter lari* (PDB : 3RCE) (233).

Les repliements des domaines catalytiques des GT les plus couramment observés sont nommés GT-A et GT-B, présents chez les GT utilisant les nucléotide-sucres comme substrats donneurs. Un troisième repliement, appelé GT-C, a été décrit plus récemment. Les GT-A et les GT-B inversent ou conservent la configuration anomérique du glucide, tandis que toutes les GT-C décrites présentent un mécanisme d'inversion (234).

i. Les GT-A

Les GT-A sont des enzymes qui contiennent un repliement de Rossmann constitué d'un feuillet β dont les brins sont séparés par des hélices α de chaque côté (235, 236). Il est estimé que ces repliements sont présents dans 15 % des protéines humaines (237). Au cours de l'étude structurale de protéines contenant des repliements de Rossmann, il est apparu que le motif le plus conservé était constitué uniquement des trois premières structures secondaires, $\beta/\alpha/\beta$. La suite de cette structure peut varier considérablement, avec plus ou moins de résidus et de structures secondaires entre deux brins du feuillet β , ou avec une succession de structures secondaires $\beta/\alpha/\beta/\alpha/\beta$ séparées de petites boucles (238) (Figure 16).

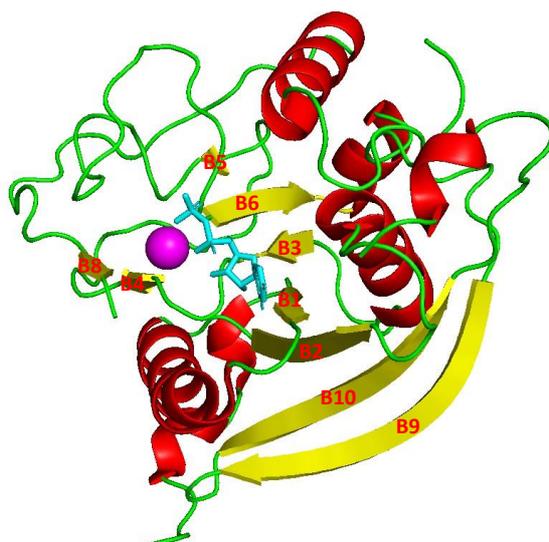


Figure 16. Structure du site actif de la β 1,4-galactosyltransférase 7 humaine en conformation ouverte. Les hélices α et les feuillets β sont représentés en rouge et en jaune, respectivement. L'uridine diphosphate (UDP) du substrat donneur est représenté en cyan et l'ion Mn^{2+} en magenta. Les brins β sont numérotés de l'extrémité N_{ter} à l'extrémité C_{ter} en rouge. Le repliement de Rossmann consiste en deux feuillets β : les brins B5-B6-B3-B1-B2-B10-B9 et B8-B4. Modèle préparé sur PyMOL (PDB : 4IRP) (239).

Ainsi, les GT-A présentent une grande diversité de structures secondaires dans ce repliement Rossmannoïde, ce qui confère une variété structurale dans cette famille. Au sein de ces structures secondaires se trouvent cependant des motifs et des résidus conservés qui jouent des rôles primordiaux dans l'activité transférase (240).

La plupart des GT de la famille GT-A requiert la fixation d'un cation divalent grâce à un motif conservé Asp-x-Asp (DxD), où x représente n'importe quel acide aminé. Le motif DxD forme également des liaisons avec le nucléotide du sucre activé. Ces deux éléments sont ainsi placés correctement dans l'espace de la poche catalytique. La mutation d'un Asp du motif peut considérablement réduire l'affinité pour le substrat ou même entraîner une perte de la fonction enzymatique (241). Cependant, la mutation d'un des deux Asp par un autre résidu acide comme le Glu au niveau de XylT-I a montré que l'activité enzymatique persiste malgré une diminution de l'affinité pour l'uridine diphosphate (UDP) (242). Certaines GT-A possèdent des particularités, comme l'existence d'un motif supplémentaire pour la coordination du cation et du nucléotide. C'est le cas la β 4GalT7, avec la séquence His-x-His (HxH) (243, 244). Enfin, une partie des GT-A ne nécessite pas de cation pour effectuer leur catalyse et ne possède aucun de ces deux motifs (245).

Les GT-A possèdent également un autre motif qui consiste en la séquence x-Glu-Asp (xED). Ce motif contient la base catalytique, un Glu ou un Asp, qui facilite la déprotonation du substrat accepteur permettant l'attaque nucléophile sur le carbone anomérique du substrat donneur (246).

La présence d'une boucle riche en Gly permet la reconnaissance du substrat donneur de plusieurs GT-A (247, 248), tandis qu'une His conservée en C_{ter} coordonne l'ion divalent (249, 250). Si l'His n'est pas présente, d'autres mécanismes de coordination sont adoptés par l'enzyme, comme l'utilisation d'une molécule d'eau ou d'autres résidus chargés (240).

Les GT-A sont des enzymes flexibles dont l'architecture subit une modification conformationnelle pour accueillir le substrat accepteur après la coordination du substrat donneur et du cation (251, 252). Des comparaisons de structures montrent que le positionnement du substrat donneur est généralement conservé chez les GT-A à mécanisme d'inversion, malgré une variabilité des résidus impliqués dans l'interaction avec le nucléotide sucre (251).

ii. Les GT-B

L'activité des GT de type GT-B est indépendante de la présence d'un cation divalent. Elles possèdent deux repliements de Rossmann, dont l'un fixe le nucléotide du substrat donneur grâce à des résidus Lys et Arg de manière similaire aux GT-A indépendante d'un cation, et l'autre le substrat accepteur (253). Les deux substrats se font face au niveau de la limite entre les deux plis de Rossmann, qui est une zone variable en fonction de la nature de l'activité enzymatique de la GT-B (56).

iii. Les GT-C

Les GT-C sont des enzymes membranaires qui utilisent un substrat donneur de type lipide-glucide. Leur découverte tardive et leur nature membranaire a considérablement limité leur étude. Il est estimé que les GT-C représentent 10 % des GT humaines (254).

c. Interactions entre des glycosyltransférases

De nombreuses recherches se sont intéressées à la capacité des GT à former des complexes multimoléculaires. Les GT sont majoritairement des enzymes transmembranaires de type II, c'est-à-dire qu'elles sont constituées d'un petit segment cytosolique, d'un segment

transmembranaire, d'un segment *stem* déstructuré et d'un domaine catalytique (255, 256). Selon l'enzyme considérée, elles peuvent former des homodimères ou des hétérodimères qui régulent la fonction enzymatique (257).

La GlcAT-I, enzyme de la biosynthèse de l'amorce tétrasaccharidique des chaînes de GAG, est capable de former des homodimères grâce à un pont disulfure impliquant la Cys³³, localisée au niveau du segment *stem*. L'homodimérisation améliore l'activité catalysée par la GlcAT-I, ce qui lui confère donc un rôle fonctionnel direct (258).

L'exotosine 1 (EXT1) et l'exotosine 2 (EXT2) sont deux GT de la biosynthèse des chaînes de GAG de type HS catalysant les ajouts successifs de GlcA et de GlcNAc lors de l'élongation de la chaîne (259). Il a été démontré que ces deux enzymes possèdent chacune deux sites actifs pour le transfert du GlcA et du GlcNAc (260). En revanche, l'EXT2 catalyse uniquement le transfert du GlcA et l'EXT1 a une faible activité catalytique lorsqu'elle est exprimée en absence d'EXT2 (261). Elles sont capables de former un hétérodimère par interaction de leurs domaines catalytiques pour obtenir une efficacité catalytique optimale (262).

d. L'organisation structurale des glycosyltransférases de la synthèse de l'amorce tétrasaccharidique

Les quatre GT de la biosynthèse de l'amorce tétrasaccharidique des GAG possèdent un repliement de type GT-A et catalysent le transfert d'un glucide en inversant la configuration de son carbone anomérique. Dans l'ordre, la XylIT-I intervient en premier, puis suivent la β 4GalT7, la β 3GalT6 et la GlcAT-I. Trois de ces enzymes, la XylIT-I (PDB : 6FOA ; (252)), la β 4GalT7 (PDB : 4IRP et 4IRQ ; (239)) et la GlcAT-I (PDB : 1FGG ; (263)) ont été cristallisées. La β 3GalT6 est la seule de ces GT dont la structure n'a pas encore été résolue. Cette section se concentre donc sur les informations structurales dont nous disposons à l'heure actuelle.

i. La xylosyltransférase I

La XylIT-I fait partie de la famille GT14, des GT qui ne requièrent pas de cation divalent. Son substrat donneur est l'uridine diphosphate xylose (UDP-Xyl) et son substrat accepteur est la Ser du motif Ser-Gly des protéines-*core* de PG.

En tant que GT ne nécessitant pas de cation divalent, la XylIT-I se lie à l'UDP-Xyl grâce à des interactions établies entre l'UDP et deux résidus chargés positivement : l'Arg⁵⁹⁸ et la Lys⁵⁹⁹. Le Xyl interagit avec plusieurs résidus de la XylIT-I. De façon notable, une liaison apolaire est

établie entre le C5 du Xyl et le Trp³⁹², proches dans l'espace. Ceci empêcherait d'autres substrats donneurs aux hexoses plus grands de s'insérer dans la poche catalytique, et donc préserverait la spécificité de substrat donneur de l'enzyme.

Le résidu Ser du substrat accepteur est le résidu 0, ceux situés plus en N_{ter} sont numérotés négativement et ceux situés plus en C_{ter} sont numérotés positivement. Le résidu 0 doit obligatoirement être une Ser pour que la réaction ait lieu. Le résidu +1, traditionnellement une Gly, peut être remplacé par d'autres petits résidus tels qu'une Ala, une Ser ou une Val, car l'espace est très resserré à cet endroit de la poche catalytique. Un autre motif accepteur que Ser-Gly n'a cependant jamais été observé. L'absence de site alternatif peut s'expliquer par la spécificité de substrat d'une autre enzyme que la XylT-I. Briggs et Hohenester (252) proposent que la FAM20B (kinase responsable de la phosphorylation du Xyl transféré par la XylT-I, modification nécessaire à la catalyse de la β 3GalT6) nécessite que le résidu +1 ne possède pas de chaîne latérale autre qu'un hydrogène. Afin de vérifier cette hypothèse et de compléter les informations dont nous disposons sur la séquence des événements enzymatiques de la formation de l'amorce tétrasaccharidique, il sera intéressant d'explorer la spécificité de substrat de la FAM20B lorsque sa structure aura été résolue.

Il est établi que les résidus -2, -3, -4 et -5 sont préférentiellement des résidus chargés négativement à cause d'une charge globalement plus positive à cet endroit de la poche catalytique. Mais dans l'ensemble, la nature de la chaîne peptidique au-delà des résidus 0 et +1 peut grandement varier. C'est un contraste intéressant avec la spécificité étroite vis-à-vis du substrat donneur de la XylT-I, l'UDP-Xyl (252).

ii. La β 1,4-galactosyltransférase 7

La β 4GalT7 fait partie de la famille GT7. Son substrat donneur est l'UDP-Gal et son substrat accepteur est le Xyl- β -O-Ser présent au niveau de la protéine-core (72). Elle coordonne le substrat donneur et le cation divalent Mn²⁺ grâce à un Asp appartenant au motif conservé DxD (Asp¹⁶³-Val¹⁶⁴-Asp¹⁶⁵) et possède également un deuxième motif conservé dans les β 1,4-galactosyltransférases, HxH (His²⁵⁷-Leu²⁵⁸-His²⁵⁹). Après coordination du cation et du substrat donneur, la β 4GalT7 passe d'une conformation ouverte à une conformation fermée grâce à une dynamique structurale permise par les nombreuses boucles retrouvées de part et d'autre du pli de Rossmann. La longue boucle (His²⁵⁹-Gly²⁸⁴) et la petite boucle (Gly²²³-Trp²²⁴) changent de conformation suite ou lors de la liaison du substrat donneur. Cela induit la

réorganisation du motif xED (Arg²²⁶-Glu²²⁷-Asp²²⁸) et du résidu Arg²⁷⁰ (264) dans la poche catalytique pour permettre la fixation du substrat accepteur et effectuer la réaction grâce à la base catalytique Asp²²⁸ (Figure 17) (239).

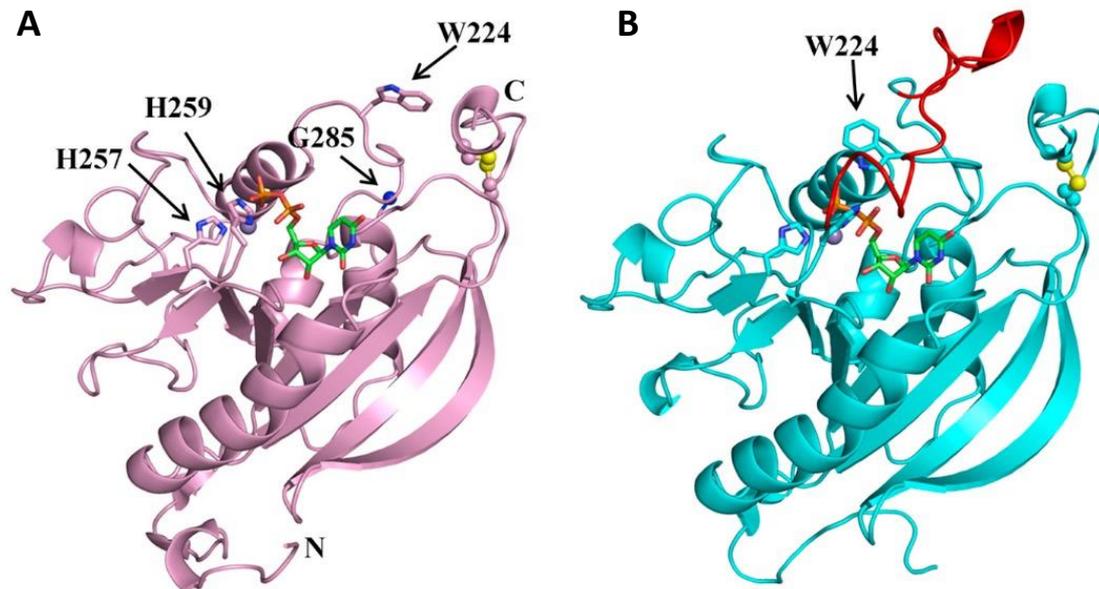


Figure 17. Les structures des deux conformations de la β1,4-galactosyltransférase humaine. L'enzyme se trouve soit en conformation ouverte (A, rose), soit en conformation fermée (B, cyan). L'ion Mn²⁺ et l'uridine diphosphate (UDP) sont représentés violet et en vert/orange, respectivement. Les résidus His²⁵⁷ et His²⁵⁹ faisant partie du motif HxH sont annotés. Les résidus His²⁵⁹ et Gly²⁸⁵ délimitant la longue boucle (rouge, visible uniquement en conformation fermée) sont représentés. Le Trp²²⁴ faisant partie de la petite boucle est annoté. De (239).

iii. La glucuronosyltransférase I

La GlcAT-I est une enzyme de la famille des GT43. Elle fixe son substrat donneur l'uridine diphosphate acide glucuronique (UDP-GlcA) grâce à un résidu Asp conservé et son motif DxD. Contrairement à la β4GalT7 dont la base catalytique est un résidu Asp, la base catalytique de la GlcAT-I est le Glu²⁸¹ (263). L'enzyme humaine a été cristallisée sous une forme soluble, tronquée de ses soixante-quinze premiers acides aminés (265). La Cys³³ semble critique dans son homodimérisation (258) : ainsi, le Gln³¹⁸ d'un monomère serait capable d'interagir avec le substrat accepteur situé dans la poche catalytique de l'autre monomère (263).

De façon intéressante, une GT-A avec un mécanisme d'inversion mais dont l'activité ne requiert par la coordination d'un cation divalent, la β1,3-galactosyl-O-glycosyl-glycoprotéine β1,6-N-acétylglucosaminyltransférase (GCNT1), utilise également un résidu Glu catalytique et

positionne son substrat donneur de la même façon que la GlcAT-I. La nature du résidu catalytique serait un paramètre important du positionnement dans l'espace du substrat donneur pour les GT-A à mécanisme d'inversion de la configuration du carbone anomérique (254).

2. La β 1,3-galactosyltransférase 6, enzyme de la famille GT31

Comme mentionné précédemment, la β 3GalT6 n'a pas été cristallisée à l'heure actuelle. L'étude de sa structure apporterait des informations importantes quant à la fixation de ses substrats, aux résidus importants pour sa structure et/ou sa fonction et aux changements conformationnels qu'elle pourrait subir. Toutefois, grâce à l'introduction de nouvelles approches, l'étude des structures protéiques non cristallisées devient plus facile et plus fiable. C'est le cas de la β 3GalT6, pour laquelle AlphaFold (266, 267) a fourni un modèle présentant un score de confiance élevé (268) (Figure 18). En l'absence d'une structure, il s'agit d'un point de départ intéressant pour formuler des hypothèses sur son fonctionnement enzymatique, sans pour autant remplacer les informations que la détermination structurale pourrait fournir (269). Nous avons eu recours à cette prédiction de structure pour des études de dynamique moléculaire (Diana *et al.*, *in external peer-review*).

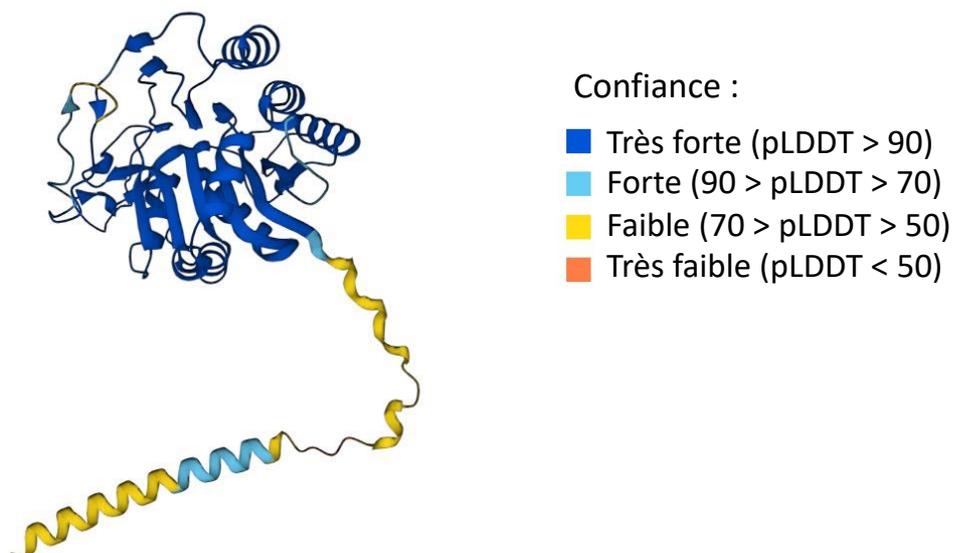


Figure 18. Prédiction de structure de la β 1,3-galactosyltransférase 6 humaine par le logiciel de prédiction AlphaFold. L'indice de confiance est mesuré en pLDDT, une unité de degré de confiance par résidu, et est représenté dans la structure par des couleurs (267).

b. Activité enzymatique

La $\beta 3\text{GalT6}$ est la troisième enzyme agissant lors de la biosynthèse de l'amorce tétrasaccharidique des chaînes de GAG. Elle catalyse la formation d'une liaison *O*-osidique de type $\beta 1,3$ entre le substrat accepteur (le disaccharide Gal-Xyl, dont le Xyl est phosphorylé sur le carbone en position C2 et lié covalentement à la Ser de la protéine-core du PG) et un Gal fourni par le substrat donneur UDP-Gal. La réaction nécessite la présence d'un cation divalent, le Mn^{2+} (67, 271) (Figure 20).

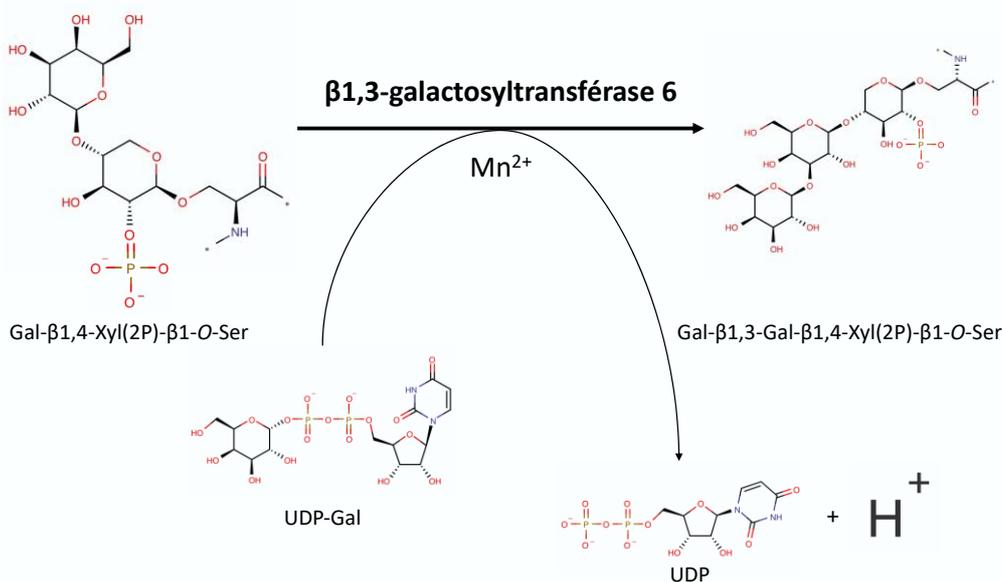


Figure 20. Réaction enzymatique catalysée par la $\beta 1,3$ -galactosyltransférase 6. L'enzyme catalyse le transfert du galactose (Gal) fourni par le substrat donneur uridine diphosphate galactose (UDP-Gal) sur le substrat accepteur Gal- $\beta 1,4$ -Xyl(2P)- $\beta 1$ -O-Ser en présence de Mn^{2+} . Le produit de la réaction est le Gal- $\beta 1,3$ -Gal- $\beta 1,4$ -Xyl(2P)- $\beta 1$ -O-Ser.

La $\beta 3\text{GalT6}$ présente plusieurs motifs conservés directement impliqués dans son activité enzymatique dont :

- Le motif DxD (Asp¹⁵⁶-Asp¹⁵⁷-Asp¹⁵⁸) qui est nécessaire à la fixation du Mn^{2+} et de l'UDP-Gal ;
- Le motif xED (Ser²⁴⁰-Glu²⁴¹-Asp²⁴²) contenant la base catalytique prédite, l'Asp²⁴² (251).

c. Structure et repliement

La β 3GalT6 est une enzyme transmembranaire de type II de l'appareil de Golgi. Elle est constituée d'un court segment cytosolique, d'un segment transmembranaire, d'un segment largement déstructuré (nommé *stem*) et d'un domaine catalytique (Figure 18).

Sa structure n'étant pas résolue, des alignements de séquences avec d'autres enzymes aident à prédire les résidus importants pour la reconnaissance des substrats et/ou pour la catalyse enzymatique. Parmi elles, les enzymes les plus pertinentes comprennent d'autres transférases, par exemple les autres β 3GalT humaines (section III.2.d) ou encore des structures résolues de la même famille, les GT31 (section III.2.e) et des galactosyltransférases comme la β 4GalT7 qui intervient dans cette même voie de biosynthèse.

d. Les β 1,3-galactosyltransférases

Les β 3GalT catalysent toutes la même réaction, le transfert d'un Gal en position β 1,3, mais les substrats accepteurs diffèrent grandement. Le glucide du substrat accepteur sur lequel est formée la liaison covalente peut-être un Gal, une GalNAc ou une GlcNAc. Il est intéressant de noter que sur les six β 3GalT connues, seule la β 3GalT6 catalyse la formation d'une liaison Gal- β 1,3-Gal.

Patel et Bajali (272) ont utilisé la structure de la *spore coat polysaccharide biosynthesis protein* (SpsA) (PDB : 1QG8 ; (231)) comme référence pour prédire les structures secondaires des β 3GalT et les résidus importants à leur structure et fonction. Sept domaines conservés sont retrouvés chez les β 3GalT. Le motif xED est situé juste devant une hélice et est totalement conservé dans le domaine conservé n°7. Le motif DxD est lui aussi totalement conservé, dans le domaine conservé n°5 (Figure 21).

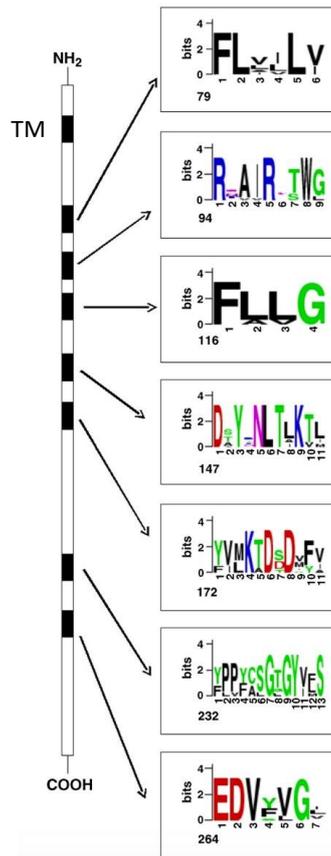


Figure 21. Séquences des motifs conservés de la famille des β 1,3-galactosyltransférases. Les motifs ont été générés en alignant les séquences des β 1,3-galactosyltransférases (β 3GalT) humaines. Les chiffres des axes des abscisses représentent la position des résidus dans chaque motif. Les ordonnées représentent le taux de présence des résidus mentionnés. Le chiffre en bas à gauche représente la position dans la séquence de la β 1,3-galactosyltransférase 1 (β 3GalT1). Adapté de (272).

Afin de vérifier l'importance fonctionnelle de certains résidus, une stratégie classique consiste à les muter notamment en Ala et à analyser l'impact sur l'activité enzymatique et l'affinité pour les substrats. Malissard *et al.* (273) ont ainsi généré des mutations sur des résidus-clés de la β 1,3-galactosyltransférase 1 (β 3GalT1) murine. Si la totalité du domaine conservé n°3 est mutée en Ala, l'activité est entièrement perdue. Ce domaine consiste en 4 résidus successifs FLLG correspondant aux résidus Phe⁹⁴-Ala⁹⁵-Val⁹⁶-Gly⁹⁷ dans la β 3GalT6 humaine. D'autres mutations conservatives ou non conservatives ont été effectuées, parfois sur des résidus totalement conservés dans la β 3GalT6. Ainsi, si l'on extrapole les résultats des mutations générées sur la β 3GalT1 murine, les résidus Trp⁸⁰ et Trp¹⁴² de la β 3GalT6 humaine pourraient être impliqués dans la reconnaissance de l'UDP-Gal et le résidu Glu²⁴¹ du motif xED jouerait un rôle dans l'affinité pour ses deux substrats.

Enfin, d'autres résidus conservés sont également retrouvés entre les différentes β 3GalT humaines. Par exemple, la boucle riche en Gly de reconnaissance du substrat accepteur (247) (Gly²¹⁵-Gly²¹⁶-Gly²¹⁷ pour β 3GalT6) ou l'His en C_{ter} qui coordonne le cation divalent et participe à la catalyse pour certaines GT-A (249) (His²⁷⁹ pour β 3GalT6).

e. Deux structures de GT31 résolues

i. La β 1,3-*N*-acétylglucosaminyltransférase 2 humaine

La seule GT31 d'*Homo sapiens* dont la structure a été cristallisée est la β 1,3-*N*-acétylglucosaminyltransférase 2 (B3GNT2), qui catalyse l'ajout d'un GlcNAc sur le substrat Gal- β 1,4-GlcNAc. Kadirvelraj *et al.* (274) ont cristallisé cette enzyme en complexe avec l'UDP et le Mg²⁺ (PDB : 7JH0).

Au sein de la B3GNT2, les résidus Lys²⁸⁸ et Tyr²⁸⁹ situés dans une longue boucle forment des interactions avec les deux phosphates de l'UDP du substrat donneur, ce qui permettrait de recruter le Mg²⁺ pour sa coordination par le motif DxD. En effet, le cation divalent interagit avec l'UDP, l'Asp²⁴⁷ du motif DxD et l'His³⁷⁶. L'Asp²⁴⁶ du motif DxD interagit avec l'uridine de l'UDP, tout comme l'Asp²¹⁵ et la Lys²²³ (Figure 22A). L'Asp³³³ du motif xED est prédit comme étant la base catalytique, et la Tyr³⁰³ interagit avec le Gal du substrat accepteur (Figure 22B).

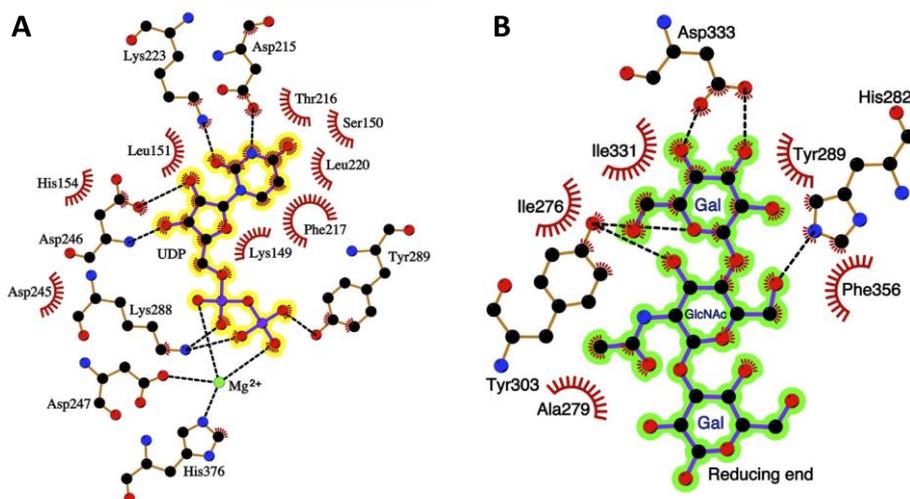


Figure 22. Schéma du site actif de la β 1,3-*N*-acétylglucosaminyltransférase 2 humaine. (A) Interactions entre les résidus du site actif, le cation divalent Mg²⁺ et l'uridine diphosphate (UDP) du substrat donneur (surligné en jaune). **(B)** Interactions entre les résidus du site actif et le substrat accepteur (surligné en vert). Les liaisons hydrogène et les encombrements sont représentés par des lignes noires pointillées et par des arcs de cercle rouges, respectivement. Gal : galactose. GlcNAc : *N*-acétyl-*D*-glucosamine. Adapté de (274).

Ces résidus sont fortement conservés au sein de la β 3GalT6 (Tableau V). Un résidu basique de la β 3GalT6 jouant un rôle comparable à la Lys²⁸⁸ de la B3GNT2, l'Arg¹⁹⁷, a été étudié au cours d'une étude portant sur un patient diagnostiqué SEDsp et porteur d'une mutation homozygote non conservative de β 3GalT6 : la mutation p.(Arg197Cys). En utilisant la structure cristallisée de la B3GNT2, l'équipe a modélisé la région impactée de la β 3GalT6 et a indiqué que l'Arg¹⁹⁷ est prédite comme ayant la même orientation que la Lys²⁸⁸ de B3GNT2 pour former une liaison avec un phosphate de l'UDP (Figure 23) (36)

| β 1,3-galactosyltransférase 6 (β 3GalT6) | β 1,3-N-acétylglucosaminyltransférase 2 (B3GNT2) | Résidu de B3GNT2 interagissant avec | Motif impliqué ou rôle du résidu |
|--|--|-------------------------------------|----------------------------------|
| His ²⁷⁹ | His ³⁷⁶ | Mg ²⁺ | |
| Asp ¹⁵⁶ Asp ¹⁵⁷ Asp ¹⁵⁸ | Asp ²⁴⁵ Asp ²⁴⁶ Asp ²⁴⁷ | Uridine Mg ²⁺ | Motif DxD |
| Arg ¹⁹⁷ | Lys ²⁸⁸ | Diphosphate | |
| Trp ¹⁹⁸ | Tyr ²⁸⁹ | Diphosphate | |
| Asp ¹²⁷ | Asp ²¹⁵ | Uridine | |
| Lys ¹³⁵ | Lys ²²³ | Uridine | |
| Asp ²⁴² | Asp ³³³ | Galactose | Base catalytique du motif xED |
| Tyr ²¹² | Tyr ³⁰³ | Galactose | |

Tableau V. Comparaison des résidus de la β 1,3-N-acétylglucosaminyltransférase 2 impliqués dans la reconnaissance des substrats avec les résidus de la β 1,3-galactosyltransférase 6. Les résidus sont comparés sur la base d'un alignement des deux séquences en acides aminés (non montré). Les interactions des résidus de la B3GNT2 ont été déterminées expérimentalement par (274).

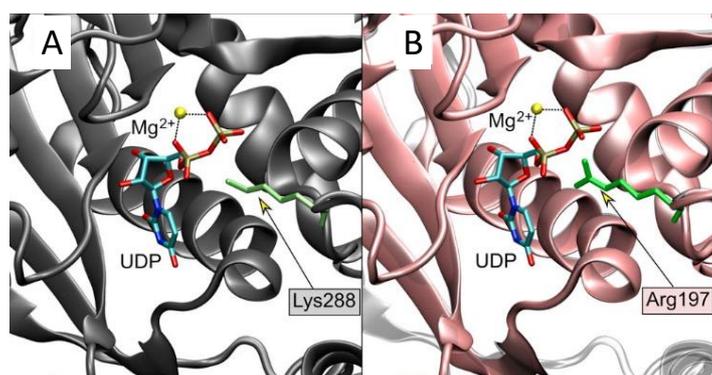


Figure 23. Visualisation des domaines catalytiques de la structure résolue de la β 1,3-N-acétylglucosaminyltransférase 2 et d'un modèle de la β 1,3-galactosyltransférase 6. (A) Structure du domaine catalytique de la β 1,3-N-acétylglucosaminyltransférase 2 (B3GNT2) déterminé par Kadirvelraj *et al.* (274). (B) Modèle du domaine catalytique de la β 1,3-galactosyltransférase 6 (β 3GalT6) généré par Leoni *et al.* (36). L'uridine diphosphate (UDP) et le cation divalent Mg²⁺ utilisé par la B3GNT2 sont représentés en cyan/jaune et jaune, respectivement. La Lys²⁸⁸ de B3GNT2 et l'Arg¹⁹⁷ de β 3GalT6 sont représentées en vert pâle et vert, respectivement. Adapté de (36).

ii. La glycoprotéine-N-acétylgalactosamine 3-β-galactosyltransférase 1 de *D. melanogaster*

La glycoprotéine-N-acétylgalactosamine 3-β-galactosyltransférase 1 (C1GalT1) de *Drosophila melanogaster* est une autre GT31 dont la structure a été résolue. Elle catalyse l'ajout d'un Gal sur un GalNAc. La fixation de la partie nucléotidique du substrat donneur UDP-Gal est considérée comme la « force motrice » qui permet la reconnaissance du substrat accepteur (275). Ce mécanisme a été directement comparé à celui d'une autre GT-A, la GalNAc-T2, qui possède deux conformations : la conformation « fermée » et la conformation « active ». La longue boucle Val³⁶⁰-Gly³⁷² va se déplacer lors de la fixation de l'UDP-Gal pour passer en conformation active et accueillir le substrat accepteur. Ce mécanisme est rendu possible grâce à un résidu en particulier, la Phe¹⁰⁴, qui ne se trouve ni dans le domaine catalytique ni dans la boucle en question, et qui est situé bien plus loin dans la séquence protéique (276). La dynamique de longues boucles est retrouvée dans d'autres GT, un exemple déjà décrit de ce manuscrit étant celui de la β4GalT7 (239).

3. Conclusion

Les GT sont des protéines flexibles subissant des modifications conformationnelles indispensables à leur activité. Des résidus-clés font partie de motifs conservés tel que le motif DxD qui coordonne le substrat donneur et le cation divalent, ou encore le motif xED qui contient la base catalytique. D'autres résidus, qui ne sont pas nécessairement présents dans des motifs particuliers, sont conservés à travers les GT pour établir des interactions avec les substrats et pour former des complexes, par exemple des homodimères nécessaires à leur activité enzymatique.

Le nombre de mutations de β3GalT6 actuellement recensé est très important. Elles conduisent à plusieurs maladies génétiques rares des TC, dont le SEDsp, avec une difficulté de diagnostic et de pronostic, des symptômes variables, une espérance de vie réduite, et globalement un manque de compréhension de la communauté scientifique. La structure tridimensionnelle de la β3GalT6 pourrait se révéler grandement utile pour comprendre les conséquences d'une mutation sur l'enzyme. Des pistes peuvent être établies provisoirement en utilisant les informations expérimentalement obtenues pour des enzymes aux mécanismes similaires et des approches de modélisation moléculaire, telles que celles mises en œuvre dans notre travail.

Problématique

Les SED représentent un groupe hétérogène de maladies génétiques rares des TC causées par des mutations de constituants de la MEC ou de protéines responsables de leur synthèse. Les PG, principaux composants des MEC, sont des macromolécules constituées d'une partie protéique et d'une ou plusieurs chaînes polysaccharidiques de GAG. Ils assurent de nombreux rôles physiopathologiques au sein des MEC, allant des interactions avec les molécules d'eau, les collagènes et de multiples constituants matriciels (277) à la régulation de l'activité de médiateurs solubles impliqués dans plusieurs cascades de signalisation (278).

Le gène *B3GALT6* code pour la β 3GalT6, une GT qui catalyse l'ajout d'un résidu Gal lors d'une des premières étapes de la synthèse des chaînes de GAG. Le SEDsp est une forme rare de SED associée à plusieurs variants pathogènes du gène *B3GALT6* présents à l'état homozygote ou hétérozygote composite. Ces mutations sont identifiées en particulier au niveau de domaines conservés du domaine catalytique de la β 3GalT6 dont les études structurales et fonctionnelles sont rares et incomplètes. Les phénotypes observés chez les patients atteints de SEDsp sont sévères, caractérisés notamment par des atteintes au niveau du système musculosquelettique. Le spectre de symptômes du SEDsp se chevauche avec l'expression clinique d'affections proches dues à d'autres mutations de *B3GALT6*, telles que la SEMD-JL1 (18) et le syndrome d'Al-Gazali (29). Les aspects cliniques se chevauchent également avec d'autres ostéochondrodysplasies d'origine génétique différente, comme le SEDk, le BCS ou le SEDmc (9). Il est donc extrêmement complexe d'établir le diagnostic de cette affection.

On dénombre aujourd'hui quarante-sept variants pathogènes de *B3GALT6* qui sont des délétions ou des substitutions entraînant l'apparition de mutations faux-sens, non-sens, ou causant un décalage du cadre de lecture (17, 18, 28, 29, 33–38, 42–52). Ces variants alléliques ont été en partie identifiés grâce à des travaux de notre équipe réalisés en collaboration avec le Dr. Fransiska Malfait (*Center of Medical Genetics, Université de Gand, Gand, Belgique*). L'équipe GlycoBio a ainsi participé à l'identification de dix-huit mutations de *B3GALT6* au sein de douze familles indépendantes. Les études réalisées grâce aux cultures de fibroblastes de peau de ces patients ont montré que les mutations de *B3GALT6* impactent à des degrés variables la glycosylation de la décorine et des PG matriciels et cellulaires. Les études de microscopie électronique de biopsies de peau de patients montrent une altération de

l'architecture du réseau de collagène. Cependant, les mécanismes par lesquels les défauts de glycosylation des PG conduisent à des défauts de l'intégrité de la MEC et de ses propriétés fonctionnelles et structurales ne sont pas établis.

Notre principal objectif de travail est de mieux cerner les mécanismes pathogéniques impliqués dans le SEDsp-*B3GALT6*, de mieux définir les relations génotype / phénotype de cette maladie et, à terme, de permettre au clinicien d'améliorer le diagnostic, le pronostic et la prise en charge des patients. Pour ce faire, nous proposons d'explorer l'impact des défauts génétiques à différentes échelles et de façon pluridisciplinaire.

- (1) Le premier axe consiste en l'**étude des conséquences d'une mutation individuelle sur la fonction de la β 3GalT6**. Nous concentrons nos efforts sur les mutations p.(Tyr182Cys) (50), p.(Asp207His) et p.(Gly217Ser) (17). Celles-ci sont retrouvées à l'état homozygote chez des patients SEDsp, ce qui nous permet de relier l'impact sur la fonction à une seule mutation. La transfection transitoire des enzymes mutées dans des cellules HeLa dont le gène *B3GALT6* a été invalidé par la technologie CRISPR-Cas9 nous permet d'analyser la biosynthèse des chaînes de GAG à partir de différents substrats. En parallèle, l'activité enzymatique spécifique des mutants est étudiée dans ce même modèle. Cette seconde partie a fait l'objet de ma première année de thèse et est décrite dans ce manuscrit, mais n'a pas permis l'obtention de résultats concluants concernant l'activité des mutants. En conséquence, nous avons modifié notre stratégie et étudié l'activité enzymatique spécifique des mutants surexprimés dans la bactérie *Escherichia coli* (*E. coli*) sous la forme de protéines tronquées, solubles et purifiées.
- (2) Le second axe consiste en l'**étude des conséquences d'une mutation individuelle sur la structure de la β 3GalT6**. La structure secondaire des protéines purifiées mentionnées est explorée par dichroïsme circulaire (DC). Leur diamètre hydrodynamique et leur distribution de taille sont étudiés par diffusion dynamique de la lumière (*dynamic light scattering*, DLS). Des modèles structuraux des mutants générés par une approche de dynamique moléculaire ont également été utilisés pour prédire des modifications de l'intensité des liaisons intramoléculaires et du taux de fluctuation de segments polypeptidiques.

(3) Le troisième axe consiste en l'**étude des conséquences de l'invalidation du gène *B3galt6* sur la morphologie et les caractéristiques biomécaniques de la MEC dans un modèle de synthèse matricielle murin.** L'invalidation de *B3galt6* dans la lignée de tératocarcinome de souris ATDC5 est réalisée par le système CRISPR-Cas9 transduit dans les cellules lors d'une infection adénovirale. Dans notre système expérimental, nous pouvons induire la chondrogenèse de ces cellules, ce qui permet l'expression de nombreuses protéines matricielles. Les cellules produisent ainsi une MEC abondante et organisée sous la forme d'un pseudo-tissu. La glycosylation des PG du sécrétome et du matrisome est étudiée. Les pseudo-tissus sont soumis à un test d'extensibilité mais également inclus en paraffine afin de réaliser des colorations histologiques. Enfin, la microscopie à force atomique (*atomic force microscopy*, AFM) fournit des informations sur la morphologie et l'élasticité des pseudo-tissus. Tous ces paramètres nous renseignent quant aux défauts d'organisation et d'assemblage de la MEC induits par la déficience en β 3GalT6.

Matériels et méthodes

I. Étude de l'impact de mutations de la β 1,3-galactosyltransférase 6 dans un modèle cellulaire invalidé pour le gène *B3GALT6*

1. Le modèle cellulaire HeLa

a. Génération d'un modèle de cellules HeLa invalidé pour le gène *B3GALT6*

Pour étudier l'effet de mutations de la β 3GalT6 sur son activité enzymatique, nous avons exprimé les enzymes mutées dans un modèle cellulaire dont le gène codant pour la β 3GalT6 a été préalablement invalidé. Ce modèle de cellules HeLa, nommé H7, a été obtenu au cours d'une thèse précédente au sein de l'équipe (279) et a été généré à partir d'une lignée HeLa obtenue à la *European Collection of Authenticated Cell Cultures* (Sigma-Aldrich, Saint-Quentin Fallavier, France). La lignée H7 porte une délétion de cinquante-deux paires de bases au niveau du gène *B3GALT6* conduisant à la mutation p.(Val61Alafs*199) avec un décalage du cadre de lecture et synthèse d'une protéine incomplète.

b. Culture cellulaire

Les cellules sont cultivées dans du milieu *Dulbecco's Modified Eagle Medium* (DMEM) à teneur élevée en glucose (4,5 g/L) (Thermo Fisher Scientific, Courtaboeuf, France) supplémenté de 10 % (v/v) de sérum de veau foetal (SVF) (Gibco, Thermo Fisher Scientific), 1 % (v/v) de L-glutamine (Sigma-Aldrich) et 1 % (v/v) de pénicilline / streptomycine (Sigma-Aldrich). Les lavages des cellules sont effectués avec du tampon phosphate salin (*phosphate buffered saline*, PBS) (Sigma-Aldrich). Lors des manipulations de passage ou d'amplification, les tapis cellulaires sont décollés grâce à un mélange trypsine - acide éthylènediaminetétraacétique (EDTA) 0,25 % (p/v) (Sigma-Aldrich). Les cellules sont cultivées dans un incubateur HERAcell™ 240i (Thermo Fisher Scientific) sous atmosphère humide à 37 °C et 5 % (v/v) de CO₂.

Les cellules H7 sontensemencées en plaques 6-puits, à raison de 3.10⁵ cellules par puits, un jour avant transfection transitoire par le vecteur permettant l'expression de la β 3GalT6 sauvage ou mutée fusionnée à une étiquette hémagglutinine (HA). Ainsi, par une combinaison de plusieurs techniques, elles seront utilisées pour l'étude de la synthèse des chaînes de GAG en cas de mutation de β 3GalT6.

2. Obtention des vecteurs d'expression des mutants de la β 1,3-galactosyltransférase 6 humaine

a. Mutagenèse dirigée

Afin de reproduire des mutations du gène *B3GALT6* décrites chez les patients atteints de SEDsp, le kit commercial de mutagenèse dirigée *QuikChange II XL* (Agilent Technologies, Les Ulis, France) a été utilisé. Dans le cadre de ma thèse, nous nous sommes concentrés sur trois mutations homozygotes retrouvées chez des patients SEDsp-*B3GALT6* : p.(Tyr182Cys) (50), p.(Asp207His) et p.(Gly217Ser) (17).

Le mélange réactionnel est réalisé selon les recommandations du fournisseur. Il consiste en 50 ng de plasmide pcDNA3.1(+)- β 3GalT6-WT disponible au laboratoire, 2,5 unités d'ADN polymérase PfuUltra™, 3 μ L de *QuikSolution* pour améliorer la qualité de la réplication, 10 mM d'un mix de désoxyribonucléotides triphosphates (dNTP) et 125 ng de chaque amorce complémentaire contenant la mutation d'intérêt (Tableau VI) dans un volume final de 50 μ L. Le programme de réaction en chaîne par polymérase (*polymerase chain reaction*, PCR) est le suivant : (i) 95 °C, 2 min (ii) 95 °C, 1 min (iii) condition variable³, 50 sec (iv) 68 °C, 10 min (v) 68 °C, 7 min. Les étapes (ii) à (iv) sont répétées 18 fois.

| Mutation générée | Brin | Séquence |
|------------------|----------|---|
| Tyr182Cys | Sens | 5'-CCGCCTCT G CTGGGGC-3' |
| | Antisens | 5'-GCCCCAGCAGAGGCGG-3' |
| Asp207His | Sens | 5'-GGCAACTCGCCACTACTACCT-3' |
| | Antisens | 5'-AGGTAGTAGT G GCAGAGTTGCC-3' |
| Gly217Ser | Sens | 5'-GCGCTGGGCGGC A GCTACGTGCTCTC-3' |
| | Antisens | 5'-GAGAGCACGTAGCTGCCGCCAGCGC-3' |

Tableau VI. Liste des amorces nucléotidiques utilisées pour la mutagenèse dirigée de la β 1,3-galactosyltransférase 6. Les nucléotides mutés sont indiqués en gras.

Le produit de PCR obtenu est ensuite incubé à 37 °C pendant une nuit avec 10 unités d'enzyme de restriction DpnI provenant du kit de mutagenèse dirigée pour cliver de façon spécifique le plasmide matrice restant.

³ Tyr182Cys : 58 °C ; Asp207His : 60 °C ; Gly217Ser : 72 °C, 2 % de DMSO.

b. Préparation des vecteurs plasmidiques

Un volume de 2 μL de produit de PCR digéré par DpnI est mis en présence de 50 μL de bactéries *E. coli* chimiocompétentes *One Shot*TM TOP10 (Invitrogen, Thermo Fisher Scientific) sur glace pendant 30 minutes. Puis, un choc thermique est réalisé en incubant successivement le mélange durant 30 secondes dans un bloc chauffant réglé à 42 °C et 3 minutes sur glace. Un volume de 150 μL de bouillon lysogène (*lysogeny broth*, LB) (Sigma-Aldrich) est ajouté à la préparation qui est incubée à 37 °C sous agitation pendant une heure.

Les précultures des bactéries transformées sont étalées sur un milieu sélectif LB-Agar supplémenté de 100 $\mu\text{g}/\text{mL}$ d'ampicilline. Les géloses sont incubées à 37 °C pendant une nuit. Après repiquage de clones et culture de ces derniers, les plasmides sont extraits et purifiés à l'aide du kit de minipréparation Qiaprep[®] (Qiagen) puis dosés au NanodropTM One (Thermo Fisher Scientific). La présence des mutations dans la séquence codante de la β3GalT6 tronquée est vérifiée par double séquençage (Eurofins Genomics, Ebersberg, Allemagne).

3. Étude de la synthèse des chaînes de glycosaminoglycanes

a. Synthèse à partir d'un substrat exogène

Les cellules H7 sont transfectées par le vecteur pcDNA3.1(+) (Thermo Fisher Scientific) permettant l'expression de la β3GalT6 sauvage ou porteuse d'une mutation Tyr182Cys, Asp207His ou Gly217Ser. Un contrôle négatif est effectué en transfectant les cellules avec le vecteur pcDNA3.1(+) vide. La transfection transitoire est réalisée le lendemain de l'ensemencement en plaques 6-puits grâce à l'agent de transfection TurbofectTM (Thermo Fisher Scientific) selon les recommandations du fournisseur, avec 3 μg d'ADN plasmidique par puits.

La méthodologie du radiomarquage des chaînes de GAG a été précédemment décrite par Bui *et al.* (264). Deux jours après la transfection, les cellules sont incubées pendant une nuit dans du milieu Fisher (Thermo Fisher Scientific) supplémenté de 2 % (v/v) de SVF, 10 $\mu\text{Ci}/\text{ml}$ de $\text{Na}_2[^{35}\text{SO}_4]$ (PerkinElmer SAS, Villebon-sur-Yvette, France) et 0, 2 ou 5 μM de 4-méthylombélliferyl- β -D-xylopyranoside (4-MOX) (Sigma-Aldrich). Le lendemain, le milieu radiomarqué est récolté et les chaînes de GAG sulfatées radiomarquées sont séparées des sulfates radiomarqués non incorporés sur une colonne Sephadex G-50 (Cytiva, Uppsala, Suède). Les chaînes de GAG radiomarquées sont quantifiées par comptage en scintillation liquide dans un cocktail de scintillation Ultima Gold (PerkinElmer SAS). En parallèle, les cellules

sont lysées et la quantité d'ADN est dosée grâce au kit *Quant-iT™ dsDNA HS Assay* (Thermo Fisher Scientific) sur un fluorimètre Qubit™ (Thermo Fisher Scientific), selon les recommandations du fournisseur. Les données de radiomarquage sont standardisées en fonction de la quantité d'ADN.

b. Synthèse à partir de la décorine recombinante

Les cellules H7 sont co-transfectées avec deux plasmides en utilisant le Turbofect™ selon les recommandations du fournisseur. Le premier plasmide, utilisé à une quantité de 1 µg, est un pcDNA3.1(+) contenant la séquence codante de la décorine humaine fusionnée à une étiquette 6His. Le second, utilisé à une quantité de 2 µg, est un pcDNA3.1(+) permettant l'expression de la β3GalT6 sauvage ou porteuse d'une mutation Tyr182Cys, Asp207His ou Gly217Ser. Un contrôle négatif est effectué en co-transfectant 1 µg du premier plasmide et 2 µg de pcDNA3.1(+) ne contenant pas la séquence codante de la β3GalT6. 48 heures après transfection, les cellules sont cultivées une nuit supplémentaire dans deux milieux différents selon l'objectif prédéfini.

Afin d'obtenir des informations sur la longueur des chaînes de GAG synthétisées, les cellules co-transfectées sont cultivées dans un milieu Fisher (Thermo Fisher Scientific) supplémenté de 10 µCi/mL de Na₂[³⁵SO₄] (PerkinElmer SAS) pour le radiomarquage (264). Les milieux radiomarqués sont récoltés le lendemain. La décorine étiquetée 6His surexprimée est purifiée à l'aide de billes magnétiques Ni-NTA (Thermo Fisher Scientific) selon les recommandations du fournisseur. Après lavages et élution de la décorine dans un volume adapté de tampon Tris-HCl (pH 7) contenant 250 mM d'imidazole, les échantillons sont mélangés avec un tampon de charge *XT Sample Buffer* (Bio-Rad, Marnes-la-Coquette, France) dans un volume final de 30 µL pour visualiser les chaînes de GAG radiomarquées séparées par électrophorèse sur gel de polyacrylamide contenant du dodécylsulfate de sodium (*sodium dodecyl sulfate polyacrylamide gel electrophoresis*, SDS-PAGE). Les échantillons sont dénaturés lors d'une incubation à 95 °C pendant 5 minutes puis déposés sur un gel à gradient 4-12 % Criterion™ XT Bis-Tris (Bio-Rad) et soumis à une tension constante de 90 V pendant deux heures. Le gel est ensuite fixé pendant une nuit dans un mélange d'eau, d'acide acétique 20 % (v/v) et de méthanol 20 % (v/v), puis après rinçage à l'eau, le gel est incubé pendant une heure dans une solution Amplify (Cytiva) permettant une amplification du signal. Le gel est

séché dans un sécheur de gel *Gel-Dry*TM (Thermo Fisher Scientific), puis mis sous film (Cytiva) durant minimum sept jours à -20°C avant d'être révélé par autoradiographie.

Afin d'obtenir des informations complémentaires sur le taux de glycosylation de la décorine, les cellules co-transfectées sont cultivées dans un milieu de culture DMEM sans SVF. Les milieux sont récoltés le lendemain et la décorine est purifiée selon la même méthodologie que pour les échantillons radiomarqués. Les échantillons sont dénaturés lors d'une incubation à 95 °C pendant 5 minutes puis mis à migrer sur un gel à gradient 4-20% Mini-PROTEAN[®] TGXTM (Bio-Rad) pendant deux heures sous une tension constante de 90 V. Les protéines sont transférées sur membrane de difluorure de polyvinylidène (PVDF) fournie dans le kit Trans-Blot[®] TurboTM Mini (Bio-Rad) au moyen du système de transfert Trans-Blot[®] TurboTM (Bio-Rad) selon les recommandations du fournisseur. La membrane est placée sous agitation pendant une heure à température ambiante dans un mélange de tampon Tris salin avec Tween (*Tris-buffered saline with Tween*, TBST) et de 5 % (p/v) d'albumine de sérum bovin (ASB). La membrane est incubée en présence d'un anticorps polyclonal de souris dirigé contre la décorine humaine (R&D Systems, Lille, France, AF143, dilution 1 / 500 dans un mélange TBST / 5 % (p/v) ASB) pendant une nuit, à 4 °C et sous agitation, puis en présence d'un anticorps polyclonal dirigé contre les IgG de souris (Cell Signaling Technology, Ozyme, Saint-Cyr-l'Ecole, France, #7076, dilution 1 / 2 000 dans un mélange TBST / 5 % (p/v) ASB) pendant une heure, à température ambiante et sous agitation. La membrane est lavée trois fois pendant 5 minutes dans du tampon TBST puis est révélée par les réactifs de chimioluminescence ECL *Western Blotting Detection* (GE Healthcare). La visualisation des signaux est permise grâce à un appareil de détection Chemidoc XRS+ (Bio-Rad).

4. Détermination de l'activité enzymatique de la β 1,3-galactosyltransférase 6 à partir d'homogénats cellulaires

a. Préparation des homogénats cellulaires et dosage des protéines totales des échantillons

Après transfection des cellules H7 par le pcDNA3.1(+) codant pour la β 3GalT6 sauvage ou l'un des mutants, les milieux de culture sont retirés des puits et les tapis cellulaires sont lavés avec du PBS. Le contrôle négatif consiste toujours en une transfection avec le pcDNA3.1(+) ne codant pas pour la β 3GalT6. Les tapis cellulaires de trois puits d'une même condition de

transfection sont récupérés par grattage dans du PBS et centrifugés durant 5 min à 16 000 *g*. Les culots cellulaires sont ensuite repris dans 100 μ L de tampon (250 mM saccharose, 5 mM HEPES à pH 7,4), puis soniqués deux fois pendant 5 secondes avec 10 secondes d'intervalle sur glace au moyen d'une sonde à ultrasons. Les protéines totales des homogénats cellulaires obtenus sont dosées par la méthode de Bradford (280) en lisant l'absorbance à 595 nm sur un spectrophotomètre UV-visible UV-1800 (Shimadzu, Noisiel, France) au moyen d'une gamme étalon d'ASB. Les lysats sont utilisés pour l'analyse de l'expression protéique de la β 3GalT6 par immunoblot et les essais enzymatiques.

b. Analyse de l'expression de la β 1,3-galactosyltransférase 6 recombinante par immunoblot

Pour chaque condition de transfection, 20 μ g de protéines totales issus des homogénats cellulaires sont mélangés à un tampon de charge composé de 200 mM de Tris-HCl à pH 6,8 supplémenté de 10 mM de dithiothréitol (DTT) (Promega, Charbonnières-lès-Bains, France), de 0,05 % (*p/v*) de bleu de bromophénol (Bio-Rad) et de 10 % (*v/v*) de glycérol (VWR Chemicals, Rosny-sous-Bois, France). Les échantillons sont dénaturés à 95 °C pendant 5 minutes, puis migrés dans un gel 4-20% Mini-PROTEAN® TGX™ (Bio-Rad), transférés et révélés selon la méthodologie décrite en section I.3.b. Les protéines produites sont détectées à l'aide d'un anticorps monoclonal de lapin est dirigé contre l'étiquette HA (Sigma-Aldrich, H6909, dilution 1 / 2 000 dans du TBST) fusionnée aux protéines surexprimées. Un anticorps polyclonal de chèvre ciblant les IgG de lapin (Cell Signaling Technology, #7074, dilution 1 / 10 000 dans du TBST) est ensuite utilisé comme anticorps secondaire.

c. Évaluation de l'activité galactosyltransférase par chromatographie liquide à haute performance

Les composés chimiques galactose- β 1,4-xylose-2-*O*-phosphate- β 1-*O*-méthoxynaphtyle (Gal-Xyl(2P)-OMN) et galactose- β 1,3-galactose- β 1,4-xylose-2-*O*-phosphate- β 1-*O*-méthoxynaphtyle (Gal-Gal-Xyl(2P)-OMN) ont été synthétisés et fournis par l'équipe du Pr. Chrystel Lopin-Bon (UMR 7311 ICOA, Université d'Orléans, France). Le premier est utilisé en tant que substrat accepteur de la réaction enzymatique tandis que le second est utilisé comme standard externe pour obtenir une gamme de produit de la réaction afin de quantifier l'activité enzymatique. Les formes déphosphorylées de ces composés, le galactose- β 1,4-

xylose- β 1-*O*-méthoxynaphtyle (Gal-Xyl-OMN) et galactose- β 1,3-galactose- β 1,4-xylose-*O*-méthoxynaphtyle (Gal-Gal-Xyl-OMN) ont également été synthétisées (281).

Le mélange réactionnel consiste en un tampon composé de 50 mM de MES à pH 6,5 avec 100 mM de NaCl et 10 mM de MnCl₂ dans lequel sont ajoutés 50 μ g de protéines totales et une concentration variable en Gal-Xyl(2P)-OMN (0 à 5 mM) en présence de 10 mM d'UDP-Gal (Sigma-Aldrich), ou une concentration variable en UDP-Gal (0 à 10 mM) en présence de 5 mM de Gal-Xyl(2P)-OMN. Les composés Gal-Xyl(2P)-OMN et Gal-Gal-Xyl(2P)-OMN étant déphosphorylés au contact des homogénats cellulaires dans les conditions utilisées, nous avons également mis en place d'autres essais enzymatiques en ajoutant 1 mM d'adénosine triphosphate (ATP) aux mélanges réactionnels, dans le but de limiter la déphosphorylation.

Le volume final de la réaction est de 50 μ L. La réaction est conduite à 37 °C pendant 1 heure puis est arrêtée par ajout de 5 μ L d'HCl 2 N. Les mélanges réactionnels sont immédiatement vortexés et placés dans la glace. Ils sont ensuite centrifugés à 4 °C à 10 000 *g* pendant 5 minutes. 10 μ L de chaque surnageant sont analysés par chromatographie liquide à haute performance (CLHP).

La phase mobile est composée de 11 % (v/v) d'acétonitrile (CARLO ERBA Reagents) et 0,02 % (v/v) d'acide trifluoroacétique (Sigma-Aldrich) dans de l'eau. La colonne de séparation en phase inverse est une colonne XBridge de type C18 (145 x 5 mm, 5 μ m) (Waters, Saint-Quentin en Yvelines, France) reliée à un module de séparation e2695 (Waters) couplé à un détecteur de fluorescence e2475 (Waters) réglé sur une longueur d'onde d'excitation de 324 nm et une longueur d'onde d'émission de 356 nm. Les échantillons sont séparés sur la colonne grâce à un flux de 1 mL/min. Les données sont analysées sur les chromatogrammes obtenus grâce au logiciel Empowers (Waters).

II. Études de la fonction des glycosyltransférases recombinantes impliquées dans la synthèse de l'amorce tétrasaccharidique

1. La β 1,3-galactosyltransférase 6 recombinante

a. Obtention des plasmides codant la 6His-MBP- β 3GalT6 Δ N_{ter}29 et la 6His-MBP- β 3GalT6 Δ N_{ter}50

Les plasmides codant pour deux formes tronquées et solubles de la β 3GalT6 recombinante humaine ont été obtenus avant le début de ce travail de thèse. Il s'agit de vecteurs d'expression pETM41 (Invitrogen) dans lesquels ont été clonées les séquences

codant la β 3GalT6 tronquée des vingt-neuf acides aminés *N*-terminaux (β 3GalT6 Δ N_{ter29}) ou des cinquante acides aminés *N*-terminaux (β 3GalT6 Δ N_{ter50}). Pour ce faire, les séquences ont été amplifiées par PCR avec les amorces appropriées (Tableau VII) (Eurofins Genomics) à partir du vecteur d'expression cellulaire pcDNA3.1(+)- β 3GalT6-WT, puis ont été insérées dans le vecteur d'expression bactérienne pETM41 (Figure 24) en utilisant les sites de restriction *Nco*I (extrémité 5') et *Xho*I (extrémité 3'). Ainsi, il est possible d'exprimer ces protéines tronquées et fusionnées à la protéine de fixation au maltose (*maltose binding protein*, MBP) et à une étiquette 6His en système bactérien. Les protéines recombinantes sauvages sont nommées 6His-MBP- β 3GalT6 Δ N_{ter29}-WT et 6His-MBP- β 3GalT6 Δ N_{ter50}-WT.

| Troncation générée | Brin | Séquence |
|--|----------|---|
| β 3GalT6 Δ N _{ter29} | Sens | 5'-CGCCATGGCAGCTGCGCGGCCGAGCCC-3' |
| | Antisens | 5'-TGCTCGAGTCAGGGGATGCCCTCCCTTCTC-3' |
| β 3GalT6 Δ N _{ter50} | Sens | 5'-CGCCATGGCAGCGCCCGCGCGGCC-3' |
| | Antisens | 5'-TGCTCGAGTCAGGGGATGCCCTCCCTTCTCTGGCA-3' |

Tableau VII. Amorces nucléotidiques utilisées pour générer les troncations de la β 1,3-galactosyltransférase 6 humaine. Les troncations obtenues sont clonées à l'aide de ces amorces dans le vecteur d'expression pETM41.

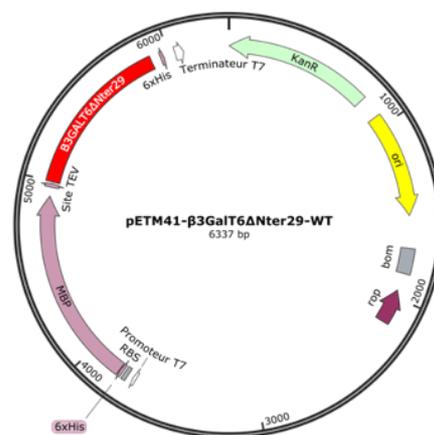


Figure 24. Carte du vecteur d'expression pETM41 dans lequel a été clonée la séquence nucléotidique de la β 1,3-galactosyltransférase 6 humaine dont les vingt-neuf premiers acides aminés ont été supprimés. Le promoteur T7 permet l'expression de la protéine d'intérêt fusionnée à la protéine de fixation au maltose (*maltose binding protein*, MBP). Les deux protéines sont séparées par un site de coupure spécifique de la protéase du virus de la gravure du tabac (*tobacco etch virus*, TEV). La protéine de fusion possède également une étiquette 6His en C_{ter}. L'enzyme sans les cinquante premiers acides aminés a été clonée de la même façon, non représentée ici. Le vecteur confère aux bactéries transformées une résistance à la kanamycine grâce au gène *KanR*.

b. Obtention des plasmides codant les mutants de la 6His-MBP-β3GalT6ΔN_{ter}29

i. Mutagenèse dirigée par PCR

Les mutations Tyr182Cys, Asp207His et Gly217Ser ont été générées à partir de la séquence codante de la 6His-MBP-β3GalT6ΔN_{ter}29 sous-clonée dans le pETM41 selon la même méthodologie que décrite précédemment (section I.2.a) avec les amorces indiquées (Tableau VI). Après digestion des brins matrices par l'enzyme *DpnI*, le produit de mutagenèse dirigée est dialysé pendant 30 minutes contre de l'eau ultrapure à travers une membrane de cellulose avec des pores de 0,025 μm (Millipore).

ii. Préparation de bactéries *Escherichia coli* électrocompétentes

Un volume de 50 mL de LB (Sigma-Aldrich) est inoculé avec 500 μL d'une préculture de bactéries *E. coli* chimiocompétentes *One Shot*TM TOP10 (Invitrogen) puis incubés à 37 °C sous agitation jusqu'à atteindre une densité optique (DO) à 600 nm de 0,6. La culture est refroidie à 4 °C pendant 10 minutes. Toutes les étapes suivantes sont réalisées à 4 °C.

La culture est centrifugée 10 minutes à 750 *g*. Le culot est resuspendu délicatement dans 50 mL d'un mélange d'eau et de 10 % (v/v) de glycérol (VWR Chemicals) puis centrifugé 10 minutes à 1 000 *g*. Cette étape est répétée deux fois, successivement avec 25 mL puis 1 mL du même mélange eau / glycérol. Enfin, le culot est resuspendu dans 200 μL du mélange eau / glycérol pour obtenir des bactéries électrocompétentes.

iii. Amplification, extraction et purification des plasmides

Toutes les étapes suivantes sont réalisées à 4 °C. 1 μL de plasmide dialysé est mélangé à 50 μL de bactéries électrocompétentes fraîchement préparées. Ce mélange est déposé dans une cuve d'électroporation préalablement conservée à - 20 °C. Le tout est placé dans l'appareil d'électroporation Eporator® (Eppendorf). L'électroporation est effectuée à 1 700 V pendant 5,7 ms. Le mélange de transformation est complété à 1 mL avec du LB (Sigma-Aldrich) préalablement conservé à 4 °C, puis les bactéries sont cultivées à 37 °C pendant 2 h sous agitation.

Les bactéries sont ensuite centrifugées à température ambiante à 2 500 *g* pendant 10 minutes puis, après élimination du surnageant, resuspendues dans 150 μL de LB. La totalité du

volume est étalée sur LB-Agar (Euromedex) supplémenté par 75 mg/L de kanamycine (Euromedex). La boîte de Pétri est placée dans un incubateur à 37 °C pendant une nuit.

Après repiquage de clones et culture de ces derniers, les plasmides sont extraits et purifiés à l'aide du kit de minipréparation Qiaprep® (Qiagen) puis dosés au Nanodrop™ One (Thermo Fisher Scientific). La présence des mutations dans la séquence codante de la β 3GalT6 Δ N_{ter}29 est vérifiée par double séquençage (Eurofins Genomics) grâce aux amorces 5'-AAGGAGTCGTCGTCGCC-3' et « T7-term » (cette dernière étant fournie par le prestataire). Les protéines recombinantes mutées sont nommées 6His-MBP- β 3GalT6 Δ N_{ter}29-Y182C, 6His-MBP- β 3GalT6 Δ N_{ter}29-D207H et 6His-MBP- β 3GalT6 Δ N_{ter}29-G217S.

c. Production et purification des enzymes sauvage et mutées recombinantes

i. Tests de surexpression des enzymes sauvages recombinantes

Les souches bactériennes chimiocompétentes *E. coli* Rosetta™ 2(DE3) (Novagen) et BL21(DE3) (Novagen) sont utilisées pour les tests d'expression des 6His-MBP- β 3GalT6 Δ N_{ter}29-WT et 6His-MBP- β 3GalT6 Δ N_{ter}50-WT.

Le premier jour, 50 μ L de bactéries sont incubées en présence de 50 ng de plasmide sur glace pendant 30 minutes. Un choc thermique est effectué disposant le tube 42 °C pendant 45 secondes, puis à 4 °C pendant 3 minutes. Le mélange de transformation est complété à 1 mL avec du LB (Sigma-Aldrich) puis est incubé à 37 °C sous agitation pendant 90 minutes. Les bactéries sont ensuite centrifugées à 1 000 *g* pendant 10 minutes à température ambiante puis resuspendues dans 200 μ L de LB (Sigma-Aldrich). La totalité est étalée sur LB-Agar (Euromedex) supplémenté par 75 mg/L de kanamycine (Euromedex). La boîte de Pétri est placée dans un incubateur à 37 °C pendant une nuit.

Le deuxième jour, une colonie bactérienne est isolée et suspendue dans 5 mL de LB contenant 75 mg/L de kanamycine. La culture bactérienne est effectuée pendant une nuit, à 37 °C et sous agitation.

Le troisième jour, 200 mL du même milieu sélectif sont inoculés avec 1 mL de la pré-culture de la veille et mis en culture selon les mêmes conditions que précédemment décrites. Lorsque la DO à 600 nm atteint 0,6 unités, la culture est placée à 4 °C pendant 1 heure sans agitation. La production de l'enzyme recombinante est alors induite par l'ajout de 1 mM d'isopropyl β -D-1-thiogalactopyranoside (IPTG) (Sigma-Aldrich) et par la poursuite de la culture à 20 °C sous agitation pendant une nuit.

Le quatrième jour, la culture est centrifugée à 5 000 *g*, à 4 °C pendant 20 minutes. Les culots bactériens sont resuspendus dans 50 mL de PBS (Sigma-Aldrich) et centrifugés de la même manière. Les culots obtenus après retrait du surnageant peuvent être conservés à -20 °C jusqu'à utilisation.

Un culot bactérien est suspendu dans 30 mL de tampon de lyse contenant 20 mM de Tris-HCl à pH 7,4, 200 mM de NaCl, 5 % (v/v) de glycérol (VWR Chemicals) et 1 mM de DTT (Promega). Les bactéries sont lysées sur glace par sonication grâce à une sonde par huit cycles de 1 minute de sonication suivie d'1 minute d'arrêt. Le lysat est centrifugé à 11 000 *g* pendant 20 minutes à 4 °C pour séparer le culot (débris cellulaires, protéines agrégées...) du surnageant (protéines solubles, matériel génomique...).

Une aliquote de chaque étape (lyse, culot et surnageant) correspondant à 1 / 2 000^e (v/v) du volume de départ est récupérée et mélangée au tampon de charge. Les échantillons sont séparés par SDS-PAGE (12 % (v/v) d'acrylamide) pour analyser l'expression et la solubilité des enzymes produites par les différentes souches bactériennes.

Le gel est coloré dans un tampon contenant 45 % (v/v) d'éthanol, 45 % (v/v) d'acide acétique (Sigma-Aldrich) et 2 g (p/v) de bleu de Coomassie R-250 (Bio-Rad). Il est ensuite décoloré par des bains successifs de tampon contenant 30 % (v/v) d'éthanol et 10 % (v/v) d'acide acétique. Les protéines séparées sont visualisées sur un appareil de détection Chemidoc XRS+ (Bio-Rad).

ii. Production et purification des enzymes recombinantes sauvages et mutées

Cette section traite des enzymes recombinantes des 6His-MBP-β3GalT6ΔN_{ter}29-WT, 6His-MBP-β3GalT6ΔN_{ter}29-Y182C, 6His-MBP-β3GalT6ΔN_{ter}29-D207H, 6His-MBP-β3GalT6ΔN_{ter}29-G217S et 6His-MBP-β3GalT6ΔN_{ter}50-WT.

Les étapes de transformation, de culture bactérienne, d'induction de l'expression, de lyse et de centrifugation sont réalisées selon la méthodologie décrite dans la section précédente.

Le surnageant contenant les protéines solubles est mis sous agitation durant 10 minutes à température ambiante après ajout de 3 μL de benzonase® (Sigma-Aldrich) et 2 mM de MgCl₂ afin de digérer l'ADN génomique. Il est ensuite passé au travers d'un filtre d'un diamètre de 0,45 μm.

La purification de la protéine est réalisée sur un système de purification à basse pression BioLogic LP (Bio-Rad), comprenant un détecteur ultra-violet couplé à un collecteur de fractions BioLogic BioFrac (Bio-Rad) et relié à un ordinateur récoltant les données grâce au logiciel LP *Data View* (Bio-Rad). La séparation est réalisée sur colonne d'affinité MBPTrap™ HP 1 mL (Cytiva) intégrée au système. La colonne est composée d'une résine de Dextrine Sepharose™ qui retiendra par affinité la protéine de fusion par la MBP.

Avant purification, le système est lavé à l'eau ultrapure puis équilibré dans le tampon de fixation (20 mM Tris-HCl pH 7,4, 200 mM NaCl, 5 % (v/v) glycérol) à un flux de 1 mL/min pendant 10 min. Le mélange protéique filtré est alors injecté dans le système avec un flux abaissé à 0,2 mL/min pour permettre la fixation de la protéine de fusion sur la colonne. Puis, la colonne est lavée avec un tampon de lavage (20 mM Tris-HCl pH 7,4, 300 mM NaCl, 5 % (v/v) glycérol) à 1 mL/min pendant 10 minutes. Enfin, le tampon d'élution (20 mM Tris-HCl pH 7,4, 200 mM NaCl, 5 % (v/v) glycérol, 10 mM maltose (Sigma-Aldrich)) est utilisé pour éluer la protéine purifiée. Enfin, la colonne est régénérée grâce à une solution composée de 500 mM de NaOH (VWR Chemicals) puis conservée dans un mélange d'eau et de 20 % (v/v) d'éthanol.

Les fractions contenant la protéine purifiée sont rassemblées puis concentrées au moyen d'un filtre à centrifuger Amicon® Ultra-15 3K (Millipore) par centrifugations successives de 15 minutes, à 4 °C et à 5 000 *g* jusqu'à obtention d'environ 1 mL de solution protéique. Celle-ci est dessalée sur une colonne PD-10 (Cytiva) dans un tampon composé de 10 mM de sodium phosphate et 100 mM de fluorure de sodium à pH 7,4. La concentration protéique de la solution finale est dosée au Nanodrop ND-1000 (Thermo Fisher Scientific) puis séparée en aliquotes de 50 µL qui seront conservées à -80 °C après une congélation rapide dans l'azote liquide.

d. Caractérisation des enzymes recombinantes purifiées

i. Dosage protéique par la méthode de Bradford

Une aliquote est prélevée à chaque étape des processus de production et de purification (lyse, surnageant, culot, surnageant traité à la benzonase®, surnageant filtré, protéines non retenues par la colonne, lavage, fractions purifiées, régénération de la colonne, conservation de la colonne). Les protéines de ces aliquotes sont dosées par la méthode de Bradford (280).

ii. Analyse des étapes de production et purification sur gel de polyacrylamide en conditions dénaturantes

Pour analyser ces étapes, 1 / 2 000^e (v/v) de chaque aliquote est séparé par SDS-PAGE selon la méthodologie proposée en section II.1.c.i. dans le but de vérifier la pureté ou la présence de contaminants et la masse moléculaire apparente de la protéine exprimée. Le gel est coloré, décoloré et visualisé comme décrit précédemment.

iii. Étude de la distribution de taille par diffusion dynamique de la lumière

Les distributions de taille de la protéine purifiée sont analysées par DLS dans le but de caractériser son état d'agrégation et sa taille estimée. Un volume de 15 µL de l'éluat de purification concentré à 20 µM est centrifugé à 16 000 *g*, à 4 °C pendant 30 minutes. L'échantillon est ensuite analysé dans un système de diffusion de lumière Zetasizer Nano ZS (Malvern Panalytical, Malvern, Worcestershire, UK) couplé au logiciel Zetasizer Nano (Malvern Panalytical).

iv. Étude de la structure secondaire par dichroïsme circulaire

Le contenu en structure secondaire des enzymes recombinantes purifiées est analysé par DC sur un spectromètre Chirascan Plus (Applied Photophysics Ltd). Des spectres dans l'UV lointain (de 180 nm à 260 nm) sont réalisés sur un volume de 30 µL de chaque protéine purifiée concentrée à 20 µM dans une cuve en quartz avec un trajet optique de 0,01 cm. Le temps d'acquisition est de 1 nm/s. Chaque spectre est la moyenne de 3 mesures desquelles le spectre du tampon seul a été soustrait.

Les spectres bruts de DC sont obtenus en absorbance, par soustraction de la polarisation circulaire gauche par la polarisation circulaire droite. Le logiciel Pro-Data (Applied Photophysics Ltd) est utilisé pour convertir les données en ellipticité molaire $[\theta]_{MRE}$, unité prenant en compte la masse moléculaire de la protéine, la longueur de la cuve ainsi que la concentration protéique utilisée.

v. Détermination des paramètres cinétiques des enzymes purifiées par chromatographie liquide à haute performance

Les paramètres cinétiques apparents de la 6His-MBP- β 3GalT6 Δ N_{ter}50-WT et de la 6His-MBP- β 3GalT6 Δ N_{ter}29-WT ou mutée sont mesurés par CLHP selon la méthodologie présentée dans la section I.4.c. 50 pmol d'enzyme purifiée sont utilisés pour les essais enzymatiques.

Les paramètres cinétiques K_M et V_{max} des enzymes recombinantes sont estimés grâce à un ajustement des données par régression non linéaire à l'équation de Michaelis-Menten en utilisant le logiciel Prism 7 (GraphPad Software, San Diego, CA, USA).

2. La β 1,3-glucuronosyltransférase 1 recombinante

a. Production et purification de l'enzyme sauvage tronquée

L'enzyme recombinante GlcAT- Δ N_{ter}76-WT a été produite, purifiée et fournie par l'équipe du Dr. Rebekka Wild (IBS UMR 5075, Université Grenoble-Alpes, France). Il s'agit de l'enzyme recombinante humaine dont les 76 premiers résidus *N*-terminaux (correspondant aux segments cytoplasmique, transmembranaire et *stem*) sont tronqués. La séquence entourée des sites de restriction *Nco*I (extrémité 5') et *Not*I (extrémité 3') a été fournie par Twist Bioscience (SF, CA, USA) et insérée dans le plasmide pETM30 (Invitrogen) pour être exprimée en fusion avec une étiquette 6His et une protéine de fusion glutathion *S*-transférase (GST) situées au niveau de l'extrémité *N*-terminale. La protéine est produite en système bactérien et purifiée par chromatographie d'affinité au moyen d'une colonne HisTrap HP™ (Cytiva). Grâce au site de clivage par la TEV situé entre la protéine d'intérêt et la GST, les deux protéines sont séparées en utilisant une TEV portant une étiquette 6His. La protéine d'intérêt est purifiée par une seconde chromatographie d'affinité retenant la protéase et la GST fusionnées à une étiquette 6His.

b. Détermination de l'activité enzymatique par bioluminescence

Les composés chimiques Gal-Xyl-OMN et Gal-Gal-Xyl-OMN (281) sont utilisés comme substrats accepteurs de la réaction de la GlcAT- Δ N_{ter}76-WT.

Le mélange réactionnel est effectué dans un tampon composé de 100 mM d'acétate de sodium à pH 6,5, 10 mM de MnCl₂ dans lequel sont ajoutés 100 pmol d'enzyme purifiée et une concentration variable de substrat accepteur (0 à 1 mM) en présence de 5 mM d'UDP-GlcA (Sigma-Aldrich). La réaction se fait à 37 °C pendant 1 heure et est arrêtée par une

incubation sur glace. La libération d'UDP à la suite de la réaction catalysée par la GlcAT-I- Δ Nter76-WT est évaluée grâce au kit de mesure d'activité GT UDP-Glo™ (Promega) selon les recommandations du fournisseur. L'UDP est converti en ATP, substrat de la réaction d'une luciférase pour générer de la bioluminescence mesurée à l'aide d'un luminomètre GloMax®-Multi+ (Promega).

III. Simulations de dynamique moléculaire de la β 1,3-galactosyltransférase 6 sauvage et mutée

La structure tridimensionnelle de la β 3GalT6 n'étant pas résolue à ce jour, nous avons établi une collaboration avec les Drs. Yasaman Karami et Hamed Khakzad (UMR 7503 LORIA, Vandoeuvre-lès-Nancy, France) afin d'étudier la dynamique moléculaire de l'enzyme sauvage et des mutants Tyr182Cys, Asp207His et Gly217Ser selon une approche préalablement décrite (282). Le modèle bioinformatique utilisé est généré par le logiciel d'intelligence artificielle AlphaFold (Figure 18).

La β 3GalT6 est une enzyme localisée dans l'appareil de Golgi. Le segment TM de la prédiction de structure de l'enzyme sauvage ou mutée est donc inséré dans un modèle de membrane biologique de l'appareil de Golgi constitué de 17 % de CHOL, 49 % de DOPC, 16 % de DOPE, 5 % de DOPS, 6 % de PSM, 6 % de POPI et 1 % de PLPA. Ces abréviations sont le système de dénomination des lipides proposés par le logiciel CHARMM-GUI⁴. Le système est hydraté par un modèle d'eau TIP3P (283) avec une concentration saline de 150 mM de NaCl. Pour chaque système, trois réplicats de simulations dynamiques de 1 μ s sont obtenus.

Le logiciel de simulation de dynamique moléculaire utilisé est GROMACS 2021.3. Toutes les expériences sont conduites à une température de 37 °C et une pression de 1 atm. Les liaisons hydrogène et les ponts salins intramoléculaires sont enregistrés selon la fréquence à laquelle ils sont observés dans des conditions définies, c'est-à-dire à une distance de 3,9 Å et 3,5 Å entre le centre de masse de l'atome donneur et celui de l'atome accepteur, respectivement. Les valeurs de fluctuations (*root-mean square fluctuations*, RMSF) des résidus sont obtenues en analysant la position dans l'espace des atomes du squelette peptidique des systèmes.

⁴ <http://www.charmm-gui.org/?doc=1?4input/membrane>

IV. Caractérisation d'un modèle de synthèse matricielle murin invalidé pour le gène

B3galt6

1. Culture cellulaire des cellules ATDC5

a. Matériel

Les cellules de la lignée cellulaire ATDC5 de tératocarcinome de souris (Riken Cell Bank, Tsukubai, Japon) ont la particularité d'être chondrogéniques, c'est-à-dire capables de réaliser un processus de synthèse cartilagineuse (284, 285). Lors de la chondrogenèse, elles produisent une MEC constituée, entre autres, de collagènes et de PG (286) qui s'organise sous la forme d'un pseudo-tissu. Les cellules sont fournies au passage 13 et sont cultivées à 37 °C dans une atmosphère humide contenant 5 % (v/v) de CO₂. Différents milieux de culture sont utilisés selon les applications :

- Le **milieu complet** est constitué de *Dulbecco's Modified Eagle Medium Nutrient Mixture F-12* (DMEM / F-12) (Gibco, Thermo Fisher Scientific), auquel est ajouté 5 % (v/v) de SVF, 1 % (v/v) de pénicilline-streptomycine (Gibco), 1 % (v/v) de glutamine, 10 µg/mL de transferrine humaine (Sigma-Aldrich) et 3.10⁻⁸ M de sélénite de sodium (Sigma-Aldrich).
- Le gène *B3galt6* est ciblé par la méthode CRISPR-Cas9 *via* une transduction adénovirale. Le **milieu d'infection** consiste en du milieu complet auquel le SVF a été substitué par la quantité adéquate d'adénovirus à utiliser pour infecter les tapis cellulaires.
- L'induction de la synthèse matricielle est réalisée en cultivant les cellules dans le **milieu d'induction**, constitué de milieu complet supplémenté par 10 µg/mL d'insuline (Sigma-Aldrich).
- Le **milieu d'induction sans SVF** est utilisé pour les analyses protéomiques afin d'éliminer les protéines issues du sérum.

Le changement du milieu de culture est effectué tous les 2 à 3 jours. Les lavages des cellules sont réalisés avec du PBS (Sigma-Aldrich) et les passages et amplifications sont effectués en décollant les tapis cellulaires à l'aide d'un mélange trypsine-EDTA (Sigma-Aldrich).

b. Stratégie d'inactivation du gène *B3galt6* par transduction virale

Deux vecteurs adénoviraux (VectorBuilder, Neu-Isenburg, Allemagne) sont utilisés pour la transduction dans les cellules ATDC5 :

- Le premier, désigné « *ΔB3galt6* », permet la transcription d'un ARN *single guide* (ARNsg) de séquence 5' - GGTTCCTCGTAGGCGTCGCGC - 3' dont la fonction est de s'hybrider au gène *B3galt6* et de recruter l'endonucléase hCas9 co-exprimée par ce même vecteur afin de couper l'ADN. Il permet également la co-expression d'une protéine fluorescente verte (*green fluorescent protein*, GFP).
- Le second, désigné « *contrôle* », permet l'expression de la GFP seule. Il s'agit donc d'un contrôle négatif de l'infection, pour lequel il n'y a pas d'édition du génome des cellules.

c. Infection adénovirale

Les cellules ATDC5 sont ensemencées dans le milieu complet à 6.10^4 cellules/cm² en plaques 12-puits, ou sur lamelles de verre ou de fluorine en plaques 24-puits. Lorsque les cellules sont à confluence, elles sont infectées avec le virus *ΔB3galt6* ou *contrôle* dans le milieu d'infection. La multiplicité d'infection étant de 240, la quantité de virus recommandée par le fournisseur est de 3 millions d'unités virales par puits. Les milieux d'infection sont retirés le lendemain pour être remplacés par du milieu complet. Pendant deux jours, les cellules sont rincées deux fois au PBS afin de s'affranchir des virus dans le milieu.

d. Induction de la synthèse matricielle et formation des pseudo-tissus

La synthèse matricielle débute lorsque le milieu complet est remplacé par le milieu d'induction qui est ensuite renouvelé tous les deux jours pendant deux semaines. Les tapis cellulaires forment alors des pseudo-tissus visibles à l'œil nu. Selon leur exploitation, les pseudo-tissus subissent différents traitements :

- Les pseudo-tissus cultivés en plaques 12-puits sont destinés à des analyses biochimiques. Étant donné que l'expression protéique de PG est analysée par immunoblot, ils sont cultivés deux jours de plus en milieu d'induction sans SVF, puis récupérés. Un test d'extensibilité leur est également appliqué.
- Les pseudo-tissus cultivés en plaques 24-puits sur lamelles de verre sont destinés à des analyses morphologiques par coloration de coupes histologiques.

- Les pseudo-tissus cultivés en plaques 24-puits sur lamelles de fluorine sont destinés à des analyses morphologiques et biomécaniques par AFM.

2. Analyses biochimiques de pseudo-tissus

a. Récupération des milieux et précipitation du sécrétome

Pour les analyses biochimiques, les milieux sans SVF (constituant la **fraction du sécrétome**) des cellules cultivées en plaques 12-puits sont récupérés puis centrifugés pendant 5 minutes à 1 000 *g* et à 4 °C. Les surnageants sont incubés avec 10 µL de désoxycholate de sodium 2 % (*v/v*), puis vortexés, centrifugés rapidement et incubés sur glace pendant 10 minutes. 200 µL d'acide trichloroacétique 100 % (*p/v*) sont alors ajoutés à chaque surnageant, qui sont ensuite vortexés et incubés à 4 °C pendant une nuit.

Le lendemain, les échantillons sont centrifugés à 15 000 *g* et à 4 °C pendant 30 minutes. Le surnageant est éliminé et les culots sont rincés deux fois dans de l'acétone conservé à – 20 °C, après deux centrifugations à 15 000 *g* et à 4 °C pendant 10 minutes. Les culots sont ensuite séchés pendant 2 minutes à température ambiante. Les culots sont directement resuspendus dans un volume de tampon de charge déterminé par le dosage par la méthode de Bradford (280) avant l'analyse par immunoblot.

b. Fractionnement protéique

Sauf mention particulière, toutes les centrifugations sont réalisées à 15 000 *g* et à 4 °C.

Les pseudo-tissus cultivés en plaques 12-puits sont récupérés dans 1 mL de PBS, Triton X-100 (Sigma-Aldrich) 1 % (*v/v*) puis sont incubés à 4 °C sous rotation pendant une heure et centrifugés pendant 15 minutes. Le surnageant constitue la fraction contenant les extraits cellulaires qui est dosée par la méthode de Bradford (280) dans le but de normaliser tous les dépôts effectués pour l'analyse en immunoblot.

Les culots sont lavés avec 1 mL de PBS, Triton X-100 1 % (*v/v*), vortexés puis centrifugés 5 minutes. Les surnageants sont éliminés. 500 µL de guanidine (6 M), Tris (50 mM), DTT (50 mM) à pH 8 sont ajoutés aux culots qui sont vortexés et incubés à 4 °C sous rotation pendant 3 jours. Après ce temps d'incubation, les échantillons sont centrifugés pendant 15 minutes. Le surnageant représente la fraction contenant les protéines de la MEC synthétisée par les

cellules ATDC5 ne formant pas de liaisons croisées (appelée **fraction guanidine**), tandis que les culots sont éliminés. A ces surnageants est ajouté 1 mL d'éthanol 100 % préalablement conservé à - 20 °C, ces derniers sont ensuite vortexés et incubés à - 20 °C pendant 3 jours.

Les échantillons sont centrifugés pendant 20 minutes et les surnageants sont éliminés. Les culots sont lavés avec 1 mL d'éthanol 90 % (v/v) conservé à - 20 °C, vortexés puis centrifugés pendant 5 minutes. Ce lavage est répété avec de l'éthanol 100 % conservé à - 20 °C. Les culots sont séchés dans un concentrateur à vide SpeedVac™ (Eppendorf) pendant 10 minutes, puis resuspendus dans un volume de tampon de charge déterminé en fonction des valeurs obtenues après dosage protéique des extraits cellulaires. Les échantillons sont soniqués au bain à sonication pendant 2 minutes puis incubés à - 80 °C pendant 1 heure. Ces étapes de sonication puis d'incubation sont répétées 3 fois pour obtenir les échantillons à déposer sur gel.

c. Analyse de l'expression protéique par immunoblot

Les protéines des échantillons obtenus lors de la précipitation du sécrétome et du fractionnement cellulaire sont séparées par SDS-PAGE sur gels 4-20 % Mini-PROTEAN® TGX Stain-Free™ (Bio-Rad). Les protéines séparées sont visualisées directement dans le gel par la technologie d'imagerie *Stain-Free*™ sur un appareil de détection Chemidoc XRS+ (Bio-Rad) pour comparer le profil de migration global et s'assurer que les quantités de protéines totales sont similaires d'un dépôt à l'autre. Les protéines sont ensuite transférées sur membrane de PVDF fournie dans le kit de transfert Trans-Blot® Turbo™ Mini (Bio-Rad) au moyen du système de transfert Trans-Blot® Turbo™ (Bio-Rad). La membrane est saturée dans un mélange de TBST et de 5 % (p/v) d'ASB sous agitation, à température ambiante, pendant 1 heure.

La décorine exprimée est révélée par un anticorps polyclonal de chèvre (R&D Systems, AF1060, dilution 1 / 400 dans un mélange TBST / 5 % (p/v) ASB) pendant une nuit, à 4 °C et sous agitation, puis par un anticorps polyclonal de lapin dirigé contre les IgG de chèvre (R&D Systems, HAF017, dilution 1 / 2 000 dans un mélange TBST / 5 % (p/v) ASB) pendant 1 heure, à température ambiante et sous agitation. Le Col XII est révélé en utilisant un anticorps polyclonal de lapin (Santa Cruz, Heidelberg, Allemagne, dilution 1 / 500 dans un mélange TBST / 5 % (p/v) ASB) et un anticorps polyclonal de chèvre ciblant les IgG de lapin (Cell Signaling Technology, #7074, dilution 1 / 10 000 dans du TBST).

La membrane est lavée trois fois pendant 5 minutes dans du tampon TBST puis est révélée par les réactifs de chimioluminescence *ECL Western Blotting Detection* (GE Healthcare). La visualisation des signaux est permise grâce à un appareil de détection Chemidoc XRS+ (Bio-Rad).

3. Analyses biomécaniques et morphologiques des pseudo-tissus

a. Coloration histologique

Les pseudo-tissus invalidés et cultivés sur lamelles de verre en plaques 24-puits sont lavés au PBS et fixés dans une solution de *paraformaldéhyde* (Sigma-Aldrich) 4 % (v/v) pendant 20 minutes sous agitation à température ambiante. Ils sont délicatement décollés et inclus dans une préparation HistoGel (Microm Microtech France, Brignais) avant leur intégration dans un bloc de paraffine. La suite des manipulations a été réalisée sur la plateforme d'Histologie, Immunomarquage, Microdissection laser (HistIM, Institut Cochin, INSERM U1016, CNRS UMR8104, Université de Paris, Paris, France). Des coupes transversales de 4 μm d'épaisseur de chaque bloc sont réalisées sur un microtome Histocore Autocut R (Leica Biosystems, Nanterre, France). Les coupes sont ensuite colorées avec le bleu Alcian (BIO-OPTICA, Microm Microtech) et le *Nuclear fast red* (BIO-OPTICA) sur un appareil de coloration Histocore Spectra (Leica) selon les recommandations du fournisseur. Le bleu Alcian colore les chaînes de GAG sulfatées en bleu clair tandis que le *Nuclear fast red* colore les noyaux des cellules ATDC5 en violet foncé. Les images sont numérisées sur un scanner de lames Lamina Multilabel (Akoya Biosciences, Perkin Elmer).

b. Évaluation de l'extensibilité

Pour mesurer leur extensibilité, les pseudo-tissus sont lavés au PBS puis délicatement décollés. Ils sont ensuite maintenus 30 s dans l'air à l'aide d'une pince. La longueur du pseudo-tissu est mesurée avec une règle graduée. Pour chaque condition, la mesure a été réalisée sur 3 pseudo-tissus provenant de deux cultures différentes.

c. Caractérisation biomécanique et morphologique des pseudo-tissus par microscopie à force atomique

Les pseudo-tissus cultivés sur lamelle de fluorine puis fixés avec du *paraformaldéhyde* selon la même méthodologie que décrite précédemment sont utilisés. Ces expériences sont

réalisées en collaboration avec le Dr. Grégory Francius (LCPME, UMR 7564 CNRS-UL, Villers-lès-Nancy, France).

Les analyses biomécaniques ont été réalisées sur un microscope à force atomique MFP3D-BIO (Asylum Research Technology, Oxford Instruments Company, Wiesbaden, Allemagne) équipé d'un microlevier de nitrure de silicium Si_3N_4 (MLCT, Bruker France SAS) par la technique de nanoindentation. Les données ont été obtenues avec une constante de raideur d'environ $0,01 \text{ nN.nm}^{-1}$ à un taux d'indentation de $1 \mu\text{m.s}^{-1}$. Chaque mesure est réalisée en triplicat en enregistrant au moins trois courbes *peak-force* à différents endroits du pseudo-tissu sur une surface de $90 \mu\text{m} \times 90 \mu\text{m}$. Les images contiennent au moins six à huit cellules et consistent en une grille de 50×50 courbes de forces. L'élasticité de l'échantillon, ou module de Young E , est évaluée en analysant les courbes de force à travers le modèle de Sneddon qui assimile la pointe à un cône. Les données de chaque image sont analysées en utilisant un algorithme automatique sur MATLAB (The MathWorks Inc., Natick, MA, USA) (287).

Les analyses morphologiques sont réalisées sur un microscope à force atomique Bioscope Resolve (Bruker Nano Surface, Bruker France SAS, Palaiseau, France). La topographie de surface des échantillons est obtenue par contact intermittent (*tapping*TM, PeakForce-HIRS-SSB, Bruker France SAS) entre le microlevier de nitrure de silicium (Si_3N_4) et l'échantillon avec une constante de raideur d'environ $0,12 \text{ nN.nm}^{-1}$. Les images obtenues ont une résolution de 512×512 pixels avec une fréquence de résonance de $0,5 \text{ Hz}$.

Résultats expérimentaux

I. Étude des conséquences fonctionnelles de mutations pathogènes de la β 1,3-galactosyltransférase 6 humaine

1. Utilisation d'un modèle cellulaire eucaryote invalidé pour le gène *B3GALT6*

Au cours de ma première année de thèse, nous souhaitons comparer les effets de différentes mutations de la β 3GalT6 sur son activité enzymatique et sur la synthèse des chaînes de GAG. Pour ce faire, nous avons surexprimé transitoirement les enzymes sauvage et mutées dans des cellules HeLa invalidées pour le gène *B3GALT6* grâce au système CRISPR-Cas9. Cette lignée cellulaire, nommée H7, a été générée par notre équipe au cours d'une thèse précédente et porte la mutation Val61Alafs*199 sur le gène *B3GALT6*, ce qui conduit à la fois aux pertes de l'expression génique et de l'activité galactosyltransférase (279). Ce modèle est intéressant car il nous permet de nous affranchir de l'activité β 3GalT6 endogène.

a. Analyse de l'expression de la β 1,3-galactosyltransférase 6 recombinante sauvage ou mutée

Les β 3GalT6 sauvage et mutées sont fusionnées à une étiquette HA en C_{ter} permettant d'analyser leur surexpression dans les cellules H7 par immunoblot. Nous les nommerons β 3GalT6-WT, β 3GalT6-Y182C, β 3GalT6-D207H et β 3GalT6-G217S, les mutations étudiées étant celles rencontrées sous forme homozygote chez des patients souffrant de SEDsp (17, 50). Un contrôle négatif est effectué en transfectant transitoirement ces mêmes cellules avec le plasmide vide. Par immunoblot, nous observons trois bandes à des masses moléculaires comprises entre 30-37 kDa, comme attendu (Figure 25). Ces bandes correspondent à différentes glycoformes de la β 3GalT6 qui disparaissent après traitement à la PNGase F (données non montrées) car c'est une protéine modifiée par une *N*-glycosylation. Les cellules H7 transfectées expriment donc la β 3GalT6 sauvage et les trois mutants dans les conditions utilisées, même si on note que le taux d'expression de la β 3GalT6-WT est moins important que celui de ses homologues mutés.

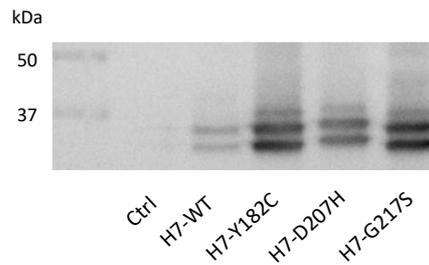


Figure 25. Expression de la β 1,3-galactosyltransférase 6 sauvage ou mutée dans les cellules HeLa invalidées pour le gène *B3GALT6*. Les cellules H7 ont été transfectées transitoirement avec un vecteur d'expression pcDNA3.1(+) qui ne contient pas la séquence codante du gène *B3GALT6* (Ctrl), qui code pour l'enzyme sauvage (H7-WT) ou qui code pour l'enzyme porteuse d'une mutation Tyr182Cys (H7-Y182C), Asp207His (H7-D207H) ou Gly217Ser (H7-G217S). Les chiffres à gauche indiquent la masse moléculaire en kDa du marqueur de taille. Un anticorps dirigé contre l'étiquette hémagglutinine (HA) a été utilisé pour cet immunoblot.

b. Mesure du taux de biosynthèse des chaînes de glycosaminoglycanes après transfection dans les cellules invalidées pour le gène *B3GALT6*

Le premier objectif était d'analyser la capacité des cellules H7 transfectées avec les enzymes recombinantes à synthétiser les chaînes de GAG *in cellulo*. Deux méthodes ont été utilisées pour mettre en évidence cette synthèse.

i. Synthèse à partir d'un xyloside exogène

Le 4-MOX est un xyloside servant de substrat à l'enzyme précédant la β 3GalT6, la β 4GalT7, et permet d'initier la synthèse de chaînes de GAG sans protéine-*core* (288). Après la transfection transitoire des plasmides d'expression codant la β 3GalT6 sauvage et mutée, le xyloside est ajouté au milieu de culture à différentes concentrations : 0, 2 ou 5 μ M en présence de soufre radiomarqué. La quantité de chaînes de GAG synthétisée, évaluée par détermination du taux d'incorporation du soufre, augmente en fonction de la concentration de 4-MOX ajoutée dans le milieu (Figure 26).

Dans toutes les conditions et en absence de 4-MOX, les cellules H7 possèdent la capacité de synthétiser les chaînes de GAG de manière résiduelle (environ 0,04 dpm/ng d'ADN). En absence d'expression de β 3GalT6, le taux d'incorporation du soufre augmente de deux fois avec 2 μ M de 4-MOX (0,08 dpm/ng) et quatre fois avec 5 μ M de 4-MOX (0,18 dpm/ng). L'expression de la β 3GalT6 permet une augmentation significative ($p < 0,001$) de la synthèse par rapport à la condition contrôle, de quatre fois avec 2 μ M de 4-MOX (0,18 dpm/ng) et de

huit fois avec 5 μM de 4-MOX (0,36 dpm/ng). L'expression des trois mutants s'accompagne d'une synthèse comparable à la condition contrôle. Les deux différences significatives ($p < 0,05$) que nous observons sont des diminutions de synthèse avec l'expression de la $\beta 3\text{GalT6-G217S}$ en absence de 4-MOX (0,026 dpm/ng) et en présence de 5 μM de 4-MOX (0,123 dpm/ng).

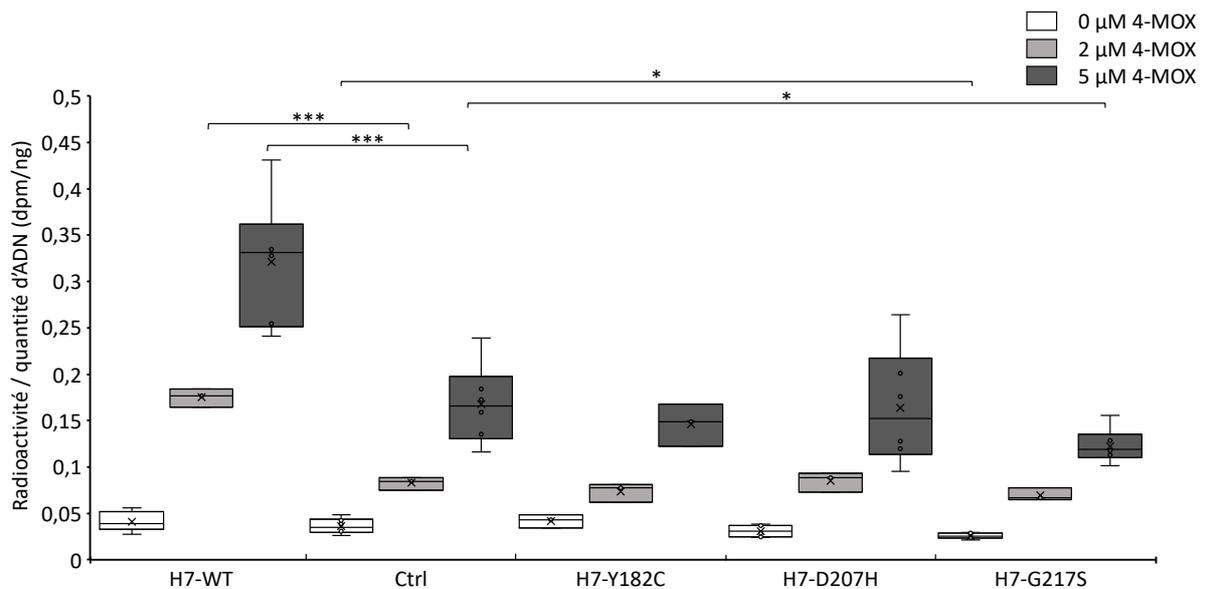


Figure 26. Radiomarquage des chaînes de glycosaminoglycanes synthétisées dans les cellules HeLa invalidées pour le gène *B3GALT6* à partir d'un substrat exogène après surexpression de la $\beta 1,3$ -galactosyltransférase 6 sauvage ou mutée. Les cellules H7 ont été transfectées transitoirement avec un vecteur d'expression pcDNA3.1(+) qui ne contient pas la séquence codante du gène *B3GALT6* (Ctrl), qui code pour l'enzyme sauvage (H7-WT) ou qui code pour l'enzyme porteuse d'une mutation Tyr182Cys (H7-Y182C), Asp207His (H7-D207H) ou Gly217Ser (H7-G217S). Le substrat exogène 4-méthylombélliferyl- β -D-xylopyranoside (4-MOX) est ajouté à différentes concentrations dans le milieu de culture. Le taux de biosynthèse est évalué par comptage de la radioactivité du S^{35} incorporé aux chaînes de glycosaminoglycanes (GAG) rapportée à la quantité d'ADN. Les valeurs sont des moyennes \pm écart-types de deux manipulations distinctes, chacune exécutée en triplicat. Les résultats des tests statistiques sont obtenus en comparant la valeur H7-WT, H7-Y182C, H7-D207H ou H7-G217S à la valeur correspondante Ctrl. * : $p < 0,05$. *** : $p < 0,001$ (Test *t* de Student). Adapté de Diana *et al.*, in external peer-review.

Ces résultats indiquent que l'expression de la $\beta 3\text{GalT6-WT}$ dans notre modèle cellulaire induit une augmentation de la synthèse des chaînes de GAG à partir de 4-MOX, de façon dose-dépendante. Les cellules H7 invalidées conservent une capacité de synthèse de chaînes de

GAG résiduelle comparable à celle mesurées lorsque les trois mutants sont exprimés, suggérant une perte de fonction des enzymes β 3GalT6-Y182C, β 3GalT6-D207H et β 3GalT6-G217S.

ii. Synthèse à partir de la protéine-core de la décorine

En complément du 4-MOX, nous avons étudié la glycosylation de la décorine, un PG extracellulaire souvent utilisé pour suivre la synthèse des chaînes de GAG. Nous avons co-transfecté transitoirement les cellules H7 pour qu'elles expriment à la fois la décorine et la β 3GalT6 sauvage ou mutée. Le contrôle négatif consiste en une transfection permettant l'expression de la décorine seule. La glycosylation de la décorine est analysée par immunoblot à partir des milieux de culture. Nous pouvons voir une bande correspondant à la protéine-core de la décorine à 42 kDa et la décorine glycosylée formant une bande diffuse de 75 à 125 kDa environ. Avec l'expression de la β 3GalT6-WT, la décorine est majoritairement glycosylée mais une fraction de protéine-core subsiste. De façon intéressante, en absence de β 3GalT6 ou en présence de β 3GalT6-D207H ou β 3GalT6-G217S, la décorine est faiblement glycosylée (cinq fois moins qu'avec la β 3GalT6-WT). Avec l'expression de la β 3GalT6-Y182C, la glycosylation est deux fois moins importante qu'avec la β 3GalT6-WT. Nous relevons également que la décorine semble être moins exprimée en cas de co-transfection avec les mutants (Figure 27A).

De façon complémentaire, l'incorporation de sulfate radiomarqué dans les chaînes de GAG de décorine est visualisée par autoradiographie (Figure 27B). Nous pouvons voir que les chaînes de GAG synthétisées par la β 3GalT6-WT ont une masse moléculaire de 75 kDa à 100 kDa environ. Nous constatons une augmentation de la masse moléculaire de ces chaînes (100 kDa à 150 kDa environ) synthétisées dans les cellules contrôles ou exprimant la β 3GalT6-Y182C, la β 3GalT6-D207H ou la β 3GalT6-G217S. La faible intensité de signal dans la piste correspondant à la β 3GalT6-G217S est due à un dépôt moins important de chaînes radiomarquées car les cellules transfectées avec ce plasmide étaient moins nombreuses.

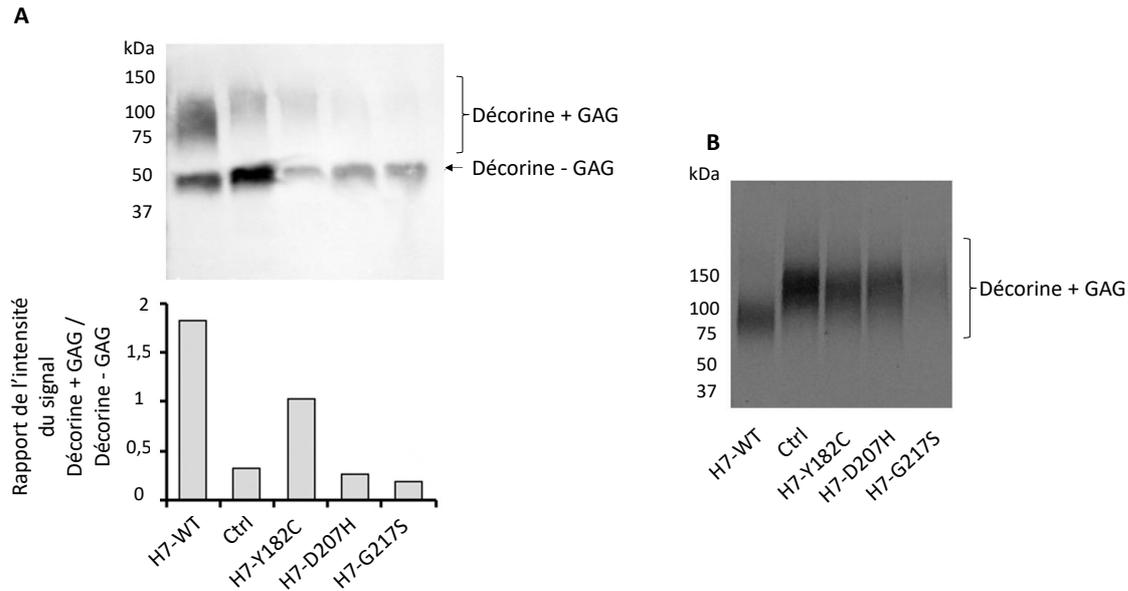


Figure 27. Analyse du taux de biosynthèse des chaînes de glycosaminoglycanes de la décorine exprimée dans les cellules HeLa invalidées pour le gène *B3GALT6* après expression de la β 1,3-galactosyltransférase 6 sauvage ou mutée. Les cellules H7 ont été co-transfectées avec deux vecteurs d'expression pcDNA3.1(+). Le premier code pour la décorine humaine fusionnée à une étiquette 6His. Le second ne code pour aucune enzyme (Ctrl), code pour l'enzyme sauvage (H7-WT) ou code pour l'enzyme porteuse d'une mutation Tyr182Cys (H7-Y182C), Asp207His (H7-D207H) ou Gly217Ser (H7-G217S). La synthèse des chaînes de glycosaminoglycanes (GAG) est analysée par (A) immunoblot ou par (B) radiomarquage au S^{35} . Dans le premier cas, le dépôt dans chaque piste correspond à un même volume de décorine purifiée par billes de nickel. Un anticorps dirigé contre l'étiquette 6His a été utilisé pour cet immunoblot. Une analyse de l'intensité des bandes par densitométrie est présentée sous la forme d'un graphique. Dans le second cas, le dépôt dans chaque piste correspond à une quantité égale de radioactivité à l'exception de la piste 5 pour laquelle une quantité insuffisante a été récupérée. Les chiffres indiquent la masse moléculaire en kDa du marqueur de taille. Adapté de Diana *et al.*, *in external peer-review*.

iii. Conclusion

De façon remarquable, nous avons observé que les cellules H7 invalidées conservent une capacité de synthèse résiduelle de chaînes de GAG pouvant être stimulée de façon dose-dépendante à la quantité de substrat exogène ajoutée au milieu. L'expression de la β 3GalT6-WT augmente le taux de synthèse des chaînes de GAG, tandis que les β 3GalT6-Y182C, β 3GalT6-D207H et β 3GalT6-G217S ne le permettent pas. Les chaînes de GAG synthétisées en absence d'enzyme sauvage ou en présence d'enzyme mutée sont donc moins nombreuses, mais également plus longues.

c. Mesure de l'activité enzymatique de la β 1,3-galactosyltransférase 6 recombinante humaine

Nous nous sommes ensuite intéressés à l'activité enzymatique de chacune des enzymes surexprimées. Pour ce faire, les cellules H7 sont transfectées avec un vecteur d'expression contenant la séquence codante de la β 3GalT6-WT, la β 3GalT6-Y182C, la β 3GalT6-D207H ou la β 3GalT6-G217S. Un contrôle négatif est effectué en transfectant les cellules H7 avec un plasmide vide, et un contrôle positif consiste en l'utilisation d'une lignée de cellules HeLa non invalidée pour le gène *B3GALT6*. Après lyse, les homogénats cellulaires sont incubés avec un substrat disaccharidique de synthèse phosphorylé, le Gal-Xyl(2P)-OMN. Ce composé mime le substrat endogène de l'enzyme, le Gal-Xyl(2P)-O-Ser (70). Les milieux d'incubation sont ensuite séparés par CLHP en phase inverse, et le produit de la réaction est détecté par fluorescence. La quantification de sa formation et l'évaluation de l'activité enzymatique sont effectuées en mesurant l'aire sous la courbe du pic correspondant.

Nous disposons au laboratoire de plusieurs composés de synthèse utiles à l'interprétation des résultats de ces essais enzymatiques. Il s'agit du Gal-Xyl(2P)-OMN (substrat de la réaction de la β 3GalT6 ajouté dans les essais enzymatiques), Gal-Gal-Xyl(2P)-OMN (produit de la réaction de la β 3GalT6 que nous utilisons pour établir une gamme étalon) et de leurs équivalents non phosphorylés, le Gal-Xyl-OMN et le Gal-Gal-Xyl-OMN. Nous avons tout d'abord établi les profils chromatographiques de chacun de ces composés afin d'en déterminer le temps de rétention et d'analyser leur degré de pureté (Figure 28).

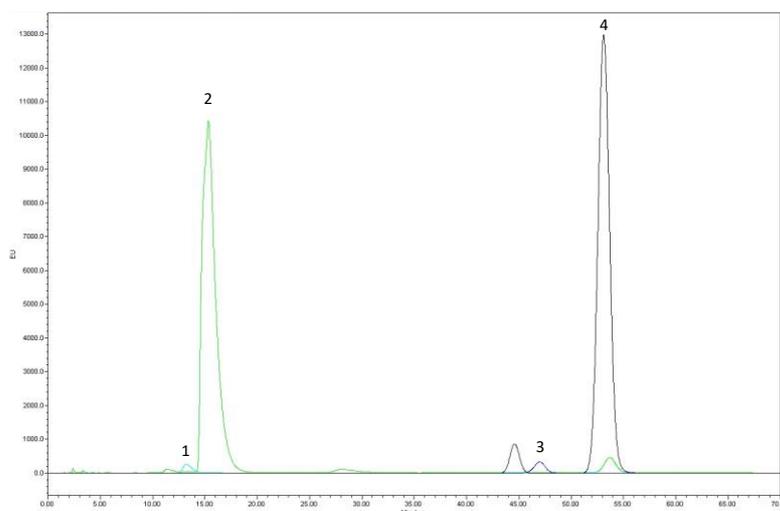


Figure 28. Profils chromatographiques des composés de synthèse utilisés pour la mesure de l'activité enzymatique de la β 1,3-galactosyltransférase 6. Les pics des composés sont désignés par des nombres (1 : Gal-

Gal-Xyl(2P)-OMN, bleu ciel ; 2 : Gal-Xyl(2P)-OMN, vert ; 3 : Gal-Gal-Xyl-OMN, bleu foncé ; 4 : Gal-Xyl-OMN, noir). Leur fluorescence est mesurée en unités d'émission (EU). Dans les conditions de cette expérience (phase mobile composée de 13 % d'acétonitrile et 0,02 % d'acide trifluoroacétique), les temps de rétention sont de 13,5 min (1), 16 min (2), 47 min (3) et 53 min (4). Les composés 2 et 4 ont été utilisés à une concentration finale de 2 mM et les composés 1 et 3 ont été utilisés à une concentration finale de 0,5 mM.

Nous souhaitons utiliser des cellules HeLa comme contrôle positif de la réaction. Malgré une optimisation des conditions de mesure, notamment en modifiant la concentration en substrat ou le temps d'incubation, l'activité spécifique est difficilement quantifiable. L'augmentation de la hauteur du pic correspondant au temps de rétention du Gal-Gal-Xyl(2P)-OMN (Figure 29, pic 1, courbe noire) indique une formation du produit par la β 3GalT6 endogène, mais elle est sujette à une grande variabilité dans nos conditions d'essai. En utilisant les cellules H7 transfectées avec la β 3GalT6-WT, nous observons également cette variabilité. De plus, un pic parasite du composé Gal-Xyl(2P)-OMN ne permet pas une mesure précise de l'aire sous la courbe car il présente le même temps de rétention que le Gal-Gal-Xyl(2P)-OMN (Figure 29, pic 1, courbe bleu foncé).

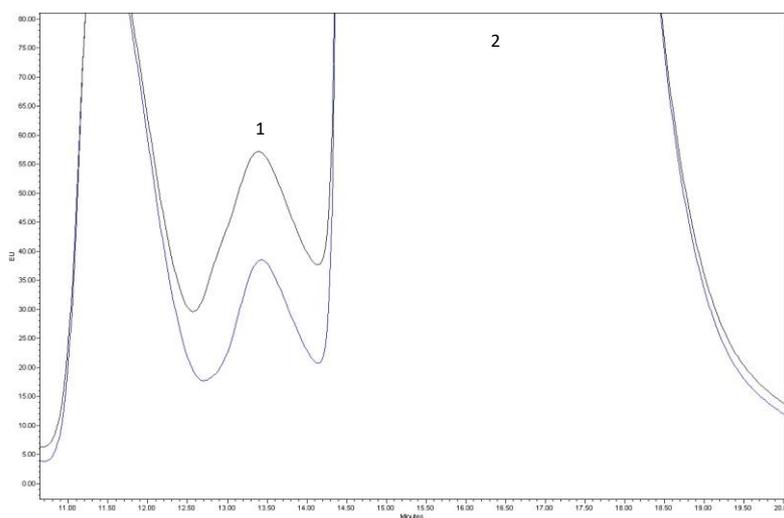


Figure 29. Exemple de chromatogramme de l'activité de la β 1,3-galactosyltransférase 6 endogène. La fluorescence est mesurée en unités d'émission (EU). Les deux pics numérotés sont le Gal-Gal-Xyl(2P)-OMN (1, produit de la réaction) et le Gal-Xyl(2P)-OMN (2, substrat de la réaction). Les deux courbes superposées représentent les essais enzymatiques en présence (noir) ou en absence (bleu foncé, contrôle négatif) d'UDP-Gal. On observe une augmentation de l'aire sous la courbe du pic 1 en présence d'UDP-Gal (11 070 880 μ V*sec, unité obtenue grâce au calcul d'intégration exécuté par le logiciel Empowers 3) par rapport à celle en absence d'UDP-Gal (7 760 531 μ V*sec).

Afin d'explorer les raisons derrière ces activités enzymatiques faibles et variables, nous avons entrepris des optimisations comprenant des modifications du pH du tampon de la réaction. Le chromatogramme du Gal-Xyl(2P)-OMN présente deux pics : celui du composé lui-même à un temps de rétention de 16 minutes, et un pic correspondant à sa forme non phosphorylée à un temps de rétention de 53 minutes. Plus le pH est élevé, plus l'aire sous la courbe du pic à 53 minutes augmente (Figure 30A). Le Gal-Xyl(2P)-OMN est déphosphorylé atteignant un maximum d'environ 45 % de déphosphorylation à pH 7,5. Les mêmes investigations ont été entreprises en utilisant le Gal-Gal-Xyl(2P)-OMN. Nous avons observé l'augmentation de l'aire sous la courbe d'un pic à un temps de rétention de 47 minutes, correspondant au Gal-Gal-Xyl-OMN (Figure 30B). Le Gal-Gal-Xyl(2P)-OMN est largement déphosphorylé à pH 6,5, atteignant quasiment 100 % de déphosphorylation.

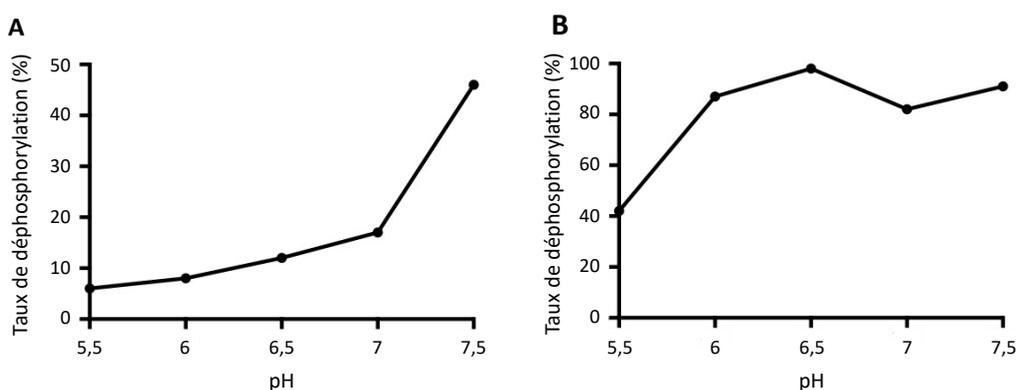


Figure 30. Taux de déphosphorylation des composés de synthèse phosphorylés en fonction du pH lors de leur incubation avec les lysats cellulaires H7. Les composés de synthèse testés sont le Gal-Xyl(2P)-OMN (A) et le Gal-Gal-Xyl(2P)-OMN (B), respectivement le substrat et le produit de la réaction catalysée par la β 1,3-galactosyltransférase 6 (β 3GalT6). Les essais contiennent 0,5 mM de Gal-Gal-Xyl(2P)-OMN ou de Gal-Xyl(2P)-OMN, 50 μ g de lysat cellulaire H7 et un tampon de réaction qui diffère en fonction du pH. A pH 5,5, 6 et 6,5, il est composé de 50 mM de MES et 100 mM de NaCl. A pH 7 et 7,5, il s'agit de 50 mM de MOPS et 100 mM de NaCl. Les mélanges réactionnels sont incubés 1 heure à 37 °C puis analysés par CLHP. Les valeurs sont obtenues en procédant au calcul suivant : (Aire forme déphosphorylée) / (Aire forme phosphorylée + aire forme déphosphorylée) x 100.

En incubant les composés en absence d'homogénat cellulaire, nous n'avons pas observé de déphosphorylation. Le phénomène de déphosphorylation des composés phosphorylés se produit donc en présence de lysats cellulaires. Ces réactions pourraient être catalysées par plusieurs phosphatases étant donné que leurs pH optimums sont différents. La

déphosphorylation expliquerait la faible activité apparente de la β 3GalT6 et la variabilité observée dans nos essais.

Afin de limiter ce phénomène, nous avons réalisé des essais enzymatiques en présence d'1 mM d'ATP, substrat donneur des kinases, dans le but de favoriser les formes phosphorylées et/ou de limiter la déphosphorylation du substrat ou du produit de la réaction. En utilisant cette méthode, il est possible de limiter leur déphosphorylation par les homogénats cellulaires (Figure 31). Néanmoins, aucune activité enzymatique spécifique n'a pu être mesurée à partir d'homogénats cellulaires de cellules HeLa qui expriment la β 3GalT6 de manière endogène.

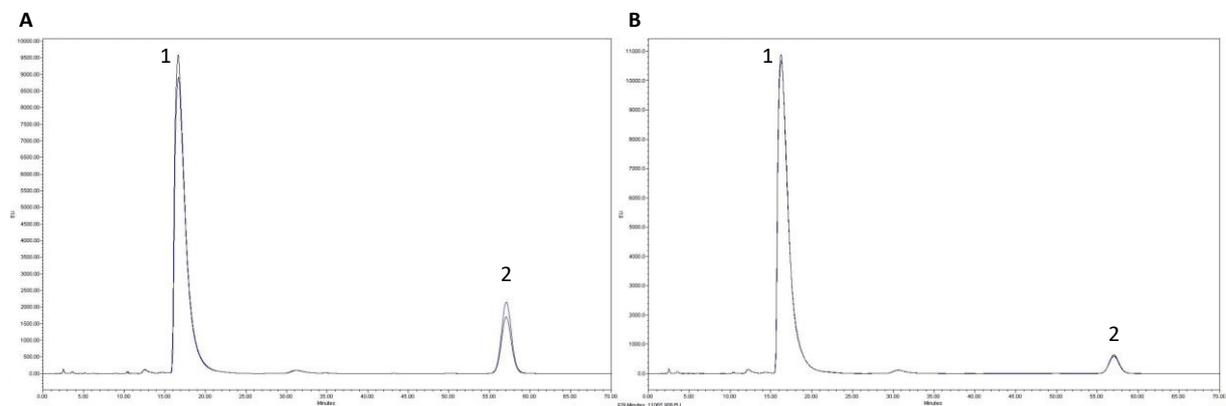


Figure 31. Effet de l'ajout d'adénosine triphosphate sur la déphosphorylation du substrat de synthèse de la réaction de la β 1,3-galactosyltransférase 6. La fluorescence est mesurée en unités d'émission (EU). L'essai consiste en l'incubation de 50 μ g de lysat de cellules HeLa non transfectées avec 2 mM de Gal-Xyl(2P)-OMN et sans ATP (A) ou avec 1 mM d'ATP (B) à pH 6,5. Le pic 1 correspond au Gal-Xyl(2P)-OMN et le pic 2 est l'impureté correspondant au Gal-Xyl-OMN. Les essais sont menés en duplicat.

Pour pallier les difficultés de mesure de l'activité enzymatique de la β 3GalT6 et des mutants dans les cellules H7, nous nous sommes orientés vers l'expression des protéines recombinantes dans la bactérie *E. coli* en vue de leur purification puis de la détermination de leurs paramètres cinétiques. Le but est de s'affranchir des protéines des lysats cellulaires, notamment des phosphatases, afin d'éviter la déphosphorylation du substrat et du produit de la réaction lors de l'incubation.

2. Utilisation d'un modèle bactérien pour la production et la purification d'enzymes recombinantes

La stratégie d'expression en système bactérien *E. coli*, classiquement utilisée pour les GT, consiste à exprimer le domaine catalytique soluble en tronquant les segments cytosolique et transmembranaire et une partie plus ou moins longue du segment *stem*. La β 3GalT6 est exprimée en fusion avec la MBP à l'extrémité N_{ter}, ce qui permet non seulement sa purification par chromatographie d'affinité, mais également d'augmenter sa solubilité. En effet, il a été difficile d'obtenir cette enzyme sous forme purifiée par le passé, et il a été montré que la MBP était un bon partenaire de fusion pour augmenter sa solubilité (279).

Nous disposons au laboratoire de deux vecteurs d'expression de la β 3GalT6 fusionnée à la MBP. La première construction, nommée 6His-MBP- β 3GalT6 Δ N_{ter}29-WT, code pour une protéine de fusion entre la MBP et la β 3GalT6 dépourvue des vingt-neuf premiers acides aminés correspondant aux segments cytosolique et transmembranaire. La seconde, nommée 6His-MBP- β 3GalT6 Δ N_{ter}50-WT, comprend la β 3GalT6 tronquée des cinquante premiers acides aminés correspondant aux mêmes segments ainsi que la majeure partie du segment *stem* (Figure 32).

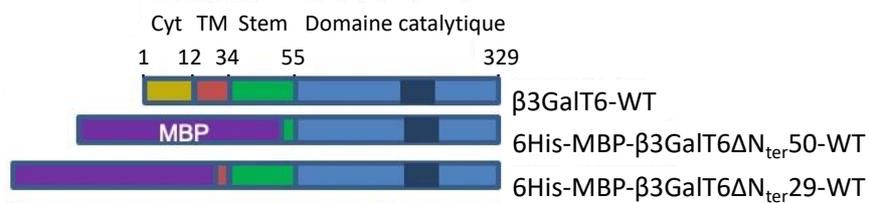


Figure 32. Représentation schématique de l'organisation de la β 1,3-galactosyltransférase 6 humaine et des protéines recombinantes produites dans le système bactérien. La β 1,3-galactosyltransférase 6 sauvage (β 3GalT6-WT) est composée d'un segment cytosolique (Cyt, jaune), un segment transmembranaire (TM, rouge), un segment *stem* (vert) et un domaine catalytique (bleu). Les nombres au-dessus de la β 3GalT6-WT correspondent aux acides aminés qui délimitent chaque domaine. Les formes tronquées de l'enzyme dans lesquelles le segment *stem* a été supprimé (6His-MBP- β 3GalT6 Δ N_{ter}50-WT) ou conservé (6His-MBP- β 3GalT6 Δ N_{ter}29-WT) ont été fusionnées à la protéine de fixation au maltose (*maltose binding protein*, MBP, violet) et exprimées dans la bactérie *E. coli*. Adapté de Diana *et al.*, *in external peer-review*.

a. Optimisation des conditions de production en système bactérien

Les enzymes ont été produites et les bactéries lysées selon les méthodes décrites dans la section « Matériels et Méthodes ». Un des paramètres importants à optimiser est le choix de la souche bactérienne à utiliser. Deux souches bactériennes *E. coli* ont été testées : BL21(DE3) et Rosetta2™(DE3). Le suivi de la surexpression de la 6His-MBP-β3GalT6ΔN_{ter}29-WT par SDS-PAGE coloré au bleu de Coomassie est présenté en Figure 33 et montre que la souche BL21(DE3) ne permet pas la surexpression de la protéine d'intérêt, contrairement à la souche Rosetta2™(DE3). La 6His-MBP-β3GalT6ΔN_{ter}29-WT présente une masse moléculaire de 75 kDa environ, ce qui correspond à la masse moléculaire attendue pour cette protéine de fusion entre la MBP à 42 kDa et la β3GalT6ΔN_{ter}29 à 34 kDa. Nous l'observons sous la forme de deux bandes distinctes de masse moléculaire proche : une analyse par spectrophotométrie de masse (Dr. Jean-Baptiste Vincourt, UMS 2008 IBSLor, Vandoeuvre-lès-Nancy, France) a confirmé qu'il s'agit bien dans les deux cas de la protéine de fusion recherchée. La bande avec un poids moléculaire plus faible pourrait correspondre à la protéine de fusion dont l'étiquette 6His a été spontanément clivée lors de la culture. La souche Rosetta2™(DE3) a donc été choisie pour poursuivre les expériences de production et de purification.

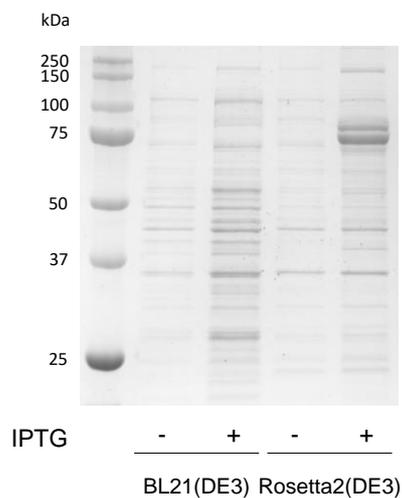


Figure 33. Analyse de l'expression de la 6His-MBP-β3GalT6ΔN_{ter}29-WT en fonction de la souche bactérienne utilisée. La production est réalisée à partir de deux souches *E. coli* : BL21(DE3) et Rosetta2™(DE3). L'induction est effectuée par ajout d'1 mM d'IPTG suivi d'une incubation pendant une nuit sous agitation à 20 °C. Les fractions analysées correspondent aux protéines totales de lysat avant (-) et après (+) induction de la production. Les chiffres à gauche indiquent la masse moléculaire en kDa du marqueur de taille.

b. Détermination des effets de la troncation sur les propriétés enzymatiques

La MBP est fusionnée aux enzymes recombinantes au niveau de l'extrémité N_{ter}. Le segment *stem* module la distance entre la MBP et le domaine catalytique de l'enzyme, ce qui demande de maîtriser sa longueur afin de choisir la troncation adéquate pour étudier l'activité enzymatique de la β 3GalT6. Pour cette raison, nous avons décidé de concentrer nos travaux sur les deux constructions décrites précédemment.

Les deux enzymes recombinantes 6His-MBP- β 3GalT6 Δ N_{ter}29-WT et 6His-MBP- β 3GalT6 Δ N_{ter}50-WT sont solubles et ont été purifiées avec un rendement similaire, de 1,5 à 2 mg de produit de purification par 200 mL de milieu de culture (Figure 34, Tableau VIII).

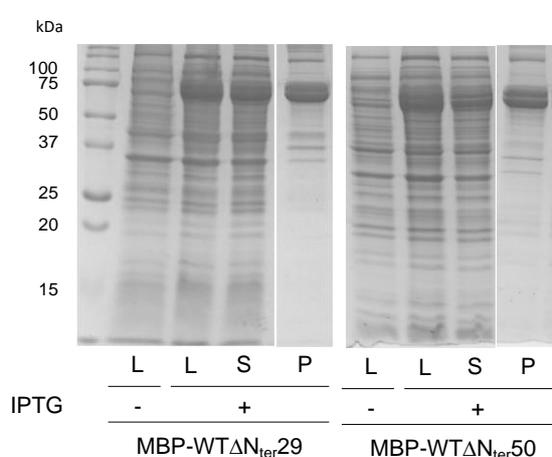


Figure 34. Analyse de l'expression, de la solubilité et de la pureté des protéines recombinantes 6His-MBP- β 3GalT6 Δ N_{ter}29-WT et 6His-MBP- β 3GalT6 Δ N_{ter}50-WT par SDS-PAGE après coloration du gel au bleu de Coomassie. Les protéines ont été exprimées dans la souche bactérienne *E. coli* Rosetta2™(DE3). L'induction est effectuée par ajout d'1 mM d'IPTG suivi d'une incubation pendant une nuit sous agitation à 20 °C. Les fractions correspondant aux protéines totales dans le lysat (L), aux protéines solubles dans le surnageant (S) et au produit de purification (P) ont été analysées. Les chiffres à gauche indiquent la masse moléculaire en kDa du marqueur de taille. Adapté de Diana *et al.*, *in external peer-review*.

| Fraction | Volume (mL) | Quantité de protéine (mg) | Concentration (mg/mL) | Facteur de concentration |
|--|-------------|---------------------------|-----------------------|--------------------------|
| Extrait brut | 200 | 41 | 0,2 | 0 |
| Chromatographie d'affinité | 10 | 4 | 0,5 | 2,5 |
| Concentration du produit de purification | 1 | 1,5 | 1,5 | 7,5 |

Tableau VIII. Bilan de purification de la β 1,3-galactosyltransférase 6 humaine tronquée et fusionnée à la MBP.

Ce tableau présente les valeurs obtenues pour la 6His-MBP- β 3GalT6 Δ N_{ter}29-WT. Les valeurs obtenues pour la 6His-MBP- β 3GalT6 Δ N_{ter}50-WT sont très similaires et non représentées. Le facteur de concentration est obtenu en divisant la concentration de l'étape de purification considérée par celle de l'extrait brut.

Les constantes cinétiques des deux enzymes recombinantes ont ensuite été évaluées selon la méthode décrite dans « Matériels et Méthodes » après optimisation de la quantité de protéine ajoutée et du temps d'incubation afin de se placer en conditions saturantes en substrats. La β 3GalT6 exprimée au moyen de la construction 6His-MBP- β 3GalT6 Δ N_{ter}29-WT présente une cinétique de type Michaelis-Menten contrairement à la 6His-MBP- β 3GalT6 Δ N_{ter}50-WT. La 6His-MBP- β 3GalT6 Δ N_{ter}29-WT possède une affinité (K_M) de 1,44 mM pour le substrat donneur, 0,38 mM pour le substrat accepteur et une V_{max} d'environ 1,2 nmol.min⁻¹.mg⁻¹. Pour la 6His-MBP- β 3GalT6 Δ N_{ter}50-WT, l'affinité pour le substrat donneur est de 2,03 mM et celle pour le substrat accepteur n'a pu être déterminée. La V_{max} calculée lors de la cinétique d'UDP-Gal est d'environ 0,5 nmol.min⁻¹.mg⁻¹. L'efficacité catalytique (V_{max}/K_M) de la 6His-MBP- β 3GalT6 Δ N_{ter}29-WT vis-à-vis du substrat donneur est donc meilleure que celle de la 6His-MBP- β 3GalT6 Δ N_{ter}50-WT (Tableau IX).

En prenant en compte ces résultats, la 6His-MBP- β 3GalT6 Δ N_{ter}29-WT a été choisie pour générer par mutagenèse dirigée les mutations d'intérêt retrouvées chez les patients.

| Enzyme | UDP-Gal | | | Gal-Xyl(2P)-OMN | | |
|--|--|---------------|--|--|---------------|--|
| | V_{max} (nmol.min ⁻¹ .mg ⁻¹) | K_m (mM) | V_{max}/K_m (mL.min ⁻¹ .mg ⁻¹) | V_{max} (nmol.min ⁻¹ .mg ⁻¹) | K_m (mM) | V_{max}/K_m (mL.min ⁻¹ .mg ⁻¹) |
| 6His-MBP- ®3GalT6 Δ N _{ter} 29- WT | 1,25 ± 0,04 | 1,44 ± 0,36 | 0,87 | 1,07 ± 0,05 | 0,38 ± 0,07 | 2,82 |
| 6His-MBP- ®3GalT6 Δ N _{ter} 50- WT | 0,51 ± 0,03 | 2,03 ± 0,34 | 0,25 | < 0,2 | N.D. | N.D. |

Tableau IX. Paramètres cinétiques des protéines recombinantes 6His-MBP- β 3GalT6 Δ N_{ter}29-WT et 6His-MBP- β 3GalT6 Δ N_{ter}50-WT purifiées. Les paramètres cinétiques envers les substrats donneur et accepteur ont été évalués à partir des valeurs de vitesse initiale de la réaction en utilisant 50 pmol d'enzyme purifiée en présence d'une concentration fixe de 5 mM de Gal-Xyl(2P)-OMN et d'une concentration variable (0 - 10 mM) d'UDP-Gal, ou en présence d'une concentration fixe de 10 mM d'UDP-Gal et d'une concentration variable (0 - 5 mM) de Gal-Xyl(2P)-OMN. Les paramètres cinétiques ont été déterminés par l'analyse des moindres carrés non-linéaires et l'ajustement de la courbe à l'équation de Michaelis-Menten en utilisant le logiciel GraphPad Prism 7. Les résultats obtenus sont les moyennes de trois expériences différentes ± l'écart-type en duplicats. Adapté de Diana *et al.*, *in external peer-review*.

c. Évaluation de l'effet de mutations sur l'activité enzymatique

Après mutagenèse dirigée réalisée sur la construction 6His-MBP- β 3GalT6 Δ N_{ter}29-WT et confirmation de la présence des mutations par séquençage, les trois mutants 6His-MBP- β 3GalT6 Δ N_{ter}29-Y182C, 6His-MBP- β 3GalT6 Δ N_{ter}29-D207H et 6His-MBP- β 3GalT6 Δ N_{ter}29-G217S ont été exprimés et purifiés selon la même méthode que l'enzyme sauvage. Le taux de production, la pureté et le rendement sont similaires pour les quatre protéines recombinantes (Figure 35, Tableau VIII).

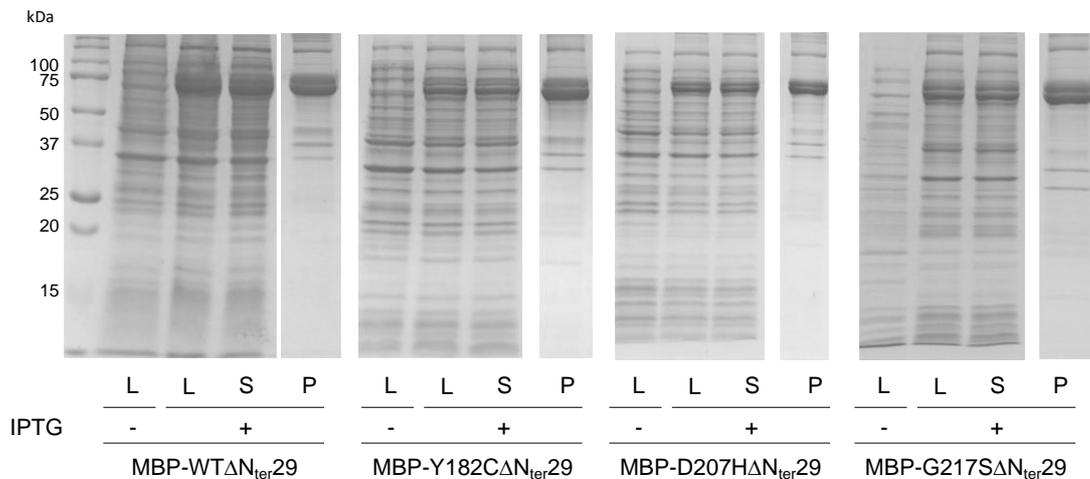


Figure 35. Analyse de l'expression, de la solubilité et de la pureté des protéines recombinantes 6His-MBP- β 3GalT6 Δ N_{ter}29-WT, 6His-MBP- β 3GalT6 Δ N_{ter}29-Y182C, 6His-MBP- β 3GalT6 Δ N_{ter}29-D207H et 6His-MBP- β 3GalT6 Δ N_{ter}29-G217S par SDS-PAGE après coloration du gel au bleu de Coomassie. Les protéines ont été exprimées dans la souche bactérienne *E. coli* Rosetta2TM(DE3). L'induction est effectuée par ajout d'1 mM d'IPTG suivi d'une incubation pendant une nuit sous agitation à 20 °C. Les fractions correspondant aux protéines totales dans le lysat (L), aux protéines solubles dans le surnageant (S) et au produit de purification (P) ont été analysées. Les chiffres à gauche indiquent la masse moléculaire en kDa du marqueur de taille. Adapté de Diana *et al.*, *in external peer-review*.

Alors que la 6His-MBP- β 3GalT6 Δ N_{ter}29-WT présente une cinétique michaelienne en conditions saturantes en substrat, aucun des trois mutants exprimés ne présente d'activité détectable, indiquant une perte totale de fonction.

3. Exploration du rôle de la β 1,3-glucuronosyltransférase 1 en cas de déficience en β 1,3-galactosyltransférase 6

Nos travaux montrent une perte de fonction enzymatique de la β 3GalT6 produites par les mutations Tyr182Cys, Asp207His ou Gly217Ser. Ce résultat est en désaccord apparent avec nos observations de la section I.1.b où nous mettons en évidence une forte diminution mais non une abolition de la synthèse des chaînes de GAG lorsque le gène *B3GALT6* est invalidé et que les cellules sont transfectées pour exprimer une β 3GalT6 mutée. En théorie, la totalité de la synthèse des chaînes de GAG de types CS, DS et HS devrait être interrompue en cas de perte de fonction de β 3GalT6, car ces chaînes débutent toutes par une amorce tétrasaccharidique.

La GlcAT-I est la quatrième enzyme de la synthèse de l'amorce tétrasaccharidique. Son activité a lieu directement après celle de la β 3GalT6, qui agit elle-même juste après la β 4GalT7. L'hypothèse d'une synthèse alternative des chaînes de GAG lorsque le gène *B3GALT6* est muté a été proposée (289). Dans ce cas, la GlcAT-I agirait à la suite de la β 4GalT7 pour former une amorce trisaccharidique et permettrait la poursuite de la synthèse des chaînes de GAG. Afin de vérifier cette hypothèse, nous avons évalué l'activité de la GlcAT-I Δ N_{ter}76 recombinante *in vitro* en présence de deux substrats accepteurs différents. Le premier est un analogue de son substrat canonique, le trisaccharide Gal-Gal-Xyl-OMN. L'activité de l'enzyme recombinante envers ce substrat à une concentration de 0,5 mM ou 1 mM est d'environ 53 $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ de protéines. Le second composé chimique est le Gal-Xyl-OMN. A 0,5 mM de Gal-Xyl-OMN, l'activité spécifique de l'enzyme est de 5,3 $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ de protéines, soit 10 fois moins que pour le substrat trisaccharidique. A 1 mM, l'activité spécifique atteint 9,8 $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ de protéines (Figure 36).

Ces résultats indiquent que la GlcAT-I est effectivement capable de catalyser, bien que moins efficacement, le transfert d'un GlcA sur un substrat disaccharidique de nature Gal-Xyl-O, permettant la formation d'une amorce trisaccharidique non canonique.

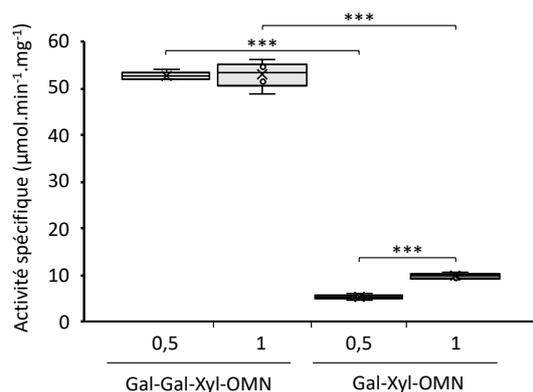


Figure 36. Activité spécifique de la β 1,3-glucuronosyltransférase I recombinante purifiée. Les deux analogues de substrats accepteurs Gal-Gal-Xyl-OMN et Gal-Xyl-OMN ont été utilisés à des concentrations de 0,5 mM ou 1 mM. Les données représentent les moyennes \pm écart-types de deux expériences distinctes réalisées en triplicat. *** : $p < 0,001$ (Test t de Student). Adapté de Diana *et al.*, in external peer-review.

II. Étude des conséquences structurales de mutations pathogènes de la β 1,3-galactosyltransférase 6

1. Étude de la distribution de taille des enzymes recombinantes par diffusion dynamique de la lumière

La perte de fonction enzymatique des β 3GalT6 recombinantes porteuses d'une mutation Tyr182Cys, Asp207His ou Gly217Ser nous a conduit à entreprendre des analyses structurales de ces enzymes. Les distributions de taille des β 3GalT6 purifiées sont analysées systématiquement à la fin de chaque purification par DLS dans le but de caractériser leur état d'agrégation. Ces données représentent un contrôle qualité avant l'utilisation des enzymes purifiées pour les essais enzymatiques. Nous voulons notamment nous assurer que les mutants ne sont pas sous forme d'agrégats, ou que la conservation à $-80\text{ }^{\circ}\text{C}$ ne cause pas de dégradation ou d'agrégation des protéines purifiées.

L'analyse des profils des protéines purifiées obtenues en sortie de purification montre que la 6His-MBP- β 3GalT6 Δ N_{ter}29-WT a un indice de polydispersité de moins de 0,2 (Figure 37). Elle est donc considérée comme monodisperse. Son diamètre hydrodynamique est d'environ 60 nm et n'indique pas d'agrégation.

La présence des mutations Tyr182Cys, Asp207His ou Gly217Ser ne cause pas de changement majeur de profil par rapport à l'enzyme sauvage. Les échantillons restent monodisperses avec un diamètre hydrodynamique similaire (Figure 37). Ces analyses

suggèrent que les mutations n'impactent pas la distribution de taille de l'enzyme recombinante.

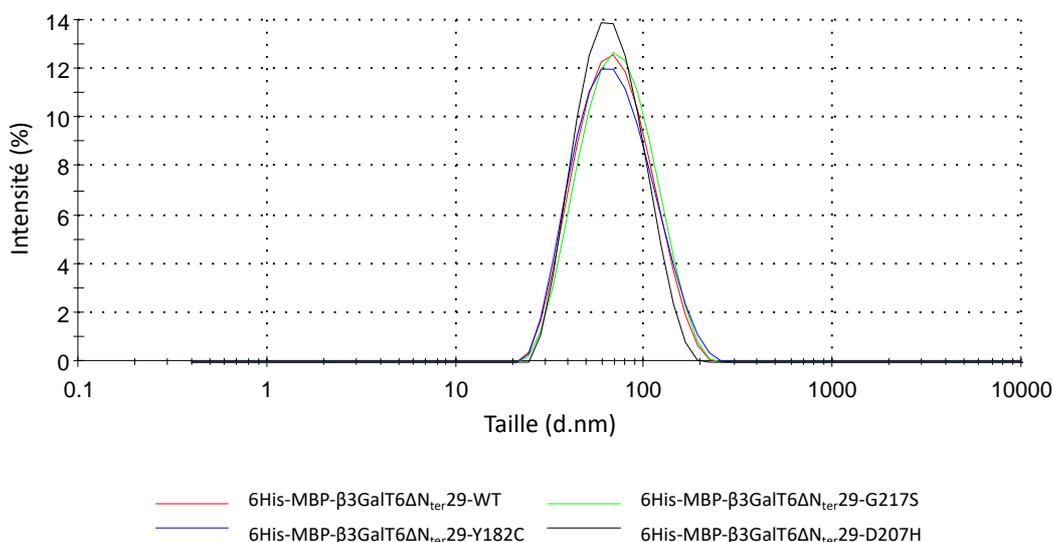


Figure 37. Profils d'intensité de diffusion dynamique de la lumière des β 1,3-galactosyltransférases 6 humaines purifiées sauvage ou mutées. L'axe des abscisses représente le rayon hydrodynamique (nm) selon une échelle logarithmique. L'axe des ordonnées représente le pourcentage total d'intensité mesuré. Chaque courbe représente la moyenne de quinze mesures distinctes sur un même échantillon.

2. Étude de la structure secondaire des enzymes recombinantes sauvage et mutées par dichroïsme circulaire

Pour compléter les études de DLS, des expériences de DC sont systématiquement entreprises en sortie de purification afin d'analyser le contenu en structures secondaires des enzymes recombinantes. Le spectre de DC de la 6His-MBP- β 3GalT6 Δ N_{ter}29-WT est caractéristique des protéines composées d'une majorité d'hélices α , avec un maximum à 195 nm et deux minima à 210 et 220 nm (Figure 38). La signature au DC des enzymes mutées est identique (Figure 38). Ces résultats confirment que les enzymes recombinantes sont correctement repliées et indiquent que la perte de fonction enzymatique relevée pour chaque mutant ne résulte pas d'une altération structurale majeure.

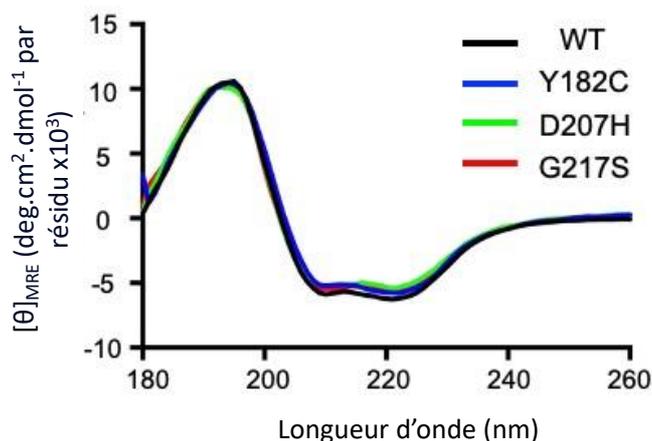


Figure 38. Spectre de dichroïsme circulaire dans l’UV lointain des β 1,3-galactosyltransférases 6 recombinantes sauvage et mutées. Une concentration de 20 μ M de 6His-MBP- β 3GalT6 Δ Nt29-WT (noir), 6His-MBP- β 3GalT6 Δ Nt29-Y182C (bleu), 6His-MBP- β 3GalT6 Δ Nt29-D207H (vert) ou 6His-MBP- β 3GalT6 Δ Nt29-G217S (rouge) a été utilisée. L’ellipticité molaire $[\theta]_{MRE}$ permet de prendre en compte le poids moléculaire, la concentration des enzymes ainsi que la longueur de la cuve. Chaque courbe représente la moyenne de trois enregistrements. Le même profil a été observé sur trois lots d’enzymes différents. Adapté de Diana *et al.*, *in external peer-review*.

3. Étude de la dynamique moléculaire de la β 1,3-galactosyltransférase 6 sauvage ou mutée

La structure tridimensionnelle de la β 3GalT6 n’est pas résolue à ce jour. Afin d’obtenir un modèle de la β 3GalT6 humaine qui permettrait d’expliquer la perte de fonction due aux mutations Tyr182Cys, Asp207His et Gly217Ser, nous nous sommes orientés vers la dynamique moléculaire (en collaboration avec les Drs. Yasaman Karami et Hamed Khakzad, UMR 7503 LORIA, Vandoeuvre-lès-Nancy, France). Le modèle de structure de la β 3GalT6 prédite par Alphafold présente un degré de confiance élevé (Figure 18) et constitue un point de départ satisfaisant pour cette approche. Ce modèle révèle trois domaines principaux : le segment transmembranaire organisé en hélice α , le segment *stem* déstructuré et le domaine catalytique constitué de feuillets β , d’hélices α et de boucles déstructurées. Le domaine catalytique adopte une conformation caractéristique des GT-A, avec un pli de Rossmann $\alpha/\beta/\alpha$. Il est constitué de deux feuillets β contenant respectivement les brins β 1- β 2- β 3- β 4- β 6- β 7- β 8 et les brins β 5- β 9 séparés par des hélices α et des boucles (Figure 39).

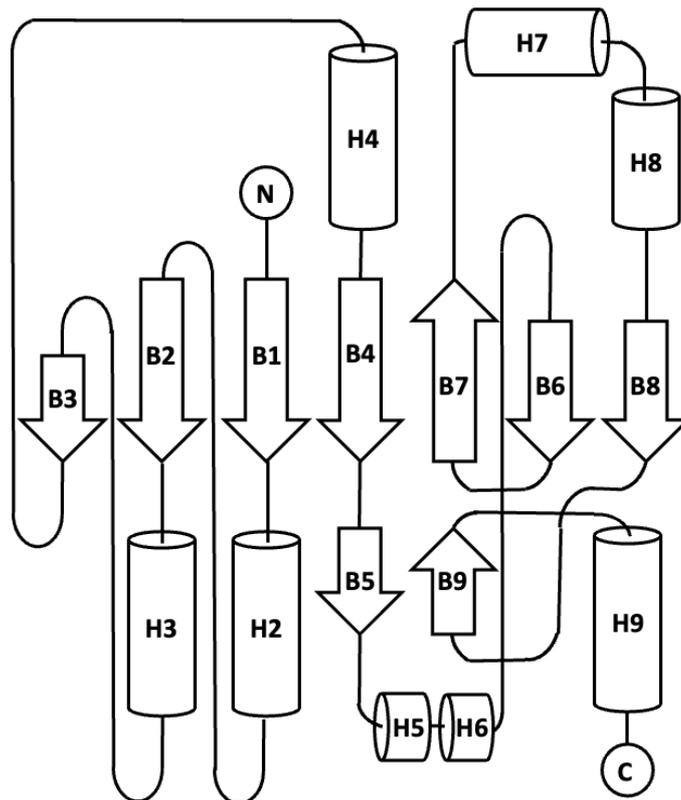


Figure 39. Diagramme de la topologie du domaine catalytique de la β1,3-galactosyltransférase 6 humaine. Les flèches et les cylindres représentent les brins β (B) et les hélices α (H), respectivement. Les structures secondaires sont numérotées selon leur ordre d'apparition, de l'extrémité N_{ter} (N) à l'extrémité C_{ter} (C). Adapté de Diana *et al.*, *in external peer-review*.

Parmi les éléments notables, on observe (Figure 40) :

- Le domaine DxD (Asp¹⁵⁶-Asp¹⁵⁷-Asp¹⁵⁸) qui coordonne le cation divalent Mn²⁺ et le substrat donneur UDP-Gal. L'Asp¹⁵⁶ se trouve sur le brin β4. L'Asp¹⁵⁷ et l'Asp¹⁵⁸ font le lien entre les brins β4 et β5 ;
- Le domaine xED (Ser²⁴⁰-Glu²⁴¹-Asp²⁴²) qui contient un résidu catalytique (prédit comme étant l'Asp²⁴²). La Ser²⁴⁰ est située dans la boucle faisant le lien entre les hélices α H7 et H8. Le Glu²⁴¹ et l'Asp²⁴² se trouve dans l'hélice α H8 ;
- Le résidu Tyr¹⁸², impliqué dans la mutation Tyr182Cys. Il est situé dans le feuillet β6 enfoui dans le site catalytique ;
- Le résidu Gly²¹⁷, impliqué dans la mutation Gly217Ser. Il est situé dans le feuillet β7 également enfoui dans le site catalytique. Dans l'espace, il se trouve à proximité du domaine DxD, les deux brins β4 et β7 étant côte-à-côte dans le plus grand des deux

feuillet β . Il se trouve également à proximité du résidu Tyr¹⁸², les deux brins β 6 et β 7 étant également adjacents dans ce même feuillet β ;

- Le résidu Asp²⁰⁷, impliqué dans la mutation Asp207His. Il est situé dans une longue boucle Gly¹⁸⁸-Pro²¹¹ entre les brins β 6 et β 7.

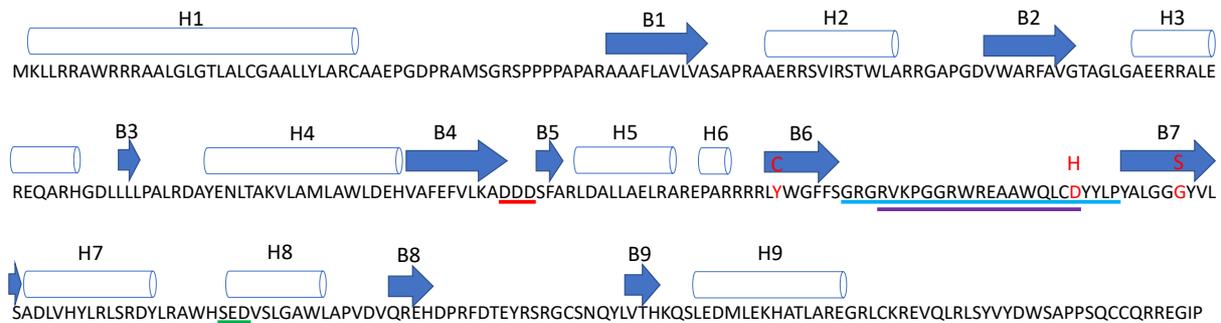


Figure 40. Séquence primaire et structures secondaires de la β 1,3-galactosyltransférase 6. Les flèches et les cylindres représentent les brins β (B) et les hélices α (H), respectivement. Les structures secondaires sont numérotées selon leur ordre d'apparition, de l'extrémité N_{ter} à l'extrémité C_{ter}. Les mutations étudiées Tyr182Cys, Asp207His et Gly217Ser sont représentées en rouge. Le domaine DxD (Asp¹⁵⁶-Asp¹⁵⁷-Asp¹⁵⁸) et le domaine xED (Ser²⁴⁰-Glu²⁴¹-Asp²⁴²) sont soulignés en rouge et en vert, respectivement. La boucle Gly¹⁸⁸-Pro²¹¹ est soulignée en cyan. Un segment présentant de hautes fluctuations dans cette boucle, Arg¹⁹¹-Asp²⁰⁷, est souligné en violet. Adapté de Diana *et al.*, *in external peer-review*.

A partir de la prédiction de structure décrite ci-dessus, les quatre modèles de la β 3GalT6 humaine sauvage ou comportant l'une des trois mutations d'intérêt ont été étudiés par dynamique moléculaire. Les mouvements des enzymes enchâssées dans un modèle de membrane lipidique de l'appareil de Golgi ont été enregistrés pendant 12 μ s. Les liaisons hydrogène et les ponts salins intramoléculaires présents dans plus de 40 % des simulations d'au moins un des quatre modèles ont été répertoriés, et leur présence a été comparée entre chaque simulation (Figure 41). Chaque mutation crée d'importantes différences d'intensité au niveau de plusieurs de ces liaisons par rapport à l'enzyme sauvage.

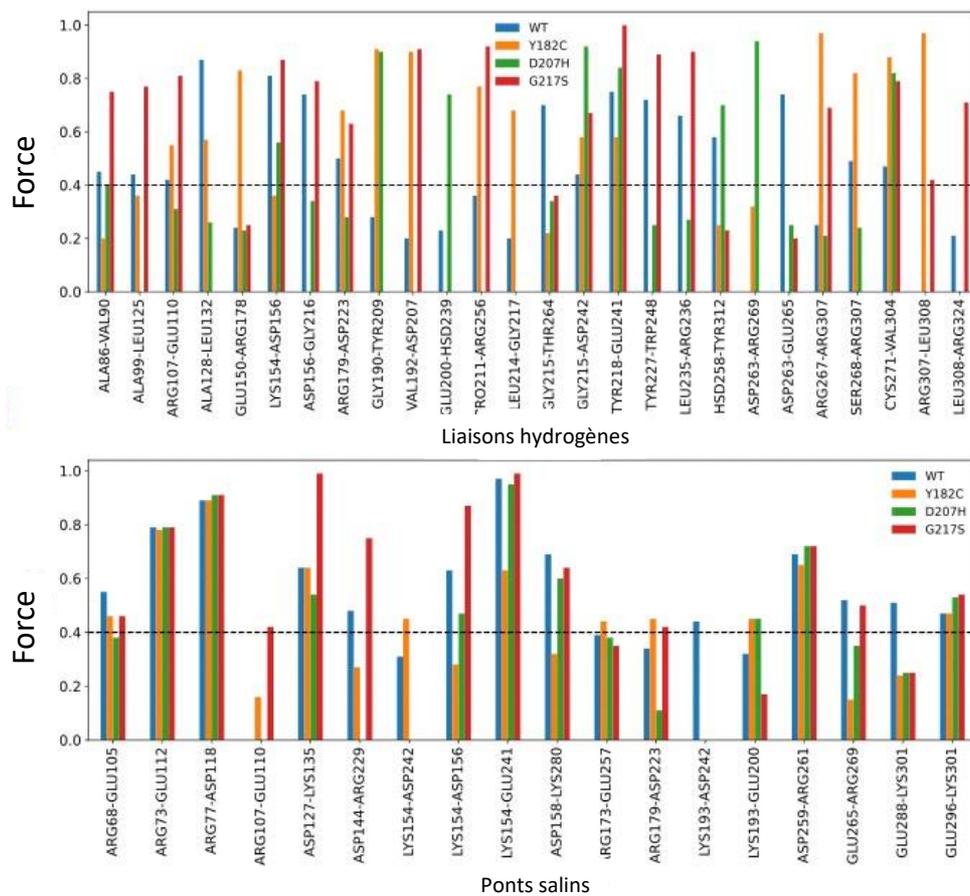


Figure 41. Forces des liaisons hydrogènes et des ponts salins enregistrées lors des simulations de dynamique moléculaire de la β 1,3-galactosyltransférase 6 sauvage et mutée. Les liaisons présentées ont un pourcentage d'intensité supérieur ou égal à 40 % dans au moins un des systèmes : β 3GalT6-WT (bleu), β 3GalT6-Y182C (jaune), β 3GalT6-D207H (vert) ou β 3GalT6-G217S (rouge). Adapté de Diana *et al.*, *in external peer-review*.

Résidus du motif xED :

Le pont salin Lys¹⁹³-Asp²⁴² est particulièrement intéressant, étant présent dans la simulation de l'enzyme sauvage mais absent dans chacun des trois mutants (Figure 41). Le résidu Lys¹⁹³ se trouve dans la longue boucle flexible Gly¹⁸⁸-Pro²¹¹. L'Asp²⁴² est la base catalytique du motif xED prédite responsable de l'attaque nucléophile, étape-clé de la catalyse enzymatique. Ce pont salin serait essentiel à la connexion de la boucle au centre catalytique. De plus, le segment Arg¹⁹¹-Asp²⁰⁷ qui se trouve dans cette boucle présente un haut taux de fluctuations qui diminue pour chaque mutant étudié (Figures 40 et 42). La rupture du pont salin Lys¹⁹³-Asp²⁴² serait donc un événement majeur contribuant à la perte de fonction enzymatique, d'une part à cause de la réduction des fluctuations de la longue boucle, d'autre part à cause de l'implication de la base catalytique Asp²⁴² dans cette liaison.

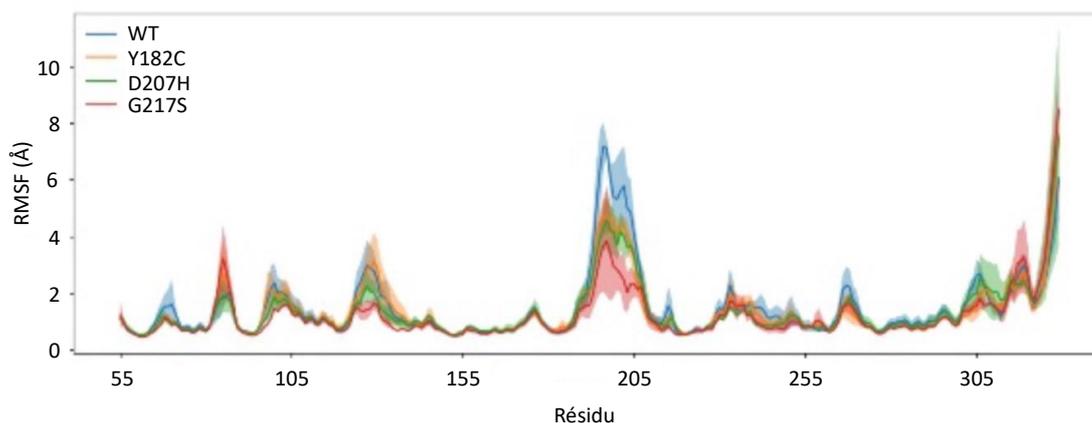


Figure 42. Fluctuations moyennes des positions des résidus du domaine catalytique de la β 1,3-galactosyltransférase 6 dans les modèles de simulations de dynamique moléculaire. Les fluctuations ont été mesurées sur les atomes du squelette peptidique par rapport à leur positionnement moyen pour chaque modèle : β 3GalT6-WT (bleu), β 3GalT6-Y182C (jaune), β 3GalT6-D207H (vert) et β 3GalT6-G217S (rouge). Les données correspondent à la moyenne quadratique des fluctuations (*root-mean-square fluctuation*, RMSF) des positions atomiques de chaque résidu, donnée en Ångström (Å). Les valeurs de fluctuations représentées en traits pleins pour chaque système sont la moyenne de trois simulations distinctes et les écart-types sont représentés par les nuances autour des traits pleins. Adapté de Diana *et al.*, *in external peer-review*.

Résidus du motif DxD :

L'intensité de liaison du pont salin Lys¹⁵⁴-Asp¹⁵⁶ est moins forte pour le mutant Asp207His et surtout pour le mutant Tyr182Cys (Figure 41). L'intensité de la liaison hydrogène Asp¹⁵⁶-Gly²¹⁶ diminue fortement chez le mutant Asp207His, et elle est rompue dans le mutant Tyr182Cys (Figure 41). Le résidu Asp¹⁵⁶ faisant partie du motif DxD, il est possible que les mutations Tyr182Cys et Asp207His aient aussi un impact sur ce motif et affectent la coordination du cation divalent Mn²⁺ et/ou la fixation ou la reconnaissance du substrat donneur de la réaction.

Nos observations indiquent que les mutants, bien qu'adoptant une structure secondaire similaire à celle de l'enzyme sauvage en simulation de dynamique moléculaire (confirmé par DC), présentent des altérations des liaisons intramoléculaires qui impliquent des éléments fonctionnels essentiels des GT, comme le motif DxD et le motif xED. Celles-ci contribueraient à la perte de fonction enzymatique observée pour ces mutants.

III. Conséquences de la déficience en β 1,3-galactosyltransférase 6 sur la matrice extracellulaire dans un modèle cellulaire de synthèse matricielle

Les MEC sont constituées d'un grand nombre de macromolécules, dont les PG, et sont la cible de dysfonctionnements dans le contexte d'une maladie génétique rare des TC telle que le SEDsp-*B3GALT6*. Afin d'étudier les conséquences de la déficience de la biosynthèse des chaînes de GAG due à l'absence de fonction de la β 3GalT6 sur la composition et les propriétés biomécaniques des MEC, nous avons développé un modèle cellulaire de synthèse matricielle en utilisant la lignée cellulaire ATDC5. Lors de la chondrogenèse et sous stimulation à l'insuline, ces cellules produisent une MEC cartilagineuse organisée sous la forme d'un pseudo-tissu, ce qui en fait un bon modèle tridimensionnel pour notre étude.

1. Évaluation de la synthèse des chaînes de glycosaminoglycanes par les pseudo-tissus invalidés pour le gène *B3galt6*

Les cellules sont invalidées pour le gène *B3galt6* puis la synthèse matricielle est induite selon les méthodes décrites. Le fractionnement protéique des échantillons permet d'analyser les chaînes de GAG des PG retrouvées dans différents compartiments de la MEC.

Le **sécrétome** consiste en la fraction de protéines sécrétées dans le milieu de culture par les cellules et contient notamment de la décorine dont nous avons analysé l'état de glycosylation par immunoblot (Figure 43). Après induction, les cellules ATDC5 non-invalidées sont capables de glycosyler la décorine comme le montre la bande diffuse de masse moléculaire de 75 kDa à 125 kDa dans le puits Ctrl. Lorsque les cellules sont invalidées, dans le puits Δ *B3galt6*, une bande nette et majoritaire correspondant à la protéine-core de la décorine apparaît à 42 kDa. Tout comme les cellules H7, invalidées pour le gène *B3GALT6*, les cellules conservent une capacité de synthèse résiduelle, comme le montre la bande diffuse de très faible intensité située entre 75 et 125 kDa. Ceci confirme que l'invalidation du gène fonctionne, réduisant fortement la capacité des cellules à synthétiser des chaînes de GAG de la décorine.

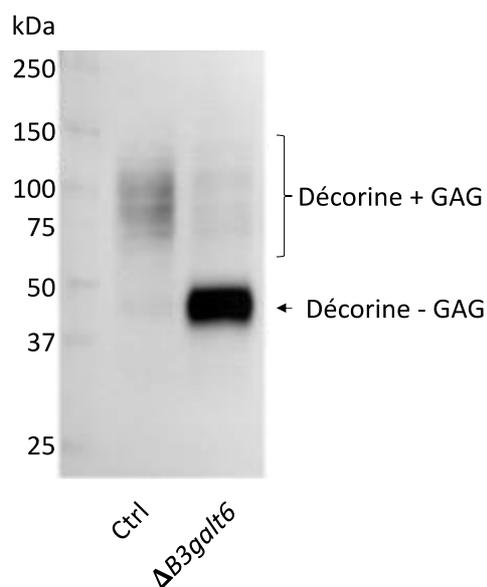


Figure 43. Immunoblot de la décorine sécrétée par les ATDC5 contrôles ou invalidées pour le gène *B3galt6*. Le gène *B3galt6* a été invalidé par CRISPR-Cas9 ($\Delta B3galt6$) ou non (Ctrl). Les chiffres à gauche indiquent la masse moléculaire en kDa du marqueur de taille. Un anticorps dirigé contre la décorine murine a été utilisé pour cet immunoblot. GAG : glycosaminoglycane. Adapté de Diana *et al.*, *in external peer-review*.

La **fraction guanidine** du pseudo-tissu est constituée des protéines matricielles dépourvues de liaisons croisées, à savoir des grands PG comme l'aggrécane, des collagènes non fibrillaires et d'autres protéines. Comme pour la fraction du sécrétome, les protéines sont séparées par SDS-PAGE. Une analyse après coloration du gel au bleu de Coomassie de la fraction guanidine des ATDC5 invalidées a révélé l'apparition d'une bande nette de haut poids moléculaire (Figure 44A), qui s'est révélée être le Col XII après analyse par spectrométrie de masse (Dr. Jean-Baptiste Vincourt, UMS 2008 IBSLor, Vandoeuvre-lès-Nancy, France). Afin de confirmer l'expression du Col XII dans les cellules invalidées, une analyse par immunoblot a été menée sur cette fraction guanidine. La forme longue du Col XII est visible sous la forme d'une bande diffuse à plus de 250 kDa dans les cellules contrôles et sous la forme d'une bande fine et nette dans les cellules invalidées. Ces résultats indiquent que l'invalidation de *B3galt6* entraîne une perte de la glycosylation de la forme longue du Col XII. La forme courte du Col XII est également révélée par immunoblot, sous la forme d'une bande nette entre 150 kDa et 250 kDa dans les cellules contrôles et invalidées (Figure 44B).

La fraction guanidine a également permis de montrer la diminution de la quantité de chaînes de GAG provenant de grands PG, principalement l'aggrécane (données non montrées)

(Diana *et al.*, *in external peer-review*), renforçant la pertinence de notre modèle cellulaire comme outil d'étude des défauts de la synthèse des chaînes de GAG au niveau matriciel.

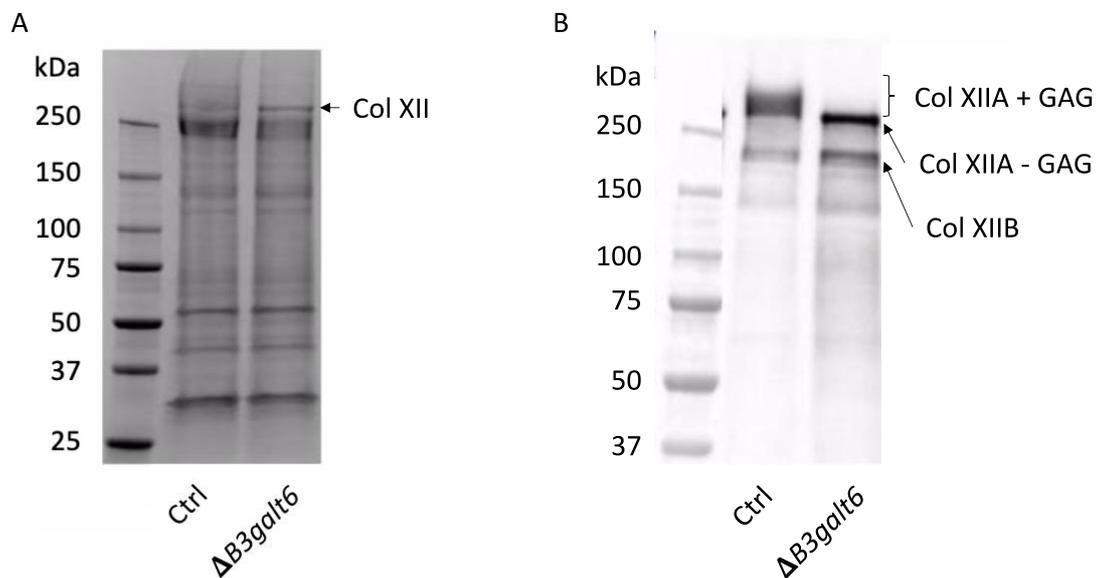


Figure 44. Analyse de la fraction guanidine extraite à partir des pseudo-tissus de cellules ATDC5 contrôles ou invalidées pour le gène *B3galt6*. (A) Profil de migration des protéines de la fraction guanidine des ATDC5 dont le gène *B3galt6* a été ciblé par CRISPR-Cas9 ($\Delta B3galt6$) ou non (Ctrl) après analyse par SDS-PAGE suivi d'une coloration du gel au bleu de Coomassie. La bande identifiée comme étant le collagène de type XII (Col XII) par spectrométrie de masse est indiquée. Les chiffres à gauche indiquent la masse moléculaire en kDa du marqueur de taille. (B) Immunoblot du Col XII de la fraction guanidine des ATDC5 invalidées ($\Delta B3galt6$) ou non (Ctrl). La forme longue du Col XII (Col XIIA) est glycosylée dans le puits Ctrl et non glycosylée dans le puits $\Delta B3galt6$. La forme courte du Col XII (Col XIIIB) est visible dans les deux puits. Les chiffres à gauche indiquent la masse moléculaire en kDa du marqueur de taille. Un anticorps dirigé contre la décorine murine a été utilisé pour cet immunoblot. GAG : glycosaminoglycane. Adapté de Diana *et al.*, *in external peer-review*.

2. Analyses des pseudo-tissus invalidés pour le gène *B3galt6*

a. Analyses histologiques

Les pseudo-tissus ont été inclus en paraffine et des coupes transversales des différents blocs ont été réalisées. Les chaînes de GAG sulfatées sont colorées au bleu Alcian. De façon globale, au grossissement x60, la MEC des pseudo-tissus invalidés est moins bien organisée, avec des lacunes montrant que l'absence des chaînes de GAG produit une altération de l'architecture matricielle (Figure 45).

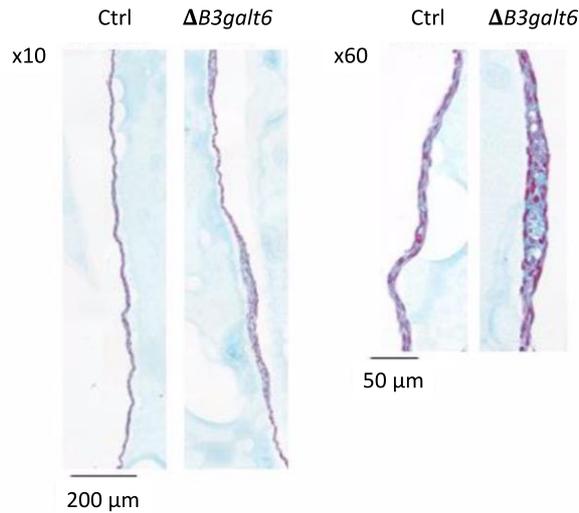


Figure 45. Analyses histologiques de coupes transversales de pseudo-tissus obtenus à partir des cellules ATDC5 contrôles ou invalidées pour le gène *B3galt6*. Le gène *B3galt6* des cellules est ciblé par CRISPR-Cas9 ($\Delta B3galt6$) ou non (Ctrl). Les chaînes de glycosaminoglycanes (GAG) sulfatées et les noyaux cellulaires sont colorés au bleu Alcian et au *Nuclear fast red*, respectivement. Les images sont obtenues par microscopie optique à des grossissements 10x (gauche) et 60x (droite). Adapté de Diana *et al.*, *in external peer-review*.

b. Tests d'extensibilité

Afin d'étudier les propriétés biomécaniques des pseudo-tissus, une mesure de leur extensibilité par la gravité a été réalisée. Les pseudo-tissus invalidés sont significativement plus extensibles que les pseudo-tissus contrôles, de l'ordre de 20 % (Figure 46). Dans le modèle de chondrogenèse, l'invalidation de *B3galt6* cause donc des défauts biomécaniques qu'il convient d'explorer plus en profondeur.

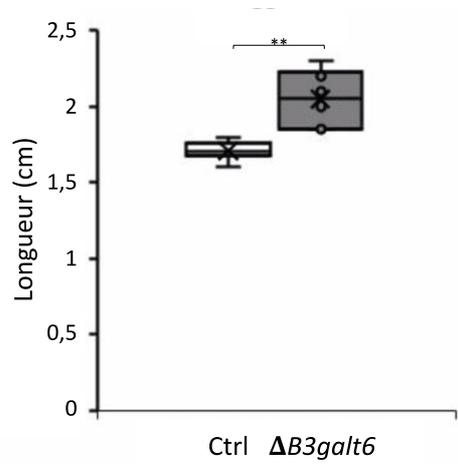


Figure 46. Mesure de l'extensibilité des pseudo-tissus de cellules ATDC5 contrôles ou invalidées pour le gène *B3galt6*. Le gène *B3galt6* des cellules est ciblé par CRISPR-Cas9 ($\Delta B3galt6$) ou non (Ctrl). Les pseudo-tissus sont

détachés et maintenus en hauteur à l'aide d'une pince pendant 30 sec. L'extensibilité est mesurée à l'aide des graduations en cm d'une règle. Les données présentées sont la moyenne \pm l'écart-type de l'extensibilité de six pseudo-tissus provenant de deux expériences distinctes. ** : $p < 0,01$ (Test t de Student). Adapté de Diana *et al.*, in external peer-review.

c. Analyses morphologiques et biomécaniques par microscopie à force atomique

L'AFM est un microscope constitué d'un levier auquel est associée une sonde nanométrique qui est mise en contact avec l'échantillon à analyser. Le levier est un élément flexible dont les déflexions lors des contacts sont enregistrées. Elles permettent d'obtenir des informations de force. C'est une technique versatile intéressante à l'interface de plusieurs disciplines. Pour notre étude (en collaboration avec le Dr. Grégory Francius, UMR 7564 LCPME, Vandoeuvre-lès-Nancy, France), cette technique de microscopie a été utilisée pour comparer les morphologies et les propriétés mécaniques des pseudo-tissus afin d'analyser les défauts causés par l'inactivation de *B3galt6*.

Grâce à ce système, la morphologie de l'échantillon peut être obtenue sous la forme d'images : les différences de déflexion du levier sont interprétées comme des informations topographiques. La morphologie des pseudo-tissus est ainsi étudiée en milieu aqueux dans du PBS ou à l'air libre après dessiccation. Comme le montrent les images en 3D, il n'y a pas de différence majeure de distribution ou de morphologie cellulaire entre les pseudo-tissus contrôles ou invalidés (Figure 47A). En revanche, la sonde nanométrique relève une plus grande profondeur d'échantillon, c'est-à-dire une différence plus importante entre le point le plus haut et le point le bas, pour les pseudo-tissus contrôles (avec une valeur médiane de 7,28 μm) que pour les pseudo-tissus invalidés (4,34 μm). Après dessiccation, cette différence est moins importante mais suit la même tendance (environ 2,1 μm contre environ 1 μm) (Figure 47B). Nous observons également que la dessiccation cause un rétrécissement plus important des pseudo-tissus invalidés (5,06 fois) que contrôles (3,33 fois) (Figure 47C).

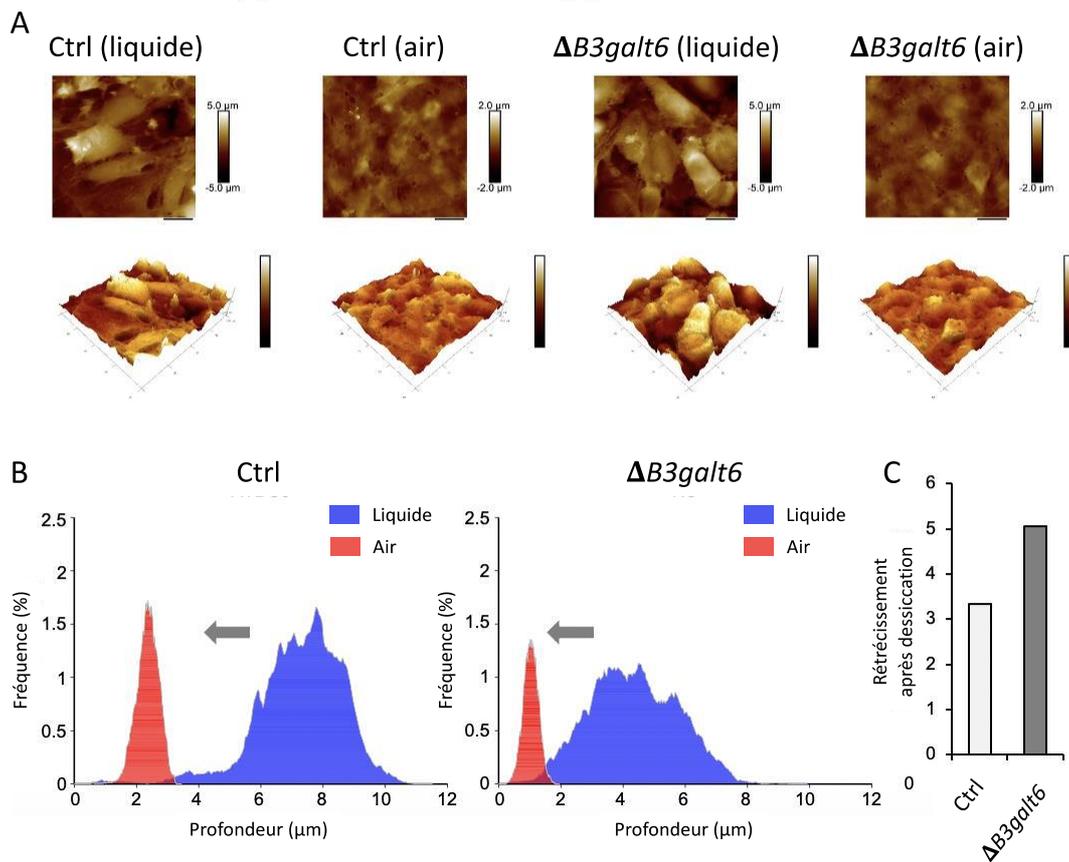


Figure 47. Analyses topographiques de pseudo-tissus ATDC5 contrôles ou invalidés pour le gène *B3galt6*. Le gène *B3galt6* des cellules est ciblé par CRISPR-Cas9 ($\Delta B3galt6$) ou non (Ctrl). **(A)** Topographie de surface et représentation tridimensionnelle des pseudo-tissus dans du PBS (liquide) ou après dessiccation (air). Les barres noires sous les images sont une échelle correspondant à une distance de 20 μm . Les barres colorées sur les côtés droits des images indiquent la hauteur mesurée en μm . **(B)** Effet de la dessiccation sur les pseudo-tissus. La distribution de mesures de profondeurs dans du PBS (bleu) ou dans l'air (rouge) représente la moyenne des mesures de quatre à six localisations pour chaque pseudo-tissu. **(C)** Rétrécissement des pseudo-tissus après dessiccation. Les valeurs sont obtenues en divisant l'aire sous la courbe des mesures de profondeurs en milieu liquide par celle des mesures de profondeurs dans l'air. Adapté de Diana *et al.*, *in external peer-review*.

L'AFM donne également accès aux propriétés biomécaniques des pseudo-tissus. Par la technique de nanoindentation, les différences de déflexion du levier sont mesurées sur une zone choisie (90 μm x 90 μm) puis sont interprétées comme des informations de hauteur et de rigidité. Les pseudo-tissus contrôles présentent de plus grandes valeurs de hauteurs (0 - 4,5 μm) que les pseudo-tissus invalidés (0 - 3 μm). Cette hétérogénéité est mise en évidence par les cartes de mesure (Figures 49A et 49B). De façon concomitante, des valeurs d'élasticité des échantillons ont été obtenues. Celles-ci sont traduites par le module de Young (E), une

valeur en Pa qui constitue une des constantes d'élasticité utilisées pour décrire la déformation d'un matériau sous une pression uniaxiale. Le test permet d'obtenir une courbe de compression conventionnelle, décrite en Figure 48.

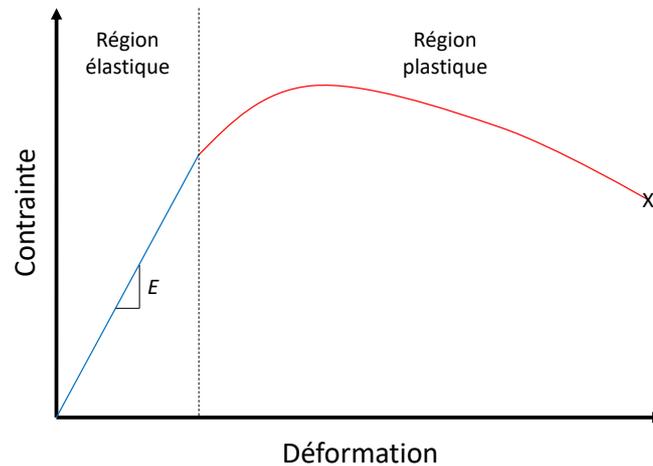


Figure 48. Courbe conventionnelle de déformation d'un matériau selon l'intensité de la contrainte appliquée.

Lors du début du test, la déformation de l'objet en fonction de la contrainte appliquée suit une fonction linéaire (bleu) dont la pente est égale au module de Young E . Cette région est dite élastique car l'objet retrouve sa forme initiale lorsque la contrainte n'est plus appliquée. Lorsque la contrainte entraîne une déformation dans la région plastique, la limite d'élasticité est dépassée : la courbe de déformation n'est plus linéaire (rouge) et la déformation de l'objet devient irréversible. Une rupture de l'objet est observée à la fin de la courbe (X).

La déformation élastique, c'est-à-dire la déformation que subit l'échantillon sans dépasser sa limite d'élasticité pour entrer dans la région de déformation plastique, est décrite sous la forme d'une courbe linéaire dont la pente est égale au module de Young. Si la valeur du module de Young est basse, le matériau est plus souple que rigide, et inversement. Le module de Young moyen obtenu pour les pseudo-tissus invalidés (13,91 kPa) est significativement plus élevé ($p < 0,001$) que celui obtenu pour les pseudo-tissus contrôles (10,89 kPa). L'hétérogénéité de la MEC des pseudo-tissus cause une disparité des modules de Young. Les zones les plus rigides sont aussi les plus hautes, ce qui correspondrait aux noyaux cellulaires. Les zones les plus souples sont les plus basses, correspondant notamment à la MEC synthétisée lors de la chondrogenèse. L'inactivation du gène *B3galt6* entraîne donc une augmentation de la rigidité des échantillons en modifiant les propriétés viscoélastiques des

MEC (Figure 49C). Ces changements biomécaniques sont certainement dus en partie à une moins bonne rétention de l'eau par une matrice possédant moins de chaînes de GAG.

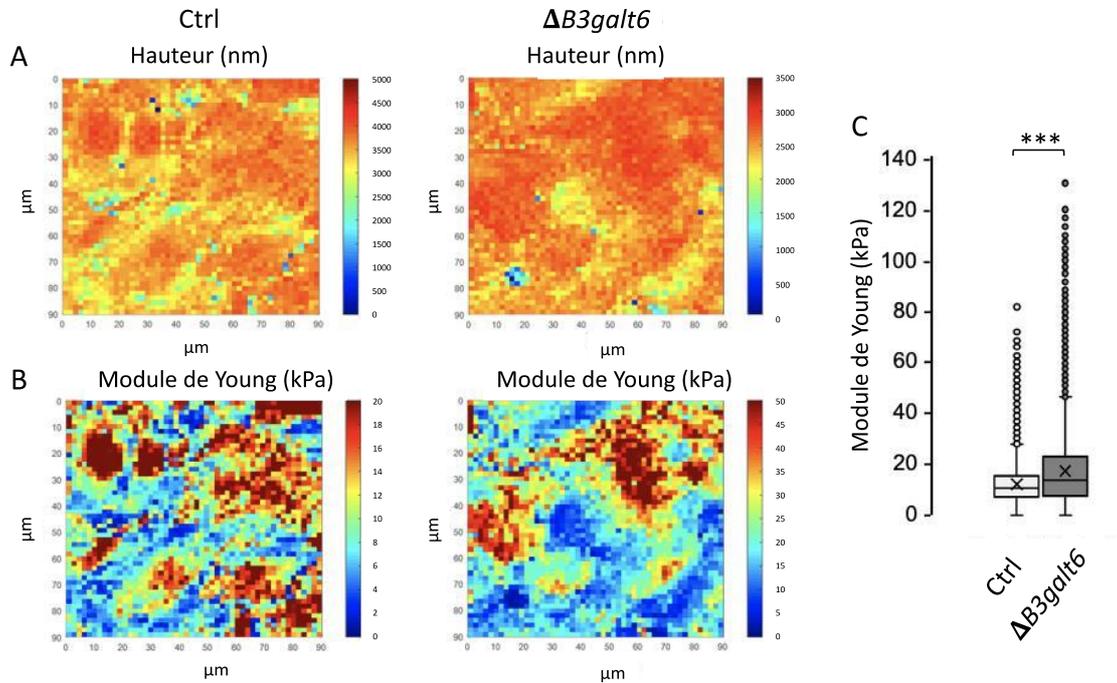


Figure 49. Analyses de la hauteur et de la rigidité des pseudo-tissus obtenus ATDC5. Le gène *B3galt6* des cellules est ciblé par CRISPR-Cas9 ($\Delta B3galt6$) ou non (Ctrl). **(A)** Hauteur en nm des pseudo-tissus mesurée en milieu liquide sur une surface de 90 μm x 90 μm . **(B)** Modules de Young en kPa des pseudo-tissus mesurés en milieu liquide sur une surface de 90 μm x 90 μm . **(C)** Distribution des modules de Young obtenus sur au moins six localisations différentes sur deux pseudo-tissus. *** : $p < 0,001$ (Test *t* de Student). Adapté de Diana *et al.*, in *external peer-review*.

Discussion

Les PG sont des macromolécules présentes au niveau des membranes cellulaires et des MEC des TC. Ils sont composés d'une partie protéique sur laquelle sont greffées une à plusieurs chaînes de GAG. La β 3GalT6 est une GT impliquée dans les étapes précoces de la biosynthèse de ces chaînes, au niveau de l'amorce tétrasaccharidique. Quarante-sept mutations du gène *B3GALT6* codant pour la β 3GalT6 sont aujourd'hui répertoriées parmi lesquelles une majorité de mutations par substitution ou délétion avec décalage du cadre de lecture, présentes à l'état homozygote ou hétérozygote composite chez les patients. Elles sont à l'origine de maladies génétiques rares des TC telles que le SEDsp (17), la SEMD-JL1 (18) et le syndrome d'Al-Gazali (29), toutes caractérisées par des atteintes musculosquelettiques multiples et sévères. Ces maladies font partie des *linkeropathies* car l'enzyme mutée catalyse l'ajout d'un glucide de l'amorce tétrasaccharidique (*tetrasaccharidic linker*) des chaînes de GAG. Les manifestations cliniques de ces affections sont de sévérité variable et ne sont pas nécessairement toutes présentes chez les patients. D'autres *linkeropathies* présentent des symptômes proches, comme le syndrome de Larsen-like de type *B3GAT3* (290) et la dysplasie de Desbuquois de type 2, due à des mutations de la XylT-I (221). De façon similaire, les symptômes d'autres ostéochondrodysplasies sont très proches de ceux dus à des mutations de *B3GALT6*, comme le SEDk, le SEDsp-*B4GALT7*, le SEDsp-*SLC39A13* ou encore le SEDmc (9). Ces chevauchements entre les symptômes sont source d'errance diagnostique. A l'échelle moléculaire, le déficit en chaînes de GAG synthétisées est plus ou moins marqué en fonction du patient considéré. L'impact des mutations de *B3GALT6* sur la fonction de la β 3GalT6 et sur la synthèse des chaînes de GAG reste donc mal défini. Mon travail vise comprendre la séquence des événements moléculaires conduisant à ces tableaux cliniques variés. Pour ce faire, nous avons réalisé une étude à différentes échelles. La première étape a été de déterminer l'impact de différents variants pathogènes sur l'activité enzymatique de la β 3GalT6.

Nous avons souhaité déterminer les paramètres cinétiques de l'enzyme sauvage et de les comparer à ceux de trois formes de β 3GalT6 mutées : Tyr182Cys, Asp207His et Gly217Ser. Ces mutations sont situées dans le domaine catalytique et ont été identifiées chez des patients à l'état homozygote (17, 50), ce qui nous permet d'étudier l'impact d'une mutation isolée. Nous avons tout d'abord utilisé un modèle cellulaire mis en place au cours d'une thèse précédente

dans notre équipe (279). Il s'agit d'une lignée cellulaire HeLa invalidée pour le gène *B3GALT6* par la technique de CRISPR-Cas9. Les cellules obtenues, nommées H7, présentent une perte de l'expression et de l'activité enzymatique de la β 3GalT6, ce qui en fait un modèle de choix pour surexprimer la β 3GalT6 sauvage ou mutée par transfection transitoire. Ce modèle a été développé pour plusieurs raisons. Tout d'abord, des tentatives d'inactivation du gène *B3GALT6* de fibroblastes au sein de l'équipe ont été infructueuses par le passé. De plus, les fibroblastes sont des cellules de lignée primaire, tandis que les cellules HeLa sont immortalisées et donc d'utilisation plus commode. Enfin, si la lignée *Chinese Hamster Ovary* (CHO) a été utilisée pour invalider une large gamme d'enzymes de la biosynthèse des chaînes de GAG dont la β 3GalT6 selon une approche similaire (291), l'origine humaine des cellules HeLa a été privilégiée.

Cette partie de mon travail a été réalisée durant ma première année de thèse. Le substrat de synthèse utilisé est le disaccharide Gal-Xyl(2P)-OMN et le produit de la réaction attendu est le Gal-Gal-Xyl(2P)-OMN, deux composés phosphorylés sur le C2 du Xyl. Le phosphate est en effet un élément important à la reconnaissance du substrat accepteur dans le cadre de la réaction enzymatique catalysée par la β 3GalT6 (70). Nous n'avons pas réussi à mesurer l'activité enzymatique de la β 3GalT6 de façon reproductible dans des lysats cellulaires malgré de nombreuses optimisations comprenant entre autres la quantité de protéines totales utilisée, la composition du tampon de réaction, la concentration en substrat ajoutée ou encore la composition de la phase mobile de la CLHP. En revanche, au cours de nos optimisations du pH du tampon de réaction, nous avons mis en évidence une déphosphorylation du substrat et du produit de la réaction. Ces déphosphorylations varient en intensité en fonction du pH et suggèrent que le substrat est plus rapidement déphosphorylé que pris en charge par la β 3GalT6 dans nos conditions enzymatiques. En incubant le substrat et le produit de la réaction en absence de lysat, la déphosphorylation n'est pas observée suggérant que des enzymes contenues dans les homogénats cellulaires sont responsables de la déphosphorylation des deux composés. Pour limiter la déphosphorylation observée, nous avons réalisé nos incubations en présence d'ATP, le substrat donneur des kinases, ce qui a permis de maintenir l'état phosphorylé de nos composés. Néanmoins, l'ajout d'ATP n'a pas permis de détecter une activité enzymatique de la β 3GalT6. Il pourrait y avoir compétition entre β 3GalT6 et phosphatases en faveur de ces dernières.

Le pH optimal de déphosphorylation de chacun des composés est différent : pH 7,5 pour le Gal-Xyl(2P)-OMN et pH 6,5 pour le Gal-Gal-Xyl(2P)-OMN, suggérant l'implication de deux phosphatases distinctes. La formation et la modification de l'amorce tétrasaccharidique fait intervenir plusieurs enzymes (GT, ST, kinase, phosphatase) mais cette séquence n'est pas totalement définie. Nous savons que la phosphorylation du Xyl est importante pour l'activité de la β 3GalT6. La kinase responsable de cette modification post-traductionnelle est la FAM20B (69), et la présence du phosphate empêche l'ajout du premier Gal sur l'amorce par la β 4GalT7 (72). La PXYLP1 est la phosphatase qui catalyse la déphosphorylation du Xyl (73), vraisemblablement après l'ajout du GlcA par la GlcAT-I (71). L'activité optimale de la PXYLP1 se situe à un pH de 5,8 (73) qui est plus acide que les pH optimaux de déphosphorylation que nous avons trouvé. Des investigations supplémentaires sont nécessaires pour déterminer si la déphosphorylation du Gal-Gal-Xyl(2P)-OMN est bien due à l'activité de la PXYLP1, en réalisant notamment des expériences de transfection avec de petits ARN interférents afin d'empêcher son expression. Il est possible que ces déphosphorylations soient liées à un mécanisme de régulation de la formation des chaînes de GAG encore inconnu. Il serait intéressant d'approfondir les mécanismes impliqués dans la déphosphorylation de nos composés de synthèse dans le but d'apporter des informations plus précises quant à la séquence des événements de la synthèse de l'amorce tétrasaccharidique.

Compte tenu des limites de l'utilisation de lysats cellulaires pour la mesure de l'activité enzymatique de la β 3GalT6, nous avons décidé de poursuivre nos investigations sur la β 3GalT6 surexprimée et purifiée à partir de bactéries. Ainsi, nous nous affranchissons des protéines présentes dans les lysats cellulaires, notamment les phosphatases. Afin d'étudier les paramètres cinétiques de l'enzyme sauvage et mutée, la production dans la bactérie et la purification d'une forme tronquée soluble de l'enzyme, comprenant son domaine catalytique, a été réalisée à la manière d'autres études de GT (262, 265, 292–294). Nous avons tout d'abord investigué quelle souche bactérienne utiliser pour surexprimer l'enzyme en grandes quantités. Nous avons déterminé que la souche *E. coli* RosettaTM2(DE3) est adéquate, la souche *E. coli* BL21(DE3) ne surexprimant pas l'enzyme dans nos conditions. Deux constructions ont été testées : la 6His-MBP- β 3GalT6 Δ N_{ter}29 (MBP fusionnée à l'extrémité N_{ter} de l'enzyme consistant en son segment *stem* et son domaine catalytique) et la 6His-MBP- β 3GalT6 Δ N_{ter}50 (protéine de fusion similaire avec un segment *stem* largement tronqué). La MBP améliore la solubilité de la protéine de fusion et facilite sa purification (279). Les deux

constructions permettent la production de protéines solubles et purifiées en quantités suffisantes pour nos expériences, environ 2 mg de protéine pour 200 mL de milieu de culture. Nous n'avons pas réussi à déterminer de paramètres cinétiques pour la construction 6His-MBP- β 3GalT6 Δ N_{ter}50 qui ne présente pas une cinétique michaelienne. En revanche, la 6His-MBP- β 3GalT6 Δ N_{ter}29 en possède une pour chacun des substrats. Le K_M de la 6His-MBP- β 3GalT6 Δ N_{ter}29 envers le Gal-Xyl(2P)-OMN est proche de celui qui a été brièvement décrit par une autre équipe : nous obtenons une valeur de $0,38 \pm 0,07$ mM lorsque Wen *et al.* (70) obtenaient une valeur de $0,25 \pm 0,03$ mM avec une enzyme purifiée et un substrat similaire, respectivement la MBP- β 3GalT6 Δ N_{ter}30 et le Gal-Xyl(2P)-O-benzyle.

La seule différence entre nos deux constructions est la longueur du segment *stem*. Le segment *stem* est déstructuré dans les GT de type II qui présentent une forte similarité de structure générale (295). Si les fonctions du domaine catalytique et du segment transmembranaire sont bien définies (activité enzymatique et ancrage membranaire/localisation subcellulaire respectivement), plusieurs hypothèses ont été émises concernant les rôles du segment *stem* des GT. Ainsi, il pourrait influencer la localisation de l'enzyme (296), son oligomérisation (297), sa flexibilité (295), son clivage (298) ou encore sa spécificité de substrat (299). Sa suppression dans la construction 6His-MBP- β 3GalT6 Δ N_{ter}50 est la cause d'une moins bonne activité enzymatique. Il est possible que le segment *stem* crée une distance entre le site actif et la MBP indispensable à l'activité et/ou qu'il joue un rôle spécifique dans l'interaction avec les substrats. Mettre en place un criblage des conditions de purification pour considérer chaque variable pourrait potentiellement permettre d'obtenir une enzyme soluble après clivage de la MBP. Une telle enzyme serait utile pour savoir si la β 3GalT6 Δ N_{ter}50 gagne en affinité pour ses substrats sans la MBP, et donc si le segment *stem* améliore l'affinité ou non.

Nos résultats ont conduit à utiliser la 6His-MBP- β 3GalT6 Δ N_{ter}29 pour générer par mutagenèse dirigée les mutations Tyr182Cys, Asp207His et Gly217Ser et étudier leur impact sur l'activité enzymatique. Dans les trois cas, nous constatons une perte totale de l'activité enzymatique *in vitro*. En conséquence, aucun paramètre cinétique n'a pu être déterminé. En accord, l'alignement de séquence des β 3GalT6 de différentes espèces montre que les résidus mutés sont tous fortement conservés, impliquant qu'ils jouent un rôle fonctionnel important, au niveau catalytique et/ou structural (Diana *et al.*, *in external peer-review*).

Faisant suite à nos résultats d'activité enzymatique, nous nous sommes intéressés à la structure de l'enzyme. La β 3GalT6 n'a, à ce jour, pas été cristallisée et nous manquons donc d'informations à ce niveau. Nous avons envisagé que la perte d'activité des mutants puisse provenir de défauts structuraux. En étudiant nos enzymes purifiées par DLS, l'indice de polydispersité est toujours d'environ 0,1, ce qui fait que nous pouvons considérer nos produits de purification comme des populations monodisperses. Nous avons observé un diamètre hydrodynamique large, de 60 ± 20 nm en moyenne. Il est possible que la flexibilité du segment *stem* exerce une influence sur la distribution des tailles en solution, car le produit de purification n'est pas constitué d'un seul mais de deux domaines globulaires (la MBP et le domaine catalytique de la β 3GalT6) reliés par un segment polypeptidique non structuré long de vingt-six acides aminés. De plus, de nombreuses GT comme la GlcAT-I peuvent s'homodimériser (258, 265). La chromatographie d'exclusion couplée à la mesure de la diffusion de la lumière sous multiples angles pourrait déterminer si nous sommes en présence de plusieurs populations en solution, monomérique et/ou multimérique (300). Une autre technique pour déterminer la masse moléculaire de nos protéines serait la photométrie de masse, qui a l'avantage de n'utiliser que très peu de matériel biologique (301). Nous avons ensuite mené des expériences de DC qui démontrent que les enzymes mutées présentent une signature en hélice α identique à celle de l'enzyme sauvage. Ce résultat indique que nos protéines sont correctement repliées. En conséquence, nous en concluons que les pertes d'activité enzymatique ne sont pas dues à des modifications majeures de la structure secondaire à la suite de la mutation de l'enzyme.

Dans l'intention d'explorer les raisons de la perte de fonction due aux mutations Tyr182Cys, Asp207His et Gly217Ser, une collaboration avec le Dr. Yasaman Karami et le Dr. Hamed Khakzad (UMR 7503 LORIA, Vandoeuvre-lès-Nancy, France) a permis d'obtenir, par une approche de simulation de dynamique moléculaire, des modèles structuraux de la β 3GalT6 sauvage et mutée. Les intensités des liaisons hydrogène et des ponts salins intramoléculaires qui stabilisent la structure secondaire ont été comparées entre les modèles générés. Les résultats suggèrent que ces liaisons qui impliquent des motifs conservés des GT sont modifiées lorsque la β 3GalT6 est mutée.

Le motif xED (Ser²⁴⁰-Glu²⁴¹-Asp²⁴² dans la séquence protéique de la β 3GalT6) contient un résidu indispensable à la catalyse, prédit comme étant l'Asp²⁴² chez la β 3GalT6. Ce résidu formerait un pont salin avec la Lys¹⁹³ qui se situe dans une longue boucle Gly¹⁸⁸-Pro²¹¹ flexible

avec de fortes fluctuations dans son segment Arg¹⁹¹-Asp²⁰⁷. Les prédictions suggèrent que les mutations Tyr182Cys, Asp207His et Gly217Ser empêcheraient la formation de la liaison Lys¹⁹³-Asp²⁴² et induiraient une baisse du taux de fluctuations du segment Arg¹⁹¹-Asp²⁰⁷. Ainsi, la perte du pont salin et la diminution de la flexibilité de la boucle Gly¹⁸⁸-Pro²¹¹ peuvent être reliées à la perte de fonction. En effet, les GT sont des enzymes dynamiques qui subissent une réorganisation du site catalytique après la fixation du substrat donneur pour accueillir le substrat accepteur. Ce mouvement est généralement permis par la flexibilité d'une longue boucle (274, 276, 302, 303). A titre d'exemple, la β 4GalT7, seconde enzyme de la synthèse de l'amorce tétrasaccharidique des chaînes de GAG et enzyme de type GT-A comme la β 3GalT6, passe d'une conformation ouverte à une conformation fermée grâce au mouvement de sa boucle His²⁵⁹-Gly²⁸³, ce qui induit la réorganisation de son motif xED (239).

La boucle Gly¹⁸⁸-Pro²¹¹ est un élément intéressant du domaine catalytique. La B3GNT2, autre GT de la famille GT31 décrite en introduction de ce manuscrit, contient une boucle similaire : Phe²⁷⁵-Gly³⁰⁷. Elle est un élément déterminant du site catalytique car certains de ses résidus forment des liaisons avec les substrats. Deux de ces résidus sont la Lys²⁸⁸ et la Tyr²⁸⁹ qui forment des liaisons hydrogène avec les phosphates de la partie UDP du substrat donneur. Ces résidus sont conservés dans la séquence protéique de la β 3GalT6, que l'on retrouve sous la forme de l'Arg¹⁹⁷ et du Trp¹⁹⁸ (274). Leoni *et al.* (36) ont identifié la mutation de β 3GalT6 p.(Arg197Cys) chez un patient et ont utilisé la structure résolue de la B3GNT2 pour créer un modèle de la β 3GalT6, mettant en lumière la même orientation dans l'espace des chaînes latérales du résidu Lys²⁸⁸ de la B3GNT2 et du résidu Arg¹⁹⁷ de la β 3GalT6. Nous émettons l'hypothèse que la boucle Gly¹⁸⁸-Pro²¹¹ joue un rôle similaire à celui de la boucle Phe²⁷⁵-Gly³⁰⁷. Ainsi, la perte de la liaison Lys¹⁹³-Asp²⁴² qui relie la boucle au cœur catalytique et la diminution de sa flexibilité pour les mutants Tyr182Cys, Asp207His et Gly217Ser affecteraient la reconnaissance de substrat et/ou la catalyse.

Le motif DxD (Asp¹⁵⁶-Asp¹⁵⁷-Asp¹⁵⁸ dans la séquence protéique de la β 3GalT6) est un élément indispensable à la coordination du cation divalent et du substrat donneur de la réaction enzymatique. La dynamique moléculaire prédit que le résidu Lys¹⁵⁴ établit une liaison hydrogène Lys¹⁵⁴-Asp¹⁵⁶ avec un Asp du motif DxD. Son intensité diminue fortement chez le mutant Tyr182Cys et diminue également chez le mutant Asp207His. L'Asp¹⁵⁶ établit aussi une liaison hydrogène avec la Gly²¹⁶ de la boucle riche en Gly (Gly²¹⁵-Gly²¹⁶-Gly²¹⁷) de

reconnaissance du substrat accepteur (247). Cette liaison n'est pas retrouvée dans le mutant Tyr182Cys et diminue en intensité dans le mutant Asp207His.

Nous proposons que les modifications des intensités de liaisons intramoléculaires qui impliquent les motifs conservés de la β 3GalT6 contribuent à la perte de fonction enzymatique que nous observons dans nos enzymes mutées. Nos données reposent sur une approche de modélisation moléculaire, avec comme point de départ le logiciel AlphaFold. Bien que le modèle proposé soit de bonne qualité selon les standards du logiciel (268), des données expérimentales sont indispensables pour confirmer nos résultats de dynamique moléculaire. Il sera également intéressant d'étudier la liaison aux substrats par des approches de docking afin de définir les rôles exacts des acides aminés dans la reconnaissance des substrats et dans la catalyse.

Par ailleurs, en dépit d'une perte de fonction de l'enzyme, il a été observé que les fibroblastes de patients porteurs de mutations de *B3GALT6* conservent la capacité de synthétiser des chaînes de GAG de façon résiduelle (17, 50). Dans le cadre de ce travail, nous l'avons confirmé sur la décorine sécrétée par des fibroblastes de patients obtenus auprès de nos collaborateurs (Dr. Fransiska Malfait et Dr. Delfien Syx, *Center of Medical Genetics*, Université de Gand, Gand, Belgique) (Diana *et al.*, *in external peer-review*). Ces constatations peuvent sembler contradictoires avec la perte de fonction de la β 3GalT6 qui, en théorie, impliquerait l'absence de formation de chaînes de GAG de types CS, DS et HS. Cependant, une absence totale de GAG serait *a priori* incompatible avec la survie de l'individu porteur de la ou des mutations. En effet, les chaînes de GAG jouent de nombreux rôles biologiques. En fonction du PG considéré, elles établissent des interactions avec les fibres de collagène (152), les FACIT (190) et d'autres protéines matricielles comme la ténascine X (193) pour réguler la fibrillogenèse (173). Les chaînes de type HS des syndécans (304), les chaînes de type CS de la décorine (305) et l'Hep interagissent également avec des facteurs de croissance (306, 307) pour réguler de multiples voies de signalisation (53). Dans leur modèle CHO invalidé pour le gène *B3galt6*, Chen *et al.* (291) ont également mis en évidence une baisse (et non suppression) de la synthèse des chaînes de GAG de type CS, DS, HS et Hep, renforçant l'hypothèse selon laquelle une perte de la fonction de la β 3GalT6 n'amène pas à l'arrêt total de la synthèse. Au contraire, ces chercheurs ont observé que l'inactivation des gènes codant pour les autres enzymes de la synthèse de l'amorce tétrasaccharidique (*Xylt2*, *B4galt7* et *B3gat1*) produit une

perte totale des chaînes de GAG. La β 3GalT6 semble être une enzyme au cœur d'un mécanisme particulier que nous avons souhaité décoder.

Nous avons mis à profit notre modèle cellulaire HeLa invalidé pour le gène *B3GALT6* qui, à défaut de nous renseigner sur l'activité de la β 3GalT6, nous est utile pour explorer la synthèse des chaînes de GAG. En utilisant un substrat accepteur exogène, le 4-MOX, nous observons une synthèse résiduelle de chaînes par les cellules invalidées, en concordance avec les résultats obtenus dans les fibroblastes de patients et les CHO invalidées. La surexpression de l'enzyme sauvage permet d'augmenter le taux de synthèse alors que la surexpression des enzymes mutées ne le permet pas, avec une synthèse comparable à la synthèse résiduelle sans surexpression. Nous observons même une faible diminution statistiquement significative de la synthèse des chaînes de GAG en absence de substrat exogène lors de la surexpression de la β 3GalT6-D207H et de la β 3GalT6-G217S. Nous avons par la suite confirmé cette observation par immunoblot, où un PG co-exprimé avec la β 3GalT6, la décorine, est glycosylé de la même façon : on note une forte glycosylation de la protéine-core en présence de la β 3GalT6 sauvage mais une glycosylation fortement réduite en absence de β 3GalT6 ou en présence de β 3GalT6 mutée. Par mesure de densitométrie, nous observons que le rapport de décorine glycosylée sur décorine non-glycosylée en absence d'enzyme ou lors de la surexpression de la β 3GalT6-D207H ou de la β 3GalT6-G217S est plus faible que lors de la surexpression de la β 3GalT6-Y182C. Il s'agit-là d'une différence entre nos deux substrats : dans notre modèle, la surexpression de la β 3GalT6-Y182C semble permettre une meilleure glycosylation de la décorine que du 4-MOX. Il est possible que la β 3GalT6-Y182C ait une meilleure affinité pour la protéine-core de la décorine que pour le 4-MOX. Globalement, nous pouvons affirmer que les mutations Tyr182Cys, Asp207His et Gly217Ser diminuent dramatiquement la synthèse des chaînes de GAG mais que la synthèse n'est pas totalement abolie.

Il faut également noter que la décorine co-exprimée avec les mutants de β 3GalT6 possède des chaînes de GAG moins nombreuses mais plus longues que celles synthétisées lors de la surexpression de la β 3GalT6-WT. La modification de la longueur des chaînes de GAG lorsqu'une enzyme de la synthèse est invalidée ou mutée a déjà été observée dans d'autres travaux. La XylT-I, première GT de la synthèse de l'amorce tétrasaccharidique, est codée par le gène *XYLT1*. Des mutations de ce gène sont reliées à une *linkeropathie*, la dysplasie de Desbuquois de type 2. Les fibroblastes de patients présentent un taux de synthèse réduit des

PG avec de longues chaînes de GAG de type CS et une accumulation de PG à chaînes de GAG de type CS plus courtes (221). Ces observations sont cohérentes avec le modèle murin invalidé pour le gène *Xylt1* pour lequel une diminution du taux de synthèse et de la taille des chaînes de GAG (308) est observée. Toutefois, ces observations diffèrent des nôtres qui mettent en évidence une augmentation de la longueur des chaînes de GAG produites. La dysplasie ostéodysplasique microcéphalique type Saul-Wilson est causée par des mutations de *COG4*, gène codant pour une sous-unité du complexe oligomérique du Golgi qui joue un rôle dans la morphologie et la fonction de ce compartiment cellulaire. Les sous-unités de ce complexe influencent la régulation, la compartimentalisation, le transport et l'activité de nombreuses enzymes golgiennes et de multiples substrats (309). Une mutation hétérozygote de *COG4* est à l'origine d'une faible synthèse de chaînes de GAG de la décorine par des fibroblastes de patients. Ces chaînes sont plus longues que dans les cellules contrôles, de manière similaire à ce que nous observons. Un défaut du trafic protéique dans l'appareil de Golgi serait à l'origine de ces profils de glycosylation (310). La compartimentalisation est indispensable au bon placement de toutes les protéines et tous les substrats pour l'homéostasie de l'appareil de Golgi, et donc à l'initiation, la polymérisation et la sulfatation des chaînes de GAG (311). Il est possible que les chaînes de GAG de décorine plus longues que nous observons en cas de mutation Tyr182Cys, Asp207His et Gly217Ser de la β 3GalT6 soient le résultat d'une altération du trafic de PG : les GAG seraient plus largement allongés dans l'appareil de Golgi par les enzymes d'élongation. De plus, la réponse à une sous-glycosylation des PG par l'appareil de Golgi entraîne une augmentation de l'expression de GT et ST de la synthèse des GAG dans une tentative de compensation (312), ce qui pourrait avoir un effet sur l'élongation des chaînes. Une autre observation est l'effet de certains facteurs de croissance sur une hyperélongation des chaînes de type CS : par exemple, le TGF β est impliqué dans cette élongation dans les cellules musculaires lisses vasculaires (313) en stimulant une synthèse *de novo* des enzymes de l'élongation des chaînes (314). La réponse au stress de l'appareil de Golgi et l'effet du TGF β tendent vers le même effet : l'hyperélongation des chaînes de GAG. La modification de la longueur des chaînes peut avoir des conséquences fonctionnelles. Par exemple, la longueur des chaînes influe directement sur la rapidité d'accumulation d'agrégats de protéines jouant un rôle direct dans certaines maladies neurodégénératives : la maladie d'Alzheimer (Tau) et la maladie de Parkinson (α -synucléine) (315). Lors de la progression de l'arthrose, les chaînes de GAG de type CS de l'aggrécane deviennent plus courtes, diminuant la quantité d'eau

retenue dans le cartilage articulaire qui perd en viscoélasticité. On observe alors une hyperélongation des chaînes de type KS de l'aggrécane, vraisemblablement pour compenser la diminution de la quantité de CS (316).

Pour résumer, il existe une synthèse résiduelle par les cellules invalidées pour le gène *B3GALT6*. Cette synthèse résiduelle n'est pas augmentée de façon significative par la surexpression des enzymes mutées, en accord avec la perte totale de fonction des enzymes mutées recombinantes que nous avons purifiées. Il semble donc que les cellules soient capables d'initier la synthèse des chaînes de GAG en absence de β 3GalT6 fonctionnelle : les chaînes de GAG synthétisées sont moins nombreuses et plus longues, raison probable pour laquelle des anomalies dans la disposition et le diamètre des fibres de collagène sont relevées par microscopie électronique à transmission dans des biopsies de patients (17).

Nous avons poursuivi notre investigation des mécanismes de synthèse des chaînes de GAG en souhaitant déterminer par quel moyen la synthèse peut être partiellement compensée lorsque la β 3GalT6 est mutée et perd totalement sa fonction enzymatique. La voie de biosynthèse conventionnelle des chaînes de GAG de types CS, DS, HS et Hep est initiée par une amorce tétrasaccharidique entre la protéine-*core* et le reste de la chaîne. Celle-ci est de composition GlcA- β 1,3-Gal- β 1,3-Gal- β 1,4-Xyl- β 1-O. Récemment, une amorce trisaccharidique de type GlcA- β 1,3-Gal- β 1,4-Xyl- β 1-O a été détectée en faibles quantités dans de l'extrait total de protéines de poisson-zèbre dont le gène *B3GALT6* a été invalidé (317) et dans l'urine de patients porteurs de mutations de *B3GALT6* (318). La formation d'une telle amorce non canonique serait le mécanisme à l'origine de l'initiation de la synthèse des chaînes de GAG en absence d'activité enzymatique de la β 3GalT6.

Dans ce contexte, nous avons exploré le rôle potentiel de la GlcAT-I (quatrième enzyme de la synthèse de l'amorce tétrasaccharidique faisant directement suite à la β 3GalT6) dans la formation d'une amorce non canonique, grâce à une collaboration avec le Dr. Rebekka Wild et Marie Bourgeais (UMR 5075 IBS, Grenoble, France). Nous avons démontré *in vitro* que la GlcAT-I est capable de catalyser l'ajout d'un GlcA sur le substrat disaccharidique pour former une amorce trisaccharidique. Si on compare cette activité par rapport à celle pour son substrat canonique trisaccharidique, elle est environ dix fois moins importante. Ce serait donc par cette activité que, même en cas de perte de fonction de la β 3GalT6, les fibroblastes de patients et les cellules invalidées conserveraient la capacité de synthétiser des chaînes de GAG. La différence de composition de l'amorce, la faible quantité de GAG synthétisée et

l'augmentation de la longueur permettraient la survie de l'individu s'accompagnant néanmoins d'atteintes musculosquelettiques et tissulaires sévères et variées. Cette synthèse alternative pourrait être la raison de l'existence d'un nombre relativement important de patients porteurs de mutations de *B3GALT6*.

Néanmoins, l'intensité de la synthèse des chaînes de GAG de la décorine et des chaînes de type HS par les cellules des patients diffère en fonction de l'individu (17, 50). Pour expliquer ces différences, nous suggérons de nous intéresser à la formation de complexes protéiques que peuvent former les GT. Il a été démontré que ces interactions sont nécessaires à l'activité enzymatique de certaines GT (257). Les GT de la synthèse des GAG ne font pas exception, dont voici quelques exemples. L'homodimérisation de la GlcAT-I de l'amorce tétrasaccharidique est observée par cristallographie. Des résidus du site actif (265) et un résidu du segment *stem* (258) sont impliqués dans cette association. Ainsi, un petit segment du site actif d'un monomère est rapproché du site actif de l'autre monomère pour se lier au substrat lors de la catalyse (71). L'hétérodimérisation de l'épimérase et la 2-O-sulfotransférase des chaînes de GAG de type HS est nécessaire pour que l'épimérase soit correctement localisée dans l'appareil de Golgi et pour son activité enzymatique (319). Selon un schéma similaire, l'hétérodimérisation des EXT1 et EXT2 (262), deux GT de l'élongation des chaînes de GAG de type HS, se produit *in vivo* (259). Enfin, de multiples enzymes de l'élongation des chaînes de GAG de type CS interagissent entre elles afin d'augmenter leur efficacité catalytique *in vitro* (320, 321). Après analyse des interactions des GT de synthèse des chaînes de GAG, un modèle hypothétique nommé GAGosome a été proposé par Esko et Selleck (322). Ces interactions seraient responsables de la fine régulation de la synthèse des chaînes de GAG, à l'image de l'interaction qui a été décrite entre la GlcAT-I et la PXYLP1, qui permet l'augmentation du taux de déphosphorylation du Xyl de l'amorce après le transfert du GlcA (73). Les interactions entre la β 3GalT6 et de potentiels partenaires n'ont pas été identifiées à ce jour. Des altérations dans l'interaction entre la β 3GalT6 et des partenaires potentiels, tels que la GlcAT-I, pourraient expliquer les différences de taux de synthèse des GAG observés chez les patients porteurs de différentes mutations. En effet, les trois mutations Tyr182Cys, Asp207His et Gly217Ser pourraient compromettre l'intégrité du GAGosome. La synthèse résiduelle *via* une amorce trisaccharidique pourrait en être différemment impactée en fonction des mutations de β 3GalT6, comme nous l'observons pour la glycosylation de la décorine par la β 3GalT6-Y182C dans les cellules H7 invalidées. L'exploration de ces complexes hypothétiques établis par la

β 3GalT6, et plus largement la caractérisation de l'entièreté du GAGosome, sont indispensables pour étayer notre hypothèse. Notre équipe étudie actuellement ces interactions.

En résumé, chez les patients, la perte de fonction de la β 3GalT6 est compensée par une prise en charge partielle de la synthèse par la GlcAT-I (Diana *et al.*, *in external peer-review*) et par la formation de chaînes de GAG à partir d'une amorce trisaccharidique et non tétrasaccharidique (318). Le faible taux de synthèse à partir de cette amorce et la modification de la longueur des chaînes seraient à l'origine d'une désorganisation matricielle qui est explorée dans la suite de ce travail.

Pour élargir nos investigations par des approches non ciblées dites « omiques », notre équipe a établi une collaboration avec le Dr. Sébastien Hergalant (UMRS 1256 NGERE, Vandoeuvre-lès-Nancy, France) afin d'étudier le transcriptome de fibroblastes de patients porteurs de mutations de *B3GALT6*. Parmi les gènes exprimés différemment par rapport aux cellules contrôles, les gènes *LOXL1*, *LOXL2*, *LOX* et *PLOD1* sont retrouvés (Diana *et al.*, *in external peer-review*). Il s'agit de gènes codant pour trois LOX et la LH1, toutes responsables de modifications post-traductionnelles des collagènes fibrillaires critiques pour la fibrillogenèse. En accord, l'analyse des biopsies de peau par microscopie électronique à transmission montre une irrégularité dans la disposition, la forme et la répartition des fibres de collagène (17). Nous avons donc cherché à étudier les conséquences d'une déficience en β 3GalT6 sur l'organisation des collagènes et des MEC. Comme décrit dans l'état de l'art, les PG et leurs chaînes glucidiques sont responsables de nombreuses fonctions physiologiques, telles que la résistance des tissus aux forces de compressions par la rétention des molécules d'eau grâce à leurs chaînes sulfatées et globalement chargées négativement (323, 324) ou l'interaction avec des médiateurs solubles (325–327). Ils participent aussi activement à la formation d'un réseau matriciel dense et ordonné lors de la fibrillogenèse (173, 328). La synthèse altérée des GAG peut avoir un impact sur chacun de ces paramètres qui sont essentiels au développement embryonnaire et post-natal et à l'homéostasie des tissus.

La lignée de tératocarcinome de souris ATDC5 a fait l'objet d'un travail de thèse au sein notre équipe (329). Ces cellules sont capables de synthétiser des composants matriciels tels que des PG et du collagène, et de former une MEC organisée sous la forme d'un pseudo-tissu après induction de la chondrogenèse (286). Nous avons invalidé l'expression du gène *B3galt6*

par édition du génome *via* le système CRISPR-Cas9 introduit par transduction virale. Cette manipulation a été un succès comme le démontrent les diminutions drastiques de glycosylation de la décorine et de l'aggrécane. Comme pour les cellules H7 et les fibroblastes de patients, la glycosylation n'est pas totalement abolie. Nous pouvons supposer que la synthèse est partiellement compensée par la GlcAT-I. Nous avons également observé pour la première fois que la glycosylation de la forme longue du Col XII est fortement impactée par l'invalidation de *B3galt6*. Le Col XII est un FACIT codé par le gène *COL12A1* et exprimé sous deux isoformes grâce à un épissage alternatif : une forme longue et une forme courte. Trois isoformes, longues ou courtes, s'associent pour former le Col XII mature. La forme longue est aussi un PG car elle est modifiée par au moins une chaîne de GAG de type CS au niveau de son domaine NC3 (128, 129).

Ayant confirmé la pertinence de ce modèle cellulaire, nous l'avons soumis à plusieurs tests pour explorer les caractéristiques morphologiques et biomécaniques des pseudo-tissus dont le gène *B3galt6* a été invalidé. Nous avons constaté que l'extensibilité des pseudo-tissus est augmentée lorsqu'ils sont invalidés, indiquant des altérations des propriétés de la MEC. Cette dernière apparaît désorganisée dans nos coupes histologiques de pseudo-tissu grâce à la coloration des chaînes de GAG au bleu Alcian. En établissant une collaboration avec le Dr. Grégory Francius (UMR 7564 LCPME, Vandoeuvre-lès-Nancy, France), des analyses par AFM ont été entreprises. L'AFM permet de fournir des informations concernant la morphologie et les paramètres biophysiques d'un échantillon. Cette technique est pertinente pour analyser les caractéristiques de pseudo-tissus invalidés pour le gène *B3galt6*. Nous avons constaté que l'invalidation des cellules cause une augmentation du module de Young, autrement dit une diminution de l'élasticité du pseudo-tissu, une caractéristique de la MEC notamment reliée aux collagènes fibrillaires (330). Ceci indique que la modification de la composition du réseau matriciel due à l'invalidation de *B3galt6* a un impact majeur sur ses propriétés viscoélastiques. La déficience de la β 3GalT6 entraînerait des altérations des interactions que les PG mettent en place avec leurs multiples partenaires, ainsi qu'une moins bonne rétention des molécules d'eau. Les chaînes de GAG alternatives, dont la GlcAT-I permet la synthèse, ne pourraient compenser pleinement ces défauts.

Un résultat original de notre étude est l'impact de l'invalidation de *B3galt6* sur le Col XII. Nos études de transcriptomique révèlent que le gène *COL12A1* fait partie des gènes dont le taux d'expression est augmenté par les fibroblastes de patients SEDsp-*B3GALT6* (Diana *et al.*,

in external peer-review), mettant en lumière le rôle potentiel du Col XII dans cette pathologie. Le Col XII établit des interactions avec le Col I *via* ses domaines collagéniques (117) et avec la ténascine X grâce à son domaine non-collagénique NC3 (191). Les domaines collagéniques du Col XII sont aussi capables de s'associer avec la chaîne de GAG de la décorine, la protéine-core de la fibromoduline (190) et la COMP (182). Ces interactions influencent l'espace interfibrillaire. De façon intéressante, des mutations du gène *COL12A1* sont responsables du SEDm (9). Certaines touchent le domaine collagénique COL2 se liant aux SLRP, à la COMP et au Col I, impliquant que les interactions entre ces partenaires sont impactées. D'autres mutations sont localisées au niveau du domaine de fixation à la thrombospondine et sont reliées à des défauts de sécrétion du Col XII. Il en résulte un mauvais assemblage des MEC avec notamment une augmentation du diamètre des fibres de collagène, qui sont moins nombreuses et plus espacées les unes des autres. Selon les mutations, on observe une diminution de l'expression de la ténascine X ou de la décorine, deux partenaires protéiques connus du Col XII (331). Le Col XII joue également un rôle biomécanique, avec une augmentation de son expression face à un stress mécanique (124, 194). Les mutations de *COL12A1* conduisent à des symptômes myopathiques comme la faiblesse musculaire que nous pouvons relier à ses propriétés mécaniques. Nous démontrons ici la forte diminution de la glycosylation de la forme longue du Col XII lorsque *B3galt6* est invalidé. Les connaissances sur la fonction exacte de cette glycosylation sont limitées, mais nous pouvons émettre l'hypothèse que cette chaîne de GAG localisée sur le domaine NC3 favorise l'interaction du Col XII avec la ténascine X.

L'impact de la déficience en Col XII a été étudiée par l'établissement de modèles murins *Col12a1*^{-/-}, montrant des défauts de cicatrisation (332) et un impact sur l'élasticité du stroma cornéen (333, 334). Récemment, il a été prouvé que le Col XII influence la voie du TGFβ en le conservant dans la MEC pour réguler sa biodisponibilité. Ainsi, le Col XII régule les propriétés structurales et fonctionnelles de la cornée en régulant l'activité TGFβ (335). Les chaînes de GAG sont impliquées dans les voies de signalisation médiées par le TGFβ (336, 337). On peut donc émettre l'hypothèse selon laquelle l'absence de chaîne de GAG due à l'invalidation de *Col12a1* serait impliquée dans la dérégulation du TGFβ.

Une limite de nos travaux réside en la nature du modèle utilisé. En effet, nos pseudo-tissus récapitulent la MEC, mais les tissus atteints chez les patients sont plus complexes et variés. Les modèles murins *B3galt6*^{-/-} ne sont malheureusement pas viables. Le

développement de modèles *knock-in* reproduisant une mutation individuelle chez l'animal serait intéressant mais reste complexe à mettre en œuvre. En l'état, notre modèle récapitulant la synthèse d'une MEC cartilagineuse est donc utile pour explorer la fonction et les interactions des macromolécules impliquées dans la fibrillogenèse dans un contexte pathologique.

Ce travail met aussi en lumière l'importance de la pluridisciplinarité de notre approche. La combinaison de différentes techniques à l'interface de domaines scientifiques, comme l'AFM comme outil d'analyse biophysique, permet d'apporter un regard nouveau sur les maladies génétiques rares des TC. L'AFM a été utilisée pour analyser les propriétés mécaniques d'une molécule individuelle (338), la ténascine X, dans le contexte du SED classique-*like* (SEDcl) (339). Relier les défauts de glycosylation de certains PG ou du Col XII et leur interactome est possible selon cette méthodologie. L'utilisation de l'AFM pour comprendre la désorganisation du réseau matriciel, plus particulièrement des collagènes fibrillaires, est de plus en plus considérée dans la communauté scientifique. Les travaux récents d'une équipe de recherche de Toronto prouvent la pertinence de l'AFM pour mesurer l'élasticité d'un tissu produit par bio-ingénierie (340, 341) ou le coupler à l'analyse de coupes histologiques (342). Cette même équipe utilise aussi l'AFM pour définir les paramètres du réseau collagénique (orientation des fibres, linéarité des fibres et périodicité axiale D ou *d-band* caractéristique de l'assemblage des fibrilles en fibres) dans le contexte d'une MEC désorganisée au cours d'un SEDc (343). Toutes ces techniques peuvent être appliquées à d'autres études, comme celle-ci qui concerne le SEDsp.

Conclusions générales et perspectives de travail

Le premier axe consistait en l'étude de conséquences d'une mutation individuelle sur la fonction de la $\beta 3\text{GalT6}$. Nous avons démontré que la synthèse de chaînes de GAG est fortement réduite et que la longueur des chaînes augmente lorsque la $\beta 3\text{GalT6}$ mutée est ré-exprimée dans un modèle cellulaire invalidé pour le gène *B3GALT6*. Ces caractéristiques sont retrouvées sur les chaînes synthétisées par les fibroblastes de patients, impliquant que les mutations sont à l'origine d'une baisse de la synthèse et d'une augmentation de la longueur des chaînes. De même, nous observons une réduction et non une abolition de la synthèse dans les cellules H7 invalidées lorsque la $\beta 3\text{GalT6}$ n'est pas ré-exprimée, et ces chaînes sont également plus longues. En parallèle, nous avons mis en évidence une perte de fonction *in vitro* de la $\beta 3\text{GalT6}$ purifiée due aux mutations étudiées. Ces résultats suggèrent que la perte de fonction n'entraîne pas un arrêt de la biosynthèse *in cellulo* mais plutôt une forte diminution de celle-ci et une régulation différente de l'élongation des chaînes. Ces modifications de la quantité et de la qualité seraient dues à une voie de biosynthèse alternative. Nous avons confirmé *in vitro* que la GlcAT-I, enzyme dont le substrat canonique est le produit de la réaction de la $\beta 3\text{GalT6}$, est capable de remplacer partiellement cette dernière dans la séquence des évènements. Ainsi, en absence de $\beta 3\text{GalT6}$, la GlcAT-I permet la synthèse d'une amorce trisaccharidique qui diffère de l'amorce tétrasaccharidique traditionnelle. L'activité de la GlcAT-I en utilisant le substrat de la $\beta 3\text{GalT6}$ est dix fois moins forte qu'en utilisant son substrat canonique. Ainsi seraient formées une faible quantité de chaînes de GAG dont la composition de l'amorce change, ce qui pourrait avoir un impact sur leur longueur. Il sera nécessaire de confirmer que la composition de l'amorce est modifiée dans les cellules H7 ré-exprimant l'enzyme mutée ou non, ainsi que dans les fibroblastes de patients. Pour ce faire, une collaboration a été établie avec l'équipe du Pr. Göran Larson (*Department of Laboratory Medicine, Université de Gothembourg, Suède*) afin d'obtenir une fine analyse de la composition de l'amorce par spectrométrie de masse. Nos travaux portent donc sur des situations de perte de fonction totale de la $\beta 3\text{GalT6}$. Nous avons étudié trois mutations par substitution présentes à l'état homozygote chez des patients SEDsp et nous en répertorions neuf de ce type dans l'état de l'art. Il sera intéressant d'étudier l'impact des six autres mutations par les mêmes techniques et de déterminer si elles mènent elles aussi à une perte de fonction.

Le second axe portait sur l'**étude de conséquences d'une mutation individuelle sur la structure de la β 3GalT6**. Nos enzymes purifiées sauvage ou mutées possèdent toutes la même signature de DC et donc la même structure secondaire générale. Par DLS, la population est monodisperse mais le diamètre hydrodynamique est assez large, il sera donc nécessaire de vérifier par diffusion de la lumière multi-angle ou par photométrie de masse si nous sommes en présence de complexes multimériques et/ou de micro-agrégats. L'enzyme sauvage présente une activité enzymatique de façon consistante et reproductible, ce qui indique que la qualité du produit de purification est compatible avec nos études fonctionnelles. Une approche de dynamique moléculaire a mis en évidence chez les mutants des modifications des intensités des liaisons intramoléculaires de la β 3GalT6 qui sollicitent plusieurs motifs conservés des GT. Les fluctuations d'une longue boucle flexible du modèle de l'enzyme sauvage diminuent pour les trois modèles des enzymes mutées. Les GT présentent une forte flexibilité de leurs séquences polypeptidiques non structurées et des modifications de conformation lors de la catalyse. Nous suggérons donc que les mutations entraînent des modifications de liaisons intramoléculaires entre des domaines clés et une rigidification de l'enzyme entraînant une perte de fonction. Pour la suite de cet axe, il sera évidemment extrêmement utile d'obtenir des informations expérimentales de structure de l'enzyme pour confirmer nos analyses et hypothèses. La cristallisation de l'enzyme en complexe avec ses substrats et/ou des études de docking seront utiles pour déterminer le rôle précis des résidus. Enfin, nous savons que les GT sont des enzymes capables de former des complexes *in vivo*, et il sera intéressant de vérifier si la β 3GalT6 possède un ou plusieurs partenaires. Si cela est confirmé, nous pourrions investiguer l'hypothèse selon laquelle les mutations affectent l'association de la β 3GalT6 avec d'autres enzymes.

Le troisième axe consistait en l'**étude des conséquences de l'inactivation du gène *B3galt6* sur la morphologie et les caractéristiques biomécaniques de la MEC dans un modèle de synthèse matricielle murin**. Nous avons établi avec succès un modèle cellulaire de chondrogenèse invalidé pour le gène *B3galt6* où nous observons à nouveau une forte diminution et non une abolition de la synthèse des chaînes de GAG. Des études de transcriptomique en cours confirment que la β 3GalT6 n'y est plus exprimée. Nous relevons une baisse de l'élasticité couplée à une diminution de rugosité à la surface des pseudo-tissus

invalidés. Ils sont également moins extensibles que les pseudo-tissus contrôles. L'invalidation de *B3galt6* entraîne donc des défauts biomécaniques de la MEC que nous réussissons à analyser notamment par AFM. Il sera intéressant de continuer ces expériences pour consolider ces informations et les relier aux résultats de transcriptomique, qui révèlent des différences d'expression de gènes-clés de la fibrillogenèse. De façon novatrice, nous avons mis l'accent sur la perte de glycosylation d'un FACIT, le Col XII, qui est également un PG impliqué dans un autre sous-type de SED. Ses rôles sont multiples, de l'interaction aux SLRP, aux fibres de collagène et à la ténascine X pour le maintien de l'architecture de la MEC, à la régulation de l'activité du TGF β et à la résistance aux forces de compression. Le rôle de la chaîne de GAG du Col XII et sa contribution dans ces capacités interactives sont peu caractérisés. Comprendre ses fonctions exactes dans la physiologie des MEC et les conséquences de mutations de *B3galt6* sur ces rôles sont de nouveaux objectifs pour caractériser le SEDsp. L'AFM sur molécule isolée pourrait permettre de vérifier que l'absence de la chaîne de GAG exerce une influence sur la structure globale du Col XII. De même, des expériences de co-immunoprécipitation permettraient de déterminer avec quelle(s) protéine(s) le Col XII interagit et quels sont les changements de son interactome lorsque *B3galt6* est invalidé.

De façon globale, nous avons apporté un nouveau regard sur les maladies génétiques rares des TC dues à des mutations de *B3GALT6*. Notre travail met en lumière de nouvelles pistes de recherche en glycobiochimie, en biologie structurale et en biophysique lorsqu'il y a une perte de fonction de la β 3GalT6. Il s'agit désormais d'explorer ces champs de recherche pour continuer à définir au mieux les maladies comme le SEDsp et relier des caractéristiques moléculaires à la symptomatologie. Ceci permettra à long terme de fournir au clinicien de nouvelles méthodes d'évaluation de la maladie dans l'objectif de réduire l'errance diagnostique et d'ouvrir de nouvelles pistes thérapeutiques.

Références bibliographiques

1. Pope FM. Ehlers-Danlos syndrome. *Baillieres Clin Rheumatol*. 1991;5(2):321–349.
2. Chernogubow N. Über einen Fall von Cutis laxa (Presentation at the first meeting of Moscow Dermatologic and Venerologic Society, Nov 13, 1891). *Monatsheft Prakt Dermatol*. 1891;14(76).
3. Denko CW. Chernogubov's syndrome: a translation of the first modern case report of the Ehlers-Danlos syndrome. *J Rheumatol*. 1978;5(3):347–352.
4. Ehlers E. Cutis laxa, Neigung zu Haemorrhagien in der Haut, Lockerung mehrerer Artikulationen. *Dermatol Z*. 1901;8:173–174.
5. Danlos H-A. Un cas de cutis laxa avec tumeurs par contusion chronique des coudes et des genoux. *Bull Soc Fr Derm Syph*. 1908;19:70–72.
6. Miget A. Le syndrome d'Ehlers-Danlos. 1933.
7. Beighton P, et al. International nosology of heritable disorders of connective tissue, Berlin, 1986. *Am J Med Genet*. 1988;29(3):581–594.
8. Beighton P, et al. Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). *Am J Med Genet*. 1998;77(1):31–37.
9. Malfait F, et al. The 2017 international classification of the Ehlers-Danlos syndromes. *Am J Med Genet C Semin Med Genet*. 2017;175(1):8–26.
10. Blackburn PR, et al. Bi-allelic Alterations in AEBP1 Lead to Defective Collagen Assembly and Connective Tissue Structure Resulting in a Variant of Ehlers-Danlos Syndrome. *Am J Hum Genet*. 2018;102(4):696–705.
11. Hernández A, et al. A distinct variant of the Ehlers-Danlos syndrome. *Clin Genet*. 1979;16(5):335–339.
12. Hernández A, et al. Third case of a distinct variant of the Ehlers-Danlos Syndrome (EDS). *Clin Genet*. 1981;20(3):222–224.
13. Hernández A, et al. Ehlers-Danlos features with progeroid facies and mild mental retardation: Further delineation of the syndrome. *Clin Genet*. 1986;30(6):456–461.
14. Kresse H, et al. Glycosaminoglycan-free small proteoglycan core protein is secreted by fibroblasts from a patient with a syndrome resembling progeroid. *Am J Hum Genet*. 1987;41(3):436–453.
15. Quentin E, et al. A genetic defect in the biosynthesis of dermatan sulfate proteoglycan: galactosyltransferase I deficiency in fibroblasts from a patient with a progeroid syndrome. *Proc*

Natl Acad Sci USA. 1990;87(4):1342–1346.

16. Okajima T, et al. Molecular Basis for the Progeroid Variant of Ehlers-Danlos Syndrome. *J Biol Chem*. 1999;274(41):28841–28844.

17. Malfait F, et al. Defective initiation of glycosaminoglycan synthesis due to B3GALT6 Mutations causes a pleiotropic Ehlers-Danlos-syndrome-like connective tissue disorder. *Am J Hum Genet*. 2013;92(6):935–945.

18. Nakajima M, et al. Mutations in B3GALT6, which encodes a glycosaminoglycan linker region enzyme, cause a spectrum of skeletal and connective tissue disorders. *Am J Hum Genet*. 2013;92(6):927–934.

19. Bin B-H, et al. Biochemical Characterization of Human ZIP13 Protein. *J Biol Chem*. 2011;286(46):40255–40265.

20. Giunta C, et al. 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*. 2008;82(6):1290–1305.

21. Fukada T, et al. The Zinc Transporter SLC39A13/ZIP13 Is Required for Connective Tissue Development; Its Involvement in BMP/TGF- β Signaling Pathways. *PLoS ONE*. 2008;3(11):e3642.

22. Brady AF, et al. The Ehlers-Danlos syndromes, rare types. *Am J Med Genet C Semin Med Genet*. 2017;175(1):70–115.

23. Raja SN, et al. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. *Pain*. 2020;161(9):1976–1982.

24. Latremoliere A, Woolf CJ. Central Sensitization: A Generator of Pain Hypersensitivity by Central Neural Plasticity. *J Pain*. 2009;10(9):895–926.

25. Malfait F, et al. Pain in the Ehlers–Danlos syndromes: Mechanisms, models, and challenges. *Am J Med Genet C Semin Med Genet*. 2021;187(4):429–445.

26. Castori M, et al. Re-writing the natural history of pain and related symptoms in the joint hypermobility syndrome/Ehlers-Danlos syndrome, hypermobility type. *Am J Med Genet A*. 2013;161(12):2989–3004.

27. Syx D, et al. Hypermobility, the Ehlers-Danlos syndromes and chronic pain. *Clin Exp Rheumatol*. 2017;35 Suppl 107(5):116–122.

28. Caraffi, et al. Severe Peripheral Joint Laxity is a Distinctive Clinical Feature of Spondylodysplastic-Ehlers-Danlos Syndrome (EDS)-B4GALT7 and Spondylodysplastic-EDS-B3GALT6. *Genes*. 2019;10(10):799.

29. Ben-Mahmoud A, et al. A B3GALT6 variant in patient originally described as Al-Gazali

syndrome and implicating the endoplasmic reticulum quality control in the mechanism of some β 3GalT6-pathway mutations. *Clin Genet*. 2018;93(6):1148–1158.

30. Beighton P, Kozlowski K. Spondylo-Epi-Metaphyseal dysplasia with joint laxity and severe, progressive kyphoscoliosis. *Skeletal Radiol*. 1980;5(4):205–212.

31. Beighton P, et al. The manifestations and natural history of spondylo-epi-metaphyseal dysplasia with joint laxity. *Clin Genet*. 1984;26(4):308–317.

32. al-Gazali LI, et al. Anterior segment anomalies of the eye associated with multiple skeletal abnormalities and early lethality: confirmation of an autosomal recessive syndrome. *Clin Dysmorphol*. 1999;8(2):87–92.

33. Coetzer KC, et al. B3GALT6-linkeropathy: Three illustrative patients spanning the disease spectrum. *Eur J Med Genet*. 2023;66(10):104829.

34. Sellars EA, et al. A Newborn With Complex Skeletal Abnormalities, Joint Contractures, and Bilateral Corneal Clouding With Sclerocornea. *Semin Pediatr Neurol*. 2014;21(2):84–87.

35. Ritelli M, et al. 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*. 2015;2:1–15.

36. Leoni C, et al. Broadening the phenotypic spectrum of Beta3GalT6-associated phenotypes. *Am J Med Genet A*. 2021;185(10):3153–3160.

37. Shen F, et al. Mutant B3GALT6 in a multiplex family: A dominant variant co-segregated with moderate malformations. *Front Genet*. 2022;13:824445.

38. Brockel M, et al. Anesthetic Considerations for a Child With Rare B3GALT6 Mutations: A Case Report. *A A Pract*. 2018;10(4):83–86.

39. Berglund B, Nordström G, Lützén K. Living a restricted life with Ehlers-Danlos Syndrome (EDS). *Int J Nurs Stud*. 2000;37(2):111–118.

40. Callewaert B, et al. Ehlers-Danlos syndromes and Marfan syndrome. *Best Pract Res Clin Rheumatol*. 2008;22(1):165–189.

41. Malfait F, Wenstrup RJ, De Paepe A. Clinical and genetic aspects of Ehlers-Danlos syndrome, classic type. *Genet Med*. 2010;12(10):597–605.

42. Alazami AM, et al. Expanding the clinical and genetic heterogeneity of hereditary disorders of connective tissue. *Hum Genet*. 2016;135(5):525–540.

43. Descartes M, et al. Keratoconus in a patient with B3GALT6-related disorder. *Clin Genet*. 2021;99(6):849–850.

44. Han S, et al. New genetic mutations in a Chinese child with Ehlers-Danlos syndrome-like spondyloepimetaphyseal dysplasia: A case report. *Front Pediatr*. 2022;10:1073748.

45. Honey EM. Spondyloepimetaphyseal dysplasia with joint laxity (Beighton type): A unique South African disorder. *S Afr Med J*. 2016;106(6):54.
46. Maddirevula S, et al. Expanding the phenome and variome of skeletal dysplasia. *Genet Med*. 2018;20(12):1609–1616.
47. Ranza E, et al. Chondrodysplasia with multiple dislocations: comprehensive study of a series of 30 cases: Chondrodysplasia with multiple dislocations. *Clin Genet*. 2017;91(6):868–880.
48. Silveira KC, et al. Molecular diagnosis in a cohort of 114 patients with rare skeletal dysplasias. *Am J Med Genet C Semin Med Genet*. 2021;187(3):396–408.
49. Trejo P, et al. Spondyloepimetaphyseal Dysplasia with Joint Laxity in Three Siblings with **B3GALT6** Mutations. *Mol Syndromol*. 2017;8(6):303–307.
50. Van Damme T, et al. Biallelic B3GALT6 mutations cause spondylodysplastic Ehlers–Danlos syndrome. *Hum Mol Genet*. 2018;27(20):3475–3487.
51. Vorster AA, Beighton P, Ramesar RS. Spondyloepimetaphyseal dysplasia with joint laxity (Beighton type); mutation analysis in eight affected South African families: South African **B3GALT6** SEMD-JL mutations. *Clin Genet*. 2015;87(5):492–495.
52. Zhang J, Huang K, Dong G. Clinical and genetic analysis of a child with spondyloepimetaphyseal dysplasia type 1 and joint laxity. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi*. 2020;37(08):887–890.
53. Theocharis AD, et al. Extracellular matrix structure. *Adv Drug Deliv Rev*. 2016;97:4–27.
54. Karamanos NK, et al. A guide to the composition and functions of the extracellular matrix. *FEBS J*. 2021;288(24):6850–6912.
55. Van Damme T, et al. The Ehlers–Danlos syndromes against the backdrop of inborn errors of metabolism. *Genes*. 2022;13(2):265.
56. Lairson LL, et al. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu Rev Biochem*. 2008;77(1):521–555.
57. Iozzo RV, Schaefer L. Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. *Matrix Biol*. 2015;42:11–55.
58. Filmus J, Capurro M, Rast J. Glypicans. *Genome Biol*. 2008;9(5):224.
59. De Rossi G, et al. Shed syndecan-2 inhibits angiogenesis. *J Cell Sci*. 2014;127(21):4788–4799.
60. Marcum JA, Thompson MA. The amino-terminal region of a proteochondroitin core protein, secreted by aortic smooth muscle cells, shares sequence homology with the pre-propeptide region of the biglycan core protein from human bone. *Biochem Biophys Res*

Commun. 1991;175(2):706–712.

61. Raman R, Sasisekharan V, Sasisekharan R. Structural Insights into Biological Roles of Protein-Glycosaminoglycan Interactions. *Chem Biol.* 2005;12(3):267–277.

62. Prydz K, Dalen KT. Synthesis and sorting of proteoglycans. *J Cell Sci.* 2000;113(2):193–205.

63. Baker JR, Cifonelli JA, Rodén L. The Linkage of Corneal Keratan Sulfate to Protein. *Connect Tissue Res.* 1975;3(2–3):149–156.

64. Couchman JR, Pataki CA. An Introduction to Proteoglycans and Their Localization. *J Histochem Cytochem.* 2012;60(12):885–897.

65. Götting C, et al. Molecular Cloning and Expression of Human UDP-d-Xylose:Proteoglycan Core Protein β -d-Xylosyltransferase and its First Isoform XT-II. *J Mol Biol.* 2000;304(4):517–528.

66. Roch C, et al. Differences in gene expression of human xylosyltransferases and determination of acceptor specificities for various proteoglycans. *Biochem Biophys Res Commun.* 2010;391(1):685–691.

67. Bai X, et al. 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.* 2001;276(51):48189–48195.

68. Kitagawa H, et al. Molecular cloning and expression of glucuronyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J Biol Chem.* 1998;273(12):6615–6618.

69. Koike T, et al. FAM20B is a kinase that phosphorylates xylose in the glycosaminoglycan–protein linkage region. *Biochem J.* 2009;421(2):157–162.

70. Wen J, et al. Xylose phosphorylation functions as a molecular switch to regulate proteoglycan biosynthesis. *Proc Natl Acad Sci USA.* 2014;111(44):15723–15728.

71. Tone Y, et al. 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.* 2008;283(24):16801–16807.

72. Gulberti S, et al. 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.* 2005;280(2):1417–1425.

73. Koike T, et al. Identification of phosphatase that dephosphorylates xylose in the glycosaminoglycan-protein linkage region of proteoglycans. *J Biol Chem.* 2014;289(10):6695–

6708.

74. Sammon D, et al. Molecular mechanism of decision-making in glycosaminoglycan biosynthesis. *Nat Commun.* 2023;14(1):6425.
75. Soares Da Costa D, Reis RL, Pashkuleva I. Sulfation of Glycosaminoglycans and Its Implications in Human Health and Disorders. *Annu Rev Biomed Eng.* 2017;19(1):1–26.
76. Annaival T, et al. Heparan Sulfate Proteoglycans Biosynthesis and Post Synthesis Mechanisms Combine Few Enzymes and Few Core Proteins to Generate Extensive Structural and Functional Diversity. *Molecules.* 2020;25(18):4215.
77. Seffouh A, et al. Expression and purification of recombinant extracellular sulfatase HSulf-2 allows deciphering of enzyme sub-domain coordinated role for the binding and 6-O-desulfation of heparan sulfate. *Cell Mol Life Sci.* 2019;76(9):1807–1819.
78. Sugaya N, et al. 6-O-Sulfation of Heparan Sulfate Differentially Regulates Various Fibroblast Growth Factor-dependent Signalings in Culture. *Journal of Biological Chemistry.* 2008;283(16):10366–10376.
79. Matsuzaka Y, Yashiro R. Classification and Molecular Functions of Heparan Sulfate Proteoglycans and Their Molecular Mechanisms with the Receptor. *Biologics.* 2024;4(2):105–129.
80. Arnold K, et al. Synthetic anticoagulant heparan sulfate attenuates liver ischemia reperfusion injury. *Sci Rep.* 2020;10(1):17187.
81. Gulberti S, et al. The role of heparan sulfate maturation in cancer: A focus on the 3O-sulfation and the enigmatic 3O-sulfotransferases (HS3STs). *Semin Cancer Biol.* 2020;62:68–85.
82. Bui C, et al. Epigenetics: methylation-associated repression of heparan sulfate 3- O -sulfotransferase gene expression contributes to the invasive phenotype of H-EMC-SS chondrosarcoma cells. *FASEB J.* 2010;24(2):436–450.
83. Mao X, et al. The heparan sulfate sulfotransferase 3-OST3A (HS3ST3A) is a novel tumor regulator and a prognostic marker in breast cancer. *Oncogene.* 2016;35(38):5043–5055.
84. Lim EH, et al. Cervical dysplasia: Assessing methylation status (Methylight) of CCNA1, DAPK1, HS3ST2, PAX1 and TFPI2 to improve diagnostic accuracy. *Gynecologic Oncology.* 2010;119(2):225–231.
85. Aspberg A. The Different Roles of Aggrecan Interaction Domains. *J Histochem Cytochem.* 2012;60(12):987–996.
86. Krusius T, Ruoslahti E. Primary structure of an extracellular matrix proteoglycan core protein deduced from cloned cDNA. *Proc Natl Acad Sci USA.* 1986;83(20):7683–7687.

87. Fisher LW, Termine JD, Young MF. Deduced protein sequence of bone small proteoglycan I (biglycan) shows homology with proteoglycan II (decorin) and several nonconnective tissue proteins in a variety of species. *J Biol Chem*. 1989;264(8):4571–4576.
88. Heinegård D, et al. Two novel matrix proteins isolated from articular cartilage show wide distributions among connective tissues. *J Biol Chem*. 1986;261(29):13866–13872.
89. Sloan AJ. Biology of the Dentin-Pulp Complex. *Stem Cell Biology and Tissue Engineering in Dental Sciences*. Elsevier; 2015:371–378.
90. Bianco P, et al. Expression and localization of the two small proteoglycans biglycan and decorin in developing human skeletal and non-skeletal tissues. *J Histochem Cytochem*. 1990;38(11):1549–1563.
91. Pulkkinen L, et al. Expression of decorin in human tissues and cell lines and defined chromosomal assignment of the gene locus (DCN). *Cytogenet Genome Res*. 1992;60(2):107–111.
92. Schaefer L, Iozzo RV. Biological Functions of the Small Leucine-rich Proteoglycans: From Genetics to Signal Transduction. *J Biol Chem*. 2008;283(31):21305–21309.
93. Matsushima N, et al. A nested leucine rich repeat (LRR) domain: The precursor of LRRs is a ten or eleven residue motif. *BMC Microbiol*. 2010;10(1):235.
94. Kram V, et al. Small leucine rich proteoglycans, a novel link to osteoclastogenesis. *Sci Rep*. 2017;7(1):12627.
95. Scott PG, et al. Crystal structure of the dimeric protein core of decorin, the archetypal small leucine-rich repeat proteoglycan. *Proc Natl Acad Sci USA*. 2004;101(44):15633–15638.
96. Iozzo RV. The Family of the Small Leucine-Rich Proteoglycans: Key Regulators of Matrix Assembly and Cellular Growth. *Crit Rev Biochem Mol Biol*. 1997;32(2):141–174.
97. Haq F, Ahmed N, Qasim M. Comparative genomic analysis of collagen gene diversity. *3 Biotech*. 2019;9(3):83.
98. Gay S, Miller EJ. *Overview: What Is Collagen, What Is Not*. *Ultrastruct Pathol*. 1983;4(4):365–377.
99. Prockop DJ, Kivirikko KI. COLLAGENS: Molecular Biology, Diseases, and Potentials for Therapy. *Annu Rev Biochem*. 1995;64(1):403–434.
100. Rossert J, De Crombrughe B. Type I Collagen. *Principles of Bone Biology*. Elsevier; 2002:189–XVIII.
101. Mayne R, et al. Isolation and characterization of the chains of type V/type XI collagen present in bovine vitreous. *J Biol Chem*. 1993;268(13):9381–9386.
102. Wu J-J, et al. Differences in Chain Usage and Cross-linking Specificities of Cartilage Type

- V/XI Collagen Isoforms with Age and Tissue. *J Biol Chem.* 2009;284(9):5539–5545.
103. Ricard-Blum S. The Collagen Family. *Cold Spring Harb Perspect Biol.* 2011;3(1):a004978–a004978.
104. Salo AM, et al. *Collagen prolyl 4-hydroxylases have sequence specificity towards different X-Pro-Gly triplets.* *Molecular Biology*; 2023.
105. Shoulders MD, Raines RT. Collagen Structure and Stability. *Annu Rev Biochem.* 2009;78(1):929–958.
106. Qi Y, Xu R. Roles of PLODs in collagen synthesis and cancer progression. *Front Cell Dev Biol.* 2018;6:66.
107. Sricholpech M, et al. Lysyl Hydroxylase 3 Glucosylates Galactosylhydroxylysine Residues in Type I Collagen in Osteoblast Culture. *J Biol Chem.* 2011;286(11):8846–8856.
108. Terajima M, et al. Role of Glycosyltransferase 25 Domain 1 in Type I Collagen Glycosylation and Molecular Phenotypes. *Biochemistry.* 2019;58(50):5040–5051.
109. Kimura T, et al. The human alpha 2(XI) collagen (COL11A2) chain. Molecular cloning of cDNA and genomic DNA reveals characteristics of a fibrillar collagen with differences in genomic organization. *J Biol Chem.* 1989;264(23):13910–13916.
110. Bourhis J-M, et al. Structural basis of fibrillar collagen trimerization and related genetic disorders. *Nat Struct Mol Biol.* 2012;19(10):1031–1036.
111. Malhotra V, Erlmann P. The Pathway of Collagen Secretion. *Annu Rev Cell Dev Biol.* 2015;31(1):109–124.
112. Tang BL, Hong W. ADAMTS: A novel family of proteases with an ADAM protease domain and thrombospondin 1 repeats. *FEBS Lett.* 1999;445(2–3):223–225.
113. Vadon-Le Goff S, Hulmes DJS, Moali C. BMP-1/tolloid-like proteinases synchronize matrix assembly with growth factor activation to promote morphogenesis and tissue remodeling. *Matrix Biol.* 2015;44–46:14–23.
114. Mäki JM. Lysyl oxidases in mammalian development and certain pathological conditions. *Histol Histopathol.* 2009;(24):651–660.
115. Vallet SD, Ricard-Blum S. Lysyl oxidases: from enzyme activity to extracellular matrix cross-links. *Essays Biochem.* 2019;63(3):349–364.
116. Canty EG, Kadler KE. Procollagen trafficking, processing and fibrillogenesis. *J Cell Sci.* 2005;118(7):1341–1353.
117. Koch M, et al. Large and small splice variants of collagen XII: differential expression and ligand binding. *J Cell Biol.* 1995;130(4):1005–1014.
118. Kania AM, et al. Structural Variation of Type XII Collagen at Its Carboxyl-terminal NC1

Domain Generated by Tissue-specific Alternative Splicing. *J Biol Chem.* 1999;274(31):22053–22059.

119. Gerecke DR, et al. Complete Primary Structure of Two Splice Variants of Collagen XII, and Assignment of $\alpha 1$ (XII) Collagen (COL12A1), $\alpha 1$ (IX) Collagen (COL9A1), and $\alpha 1$ (XIX) Collagen (COL19A1) to Human Chromosome 6q12–q13. *Genomics.* 1997;41(2):236–242.

120. Mazzorana M, et al. Collagenous Sequence Governs the Trimeric Assembly of Collagen XII. *J Biol Chem.* 2001;276(30):27989–27998.

121. Chiquet M, et al. Collagen XII: Protecting bone and muscle integrity by organizing collagen fibrils. *Int J Biochem Cell Biol.* 2014;53:51–54.

122. Wälchli C, et al. Tissue-specific expression of the fibril-associated collagens XII and XIV. *J Cell Sci.* 1994;107(2):669–681.

123. Zvackova I, Matalova E, Lesot H. Regulators of Collagen Fibrillogenesis during Molar Development in the Mouse. *Front Physiol.* 2017;8:554.

124. Karimbux NY, Nishimura I. Temporal and Spatial Expressions of Type XII Collagen in the Remodeling Periodontal Ligament during Experimental Tooth Movement. *J Dent Res.* 1995;74(1):313–318.

125. Anderson S, et al. Developmentally regulated appearance of spliced variants of type XII collagen in the cornea. *Invest Ophthalmol Vis Sci.* 2000;41(1):55–63.

126. Farhat YM, et al. Gene Expression Analysis of the Pleiotropic Effects of TGF- $\beta 1$ in an In Vitro Model of Flexor Tendon Healing. *PLoS ONE.* 2012;7(12):e51411.

127. Izu Y, Birk DE. Collagen XII mediated cellular and extracellular mechanisms in development, regeneration, and disease. *Front Cell Dev Biol.* 2023;11:1129000.

128. Koch M, Bernasconi C, Chiquet M. A major oligomeric fibroblast proteoglycan identified as a novel large form of type-XII collagen. *Eur J Biochem.* 1992;207(3):847–856.

129. Toledo AG, et al. An affinity chromatography and glycoproteomics workflow to profile the chondroitin sulfate proteoglycans that interact with malarial VAR2CSA in the placenta and in cancer. *Glycobiology.* 2020;30(12):989–1002.

130. Birk DE, Brückner P. Collagens, Suprastructures, and Collagen Fibril Assembly. In: Mecham RP, ed. *The Extracellular Matrix: an Overview.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2011:77–115.

131. Cooper A. Thermodynamic studies of the assembly *in vitro* of native collagen fibrils. *Biochem J.* 1970;118(3):355–365.

132. Kadler KE, et al. Collagen fibril formation. *Biochem J.* 1996;316(1):1–11.

133. Shaw LM, Olsen BR. FACIT collagens: diverse molecular bridges in extracellular

- matrices. *Trends Biochem Sci.* 1991;16:191–194.
134. Scott JE. Proteoglycan-fibrillar collagen interactions. *Biochem J.* 1988;252(2):313–323.
135. Naomi R, Ridzuan PM, Bahari H. Current Insights into Collagen Type I. *Polymers.* 2021;13(16):2642.
136. Birk DE. Type V collagen: heterotypic type I/V collagen interactions in the regulation of fibril assembly. *Micron.* 2001;32(3):223–237.
137. Wenstrup RJ, et al. Type V Collagen Controls the Initiation of Collagen Fibril Assembly. *J Biol Chem.* 2004;279(51):53331–53337.
138. Birk DE, et al. Collagen fibrillogenesis *in vitro*: interaction of types I and V collagen regulates fibril diameter. *J Cell Sci.* 1990;95(4):649–657.
139. Linsenmayer TF, et al. Type V collagen: molecular structure and fibrillar organization of the chicken alpha 1(V) NH2-terminal domain, a putative regulator of corneal fibrillogenesis. *J Cell Biol.* 1993;121(5):1181–1189.
140. Wenstrup RJ, et al. Regulation of Collagen Fibril Nucleation and Initial Fibril Assembly Involves Coordinate Interactions with Collagens V and XI in Developing Tendon. *J Biol Chem.* 2011;286(23):20455–20465.
141. Kadler KE, Hill A, Canty-Laird EG. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. *Curr Opin Cell Biol.* 2008;20(5):495–501.
142. Kleinman HK, et al. Localization of the binding site for cell attachment in the alpha1(I) chain of collagen. *J Biol Chem.* 1978;253(16):5642–5646.
143. Dzamba BJ, et al. Fibronectin binding site in type I collagen regulates fibronectin fibril formation. *J Cell Biol.* 1993;121(5):1165–1172.
144. Paten JA, et al. Molecular Interactions between Collagen and Fibronectin: A Reciprocal Relationship that Regulates De Novo Fibrillogenesis. *Chem.* 2019;5(8):2126–2145.
145. Menzel J, Borth W. Influence of Plasma Fibronectin on Collagen Cleavage by Collagenase. *Coll Relat Res.* 1983;3(3):217–230.
146. Shi F, et al. Collagen I matrix turnover is regulated by fibronectin polymerization. *Am J Physiol Cell Physiol.* 2010;298(5):C1265–C1275.
147. Varani J, et al. Decreased Collagen Production in Chronologically Aged Skin. *Am J Pathol.* 2006;168(6):1861–1868.
148. Velling T, et al. Polymerization of Type I and III Collagens Is Dependent On Fibronectin and Enhanced By Integrins $\alpha 11\beta 1$ and $\alpha 2\beta 1$. *J Biol Chem.* 2002;277(40):37377–37381.
149. Li S, et al. Vascular Smooth Muscle Cells Orchestrate the Assembly of Type I Collagen via $\alpha 2\beta 1$ Integrin, RhoA, and Fibronectin Polymerization. *Am J Pathol.* 2003;163(3):1045–

1056.

150. Chen S, Birk DE. The regulatory roles of small leucine-rich proteoglycans in extracellular matrix assembly. *FEBS J.* 2013;280(10):2120–2137.
151. Kalamajski S, Oldberg Å. The role of small leucine-rich proteoglycans in collagen fibrillogenesis. *Matrix Biol.* 2010;29(4):248–253.
152. Douglas T, et al. Fibrillogenesis of Collagen Types I, II, and III with Small Leucine-Rich Proteoglycans Decorin and Biglycan. *Biomacromolecules.* 2006;7(8):2388–2393.
153. Kalamajski S, Aspberg A, Oldberg Å. The Decorin Sequence SYIRIADTNIT Binds Collagen Type I. *J Biol Chem.* 2007;282(22):16062–16067.
154. Orgel JPRO, et al. Decorin Core Protein (Decoron) Shape Complements Collagen Fibril Surface Structure and Mediates Its Binding. *PLoS ONE.* 2009;4(9):e7028.
155. Henninger HB, et al. Spatial distribution and orientation of dermatan sulfate in human medial collateral ligament. *J Struct Biol.* 2007;158(1):33–45.
156. Lewis PN, et al. Structural Interactions between Collagen and Proteoglycans Are Elucidated by Three-Dimensional Electron Tomography of Bovine Cornea. *Structure.* 2010;18(2):239–245.
157. Raspanti M, et al. Glycosaminoglycans show a specific periodic interaction with type I collagen fibrils. *J Struct Biol.* 2008;164(1):134–139.
158. Rühland C, et al. The glycosaminoglycan chain of decorin plays an important role in collagen fibril formation at the early stages of fibrillogenesis: Decorin and fibrillogenesis. *FEBS J.* 2007;274(16):4246–4255.
159. Bengtsson E, et al. The Leucine-rich Repeat Protein PRELP Binds Perlecan and Collagens and May Function as a Basement Membrane Anchor. *J Biol Chem.* 2002;277(17):15061–15068.
160. Krishnan P, et al. Distinct Secondary Structures of the Leucine-rich Repeat Proteoglycans Decorin and Biglycan. *J Biol Chem.* 1999;274(16):10945–10950.
161. Kresse H, et al. Critical Role of Glutamate in a Central Leucine-rich Repeat of Decorin for Interaction with Type I Collagen. *J Biol Chem.* 1997;272(29):18404–18410.
162. Nareyeck G, et al. Differential interactions of decorin and decorin mutants with type I and type VI collagens: Decorin interacting with type I and VI collagen. *Eur J Biochem.* 2004;271(16):3389–3398.
163. Kalamajski S, Oldberg Å. Fibromodulin Binds Collagen Type I via Glu-353 and Lys-355 in Leucine-rich Repeat 11. *J Biol Chem.* 2007;282(37):26740–26745.
164. Kalamajski S, Oldberg Å. Homologous Sequence in Lumican and Fibromodulin Leucine-

- rich Repeat 5-7 Competes for Collagen Binding. *J Biol Chem.* 2009;284(1):534–539.
165. Kalamajski S, et al. Fibromodulin Interacts with Collagen Cross-linking Sites and Activates Lysyl Oxidase. *J Biol Chem.* 2016;291(15):7951–7960.
166. Tashima T, et al. Molecular basis for governing the morphology of type-I collagen fibrils by Osteomodulin. *Commun Biol.* 2018;1(1):33.
167. Redaelli A, et al. Possible role of decorin glycosaminoglycans in fibril to fibril force transfer in relative mature tendons—a computational study from molecular to microstructural level. *J Biomech.* 2003;36(10):1555–1569.
168. Al Makhzoomi AK, et al. Contribution of glycosaminoglycans to the structural and mechanical properties of tendons – A multiscale study. *J Biomech.* 2021;128:110796.
169. Silver FH, Kelkar N, Deshmukh T. Molecular Basis for Mechanical Properties of ECMs: Proposed Role of Fibrillar Collagen and Proteoglycans in Tissue Biomechanics. *Biomolecules.* 2021;11(7):1018.
170. Robinson KA, et al. Decorin and biglycan are necessary for maintaining collagen fibril structure, fiber realignment, and mechanical properties of mature tendons. *Matrix Biol.* 2017;64:81–93.
171. Han B, et al. Decorin Regulates the Aggrecan Network Integrity and Biomechanical Functions of Cartilage Extracellular Matrix. *ACS Nano.* 2019;13(10):11320–11333.
172. Tatara Y, et al. Chondroitin sulfate cluster of epiphykan from salmon nasal cartilage defines binding specificity to collagens. *Glycobiology.* 2015;25(5):557–569.
173. Chen D, et al. Distinct effects of different matrix proteoglycans on collagen fibrillogenesis and cell-mediated collagen reorganization. *Sci Rep.* 2020;10(1):19065.
174. Takagi J, et al. Structure of integrin $\alpha 5 \beta 1$ in complex with fibronectin. *EMBO J.* 2003;22(18):4607–4615.
175. Takagi J. Structural basis for ligand recognition by RGD (Arg-Gly-Asp)-dependent integrins. *Biochem Soc Trans.* 2004;32(3):403–406.
176. Hynes RO. Integrins: bidirectional, allosteric signaling machines. *Cell.* 2002;110(6):673–687.
177. Colombatti A, Bonaldo P, Doliana R. Type A Modules: Interacting Domains Found in Several Non-Fibrillar Collagens and in Other Extracellular Matrix Proteins. *Matrix.* 1993;13(4):297–306.
178. Cruz MA, Handin RI, Wise RJ. The interaction of the von Willebrand factor-A1 domain with platelet glycoprotein Ib/IX. The role of glycosylation and disulfide bonding in a monomeric recombinant A1 domain protein. *J Biol Chem.* 1993;268(28):21238–21245.

179. Hoylaerts MF, et al. von Willebrand factor binds to native collagen VI primarily via its A1 domain. *Biochem J.* 1997;324(1):185–191.
180. Lankhof H, et al. A3 domain is essential for interaction of von Willebrand factor with collagen type III. *Thromb Haemost.* 1996;75(6):950–958.
181. Blumbach K, et al. Combined role of type IX collagen and cartilage oligomeric matrix protein in cartilage matrix assembly: Cartilage oligomeric matrix protein counteracts type IX collagen–induced limitation of cartilage collagen fibril growth in mouse chondrocyte cultures. *Arthritis Rheum.* 2009;60(12):3676–3685.
182. Agarwal P, et al. Collagen XII and XIV, New Partners of Cartilage Oligomeric Matrix Protein in the Skin Extracellular Matrix Suprastructure. *J Biol Chem.* 2012;287(27):22549–22559.
183. Posey KL, Coustry F, Hecht JT. Cartilage oligomeric matrix protein: COMPopathies and beyond. *Matrix Biol.* 2018;71–72:161–173.
184. Rosenberg K, et al. Cartilage Oligomeric Matrix Protein Shows High Affinity Zinc-dependent Interaction with Triple Helical Collagen. *J Biol Chem.* 1998;273(32):20397–20403.
185. Halász K, et al. COMP Acts as a Catalyst in Collagen Fibrillogenesis. *J Biol Chem.* 2007;282(43):31166–31173.
186. Keene DR, et al. Two type XII-like collagens localize to the surface of banded collagen fibrils. *J Cell Biol.* 1991;113(4):971–978.
187. Font B, et al. Binding of collagen XIV with the dermatan sulfate side chain of decorin. *J Biol Chem.* 1993;268(33):25015–25018.
188. Ehnis T, et al. Localization of a Binding Site for the Proteoglycan Decorin on Collagen XIV (Undulin). *J Biol Chem.* 1997;272(33):20414–20419.
189. Ansorge HL, et al. Type XIV Collagen Regulates Fibrillogenesis. *J Biol Chem.* 2009;284(13):8427–8438.
190. Font B, et al. Characterization of the interactions of type XII collagen with two small proteoglycans from fetal bovine tendon, decorin and fibromodulin. *Matrix Biol.* 1996;15(5):341–348.
191. Veit G, et al. Collagen XII Interacts with Avian Tenascin-X through Its NC3 Domain. *J Biol Chem.* 2006;281(37):27461–27470.
192. Lethias C, et al. A model of tenascin-X integration within the collagenous network. *FEBS Lett.* 2006;580(26):6281–6285.
193. Elefteriou F, et al. Binding of tenascin-X to decorin. *FEBS Lett.* 2001;495(1–2):44–47.
194. Flück M, Tunç-Civelek V, Chiquet M. Rapid and reciprocal regulation of tenascin-C and

- tenascin-Y expression by loading of skeletal muscle. *J Cell Sci.* 2000;113(20):3583–3591.
195. Danielson KG, et al. Targeted Disruption of Decorin Leads to Abnormal Collagen Fibril Morphology and Skin Fragility. *J Cell Biol.* 1997;136(3):729–743.
196. Reed CC, Iozzo RV. The role of decorin in collagen fibrillogenesis and skin homeostasis. *Glycoconj J.* 2002;19(4/5):249–255.
197. Chakravarti S, et al. Lumican Regulates Collagen Fibril Assembly: Skin Fragility and Corneal Opacity in the Absence of Lumican. *J Cell Biol.* 1998;141(5):1277–1286.
198. Svensson L, et al. Fibromodulin-null Mice Have Abnormal Collagen Fibrils, Tissue Organization, and Altered Lumican Deposition in Tendon. *J Biol Chem.* 1999;274(14):9636–9647.
199. Zou Y, et al. Recessive and dominant mutations in COL12A1 cause a novel EDS/myopathy overlap syndrome in humans and mice. *Hum Mol Genet.* 2014;23(9):2339–2352.
200. Izu Y, et al. Collagen XII mediated cellular and extracellular mechanisms regulate establishment of tendon structure and function. *Matrix Biol.* 2021;95:52–67.
201. Izu Y, et al. Type XII collagen regulates osteoblast polarity and communication during bone formation. *J Cell Biol.* 2011;193(6):1115–1130.
202. Zhu M, et al. Ablation of collagen XII disturbs joint extracellular matrix organization and causes patellar subluxation. *iScience.* 2023;26(7):107225.
203. Marini JC, et al. Osteogenesis imperfecta. *Nat Rev Dis Primers.* 2017;3(1):17052.
204. Bonafe L, et al. Nosology and classification of genetic skeletal disorders: 2015 revision. *Am J Med Genet A.* 2015;167(12):2869–2892.
205. Malfait F, et al. The Ehlers–Danlos syndromes. *Nat Rev Dis Primers.* 2020;6(1):64.
206. Lamandé SR, Bateman JF. Genetic Disorders of the Extracellular Matrix. *Anat Rec (Hoboken).* 2020;303(6):1527–1542.
207. Halper J, ed. *Progress in Heritable Soft Connective Tissue Diseases.* Dordrecht: Springer Netherlands; 2014.
208. Wenstrup RJ, et al. Reduced type I collagen utilization: A pathogenic mechanism in COL5A1 haplo-insufficient Ehlers-Danlos syndrome. *J Cell Biochem.* 2004;92(1):113–124.
209. Wenstrup RJ, et al. COL5A1 Haploinsufficiency Is a Common Molecular Mechanism Underlying the Classical Form of EDS. *Am J Hum Genet.* 2000;66(6):1766–1776.
210. Duong J, et al. A family with Classical Ehlers-Danlos Syndrome (cEDS), mild bone fragility and without vascular complications, caused by the p.Arg312Cys mutation in COL1A1. *Eur J Med Genet.* 2020;63(2):103730.

211. Malfait F, et al. 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.* 2007;28(4):387–395.
212. Schwarze U, et al. 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.* 2004;74(5):917–930.
213. Byers PH, et al. Ehlers-Danlos syndrome type VIIA and VIIB result from splice-junction mutations or genomic deletions that involve exon 6 in the COL1A1 and COL1A2 genes of type I collagen. *Am J Med Genet.* 1997;72(1):94–105.
214. Gnoli M, et al. COL1-Related Disorders: Case Report and Review of Overlapping Syndromes. *Front Genet.* 2021;12:640558.
215. Van Damme T, et al. Expanding the clinical and mutational spectrum of the Ehlers–Danlos syndrome, dermatosparaxis type. *Genet Med.* 2016;18(9):882–891.
216. Bekhouche M, Colige A. The procollagen N-proteinases ADAMTS2, 3 and 14 in pathophysiology. *Matrix Biol.* 2015;44–46:46–53.
217. Bekhouche M, et al. Determination of the substrate repertoire of ADAMTS2, 3, and 14 significantly broadens their functions and identifies extracellular matrix organization and TGF- β signaling as primary targets. *FASEB J.* 2016;30(5):1741–1756.
218. Colige A, et al. 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.* 2004;123(4):656–663.
219. Hicks D, et al. Mutations in the collagen XII gene define a new form of extracellular matrix-related myopathy. *Hum Mol Genet.* 2014;23(9):2353–2363.
220. Coppens S, et al. EHLERS-DANLOS /myopathy overlap syndrome caused by a large de novo deletion in *COL12A1*. *Am J Med Genet A.* 2022;188(5):1556–1561.
221. Bui C, et al. XYLT1 Mutations in Desbuquois Dysplasia Type 2. *Am J Hum Genet.* 2014;94(3):405–414.
222. Rajabi F, et al. Novel compound heterozygous variants in XYLT1 gene caused Desbuquois dysplasia type 2 in an aborted fetus: a case report. *BMC Pediatr.* 2022;22(1):63.
223. Seidler DG, et al. 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 (β 4GalT-7). *J Mol Med.* 2006;84(7):583–594.
224. Ritelli, et al. Further Defining the Phenotypic Spectrum of B3GAT3 Mutations and

Literature Review on Linkeropathy Syndromes. *Genes*. 2019;10(9):631.

225. Minatogawa M, et al. Clinical and molecular features of 66 patients with musculocontractural Ehlers–Danlos syndrome caused by pathogenic variants in *CHST14* (mcEDS- *CHST14*). *J Med Genet*. 2022;59(9):865–877.

226. Schirwani S, et al. DSE associated musculocontractural EDS, a milder phenotype or phenotypic variability. *Eur J Med Genet*. 2020;63(4):103798.

227. Lombard J. The multiple evolutionary origins of the eukaryotic N-glycosylation pathway. *Biol Direct*. 2016;11(1):36.

228. Drula E, et al. The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res*. 2022;50(D1):D571–D577.

229. Rini JM, et al. Glycosyltransferases and Glycan-Processing Enzymes. In: Varki A, et al., eds. *Essentials of Glycobiology*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022.

230. Breton C, Fournel-Gigleux S, Palcic MM. Recent structures, evolution and mechanisms of glycosyltransferases. *Curr Opin Struct Biol*. 2012;22(5):540–549.

231. Charnock SJ, Davies GJ. Structure of the Nucleotide-Diphospho-Sugar Transferase, SpsA from *Bacillus subtilis*, in Native and Nucleotide-Complexed Forms. *Biochemistry*. 1999;38(20):6380–6385.

232. Vrielink A, et al. Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J*. 1994;13(15):3413–3422.

233. Lizak C, et al. X-ray structure of a bacterial oligosaccharyltransferase. *Nature*. 2011;474(7351):350–355.

234. Moremen KW, Haltiwanger RS. Emerging structural insights into glycosyltransferase-mediated synthesis of glycans. *Nat Chem Biol*. 2019;15(9):853–864.

235. Medvedev KE, et al. Functional analysis of Rossmann-like domains reveals convergent evolution of topology and reaction pathways. *PLoS Comput Biol*. 2019;15(12):e1007569.

236. Rao ST, Rossmann MG. Comparison of super-secondary structures in proteins. *J Mol Biol*. 1973;76(2):241–256.

237. Medvedev KE, et al. A Fifth of the Protein World: Rossmann-like Proteins as an Evolutionarily Successful Structural unit. *J Mol Biol*. 2021;433(4):166788.

238. Hanukoglu I. Proteopedia: Rossmann fold: A beta-alpha-beta fold at dinucleotide binding sites: Rossmann Fold in FAD, NAD and NADP Binding Domains. *Biochem Mol Biol Educ*. 2015;43(3):206–209.

239. Tsutsui Y, Ramakrishnan B, Qasba PK. Crystal Structures of β -1,4-Galactosyltransferase 7 Enzyme Reveal Conformational Changes and Substrate Binding. *J Biol Chem.* 2013;288(44):31963–31970.
240. Taujale R, et al. Deep evolutionary analysis reveals the design principles of fold A glycosyltransferases. *Elife.* 2020;9:e54532.
241. Gulberti S, et al. The Functional Glycosyltransferase Signature Sequence of the Human β 1,3-Glucuronosyltransferase Is a XDD Motif. *J Biol Chem.* 2003;278(34):32219–32226.
242. Götting C, et al. Analysis of the DXD Motifs in Human Xylosyltransferase I Required for Enzyme Activity. *Journal of Biological Chemistry.* 2004;279(41):42566–42573.
243. Ramakrishnan B, Qasba PK. Crystal Structure of the Catalytic Domain of Drosophila β 1,4-Galactosyltransferase-7. *J Biol Chem.* 2010;285(20):15619–15626.
244. Gao J, Huang X. Recent advances on glycosyltransferases involved in the biosynthesis of the proteoglycan linkage region. *Advances in Carbohydrate Chemistry and Biochemistry.* Elsevier; 2021:95–119.
245. Pak JE, et al. Structural and Mechanistic Characterization of Leukocyte-Type Core 2 β 1,6-N-Acetylglucosaminyltransferase: A Metal-Ion-Independent GT-A Glycosyltransferase. *J Mol Biol.* 2011;414(5):798–811.
246. Gloster TM. Advances in understanding glycosyltransferases from a structural perspective. *Curr Opin Struct Biol.* 2014;28:131–141.
247. Patenaude SI, et al. The structural basis for specificity in human ABO(H) blood group biosynthesis. *Nat Struct Biol.* 2002;9(9):685–690.
248. Empadinhas N, et al. Functional and structural characterization of a novel mannosyl-3-phosphoglycerate synthase from *Rubrobacter xylanophilus* reveals its dual substrate specificity: Mannosyl-3-phosphoglycerate synthase from *R. xylanophilus*. *Mol Microbiol.* 2011;79(1):76–93.
249. Fritz TA, et al. The beginnings of mucin biosynthesis: The crystal structure of UDP-GalNAc:polypeptide α -N-acetylgalactosaminyltransferase-T1. *Proc Natl Acad Sci USA.* 2004;101(43):15307–15312.
250. Persson K, et al. Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs. *Nat Struct Biol.* 2001;8(2):166–175.
251. Kadirvelraj R, et al. Human N-acetylglucosaminyltransferase II substrate recognition uses a modular architecture that includes a convergent exosite. *Proc Natl Acad Sci USA.* 2018;115(18):4637–4642.

252. Briggs DC, Hohenester E. Structural Basis for the Initiation of Glycosaminoglycan Biosynthesis by Human Xylosyltransferase 1. *Structure*. 2018;26(6):801-809.e3.
253. Li Z, et al. Recognition of EGF-like domains by the Notch-modifying O-fucosyltransferase POFUT1. *Nat Chem Biol*. 2017;13(7):757–763.
254. Bohl H, Bai L, Li H. Recent Progress in Structural Studies on the GT-C Superfamily of Protein Glycosyltransferases. In: Harris JR, Marles-Wright J, eds. *Macromolecular Protein Complexes III: Structure and Function*. Cham: Springer International Publishing; 2021:259–271.
255. Colley KJ. Golgi localization of glycosyltransferases: more questions than answers. *Glycobiology*. 1997;7(1):1–13.
256. Breton C, et al. Structures and mechanisms of glycosyltransferases. *Glycobiology*. 2006;16(2):29R-37R.
257. Kellokumpu S, Hassinen A, Glumoff T. Glycosyltransferase complexes in eukaryotes: long-known, prevalent but still unrecognized. *Cell Mol Life Sci*. 2016;73(2):305–325.
258. Ouzzine M, et al. Structure/Function of the Human Gal β 1,3-glucuronosyltransferase. *J Biol Chem*. 2000;275(36):28254–28260.
259. McCormick C, et al. 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 USA*. 2000;97(2):668–673.
260. Lind T, et al. The Putative Tumor Suppressors EXT1 and EXT2 Are Glycosyltransferases Required for the Biosynthesis of Heparan Sulfate. *J Biol Chem*. 1998;273(41):26265–26268.
261. Busse M, Kusche-Gullberg M. In Vitro Polymerization of Heparan Sulfate Backbone by the EXT Proteins. *J Biol Chem*. 2003;278(42):41333–41337.
262. Leisico F, et al. Structure of the human heparan sulfate polymerase complex EXT1-EXT2. *Nat Commun*. 2022;13(1):7110.
263. Pedersen LC, Darden TA, Negishi M. Crystal Structure of β 1,3-Glucuronyltransferase I in Complex with Active Donor Substrate UDP-GlcUA. *J Biol Chem*. 2002;277(24):21869–21873.
264. Bui C, et al. Molecular characterization of β 1,4-galactosyltransferase 7 genetic mutations linked to the progeroid form of Ehlers-Danlos syndrome (EDS). *FEBS Lett*. 2010;584(18):3962–3968.
265. Pedersen LC, et al. Heparan/Chondroitin Sulfate Biosynthesis. *J Biol Chem*. 2000;275(44):34580–34585.
266. Jumper J, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*.

2021;596(7873):583–589.

267. Varadi M, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 2022;50(D1):D439–D444.

268. Porta-Pardo E, et al. The structural coverage of the human proteome before and after AlphaFold. *PLoS Comput Biol.* 2022;18(1):e1009818.

269. Terwilliger TC, et al. AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination. *Nat Methods.* [published online ahead of print: November 30, 2023]. <https://doi.org/10.1038/s41592-023-02087-4>.

270. Madeira F, et al. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* 2022;50(W1):W276–W279.

271. Zhou D, et al. A beta -1,3-N-acetylglucosaminyltransferase with poly-N-acetyllactosamine synthase activity is structurally related to beta -1,3-galactosyltransferases. *Proc Natl Acad Sci USA.* 1999;96(2):406–411.

272. Patel RY, Balaji PV. Fold-recognition and comparative modeling of human β 3GalT I, II, IV, V and VI and β 3GalNAcT I: Prediction of residues conferring acceptor substrate specificity. *J Mol Graph Model.* 2007;26(1):255–268.

273. Malissard M, et al. Functional assignment of motifs conserved in β 1,3-glycosyltransferases: A mutagenesis study of murine UDP-galactose: β -N-acetylglucosamine β 1,3-galactosyltransferase-I. *Eur J Biochem.* 2002;269(1):233–239.

274. Kadirvelraj R, et al. Comparison of human poly-N-acetyl-lactosamine synthase structure with GT-A fold glycosyltransferases supports a modular assembly of catalytic subsites. *J Biol Chem.* 2021;296:100110.

275. González-Ramírez AM, et al. Structural basis for the synthesis of the core 1 structure by C1GalT1. *Nat Commun.* 2022;13(1):2398.

276. De Las Rivas M, et al. Structural Analysis of a GalNAc-T2 Mutant Reveals an Induced-Fit Catalytic Mechanism for GalNAc-Ts. *Chem Eur J.* 2018;24(33):8382–8392.

277. Iozzo RV, et al. Small Leucine-Rich Proteoglycans. In: Mecham RP, ed. *The Extracellular Matrix: an Overview.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2011:197–231.

278. Iozzo RV, et al. Complexity of matrix phenotypes. *Matrix Biol Plus.* 2020;6–7:100038.

279. Jolivet B. Rôle de la β 1,3-Galactosyltransférase 6 (β 3GalT6) dans la pathogénie d'une maladie génétique rare, les syndromes d'Ehlers-Danlos (SED). 2020. <https://www.theses.fr/2020LORR0085>. Accessed March 12, 2024.

280. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities

of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 1976;72(1–2):248–254.

281. Jacquinet J-C. Synthesis of a set of sulfated and/or phosphorylated oligosaccharide derivatives from the carbohydrate–protein linkage region of proteoglycans. *Carbohydr Res.* 2006;341(10):1630–1644.

282. Karami Y, et al. Computational and biochemical analysis of type IV pilus dynamics and stability. *Structure.* 2021;29(12):1397-1409.e6.

283. Jorgensen WL, et al. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 1983;79(2):926–935.

284. Atsumi T, et al. A chondrogenic cell line derived from a differentiating culture of AT805 teratocarcinoma cells. *Cell Differ Dev.* 1990;30(2):109–116.

285. Yao Y, Wang Y. ATDC5: An excellent in vitro model cell line for skeletal development. *J Cell Biochem.* 2013;114(6):1223–1229.

286. Wilhelm D, et al. ATDC5 cells as a model of cartilage extracellular matrix neosynthesis, maturation and assembly. *J Proteomics.* 2020;219:103718.

287. Polyakov P, et al. Automated Force Volume Image Processing for Biological Samples. *PLoS ONE.* 2011;6(4):e18887.

288. Talhaoui I, et al. Identification of key functional residues in the active site of human β 1,4-galactosyltransferase 7: A major enzyme in the glycosaminoglycan synthesis pathway. *J Biol Chem.* 2010;285(48):37342–37358.

289. Persson A, et al. Identification of a non-canonical chondroitin sulfate linkage region trisaccharide. *Glycobiology.* 2019;29(5):366–371.

290. Baasanjav S, et al. Faulty Initiation of Proteoglycan Synthesis Causes Cardiac and Joint Defects. *Am J Hum Genet.* 2011;89(1):15–27.

291. Chen Y-H, et al. The GAGOme: a cell-based library of displayed glycosaminoglycans. *Nat Methods.* 2018;15(11):881–888.

292. Daligault F, et al. Thermodynamic insights into the structural basis governing the donor substrate recognition by human β 1,4-galactosyltransferase 7. *Biochem J.* 2009;418(3):605–614.

293. Rahuel-Clermont S, et al. Biochemical and thermodynamic characterization of mutated β 1,4-galactosyltransferase 7 involved in the progeroid form of the Ehlers–Danlos syndrome. *Biochem J.* 2010;432(2):303–311.

294. Lauber J, et al. Expression of the functional recombinant human glycosyltransferase GalNAcT2 in *Escherichia coli*. *Microb Cell Fact.* 2015;14(1):3.

295. Paulson JC, Colley KJ. Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. *J Biol Chem.* 1989;264(30):17615–17618.

296. Masibay AS, et al. Mutational analysis of the Golgi retention signal of bovine beta-1,4-galactosyltransferase. *J Biol Chem*. 1993;268(13):9908–9916.
297. Sasai K, et al. The Critical Role of the Stem Region as a Functional Domain Responsible for the Oligomerization and Golgi Localization of N-Acetylglucosaminyltransferase V. *Journal of Biological Chemistry*. 2001;276(1):759–765.
298. Cho SK, Cummings RD. A Soluble Form of α 1,3-Galactosyltransferase Functions within Cells to Galactosylate Glycoproteins. *J Biol Chem*. 1997;272(21):13622–13628.
299. Cid E, et al. The stem region of group A transferase is crucial for its specificity, and its alteration promotes heterologous Forssman synthase activity. *Sci Rep*. 2023;13(1):13996.
300. Some D, et al. Characterization of Proteins by Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS). *J Vis Exp*. 2019;(148):59615.
301. Wu D, Piszczek G. Standard protocol for mass photometry experiments. *Eur Biophys J*. 2021;50(3–4):403–409.
302. De Las Rivas M, et al. Structural and Mechanistic Insights into the Catalytic-Domain-Mediated Short-Range Glycosylation Preferences of GalNAc-T4. *ACS Cent Sci*. 2018;4(9):1274–1290.
303. García-García A, et al. Structural basis for substrate specificity and catalysis of α 1,6-fucosyltransferase. *Nat Commun*. 2020;11(1):973.
304. Filla MS, Dam P, Rapraeger AC. The cell surface proteoglycan syndecan-1 mediates fibroblast growth factor-2 binding and activity. *J Cell Physiol*. 1998;174(3):310–321.
305. Yamaguchi Y, Mann DM, Ruoslahti E. Negative regulation of transforming growth factor- β by the proteoglycan decorin. *Nature*. 1990;346(6281):281–284.
306. Huynh MB, et al. Glycosaminoglycans from Alzheimer's disease hippocampus have altered capacities to bind and regulate growth factors activities and to bind tau. *PLoS ONE*. 2019;14(1):e0209573.
307. Lee J, et al. Structural determinants of heparin–transforming growth factor- β 1 interactions and their effects on signaling. *Glycobiology*. 2015;25(12):1491–1504.
308. Taieb M, et al. Xylosyltransferase I mediates the synthesis of proteoglycans with long glycosaminoglycan chains and controls chondrocyte hypertrophy and collagen fibers organization of in the growth plate. *Cell Death Dis*. 2023;14(6):355.
309. Ungar D, et al. Characterization of a mammalian Golgi-localized protein complex, COG, that is required for normal Golgi morphology and function. *J Cell Biol*. 2002;157(3):405–415.
310. Ferreira CR, et al. A Recurrent De Novo Heterozygous COG4 Substitution Leads to Saul-Wilson Syndrome, Disrupted Vesicular Trafficking, and Altered Proteoglycan Glycosylation.

Am J Hum Genet. 2018;103(4):553–567.

311. Fernández CJ, Warren G. In Vitro Synthesis of Sulfated Glycosaminoglycans Coupled to Inter-compartmental Golgi Transport. *J Biol Chem.* 1998;273(30):19030–19039.

312. Sasaki K, et al. PGSE Is a Novel Enhancer Regulating the Proteoglycan Pathway of the Mammalian Golgi Stress Response. *Cell Struct Funct.* 2019;44(1):1–19.

313. Dadlani H, et al. Smad and p38 MAP Kinase-mediated Signaling of Proteoglycan Synthesis in Vascular Smooth Muscle. *J Biol Chem.* 2008;283(12):7844–7852.

314. Little PJ, et al. Biosynthesis of Natural and Hyperelongated Chondroitin Sulfate Glycosaminoglycans: New Insights into an Elusive Process. *Open Biochem J.* 2008;2:135–142.

315. Stopschinski BE, et al. Specific glycosaminoglycan chain length and sulfation patterns are required for cell uptake of tau versus α -synuclein and β -amyloid aggregates. *J Biol Chem.* 2018;293(27):10826–10840.

316. Roughley PJ, Mort JS. The role of aggrecan in normal and osteoarthritic cartilage. *J Exp Orthop.* 2014;1(1):8.

317. Delbaere S, et al. *b3galt6* knock-out zebrafish recapitulate β 3GalT6-deficiency disorders in human and reveal a trisaccharide proteoglycan linkage region. *Front Cell Dev Biol.* 2020;8:597857.

318. Nikpour M, et al. Glycosaminoglycan linkage region of urinary bikunin as a potentially useful biomarker for β 3GalT6-deficient spondylodysplastic Ehlers–Danlos syndrome. *JIMD Rep.* 2022;jmd2.12311.

319. Pinhal MAS, et al. Enzyme interactions in heparan sulfate biosynthesis: Uronosyl 5-epimerase and 2-O-sulfotransferase interact in vivo. *Proc Natl Acad Sci USA.* 2001;98(23):12984–12989.

320. Izumikawa T, et al. Involvement of chondroitin sulfate synthase-3 (chondroitin synthase-2) in chondroitin polymerization through its interaction with chondroitin synthase-1 or chondroitin-polymerizing factor. *Biochem J.* 2007;403(3):545–552.

321. Izumikawa T, et al. Identification of Chondroitin Sulfate Glucuronyltransferase as Chondroitin Synthase-3 Involved in Chondroitin Polymerization. *J Biol Chem.* 2008;283(17):11396–11406.

322. Esko JD, Selleck SB. Order Out of Chaos: Assembly of Ligand Binding Sites in Heparan Sulfate. *Annu Rev Biochem.* 2002;71(1):435–471.

323. Han E, et al. Contribution of Proteoglycan Osmotic Swelling Pressure to the Compressive Properties of Articular Cartilage. *Biophys J.* 2011;101(4):916–924.

324. Wang X, et al. Coupling Effect of Water and Proteoglycans on the In Situ Toughness of

Bone. *J Bone Miner Res.* 2016;31(5):1026–1029.

325. Andres JL, et al. Membrane-anchored and soluble forms of betaglycan, a polymorphic proteoglycan that binds transforming growth factor-beta. *J Cell Biol.* 1989;109(6):3137–3145.

326. Andres JL, et al. Binding of two growth factor families to separate domains of the proteoglycan betaglycan. *J Biol Chem.* 1992;267(9):5927–5930.

327. Ornitz DM, Marie PJ. Fibroblast growth factor signaling in skeletal development and disease. *Genes Dev.* 2015;29(14):1463–1486.

328. Reese SP, Underwood CJ, Weiss JA. Effects of decorin proteoglycan on fibrillogenesis, ultrastructure, and mechanics of type I collagen gels. *Matrix Biol.* 2013;32(7–8):414–423.

329. Wilhelm D. Utilisation du modèle cellulaire ATDC5 pour la caractérisation des mécanismes de maturation des procollagènes et de leurs relations avec le processus de minéralisation matricielle. 2019. <https://www.theses.fr/2019LORR0286>. Accessed March 12, 2024.

330. Tang VW. Collagen, stiffness, and adhesion: the evolutionary basis of vertebrate mechanobiology. *Mol Biol Cell.* 2020;31(17):1823–1834.

331. Delbaere S, et al. 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.* 2020;22(1):112–123.

332. Donovan C, et al. Collagen XII regulates stromal wound closure. *Exp Eye Res.* 2023;230:109456.

333. Sun M, et al. Collagen XII is a regulator of corneal stroma structure and function. *Invest Ophthalmol Vis Sci.* 2020;61(5):61.

334. Nair A, et al. Multiple Optical Elastography Techniques Reveal the Regulation of Corneal Stiffness by Collagen XII. *Invest Ophthalmol Vis Sci.* 2022;63(12):24.

335. Sun M, et al. Collagen XII Regulates Corneal Stromal Structure by Modulating Transforming Growth Factor- β Activity. *Am J Pathol.* 2022;192(2):308–319.

336. Lin X. Functions of heparan sulfate proteoglycans in cell signaling during development. *Development.* 2004;131(24):6009–6021.

337. Goetsch KP, Niesler CU. The extracellular matrix regulates the effect of decorin and transforming growth factor beta-2 (TGF- β 2) on myoblast migration. *Biochem Biophys Res Commun.* 2016;479(2):351–357.

338. Scholl ZN, Marszalek PE. AFM-Based Single-Molecule Force Spectroscopy of Proteins. In: Lyubchenko YL, ed. *Nanoscale Imaging*. New York, NY: Springer New York; 2018:35–47.

339. Zhuang S, Linhananta A, Li H. Phenotypic effects of Ehlers-Danlos syndrome-associated mutation on the FnIII domain of tenascin-X: Effects of EDS-Associated Mutation. *Protein Sci.* 2010;19(11):2231–2239.
340. Norman MDA, et al. Measuring the elastic modulus of soft culture surfaces and three-dimensional hydrogels using atomic force microscopy. *Nat Protoc.* 2021;16(5):2418–2449.
341. Vaez M, et al. Modulation of the biophysical and biochemical properties of collagen by glycation for tissue engineering applications. *Acta Biomater.* 2023;155:182–198.
342. Huang S, et al. Quantitative nanohistology of aging dermal collagen. *Front Aging.* 2023;4:1178566.
343. Khattignavong E, et al. Development of a facile method to compute collagen network pathological anisotropy using AFM imaging. *Sci Rep.* 2023;13(1):20173.

Annexes

Publication dans une revue internationale à comité de lecture (en cours de révision)

Diana RM, Jolivet B, Vincourt JB, Hergalant S, Francius G, Karami Y, Khakzad H, Wild R, Bourgeois M, Robert A, Wurtz A, Barreto G, Ramalanjaona N, Onifarasoaniaina R, Font S, Lopin-Bon C, Syx D, Malfait F, Gulberti S, Bui C, Fournel-Gigleux S. *B3GALT6* mutations lead to compromised connective tissue biomechanics in Ehlers-Danlos syndrome. In external peer-review.

Communication orale internationale

Diana RM, Robert A, Vincourt JB, Bui C, Fournel-Gigleux S. Impact of β 1,3-galactosyltransferase 6 deficiency in a rare connective tissue genetic disorder. 8th FEBS Advanced Lecture Course “Matrix pathobiology, signaling and molecular targets”. May 2022, Liménas Chersoníssou, Greece.

Communication affichée nationale

Diana RM, Robert A, Vincourt JB, Fournel-Gigleux S, Bui C. Impact of β 1,3-galactosyltransferase 6 deficiency in a rare connective tissue genetic disorder. Journée scientifique de l'ED BioSE. Mai 2023. Vandoeuvre-lès-Nancy, France.

Les protéoglycanes (PG) sont des macromolécules présentes au niveau des membranes plasmiques cellulaires et abondantes au sein des matrices extracellulaires (MEC) des tissus conjonctifs (TC). Ils sont composés d'une partie protéique sur laquelle sont greffées une à plusieurs chaînes de glycosaminoglycanes (GAG) qui participent activement à la fibrillogenèse des collagènes, aux propriétés biomécaniques des TC et à de nombreux autres processus biologiques tels que la régulation des cascades de signalisation intra- et extracellulaires. La β 1,3-galactosyltransférase 6 (β 3GalT6) est une glycosyltransférase (GT) impliquée de façon précoce dans la biosynthèse des chaînes de GAG. Le syndrome d'Ehlers-Danlos spondylodysplasique (SEDsp) est une maladie génétique rare et sévère des TC due à des mutations du gène *B3GALT6*. Le spectre de symptômes musculo-squelettiques est large et associé à des difficultés de diagnostic. Ce travail de thèse vise à élucider les mécanismes moléculaires sous-jacents au SEDsp par une approche pluridisciplinaire à plusieurs échelles. Une forme tronquée et soluble de la β 3GalT6 ainsi que trois variants pathogènes ont été purifiés dans le but de caractériser leurs paramètres cinétiques : les trois variants causent une perte totale de fonction. Par dynamique moléculaire, nous avons mis en évidence que les trois mutations sont à l'origine de variations d'intensités de liaisons hydrogène et de ponts salins entre des résidus faisant partie de motifs conservés des GT. Une diminution du taux de fluctuations d'une boucle est également observée lorsque l'enzyme est mutée. Nous proposons que ces paramètres contribuent à la perte de fonction. Dans un modèle cellulaire eucaryote invalidé pour le gène *B3GALT6* et dans les fibroblastes de patients, nous avons observé une synthèse résiduelle des chaînes de GAG. Nous avons démontré *in vitro* qu'une autre GT de la synthèse, la glucuronosyltransférase I, est responsable de cette synthèse résiduelle. Toutefois, cette synthèse est insuffisante et conduit à la formation d'une MEC défectueuse. Enfin, grâce à un modèle cellulaire ATDC5 chondrogénique murin, nous avons étudié les conséquences d'une invalidation de *B3galt6* sur les propriétés biomécaniques de la MEC. Nous avons relevé la forte diminution de la glycosylation du collagène de type XII, un PG matriciel, jouant un rôle potentiel dans la pathogénie du SEDsp. Nous relierons également par microscopie à force atomique la perte de fonction de β 3GalT6 à la diminution de l'élasticité de la MEC. Nos travaux apportent un nouveau regard sur la pathogénie du SEDsp et sont un pas vers la compréhension des événements responsables de cette maladie.

Mots-clefs : Syndrome d'Ehlers-Danlos, β 1,3-galactosyltransférase 6, glycosaminoglycane, matrice extracellulaire, fibrillogenèse.

Proteoglycans (PG) are macromolecules found in cell plasma membranes and which are abundant in extracellular matrix (ECM) of connective tissues (CT). They are composed of a core-protein onto which one or more glycosaminoglycan (GAG) chains are attached. They play multiple roles in collagen fibrillogenesis, the biomechanical properties of CT, and in other crucial biological processes such as cell signaling regulation. β 1,3-galactosyltransferase 6 (β 3GalT6) is a glycosyltransferase (GT) involved in the early steps of GAG chains biosynthesis. Spondylodysplastic Ehlers-Danlos syndrome (spEDS) is a rare and severe CT disorder caused by mutations in *B3GALT6*. It presents a wide range of musculoskeletal symptoms, overlapping with other osteochondrodysplasia making diagnosis challenging. The aim of this thesis is to gain new insights into the molecular mechanisms underlying spEDS, using a multiscale and multidisciplinary approach. A truncated, soluble form of β 3GalT6 as well as three pathogenic variants were purified to characterize their kinetic parameters: the three variants cause a total loss of function. Using molecular dynamics, we have shown that mutations cause variations in the intensity of hydrogen bonds and salt bridges between residues belonging to conserved motifs. A decrease in the fluctuations of a loop is also observed when the enzyme is mutated. We propose that these parameters contribute to the loss of function. In an eucaryotic cell model invalidated for the *B3GALT6* gene and in patients' fibroblasts, we observed a drastic reduction but residual GAG chains synthesis. For the first time, we have shown *in vitro* that another enzyme of the pathway is responsible for this residual synthesis. However, the residual GAG synthesis is insufficient and leads to defective ECM synthesis. Finally, using a murine chondrogenic ATDC5 cell model, we studied the consequences of *B3galt6* invalidation on the biomechanical properties of the ECM. A marked reduction in the glycosylation of type XII collagen (Col XII), was observed. This result highlights a possible pathogenic role of this matrix PG in spEDS. Using atomic force microscopy, we also associated the loss of function of β 3GalT6 to the diminution of the matrix elasticity. Our work provides new insight into the pathogenesis of spEDS and is a step towards dissecting the molecular events responsible for this disease.

Key words: Ehlers-Danlos, β 1,3-galactosyltransferase 6, glycosaminoglycan, extracellular matrix, fibrillogenesis.