



**HAL**  
open science

**Les enzymes de biosynthèse des glycosaminoglycanes :  
étude structurale et fonctionnelle de la  $\beta$ 4GalT7  
humaine et caractérisation moléculaire des mutations  
responsables du syndrome progéroïde d'Ehlers-Danlos**

Ibtissam Talhaoui

► **To cite this version:**

Ibtissam Talhaoui. Les enzymes de biosynthèse des glycosaminoglycanes : étude structurale et fonctionnelle de la  $\beta$ 4GalT7 humaine et caractérisation moléculaire des mutations responsables du syndrome progéroïde d'Ehlers-Danlos. Médecine humaine et pathologie. Université Henri Poincaré - Nancy 1, 2010. Français. NNT : 2010NAN10126 . tel-01746345

**HAL Id: tel-01746345**

**<https://hal.univ-lorraine.fr/tel-01746345>**

Submitted on 29 Mar 2018

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

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



## AVERTISSEMENT

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

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

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

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

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

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

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

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

---

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

**Thèse**

Présentée et soutenue publiquement pour l'obtention du titre de

**DOCTEUR DE L'UNIVERSITE HENRI POINCARÉ**

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

par Ibtissam TALHAOUI

**Les enzymes de biosynthèse des glycosaminoglycanes:  
Etude structurale et fonctionnelle de la  $\beta$ 4GalT7 humaine et  
caractérisation moléculaire des mutations responsables du  
syndrome progéroïde d'Ehlers-Danlos**

**Le 10 Décembre 2010**

**Membres du jury :**

**Rapporteurs :** Madame Christelle BRETON, Professeur, CERMAV-CNRS, Grenoble

Monsieur Philippe LASSALLE, Directeur de Recherche INSERM, INSERM U774, Lille

**Examineurs :** Monsieur David BONNAFFE, Professeur, UMR-CNRS-UPS 8182, Orsay

Madame Sylvie FOURNEL-GIGLEUX, Directeur de Recherche INSERM, UMR 7561  
CNRS-Nancy Université, Vandœuvre-lès-Nancy (Directeur de thèse)

Monsieur Mohamed OUZZINE, Directeur de Recherche INSERM, UMR 7561 CNRS-  
Nancy Université, Vandœuvre-lès-Nancy

Madame Brigitte LEININGER-MULLER, Professeur, UMR-S 954 INSERM,  
Vandœuvre-lès-Nancy

Ce travail a été réalisé au sein du  
Laboratoire de Pharmacologie, Physiopathologie et Ingénierie Articulaires  
UMR 7561 - Université Henri Poincaré - Nancy I  
(Directeur: Dr. Jacques Magdalou; Directeur adjoint: Pr. Patrick Netter)

Equipe Pharmacologie Moléculaire, Structure-Fonction  
Responsables: Dr. Sylvie Fournel-Gigleux, Dr. Mohamed Ouzzine

Les études phylogénétiques ont été effectuées en collaboration avec le Dr. Rafael ORIOL  
(INSERM U1004, Villejuif, France).

Les études de modélisation de la  $\beta$ 1,4-galactosyltransférase ont été réalisées en collaboration  
avec le Dr. Guillermo MULLIERT (UMR CNRS 7036 - UHP Nancy 1).

Nous les remercions vivement pour leur contribution.

Ce travail a bénéficié du soutien financier de l'Institut National de la Santé et  
de la Recherche Médicale (INSERM), de la Région Lorraine, de la Communauté  
Urbaine du Grand Nancy, du Ministère de l'Education Nationale de la Recherche  
et de la Technologie, de l'Université Henri-Poincaré et du Centre National de la  
Recherche Scientifique (CNRS)

Il a bénéficié de financement de l'Agence Nationale pour la Recherche  
(GAGNetwork ANR-08-PCVI-0023-01), de la Ligue Régionale contre le Cancer et du  
Laboratoire Européen Associé CNRS-UHP-Nancy I-Université de Dundee (SFGEN)

## *Remerciements*

Je souhaite exprimer tout d'abord mes remerciements aux membres du jury qui ont accepté d'évaluer mon travail de thèse.

A Madame le Professeur, Christelle BRETON, pour avoir accepté la charge de juger cette thèse en qualité de rapporteur. Qu'elle trouve ici l'expression de ma sincère reconnaissance et l'expression de mes remerciements les plus respectueux.

A Monsieur le Docteur Philippe LASSALLE, pour avoir accepté également de juger ce travail de thèse en qualité de rapporteur. Qu'il trouve ici l'assurance de ma sincère reconnaissance et de mon profond respect.

A Monsieur le professeur David BONNAFFE, pour avoir accepté de juger ce travail en qualité d'examineur. Veuillez trouver ici l'expression de ma profonde gratitude.

Mes plus vifs remerciements à Monsieur le Docteur Mohamed OUZZINE, pour ses précieux conseils et pour ses idées scientifiques toujours efficaces et originales. Merci également pour avoir accepté de juger ce travail en tant qu'examineur.

A Madame le Professeur Brigitte LEININGER-MULLER, pour avoir accepté d'être membre de ce jury de thèse. Merci de m'avoir introduit à l'univers passionnant de l'enseignement supérieur et de m'avoir permis de découvrir le métier d'enseignant-chercheur au cours de mon expérience d'ATER.

A Madame le Docteur Sylvie FOURNEL-GIGLEUX, ma Directrice de thèse, qui a su me laisser la liberté nécessaire à l'accomplissement de mes travaux, tout en y gardant un œil critique et avisé. Merci pour avoir cru en moi, même si ça n'a pas été toujours facile, pour avoir orientée mes recherches au bon moment, pour le soutien permanent et pour tout le reste, un grand merci.

A Monsieur le Docteur Jacques MAGDALOU, pour m'avoir accueillie au sein de son Laboratoire, et pour ses conseils, qu'il trouve ici l'expression de ma gratitude.

A Monsieur le Professeur Patrick NETTER, qu'il trouve ici le témoignage de mes sincères remerciements.

Un grand merci à Sandrine pour le temps consacré à la correction de ce manuscrit, j'espère que je n'ai pas trop bousculé ton emploi du temps.

Merci à Lydia pour sa bonne humeur, tu as toujours su mettre une bonne ambiance même quand on était en pleines manip, ça ne t'a jamais empêcher de chanter. Continue comme ça !

Merci à Matthieu pour ton aide précieuse, heureusement que tu as été là. Ça ne fait pas très longtemps depuis que tu appartiens à l'équipe, et pourtant tu as su gagner la confiance et le grand respect de tout le monde. Je n'oublierai pas ta gentillesse, je te souhaite bonne continuation et bonne chance dans la vie.

Je remercie tout particulièrement Dong, pour avoir été là pour moi, pour les nombreuses discussions (de tout et de rien), qu'on a eu dans le *fameux Tram Nancéen*, pour tous les moments qu'on a partagé depuis plus de 5 ans déjà au R... Je te souhaite bon courage et bonne chance pour ton post-doc.

Au *grand* Moustapha, merci pour ta joie de vivre et ta philosophie, tu es quelqu'un qui sait prendre du plaisir en aidant les autres, je te dis tout simplement, chapeau. Pour tes talents de bricoleur avec un peu plus de pratique et ton "*ventilo*" marchera parfaitement la prochaine fois.

A Catherine pour tous les moments qu'on a partagé ensemble. Au *petit* Moustapha, le doctorant de l'équipe, je te souhaite bonne chance pour ta thèse. A Nick, le dernier arrivé dans l'équipe, merci pour tes conseils et surtout pour tes fous rires contagieux, même si cela m'empêchait de manger. A Xianqing, je te souhaite beaucoup de chance et de courage, je sais que ce n'est pas facile car je suis déjà passée par là, mais dis toi que la prochaine soutenance sera la tienne. Je crois que je vais faire bref car la liste est interminable, un grand merci à tous les étudiants et personnels que j'ai côtoyé à l'UMR7561.

Merci à mes amis de *fac* et d'enfance, en France, Allemagne, Belgique, Espagne et Maroc, pour avoir toujours pris de mes nouvelles, pour avoir toujours été là pour moi et pour leurs encouragements, je sais que nos chemins se sont séparés mais vive l'internet !

Enfin, à mes parents qui m'ont toujours soutenu et encouragé pour aller jusqu'au bout. Les mots ne suffiront pour vous remercier. A ma petite famille Hicham, Mohamed, Souhayla, Houssam et la petite Sarah, vous me manquez beaucoup.

A tous ceux que je n'ai pas cité et qui se reconnaîtront dans ces quelques lignes, tout simplement, merci.

*A mes parents*

*A mes frères et sœurs*

*A tous ceux qui me sont chers*

*Dans un voyage, le plus long est de franchir le seuil*

VARRON

# *Sommaire*



<b>Sommaire</b> .....	<b>1</b>
Liste des illustrations.....	7
Liste des abréviations.....	11

## **Introduction bibliographique**

<b>Chapitre 1: Protéoglycanes et glycosaminoglycanes</b> .....	<b>14</b>
I. INTRODUCTION.....	14
II. STRUCTURE ET ORGANISATION DES PGs.....	14
II.1. La protéine "core".....	14
II.2. Les chaînes de GAGs.....	15
II.2.1. L'héparine et les héparane sulfates (Hp/HS).....	17
II.2.2. Les kératane sulfates (KS).....	17
II.2.2.1. La classe des kératane sulfates de type I (KSI).....	17
II.2.2.2. La classe des kératane sulfates de type II (KSII).....	18
II.2.2.3. La classe des kératane sulfates de type III (KSIII).....	19
II.2.3. Les chondroïtine sulfates (CS).....	19
II.2.4. Les dermatane sulfates (DS).....	19
II.2.5. L'acide hyaluronique (AH).....	20
II.3. Les principales fonctions biologiques des GAGs.....	20
III. CLASSIFICATION DES PGs.....	22
III.1. Les PGs membranaires.....	22
III.1.1. La famille des syndécans.....	22
III.1.1.1. L'organisation des syndécans.....	22
III.1.1.2. Distribution et fonctions biologiques des syndécans.....	24
<i>La signalisation cellulaire du syndécan via ses chaînes d'HS</i> .....	24
<i>La signalisation du syndécan via la protéine "core", exemple de l'implication du syndécan-2 dans la voie de signalisation du TGF-<math>\beta</math></i> .....	26
III.1.2. La famille des CD44.....	27
III.1.2.1. L'organisation des CD44.....	27
III.1.2.2. Les fonctions des CD44.....	28
III.1.3. La famille des glypicans.....	29
III.1.3.1. L'organisation des glypicans.....	29

III.1.3.2. Les fonctions des glypicans.....	30
III.2. Les PGs intracellulaires.....	31
III.3. Les PGs extracellulaires.....	31
III.3.1. L'agrécane.....	32
III.3.1.1. L'organisation structurale de l'agrécane.....	33
III.3.1.2. Les fonctions de l'agrécane.....	33
III.3.2. Les PGs de la membrane basale.....	34
III.3.2.1. Le perlecan.....	34
III.3.2.1.a. L'organisation structurale du perlecan.....	35
III.3.2.1.b. Les fonctions du perlecan.....	35
III.3.2.2. Le bamacan.....	36
III.3.2.3. L'agrine.....	37
III.3.2.4. Le collagène XV et le collagène XVIII.....	38
III.3.3. La famille des <i>Small Leucine Rich proteoglycans</i> (SLRPs).....	40
III.3.3.1. Les principaux SLRPs.....	43
III.3.3.1.1. La fibromoduline.....	43
III.3.3.1.2. Le lumican.....	43
III.3.3.1.3. Le biglycan.....	44
III.3.3.1.4. La décorine.....	45
III.3.4. Les PGs circulants: exemple de l'endocan.....	47
<b>Chapitre 2 : Les glycosyltransférases (GTs).....</b>	<b>49</b>
Généralités.....	49
I. CLASSIFICATION DES GTs.....	49
II. LE REPLIEMENT DES GTs, "FOLDS".....	50
II.1. La structure de type GT-A.....	50
II.2. La topologie de type GT-B.....	51
III. LES GALACTOSYLTRANSFERASES (GalTs).....	52
III.1. La famille des $\beta$ 4GalTs.....	53
III.2. La $\beta$ 4GalT1 bovine.....	55
III.3. Les $\beta$ 4GalT7s.....	57
III.3.1. Les propriétés enzymatiques.....	57
III.3.2. La $\beta$ 4GalT7 de drosophile ( <i>Drosophila melanogaster</i> ).....	61

<b>Chapitre 3 : Les voies de biosynthèse des PGs.....</b>	<b>64</b>
I. LA BIOSYNTHESE DES PGS.....	64
II. LA FORMATION DE L'AMORCE TETRASACCHARIDIQUE.....	65
III. LA POLYMERISATION DES CHAINES DE GAGs.....	67
III.1. La polymérisation des chaînes de CS/DS.....	67
III.2. La polymérisation des chaînes d'Hp/HS.....	68
IV. LES REACTIONS DE MODIFICATION DES CHAINES DE GAGs DES PGs	69
IV.1. Les réactions de modification de l'amorce tétrasaccharidique.....	69
IV.1.1. La phosphorylation en C2 du xylose de l'amorce tétrasaccharidique.....	70
IV.1.2. La sulfatation des galactoses de l'amorce tétrasaccharidique.....	71
IV.2. Les modifications des différentes chaînes de GAGs.....	72
IV.2.1. La sulfatation des Hp/HS.....	73
IV.2.1.1. La N-déacétylation/N-sulfatation des HS.....	73
IV.2.1.2. La O-sulfatation des Hp/HS.....	73
IV.2.2. La sulfatation des CS/DS.....	74
V. LA DEGRADATION ET LE RECYCLAGE OU "TURNOVER" DES PGs.....	75

## **Présentation de l'étude**

Présentation de l'étude.....	77
------------------------------	----

## **Matériels et méthodes..... 82**

### **Matériels..... 82**

1. Les lignées cellulaires et réactifs de culture cellulaire.....	82
2. Les réactifs de biologie moléculaire.....	82
3. Les réactifs chimiques.....	82
4. Les réactifs radiomarqués.....	83
5. Les réactifs biochimiques.....	83

### **Méthodes**

#### **expérimentales..... 83**

I. CLONAGE DE LA $\beta$ 4GalT7 HUMAINE ET CONSTRUCTION DES VECTEURS D'EXPRESSION.....	83
1. Construction du vecteur d'expression pcDNA- $\beta$ 4GalT7.....	84
2. Construction du vecteur d'expression pET- $\beta$ 4GalT7.....	84

II. CLONAGE DE LA DECORINE HUMAINE ET CONSTRUCTION DU VECTEUR D'EXPRESSION pcDNA-DECORINE.....	85
III. MUTAGENESE DIRIGEE PAR PCR.....	85
IV. EXPRESSION HETEROLOGUE DE LA $\beta$ 4GalT7 RECOMBINANTE HUMAINE.....	86
1. Expression de la $\beta$ 4GalT7 dans des cellules eucaryotes.....	86
1.1. Transfection transitoire des cellules eucaryotes (HeLa et CHO-pgsB618) avec la $\beta$ 4GalT7.....	86
2. Expression stable de la décorine dans des cellules CHO-pgsB618 .....	86
a. Purification de la décorine par chromatographie d'affinité ( <i>Immobilized Metal Affinity Chromatography</i> , IMAC).....	87
<u>Principe</u> .....	87
<u>Protocole de purification</u> .....	87
1.3. Analyse de l'expression protéique par immunoblot.....	87
<u>Séparation des protéines par électrophorèse</u> .....	87
<u>Transfert des protéines sur membrane de polyvinylidène difluorure (PVDF) et immunodétection</u> .....	87
2. Expression de la $\beta$ 4GalT7 sous forme soluble dans les bactéries BL21.....	89
2.1. Purification de la protéine par chromatographie d'affinité.....	89
<u>Principe</u> .....	89
<u>Protocole de purification</u> .....	89
3. Dosage de protéines.....	90
4. Mesure de l'activité enzymatique.....	90
4.1. Mesure de l'activité enzymatique par CLHP en phase inverse.....	91
4.2. Mesure de l'activité enzymatique par CCM selon la méthode de Bansal et Gessner (1980).....	91
5. Détermination des constantes enzymatiques $V_{max}$ , $K_m$ et $k_{cat}$ .....	91
6. Exploitation des données.....	92
V. EXTRACTION ET ANALYSE DES GAGs PAR INCORPORATION DU $[^{35}\text{S}]\text{SO}_4$ .....	93
1. Analyse du taux de biosynthèse des GAGs endogènes.....	93
2. Analyse des chaînes de GAGs synthétisées sur la protéine "core" de la décorine par SDS-PAGE et autoradiographie.....	94

3. Analyse de l'initiation de la biosynthèse des chaînes de GAGs sur les xylosides exogènes par SDS-PAGE.....	94
4. Marquage des cellules CHO-pgsB618 par immunofluorescence indirecte.....	95
4.1. Analyse de l'expression des chaînes d'HS au niveau de la membrane cellulaire.....	95
4.2. Analyse de l'expression et de la localisation cellulaire de la $\beta$ 4GalT7 par immunofluorescence indirecte.....	95

## Résultats

### Publication n°1

<b>Caractérisation structurale et fonctionnelle de la <math>\beta</math>4GalT7 humaine, et exploration du site actif de l'enzyme.....</b>	<b>96</b>
I. ANALYSE PHYLOGENETIQUE DE LA FAMILLE DES $\beta$ 4GalT7s.....	96
II. EXPLORATION DES ACIDES AMINES DES MOTIFS <sup>163</sup> DVD <sup>165</sup> ET <sup>221</sup> FWGWGREDDE <sup>23</sup> par mutagenèse dirigée.....	97
II.1. Activité galactosyltransférase <i>in vitro</i> de la $\beta$ 4GalT7 recombinante humaine et propriétés cinétiques de l'enzyme sauvage et des mutants.....	97
III.2. Etude de l'effet de la mutation du Trp224 sur la spécificité de substrat de la $\beta$ 4GalT7 humaine.....	103
II.2.1. Effet de la mutation W224H sur la spécificité de la $\beta$ 4GalT7 vis-à-vis su substrat donneur.....	103
II.2.2. Effet de la mutation W224H sur la spécificité de la $\beta$ 4GalT7 vis-à-vis du substrat accepteur.....	104
II.3. Activité galactosyltransférase <i>ex vivo</i> de la $\beta$ 4GalT7 humaine sauvage et mutée.....	107
II.3.1. Substrat accepteur xylosidique exogène: 4-MOX.....	107
II.3.2. Substrat accepteur endogène: protéine "core" de la décorine.....	108
II.4. Caractérisation <i>ex vivo</i> des propriétés fonctionnelles de la $\beta$ 4GalT7 sauvage et des mutants.....	108
III. ETUDE DU ROLE DU PREMIER RESIDU HIS (His257) DU MOTIF CONSERVE HXH DANS LA FORMATION DE LIAISONS DE COORDINATIONS AVEC LES CATIONS DIVALENTS.....	111
III.1. Etudes <i>in vitro/ex vivo</i> .....	111
III.2. Caractérisation de l'activation de la $\beta$ 4GalT7 sauvage et du mutant H257A	

par les cations métalliques.....	113
<b><u>Publication n°2</u></b>	
<b>Caractérisation enzymatique des mutations de la <math>\beta</math>4GalT7 humaine associées au syndrome progéroïde d'Ehlers-Danlos.....</b>	<b>115</b>
1. Analyse des effets des mutations du syndrome progéroïde d'Ehlers-Danlos sur l'activité enzymatique de la $\beta$ 4GalT7 humaine et ses propriétés enzymatiques.....	115
2. Modélisation moléculaire de la $\beta$ 4GalT7 humaine.....	119
3. Analyse des conséquences fonctionnelles des mutations associées à la forme progéroïde du syndrome d'Ehlers-Danlos, sur la synthèse des GAGs <i>ex vivo</i> dans les cellules CHO-pgsB618.....	119
<b><u>Publication n°3</u></b>	
<b>Etude des mécanismes de régulation des voies de biosynthèse des GAGs.....</b>	<b>122</b>
1. Analyse du profil de méthylation des gènes codant pour les GTs et les STs impliquées dans la synthèse et la modification des HS, dans les cellules HEMC.....	122
2. Analyse des conséquences de la déméthylation des gènes d'intérêt par évaluation de leur taux d'expression dans des cellules de chondrosarcome humain après traitement à la 5-Aza-dc.....	123
3. Analyse de l'effet du traitement des cellules HEMC par la 5-Aza-dc sur la sulfatation des chaînes de GAGs des PGs cellulaires.....	124
4. Comparaison entre l'effet de la surexpression des 3-OSTs et le traitement par la 5-Aza-dc sur l'expression des HSPGs dans les cellules HEMC.....	125
5. Effet des modifications de la sulfatation des HSPGs sur le comportement cellulaire des HEMC.....	126
<b>Discussion.....</b>	<b>128</b>
<b>Conclusions et perspectives.....</b>	<b>141</b>
<b>Références bibliographiques.....</b>	<b>145</b>
<b>Annexes.....</b>	<b>171</b>

*Liste des  
illustrations*

## Liste des illustrations

### Figures

- Figure 1 (page 16): Représentation schématique des motifs disaccharidiques répétitifs des différents types de GAGs
- Figure 2 (page 18): Organisation structurale schématique des kératane sulfates de type I (KSI)
- Figure 3 (page 18): Organisation structurale schématique des kératane sulfates de type II (KSII)
- Figure 4 (page 19): Organisation structurale schématique des kératane sulfates de type III (KSIII)
- Figure 5 (page 22): Représentation schématique des membres de la famille des syndécans
- Figure 6 (page 23): Représentation schématique des domaines fonctionnels des syndécans
- Figure 7 (page 25): Présentation schématique de la signalisation du syndécan *via* la fixation des facteurs de croissance par les chaînes d'HS
- Figure 8 (page 26): Présentation schématique de la modulation de la signalisation du TGF- $\beta$  par le syndécan-2
- Figure 9 (page 27): Présentation schématique de la structure de la protéine CD44
- Figure 10 (page 29): Présentation schématique de la structure des glypicans
- Figure 11 (page 32): Représentation schématique de la MEC du cartilage articulaire
- Figure 12 (page 33): Représentation schématique de l'organisation d'un monomère d'agrécane et de ses domaines structuraux
- Figure 13 (page 35): Présentation schématique de la structure du perlecan
- Figure 14 (page 36): Présentation schématique du bamacan
- Figure 15 (page 37): Présentation schématique de l'agrine
- Figure 16 (page 39): Présentation schématique de l'organisation du Col XV chez l'homme
- Figure 17 (page 42): Présentation schématique de la famille des SLRPs
- Figure 18 (page 45): Représentation schématique de la superposition des dimères de la décorine et du biglycan



- Figure 19 (page 46): Représentation schématique de l'interaction entre les SLRPs et les fibres de collagène
- Figure 20 (page 50): Mécanismes réactionnels enzymatique de type glycosyltransférase avec rétention ou inversion de la configuration anomérique du sucre donneur
- Figure 21 (page 51): Principales structures 3D des GTs
- Figure 22 (page 56): Structure cristallographique du domaine catalytique de la  $\beta$ 4GalT1 bovine en présence de l'UDP-Gal et d'ions  $Mn^{2+}$
- Figure 23 (page 57): Représentation schématique de la réaction enzymatique catalysée par la  $\beta$ 4GalT7
- Figure 24 (page 58): Comparaison schématique de l'organisation des membres de la famille des  $\beta$ 4GalTs chez l'homme
- Figure 25 (page 59): Analyse de la relation phylogénétique entre les  $\beta$ 4GalTs humaines, la  $\beta$ 4GalT7 de drosophile et de *sgv-3*
- Figure 26 (page 62): Organisation cristallographique de la fixation du manganèse et de l'UDP sur la  $\beta$ 4GalT7 de drosophile
- Figure 27 (page 63): Structure cristallographique de la  $\beta$ 4GalT7 de drosophile en présence d'ions  $Mn^{2+}$  et d'une molécule d'UDP
- Figure 28 (page 65): Les enzymes impliquées dans la biosynthèse de l'amorce tétrasaccharidique commune aux CS/DS et Hp/HS
- Figure 29 (page 68): Représentation schématique des voies d'élongation des chaînes CS/DS et des enzymes impliquées dans ces réactions
- Figure 30 (page 69): Représentation schématique de la polymérisation des chaînes d'Hp/HS et enzymes impliquées dans ces réactions
- Figure 31 (page 70): Modification de la région de liaison des chaînes de GAGs
- Figure 32 (page 72): Réactions de modifications des chaînes d'HS et de CS/DS
- Figure 33 (page 84): Schéma de la construction GST- $\beta$ 4GalT7 recombinante, exprimée dans les cellules BL21
- Figure 34 (page 88): Schéma du protocole de purification de la décorine par le kit MagnHis Protein Purification System de Promega
- Figure 35 (page 102): Mesure de l'activité enzymatique de la  $\beta$ 4GalT7 sauvage et mutée en position 224, 227 ou 228, exprimées dans la bactérie et purifiées par chromatographie d'affinité

- Figure 36 (page 106): Structure chimique des substrats accepteurs testés
- Figure 37 (page 110): Effet de la surexpression de la  $\beta$ 4GalT7 sauvage et mutée sur l'initiation de la biosynthèse des GAGs dans des cellules CHO-pgsB618 traitées au 4-MOX
- Figure 38 (page 111): Effet de la surexpression de la  $\beta$ 4GalT7 sauvage et mutée sur la biosynthèse des chaînes de GAGs sur la décorine exprimée dans des cellules CHO-pgsB618
- Figure 39 (page 112): Détermination de l'effet de la mutation conservative et non conservative du résidu His257 sur l'activité GalT *in vitro*
- Figure 40 (page 113): Détermination de l'effet de la mutation l'His257 sur l'initiation de la synthèse des chaînes de GAGs *ex vivo*
- Figure 41 (page 114): Activation de la  $\beta$ 4GalT7 sauvage humaine et du mutant H257A par les cations divalents
- Figure 42 (page 116): Analyse du taux d'expression de la  $\beta$ 4GalT7 sauvage et des mutants Ehlers-Danlos, par immunoblot
- Figure 43 (page 118): Mesure de l'activité enzymatique, *in vitro*, de la  $\beta$ 4GalT7 sauvage et des mutants Ehlers-Danlos exprimés dans la bactérie *E. coli*

## Tableaux

- Tableau 1 (page 21): Les principales fonctions biologiques des chaînes de GAGs
- Tableau 2 (page 30): La famille des glypicans chez l'homme
- Tableau 3 (page 39): Exemples de fonctions du Col XV
- Tableau 4 page (41): Classification de la famille des SLRPs
- Tableau 5 (page 54): Principales caractéristiques fonctionnelles et distribution tissulaire des  $\beta$ 4GalTs
- Tableau 6 (page 76): Les différentes sulfotransférases impliquées dans les réactions de modification des chaînes de GAs
- Tableau 7 (page 101): Tableau de purification de la  $\beta$ 4GalT7 recombinante humaine exprimée dans les cellules *E. coli*
- Tableau 8 (page 117): Tableau de purification de la protéine GST-R270C recombinante humaine exprimée dans les cellules *E. coli*

*Liste des  
abréviations*

## Liste des abréviations

5-Aza-dc	5-Aza-2'-déoxycytidine
ACN	ACétoNitrile
ADN	Acide DésoxyriboNucléique
ADNc	Acide DésoxyriboNucléique complémentaire
AH	Acide Hyaluronique
ARN	Acide RiboNucléique
ARNm	Acide RiboNucléique messenger
ASB	Albumine de Sérum Bovin
BGN	Biglycan
CASK	<i>Calcium/Calmodulin-dependent Serine Protein Kinase</i>
CAZy	Carbohyrate Active enZymes
Col	Collagène
CPC	Chlorure de CétylPyridinium
CpG	Cytidine phosphodiester Guanosine
CSPG	ProtéoGlycanes à Chondroïtine Sulfates
3D	Tridimensionnel
DEAE	DiEthylAminoEthyl
DCN	Décorine
DDSH	Dysplasie Dysségmentale de type Silverman-Handmaker
DMEM	<i>Dulbecco's Modified Eagle's Medium</i>
DNA	DeoxyRibonucleic Acid
DNMT	DNA MethylTransferase
<i>E. coli</i>	<i>Escherichia coli</i>
ED	Ehlers-Danlos
EDS	<i>Ehlers-Danlos Syndrome</i>
EGF	<i>Epidermal Growth Factor</i>
EST	<i>Expression Sequence Tag</i>
EHS	Engelbreth-Holm-Swarm
EXT	Exostose
FAK	<i>Focal Adhesion Kinase</i>
FGF	<i>Fibroblast Growth Factor</i>
FGFR	<i>Fibroblast Growth Factor Receptor</i>
FM	Fibromoduline
Gal	Galactose
GalNAc	N-AcétylGalactosamine
$\beta$ 3GalT	$\beta$ 1,3-GalactosylTransfêrase
$\beta$ 4GalT	$\beta$ 1,4-GalactosylTransfêrase
gB	glycoprotéine de type B
gC	glycoprotéine de type C
gD	glycoprotéine de type D
Glc	Glucose

GlcA	Acide Glucuronique
GlcAT-I	$\beta$ 1,3-GlucuronosylTransférase
GlcNAc	N-AcétylGlucosamine
GST	Glutathion-S-Transférase
GPC	Glypican
GPI	Glycosyl-PhosphatidylInositol
HB-GAM	<i>Heparin-Binding Growth-Associated-Molecule</i>
HEPES	Tris [hydroxyméthyl]aminoéthane, acide N-(2 HydroxyEthyl)Pipérazine-N'-(2-Ethane) Sulfonique
Hp	Héparine
HS	Héparane Sulfates
HSPG	ProtéoGlycanes à Héparane Sulfate
HSV	<i>Herpes Simplex Virus</i>
IdoA	Acide Iduronique
IgG	Immunoglobuline de classe G
IPTG	Isopropyl $\beta$ -D-galactopyranoside
KS	Kératane sulfates
KO	<i>Knock-out</i>
LA	$\alpha$ -Lactalbumine
LDL	<i>Low Density Lipoprotein</i>
LP	<i>Link protein</i>
LRR	<i>Leucine Rich Repeat</i>
LUM	Lumican
MB	Membrane Basale
MEC	Matrice ExtraCellulaire
MOPS	acide N-Morpholino-3-Propane Sulfonique
4-MOX	4-MéthylOmbélliférone-7- $\beta$ -D-Xyloside
NC	Non collagénique
NCAM	<i>Neural Cell Adhesion Molecule</i>
NDST	N-Déacétylase/N-SulfoTransférase
NS	Non-Sulfaté
PBS	<i>Phosphate Buffered Saline</i>
PCR	<i>Polymerase Chain Reaction</i>
Pgp-1	<i>Phagocytic glycoprotein-1</i>
PG	ProtéoGlycane
PKC	Protéine Kinase C
PMSF	PhénylMéthylSulfonylFluoride
PRELP	Prolargine
PSD	<i>Post-Synaptic Density</i>
PVDF	PolyVinylidène DiFluorure
RE	Reticulum Endoplasmique
RNA	RiboNucleic Acid
SDS-PAGE	<i>Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis</i>
SEA	<i>Sperm protein, Enterokinase, Agrin</i>

SGBS	Simpson-Golabi-Behmel
SLRP	<i>Small Leucine Rich Proteoglycan</i>
SSJI	Syndrome de Schwartz-Jampel de type I
ST	SulfoTransférase
SVF	Sérum de Veau Foetal
TBST	<i>Tris Buffer Saline Tween</i>
TGF	<i>Transforming Growth Factor</i>
TNF	<i>Tumor Necrosis Factor</i>
VEGF	<i>Vascular Epidermal Growth Factor</i>
Xyl	Xylose
XT	XylosylTransférase

*Introduction*  
*bibliographique*



# ***Chapitre 1: Protéoglycanes et glycosaminoglycanes***

## **I. INTRODUCTION**

Les protéoglycanes (PGs) sont des macromolécules hydrophiles qui se composent d'une protéine porteuse (protéine "core") sur laquelle sont fixées une ou plusieurs chaînes de carbohydrates appelées glycosaminoglycanes (GAGs). L'étude des PGs a débuté au cours des années 1920 et s'est accélérée à partir des années 1950. Ce sont des molécules remarquablement complexes du fait de leur double nature biochimique, de leur localisation cellulaire variable et des nombreuses modifications qu'elles peuvent subir. La fonction des PGs dépend à la fois de la protéine "core" et des chaînes de GAGs. Les études réalisées au cours des années 1970 ont démontré que les PGs sont situés majoritairement à la surface des cellules et dans les matrices extracellulaires (MEC). Cette localisation stratégique leur permet de jouer un rôle clé dans de multiples processus physiologiques et pathologiques. En effet, ils interagissent avec de nombreuses molécules telles que les cytokines et différents facteurs de croissance. Par conséquent, les PGs contrôlent des fonctions vitales telles que la migration, l'adhésion, la prolifération et la signalisation cellulaires, et sont également impliqués dans l'angiogenèse et l'inflammation. Ces molécules suscitent un intérêt grandissant. En effet, de nombreuses études ont mis en évidence l'implication des PGs et/ou des GAGs dans de multiples pathologies tels que les cancers, l'athérosclérose, les néphropathies, la maladie d'Alzheimer et les arthropathies, d'où l'importance de la compréhension des mécanismes d'action de ces molécules, de leurs voies de biosynthèse et de dégradation.

## **II. STRUCTURE ET ORGANISATION DES PGs**

### **II.1. La protéine "core"**

La protéine "core" présente une taille variant de 30 à plus de 500 kDa (Poole, 1986; Ruoslahti, 1988). Elle possède des sites d'ancrage pour les chaînes de GAGs et est responsable de la localisation des PGs. Des études récentes ont identifié plus de trente protéines "core" différentes. Ces protéines contiennent également des domaines qui exercent des activités biologiques particulières (Iozzo, 1998), participant aux multiples fonctions des PGs.

## II.2. Les chaînes de GAGs

Les chaînes de GAGs sont présentes en nombre variable, généralement entre 1 et 100 (Kjellén and Lindahl, 1991; Silbert and Sugumaran, 1995). Ce sont des polysaccharides linéaires et hétérogènes, constitués d'unités disaccharidiques répétitives. Ce motif disaccharidique est généralement composé (sauf pour les kératane sulfates (KS)) d'un acide hexuronique (acide glucuronique (GlcA) ou acide iduronique (IdoA)) et d'une hexosamine (N-acétylglucosamine (GlcNAc) ou N-acétylgalactosamine (GalNAc)). Le nombre et la longueur des chaînes de GAGs sont très variables, auxquels s'ajoutent des modifications structurales telles que des N-sulfatations, des O-sulfatations et des épimérisations. La présence, au sein des chaînes de GAGs, d'acide uronique et de groupements sulfates confèrent aux PGs une forte charge négative.

Les GAGs sont classés en cinq grandes familles selon la nature de l'unité disaccharidique. On distingue les héparine/héparane sulfates (Hp/HS), les KS, les chondroïtine/dermatane sulfates (CS/DS) et l'acide hyaluronique (AH).

A l'exception de l'AH, tous les GAGs sont attachés par une liaison covalente à une protéine "core" pour former un PG.

La Figure 1 montre deux types de GAGs, les GAGs sulfatés: Hp/HS, CS/DS et KS; et les GAGs non sulfatés: AH.

Type de GAG	Acide hexuronique ou iduronique	Galactose	Hexosamine	Unité disaccharidique
Héparine /Héparane sulfates (Hp/HS)	Acide-D-glucuronique (GlcA)	-	N-acétylglucosamine (GlcNAc)	<p>GlcA <math>\beta(1\rightarrow4)</math> GlcNAc <math>\alpha(1\rightarrow4)</math></p>
	Acide-L-iduronique (IdoA)			<p>IdoA <math>\alpha(1\rightarrow4)</math> GlcNAc <math>\alpha(1\rightarrow4)</math></p>
Kératane sulfates (KS)	-	Galactose (Gal)	N-acétylglucosamine (GlcNAc)	<p>Gal <math>\beta(1\rightarrow4)</math> GlcNAc <math>\beta(1\rightarrow3)</math></p>
Chondroïtine sulfates (CS)	Acide-D-glucuronique (GlcA)	-	N-acétylgalactosamine (GalNAc)	<p>GlcA <math>\beta(1\rightarrow3)</math> GalNAc <math>\beta(1\rightarrow4)</math></p>
Dermatane sulfates (DS)	Acide-D-glucuronique (GlcA) Acide-L-iduronique (IdoA)	-	N-acétylgalactosamine (GalNAc)	<p>IdoA <math>\alpha(1\rightarrow3)</math> GalNAc <math>\beta(1\rightarrow4)</math></p>
Acide hyaluronique (AH)	Acide-D-glucuronique (GlcA)	-	N-acétylglucosamine (GlcNAc)	<p>GlcA <math>\beta(1\rightarrow3)</math> GlcNAc <math>\beta(1\rightarrow4)</math></p>

**Figure 1. Représentation schématique des motifs disaccharidiques répétitifs des différents types de GAGs**

### **II.2.1. L'héparine et les héparane sulfates (Hp/HS)**

L'Hp a été découverte en 1916 par Mc Lean. C'est le GAG le plus sulfaté, et par conséquent, celui qui porte la plus forte charge négative. Il est exclusivement présent dans les mastocytes au niveau des granules de stockage cytoplasmiques (Nader *et al.*, 1980, 1999; Straus *et al.*, 1982). En revanche, les HS sont généralement moins sulfatés et sont distribués de façon ubiquitaire, chez les vertébrés et les invertébrés, à la surface cellulaire et dans les MEC (Cassaró and Dietrich, 1977). La protéine "core" de l'Hp est la serglycine, tandis que les PGs à HS (HSPGs) présentent différents types de squelettes peptidiques (Kjellén, 1991; David, 1993; Iozzo, 1994). L'Hp et les HS sont des polysaccharides linéaires qui présentent des similarités structurales. Ils sont composés d'une D-glucosamine (GlcN) liée à un acide D-glucuronique (GlcA) ou un acide L-iduronique (IdoA), par la formation d'une liaison de type  $\alpha$ 1,4. Au cours des processus de biosynthèse des GAGs, ce motif disaccharique va connaître des modifications très diverses. En effet, la GlcN peut être N-acétylée (GlcNAc) ou sulfatée (GlcNS), ou encore O-sulfatée en position 6 (GlcNAc(6S) et GlcN(6S)), et plus rarement O-sulfatée en position 3 (GlcNS(3S) et GlcNS(3,6S)). Il est plus rare, mais possible, de trouver la GlcN sous forme modifiée (GlcNH<sub>2</sub>). Quant au GlcA, il peut être épimérisé en position C5 pour donner l'IdoA, ou O-sulfaté en position 2 (GlcA2S et IdoA2S). La charge négative des groupements sulfates des Hp/HS constitue un élément clé de l'interaction entre les chaînes de GAGs et d'autres macromolécules et médiateurs solubles.

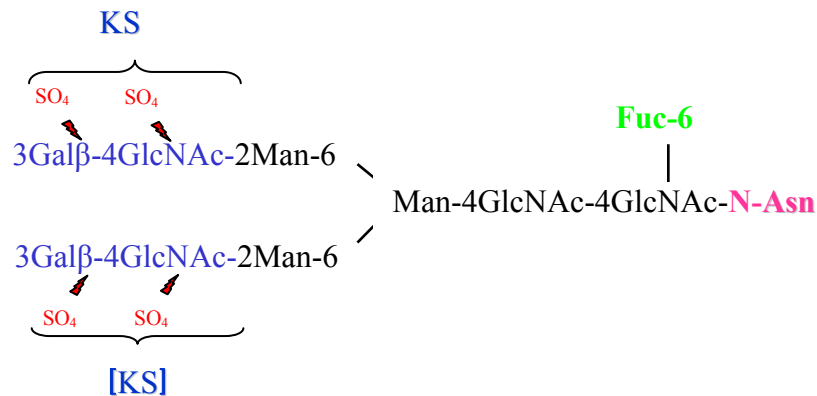
### **II.2.2. Les kératane sulfates (KS)**

Les chaînes de KS ont été extraites à partir de la cornée et identifiées pour la première fois en 1939 (Suzuki *et al.*, 1939). La structure de ces GAGs n'a été élucidée que beaucoup plus tard en 1953 par Karl Meyer, qui leur a attribué le nom de « kératosulfates » (Meyer *et al.*, 1953). L'unité disaccharidique répétitive des KS est constituée d'un galactose (Gal) lié à une GlcNAc pour former le disaccharide: 3Gal $\beta$ 1,4GlcNAc. Chaque sucre peut être sulfaté en position C6. Selon la structure de l'oligosaccharide qui permet la liaison des KS à la protéine, on distingue trois classes de KS: KSI, KSII et KSIII.

#### **II.2.2.1. La classe des kératane sulfates de type I (KSI)**

Les KS de type I sont attachées à une asparagine (Asn) de la protéine "core" par une liaison N-glycosidique. On les trouve aussi bien dans la cornée que dans le cartilage. En effet, la fibromoduline, la prolargine (PRELP) et l'ostéo-cadhérine localisées dans le cartilage et dans

l'os portent des chaînes de KSI. Un fucose peut éventuellement se fixer sur le premier résidu GlcNAc lié à l'Asn comme indiqué dans la Figure 2.

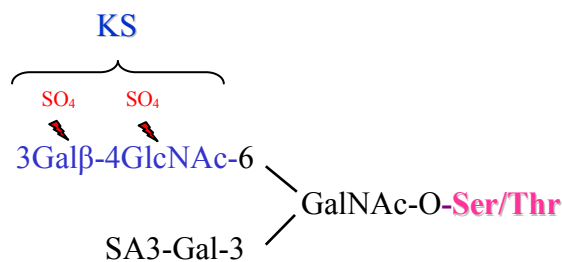


**Figure 2. Organisation structurale schématique des kératane sulfates de type I (KSI)**

L'oligosaccharide qui lie les chaînes de KSI à l'Asn de la protéine "core" peut porter un fucose lié à la première GlcNAc. Les parenthèses indiquent que ce KS peut être absent. Les deux sucres du motif disaccharidique peuvent subir des sulfatations en C6. Asparagine (Asn), Galactose (Gal), N-acétylglucosamine (GlcNAc), Mannose (Man), Fucose (Fuc).

### II.2.2.2. La classe des kératane sulfates de type II (KSII)

Ce groupe de GAGs est toujours lié à une sérine ou une thréonine (Ser/Thr) de la protéine "core" par une liaison O-glycosidique (O-GalNAc) (Figure 3). Les KSII se lient exclusivement à la protéine "core" de l'agrécan, qui représente le PG majoritaire du cartilage. Les KSII sont des GAGs hautement sulfatés (Nieduszynski *et al.*, 1990) et peuvent porter des acides sialiques (Funderburgh, 2000).

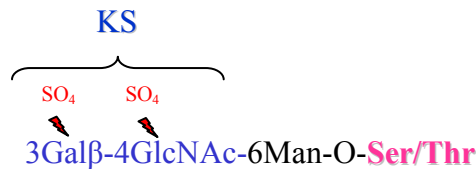


**Figure 3. Organisation structurale schématique des kératane sulfates de type II (KSII)**

Ces KS se fixent sur un acide aminé de type sérine (Ser) ou thréonine (Thr) de la protéine "core", *via* la formation d'une liaison O-glycosidique. Un acide sialique (SA) peut être lié au galactose (Gal) de la région de liaison des KSII.

### II.2.2.3. La classe des kératane sulfates de type III

Les KS de type III ont été identifiés plus récemment dans le cerveau (Krusius *et al.*, 1986). Ils présentent la structure la plus simple des chaînes de KS. Les KSIII sont attachés au mannose lui-même lié à une Ser (ou Thr) de la protéine "core" par une liaison O-glycosidique (Figure 4).



**Figure 4. Organisation structurale schématique des kératane sulfates de type III (KSIII)**

Les chaînes de KS se fixent sur la sérine (Ser) ou la thréonine (Thr) de la protéine "core" par une liaison O-glycosidique de type O-mannose (Man). Galactose (Gal), N-acétylglucosamine (GlcNAc).

### II.2.3. Les chondroïtine sulfates (CS)

L'unité disaccharidique répétitive des CS est composée d'un GlcA attaché à une GalNAc, *via* une liaison de type β1,3 (Figure 1) formant un motif GlcAβ(1,3)GalNAcβ(1,4). Les CS sont des GAGs sulfatés majeurs. Ils sont présents au niveau de la surface cellulaire et surtout dans les MEC des tissus. Au niveau du cartilage, le PG majoritaire est l'agrécan qui présente principalement des chaînes de CS et un plus petit nombre de chaînes de KS. Au cours des pathologies articulaires, la perte d'anabolisme des CS et l'altération de leur structure est un événement majeur, contribuant à la dégradation de la matrice cartilagineuse et à la perte de la fonction articulaire. Les CS peuvent subir différentes modifications, ainsi le GlcA peut être sulfaté en C2 et le GalNAc peut être sulfaté en C4 et/ou C6 pour produire respectivement des CS de type A (βGlcA(1,3)βGalNAc4S) et des CS de type C (βGlcA(1,3)βGalNAc6S).

### II.2.4. Les dermatane sulfates (DS)

Les DS sont les GAGs exprimés de façon prédominante dans la peau. Ils sont aussi connus sous le nom de CS de type B (CS-B) et présentent des activités anti-coagulante et anti-thrombotique importantes. Un DS peut être considéré comme un "dérivé" de CS, le motif disaccharique répétitif est composé d'un IdoA lié à une GalNAc par une liaison α1,3 pour former le motif IdoAα(1,3)GalNAcβ(1,4).

L'IdoA résulte de l'épimérisation en C5 du GlcA, présent sur les chaînes de CS (Malnstron *et al.*, 1975; 1984). Ainsi la présence de l'IdoA permet de distinguer les DS (ou CS-B) des CS-A

(chondroïtine 4-O-sulfatés) et des CS-C (chondroïtine 6-O-sulfatés). Ce sucre permet l'acquisition d'une flexibilité conformationnelle facilitant les interactions entre protéines (Ferro *et al.*, 1990). Le poids moléculaire des DS varie entre 12 et 45 kDa (Linhardt *et al.*, 1995). Ce sont des chaînes qui subissent différentes modifications de l'acide uronique et/ou de l'hexosamine. Par conséquent, une chaîne de DS peut contenir: 1) IdoA et GlcA, ou IdoA 2-O-sulfaté, 2) GalNAc non sulfatée, 3) GalNAc 4-O-sulfatée, 4) GalNAc 6-O-sulfatée, 5) GalNAc 4-O- et 6-O-disulfatée.

### **II.2.5. L'acide hyaluronique (AH)**

L'AH a été décrit pour la première fois en 1934 (Meyer *et al.*, 1934). C'est un hétéropolysaccharide linéaire composé d'une seule unité disaccharidique répétitive, constituée de GlcA et de GlcNAc ((1,4)GlcA $\beta$ (1,3)GlcNAc $\beta$ , Figure 1). L'AH présente une taille moyenne de 5,000 kDa (entre 10<sup>2</sup> et 10<sup>4</sup> kDa) et est composé de 10,000 unités disaccharidiques en moyenne. C'est un GAG qui a la particularité de ne jamais être sulfaté et de ne pas être lié à une protéine "core".

On trouve l'AH dans les tissus de tous les vertébrés et à la surface de bactéries pathogènes, telles que *Streptococcus* et *Pasteurella* (Fraser *et al.*, 1997), ainsi qu'à l'intérieur des cellules de mammifères. La forte charge anionique de l'AH fait de cette molécule un déterminant important de l'hydratation tissulaire (Turino and Cantor, 2003).

### **II.3. Les principales fonctions biologiques des GAGs**

Les GAGs sont des molécules qui ont la capacité d'interagir avec un grand nombre de protéines telles que les protéases, les facteurs de croissances et les récepteurs situés à la surface des cellules. Ainsi, ces molécules sont capables de moduler la transduction du signal associé à des processus physiologiques tel que le développement, la prolifération cellulaire et l'angiogenèse. De plus, il a été démontré que les GAGs sont utilisés par plusieurs agents pathogènes pour pénétrer dans la cellule (Handel, 2005).

Les principales fonctions biologiques des six types de chaînes de GAGs sont présentées dans le Tableau 1.

Type de GAGs	Fonction biologique
<b>Hp</b>	Effet anti-coagulant (Whitelock <i>et al.</i> , 2005) Stabilisation de certaines tryptases exprimées par les mastocytes (Ruoss <i>et al.</i> , 1991) Modulation de l'activité de diverses chymases (Green <i>et al.</i> , 1993) Régulation de la réponse inflammatoire (Page <i>et al.</i> , 1997) Remodelage des voies respiratoires dans le cas de l'asthme (Tyrell <i>et al.</i> , 1995)
<b>HS</b>	Interaction avec les cytokines, les chimiokines et les interleukines (Handel <i>et al.</i> , 2005, Aviezer, <i>et al.</i> , 1999) Morphogenèse, développement et organogenèse (Whitelock <i>et al.</i> , 2005) Corécepteurs de différents récepteurs à tyrosine kinases (Miserocchi <i>et al.</i> , 2001)
<b>KS</b>	Hydratation des tissus (Monzon <i>et al.</i> , 2006) Biologie cellulaire (Monzon <i>et al.</i> , 2006) GAG le plus abondant des voies respiratoires (De Merdios <i>et al.</i> , 2005)
<b>CS</b>	Prévention de l'inflammation (Takagaki <i>et al.</i> , 2002) Modulation de la réponse immunitaire (Raman <i>et al.</i> , 2005) Maintien de la structure et des propriétés biomécaniques du cartilage (Takagaki <i>et al.</i> , 2002) Régulation de l'adhésion des cellules à la MEC (Takagaki <i>et al.</i> , 2002)
<b>DS</b>	Organisation du collagène (Handel <i>et al.</i> , 2005) Régulation de l'activité du TGF- $\beta$ (Turino <i>et al.</i> , 2003) Stabilisation de la membrane basale (Handel <i>et al.</i> , 2005) Régulation des interactions cellule-cellule et cellule-matrice (Turino <i>et al.</i> , 2003)
<b>AH</b>	Stabilisation des tissus conjonctifs (Gerdin <i>et al.</i> , 1997) Organisation de la MEC (Gerdin <i>et al.</i> , 1997) Hydratation et homéostasie de l'eau (Gerdin <i>et al.</i> , 1997) Morphogenèse et homéostasie des tissus (Skold <i>et al.</i> , 1996; Li <i>et al.</i> , 2000) Régulation de la réponse inflammatoire (Cantor <i>et al.</i> , 2000) Organisation et remodelage tissulaire (Toole 1990) Phagocytose et migration cellulaire (Turino <i>et al.</i> , 2003)

**Tableau 1. Principales fonctions biologiques des chaînes de GAGs**

Hp, héparine; HS, héparane sulfates; CS, chondroïtine sulfates; DS, dermatane sulfates; AH, acide hyaluronique; TGF- $\beta$ , *transforming growth factor- $\beta$* .



### III. CLASSIFICATION DES PGs

Il existe différents critères de classification des PGs. En se basant sur leur localisation, on peut distinguer trois grands groupes: les PGs membranaires, les PGs intracellulaires et les PGs de la MEC.

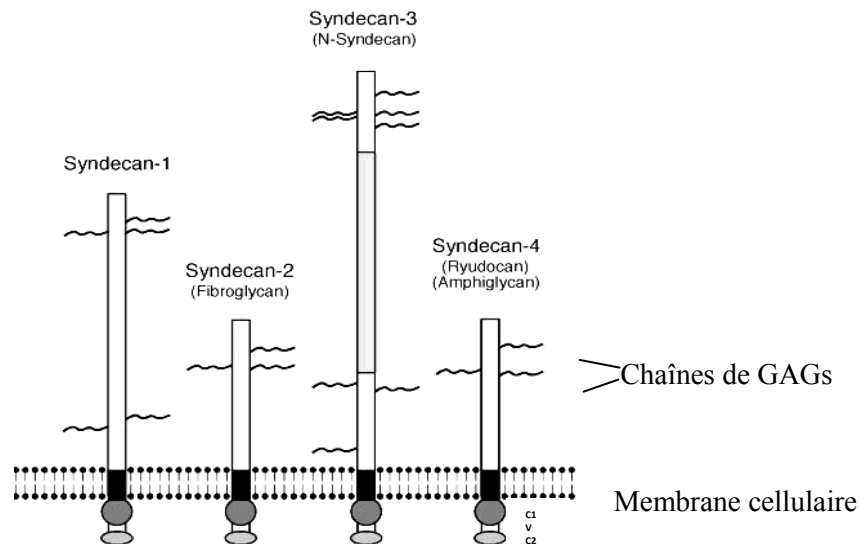
#### III.1. Les PGs membranaires

A leur découverte, certains PGs localisés à la surface des cellules ont été envisagés comme des glycoprotéines, avant d'être considérés comme de "vrais" PGs composés d'une protéine "core" et d'une ou plusieurs chaînes de GAGs. Il existe trois grandes familles de PGs membranaires: les syndécans, les CD44s et les glypicans.

##### III.1.1. La famille des syndécans

###### III.1.1.1. L'organisation des syndécans

Les syndécans sont les principaux HSPGs membranaires (Bernfield *et al.*, 1999; Carey, 1997; Couchman and Woods, 1996; Park *et al.*, 2000 b; Rapraeger and Ott, 1998). Cette famille est composée de quatre membres: le syndécan-1 (Saunders *et al.*, 1989), le syndécan-2 ou fibroglycan (Marynen *et al.*, 1989), le syndécan-3 ou N-syndécan (Carey *et al.*, 1992; Gould *et al.*, 1992), et le syndécan-4 appelé aussi amphiglycan ou ryudocan (David *et al.*, 1992; Kojima *et al.*, 1992) (Figure 5).



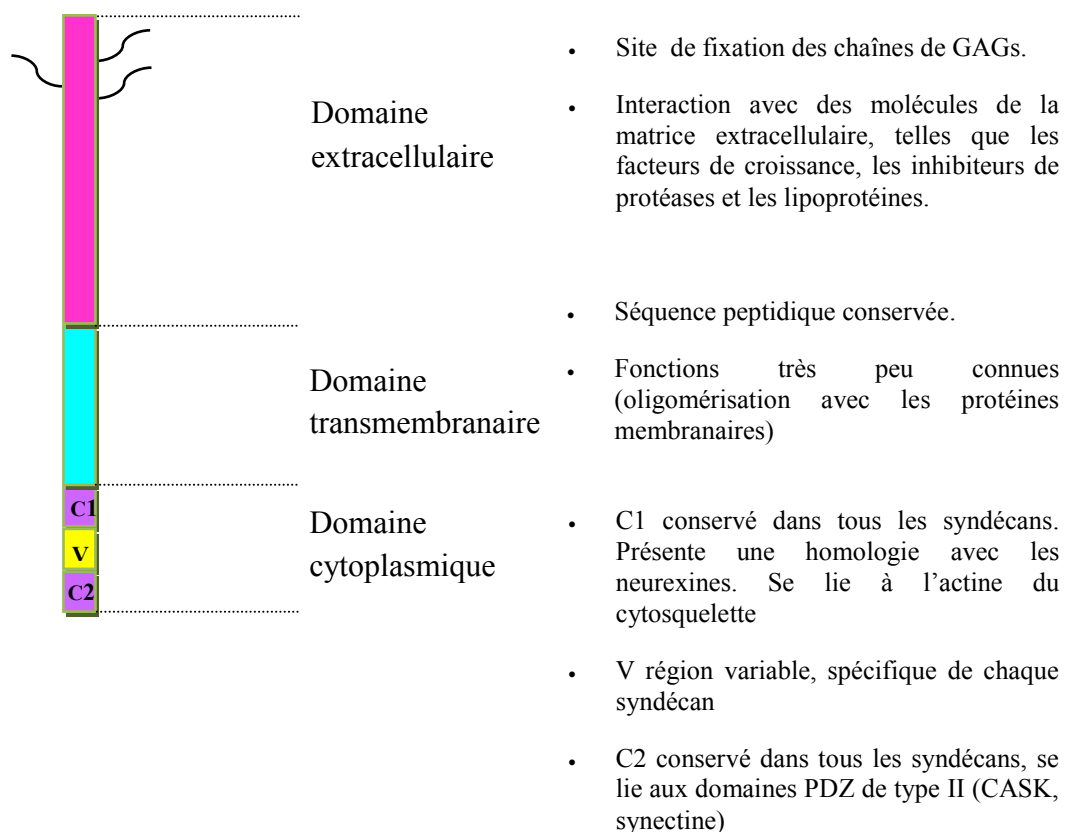
**Figure 5. Représentation schématique des membres de la famille des syndécans**

Les syndécans sont des glycoprotéines qui présentent une topologie de type I, avec un domaine extracellulaire qui porte les chaînes de GAGs, de nombre et de nature variable; un domaine transmembranaire et un court domaine cytoplasmique, divisé à son tour en trois régions: C1, V et C2.

D'après Bandtlow and Zimmermann, 2000

Toutes les cellules adhérentes expriment un ou plusieurs types de syndécans, localisés à la surface des cellules. Les syndécans sont fixés à la membrane plasmique par un segment transmembranaire et présentent une topologie de type I et une organisation structurale similaire comprenant trois domaines majeurs: le domaine extracellulaire ou ectodomaine, le domaine transmembranaire et le domaine cytoplasmique (Figure 6).

- Le domaine extracellulaire ou ectodomaine: Il présente une faible homologie de séquence entre les membres de la famille des syndécans et se situe à l'extrémité N-terminale de la protéine. Il contient des séquences consensus de type Ser-Gly, entourées d'acides aminés de nature acide qui permettent l'accrochage des chaînes de GAGs. Il porte également des sites peptidiques sensibles à l'action des protéases péricellulaires. En effet, les syndécans peuvent exister sous forme soluble dans le milieu extracellulaire suite au clivage de ce domaine.



### Figure 6. Représentation schématique des domaines fonctionnels des syndécans

Les syndécans sont des HSPGs localisés à la surface des cellules. Ce sont des glycoprotéines qui présentent une topologie de type I, avec trois domaines distincts: l'ectodomaine, le domaine transmembranaire et le domaine cytoplasmique. CASK: calcium/calmodulin-dependent serine protein kinase; PDZ: Post-synaptic density 95, Discs large, Zonula occludens-1.

D'après Multhaupt *et al.*, 2009

- Le domaine transmembranaire: constitué de 24 à 25 acides aminés hydrophobes, il présente une forte homologie de séquence au sein de la famille des syndécans et permet l'ancrage membranaire.
- Le domaine cytoplasmique: c'est un segment très conservé et relativement court, constitué de 28 à 34 acides aminés, dont 4 résidus tyrosine (Tyr) qui sont strictement conservés au sein de la famille des syndécans. Ces résidus peuvent être phosphorylés par des tyrosine kinases (TK).
- Le domaine cytoplasmique est situé à l'extrémité C-terminale de la protéine "core" et présente des séquences hautement conservées entre espèces, et entre différents syndécans au sein de la même espèce. Il est divisé en trois régions: C1, V et C2.

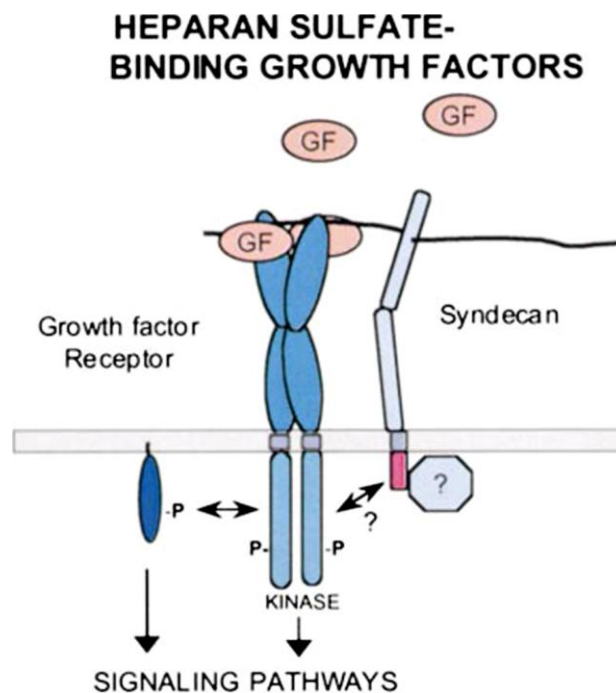
### **III.1.1.2. Distribution et fonctions biologiques des syndécans**

La localisation et le taux d'expression des syndécans varient selon les conditions physiologiques et pathologiques. Leur distribution dépend également du stade de développement et du type cellulaire. Ainsi, le syndécan-1 est exprimé à la surface des cellules épithéliales, le syndécan-2 est exprimé à la surface des cellules mésenchymateuses (Marynen *et al.*, 1989), d'où le nom de fibroglycan, le syndécan-3 est exprimé à la surface des cellules du système nerveux, d'où le nom de N-syndécan ou *Neuronal-syndecan* (Cary *et al.*, 1992). Quant au syndécan-4, il présente une distribution ubiquitaire (David *et al.*, 1992).

Les syndécans jouent un rôle essentiel dans la modulation de la signalisation cellulaire au cours de nombreux processus biologiques tels que le développement embryonnaire, la tumorigenèse et l'angiogenèse. Ce sont des corécepteurs ou récepteurs "classiques" qui interagissent avec les composants de la MEC, les facteurs de croissance et autres molécules situées à la surface des cellules.

#### ***La signalisation cellulaire du syndécan via ses chaînes d'HS***

De nombreuses fonctions assurées par les syndécans sont attribuées à leurs chaînes d'HS qui interagissent avec les facteurs de croissance, les cytokines et les molécules de la MEC (Figure 7). Les HS situés à la surface des cellules endothéliales, par exemple, exercent un effet direct sur la croissance de la tumeur, en modulant l'angiogenèse. Ce phénomène constitue une étape clé dans le processus du développement tumoral. Il a été démontré que les HS peuvent à la fois promouvoir et inhiber la néovascularisation, en régulant la signalisation associée aux récepteurs cellulaires des facteurs de croissance FGF2 (Sasisekharan *et al.*, 1994) ou VEGF (Iozzo and San Antonio, 2001).



**Figure 7. Présentation schématique de la signalisation du syndécan *via* la fixation des facteurs de croissance par les chaînes d’HS**

Les HS des syndécans jouent un rôle important dans la régulation de la signalisation cellulaire impliquée dans la croissance cellulaire et l’angiogenèse, en interagissant avec les facteurs de croissance et leur récepteurs. GF, *growth factor*.

D’après Rapraeger, 2000

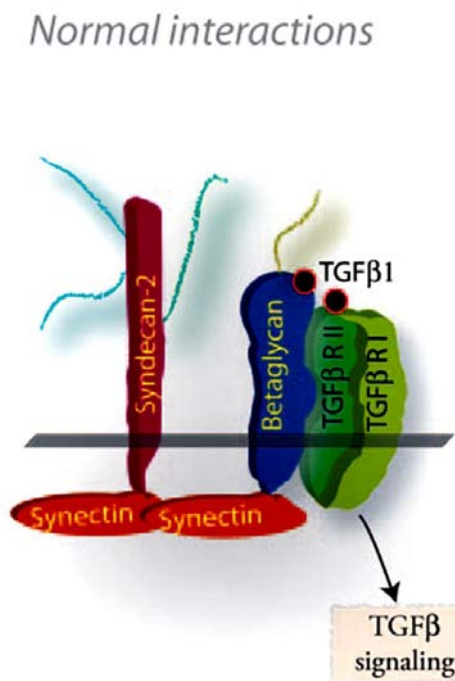
De façon intéressante, il a été démontré que chez les patients atteints de myélome, la forte augmentation du taux sérique du syndécan-1, provoquée par le "shedding" ou clivage de ce PG, est considérée comme indicateur de mauvais pronostic (Ridley *et al.*, 1993; Wijdenes *et al.*, 1996; Bayer-Garner *et al.*, 2001). Ce phénomène provoque une augmentation considérable de la quantité des chaînes d’HS présentes dans le microenvironnement tumoral. Il en résulte une stimulation de la croissance des cellules cancéreuses due à l’interaction entre ces chaînes de GAGs et certains facteurs de croissance tels que le FGF-2 et le VEGF (Filla *et al.*, 1998; Jakobsson *et al.*, 2006).

En outre, des études récentes montrent que la surexpression d’endosulfatases HSULF-1 ou HSULF-2 (enzymes qui clivent de façon spécifique les groupes 6-O-sulfates des chaînes d’HS) dans une lignée cellulaire de myélome, diminue la croissance tumorale *in vivo* de façon significative (Dai *et al.*, 2005). Il a été suggéré que les chaînes d’HS du syndécan-1

constituent une cible pharmacologique potentielle pour le traitement de ce type de cancer (Yang *et al.*, 2007).

***La signalisation du syndécan via la protéine "core", exemple de l'implication du syndécan-2 dans la voie de signalisation du TGF- $\beta$***

Par ailleurs, le TGF- $\beta$  peut se fixer directement sur le syndécan-2, *via* une interaction de type protéine-protéine. En conditions normales, le syndécan-2 se fixe sur le bétaglycan qui est transféré sur le récepteur de type II du TGF- $\beta$  (TGF- $\beta$ RII). Ce dernier va subir une autophosphorylation et va à son tour induire la phosphorylation du récepteur de type I du TGF- $\beta$  (TGF- $\beta$ RI). Ainsi le complexe TGF- $\beta$ RI/TGF- $\beta$ RII phosphorylé va permettre la transduction du signal TGF- $\beta$  (Figure 8).



**Figure 8. Présentation schématique de la modulation de la signalisation du TGF- $\beta$  par le syndécan-2**

Le syndécan-2 interagit avec le bétaglycan et conduit à la phosphorylation du complexe TGF- $\beta$ RI/TGF- $\beta$ RII situé en amont de la cascade de la signalisation du TGF- $\beta$ .

D'après Tkachenko *et al.*, 2005

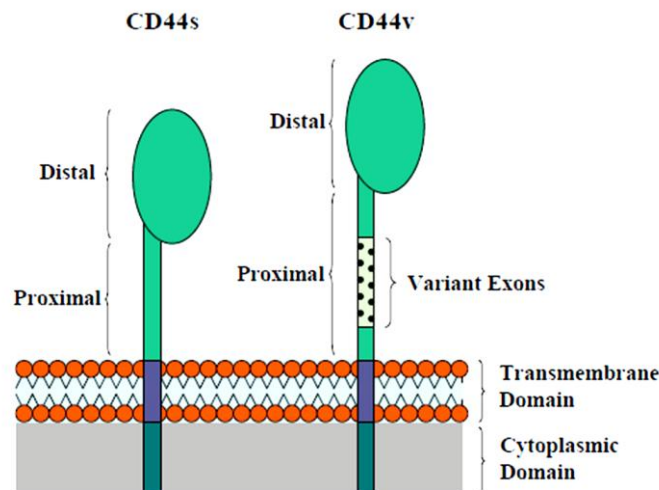
### III.1.2. La famille des CD44

#### III.1.2.1. L'organisation des CD44

Les CD44 constituent une famille de glycoprotéines membranaires de type I, qui présentent une expression ubiquitaire. Ils jouent un rôle majeur dans les interactions cellule-cellule et cellule-MEC.

Le nom de CD44 a été attribué pour la première fois à cette famille de molécules en 1987 (Cobbold, 1987). Ce terme a remplacé toutes les dénominations précédentes notamment, "lymphocyte homing receptor" (pg90<sup>Hermes</sup>), "phagocytic glycoprotein" (Pgp-1), "extracellular matrix protein III" (ECMP III) et "hyaluronate receptors" (H-CAM) et HUTCH-1.

Tous les membres de la famille des CD44 sont codés par un seul gène. Il s'agit d'isoformes issues d'un mécanisme d'épissage alternatif.



**Figure 9. Présentation schématique de la structure de la protéine CD44**

A gauche, les CD44s "standards" avec un ectodomaine distal qui reconnaît l'AH, un ectodomaine proximal, suivi d'un domaine transmembranaire et d'un domaine cytoplasmique intracellulaire, capable d'interagir avec l'ankyrine et la famille ERM (ezrine, radixine, moésine). A droite, les formes de CD44v "variants" produites par épissage alternatif. Elles sont caractérisées par la présence d'un domaine proximal allongé.

D'après Jianing and Guoshheng, 2006

Les CD44 sont virtuellement exprimés par toutes les cellules d'origine neuroectodermale. La molécule la plus petite est appelée CD44 "standard" ou CD44s. Ce PG est dépourvu de la région variable et est essentiellement exprimé à la surface des cellules d'origine lymphohématopoïétique (CD44H). On trouve aussi le CD44v8-v10 exprimé à la surface des cellules

épithéliales (CD44E) et le CD44v3-v10 détecté dans les kératinocytes. Le pMeta-1(CD44v4-v7) et le pMeta-2 (CD44v6,v7) ont été identifiés comme deux PGs « métastatiques » (Naor *et al.*, 2003). Quant aux isoformes CD44-6v et CD44-9v, elles sont détectées à la surface des monocytes, des macrophages, des lymphocytes et des cellules dendritiques (Legras *et al.*, 1998).

Le domaine extracellulaire du CD44 est situé à l'extrémité N-terminale de la protéine "core" et présente des sites potentiels de N-glycosylation, de O-glycosylation (Bennet *et al.*, 1995), et un site de fixation de l'AH qui est le principal ligand du CD44. Il possède aussi des sites consensus pour l'attachement des chaînes de GAGs de type CS, HS et KS. L'ectodomaine est divisé en deux régions: la région extracellulaire distale largement conservée entre les espèces, et la région extracellulaire proximale. Cette dernière est directement liée au domaine transmembranaire qui est suivi du domaine cytoplasmique situé au niveau de l'extrémité C-terminale de la protéine (Figure 9). De plus, il a été démontré que la séquence cytoplasmique peut être phosphorylée par la protéine kinase C (PKC), sur deux résidus lysine et arginine (Lys et Arg).

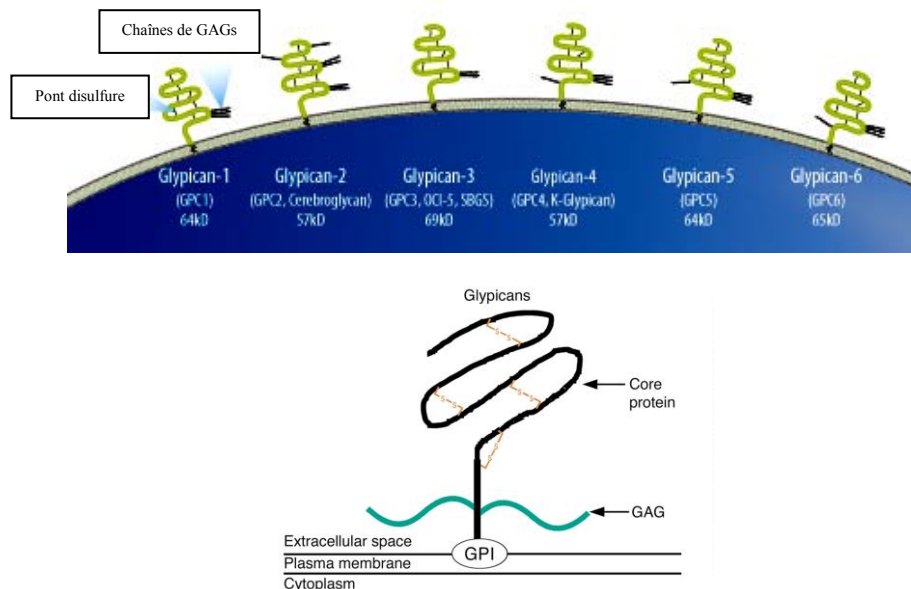
#### **III.1.2.2. Les fonctions des CD44**

Les différentes isoformes de CD44 sont impliquées dans de nombreuses fonctions biologiques vitales, notamment la migration des cellules dans les vaisseaux sanguins. Elles interagissent avec les cellules et de nombreuses molécules, principalement les facteurs de croissance, les chimiokines et des enzymes. Les CD44 assurent la transduction du signal de l'extérieur vers l'intérieur de la cellule. Ces processus sont liés à la prolifération, la survie ou encore l'induction de l'apoptose cellulaire. Les principales fonctions biologiques des CD44 sont liées à la fixation de l'AH. Il a été montré que cette protéine joue un rôle majeur dans la biologie des cellules cancéreuses et cela depuis les tous premiers stades du processus de transformation maligne des cellules. En effet, l'analyse du taux d'expression de CD44 est utilisée comme un moyen d'étude de la progression tumorale dans de nombreux cancers chez l'homme. Ainsi, dans le cas de carcinome gastrique le CD44v8-v10 est corrélé avec la présence de tumeurs et l'invasion métastatique (Yamaguchi *et al.*, 1995; Miwa *et al.*, 1996). De la même façon, dans le cas du mélanome, le CD44v5 est un indicateur de cancer agressif (Manten-Hort *et al.*, 1995).

### III.1.3. La famille des glypicans

#### III.1.3.1. L'organisation des glypicans

Le nom de glypican est attribué à une famille de PGs à HS. La protéine "core" est située sur la face externe de la membrane plasmique de la cellule, liée à celle-ci grâce à une liaison glycosyl-phosphatidyl-inositol (GPI) (David, 1993) (Figure 10). Cette famille de HSPGs est composée de six membres chez les mammifères (GPC1-GPC6). Tous les glypicans possèdent une séquence signal du côté N-terminal qui permet l'initiation de la translocation vers le lumen du reticulum endoplasmique (RE) et le trafic de la protéine vers la membrane plasmique.



**Figure 10. Présentation schématique de la structure des glypicans**

La famille des glypicans contient 6 membres de poids moléculaire proche (57 à 64 kDa). La protéine "core" est fixée à la membrane plasmique *via* une ancre GPI (glycosyl-phosphatidyl-inositol). La formation des ponts disulfures confère à la protéine une structure 3D globulaire. Les chaînes de GAGs sont principalement des HS, localisées au niveau de la partie C-terminale, à proximité de la membrane plasmique de la cellule.

D'après Filmus and Selleck, 2001

L'homologie de séquence entre les membres de la famille des glypicans est faible. En revanche, l'analyse de leur structure 3D montre une organisation similaire due, en partie, à la présence de 14 résidus cystéine (Cys) conservés dans tous les glypicans (Veugelers *et al.*, 1999; Topczewsky *et al.*, 2001). Une autre caractéristique fonctionnelle importante des



glypicans est la localisation des sites (Ser-Gly) de fixation des chaînes d'HS à proximité de l'extrémité C-terminale de la protéine "core". Cette localisation stratégique suggère que les HS peuvent jouer le rôle de médiateurs au cours de l'interaction entre le glypican et les autres molécules situées à la surface de la cellule, tels que les récepteurs des facteurs de croissance.

### III.1.3.2. Les fonctions des glypicans

De nombreuses études ont montré que la principale fonction des glypicans *in vivo* est la régulation de la signalisation de nombreux facteurs de croissance, tels que les FGFs, les Wnts (Yayon *et al.*, 1991; Schlessinger *et al.*, 1995; Reichsman *et al.*, 1996; Binari *et al.*, 1997), ainsi que des protéines de la famille des BMPs (*Bone Morphogenic Proteins*) et des Hhs (*Hedgehogs*) (Topczewsky *et al.*, 2001; Ohkawara *et al.*, 2003).

Nom du gène	Synonyme	Expression chez l'embryon	Expression chez l'adulte	Localisation chromosomique
<b>GPC1</b>	Glypican	Os, moelle osseuse, muscles, reins	La plupart des tissus	2q35-37
<b>GPC2</b>	Cérebroglycan	Système nerveux	Non détecté	7q22.1
<b>GPC3</b>	OCI-5, MXR7	La plupart des tissus	Ovaires, glandes mammaires	Xq26
<b>GPC4</b>	K-glypican	Cerveau, reins, poumons	La plupart des tissus	Xq26.1
<b>GPC5</b>		Cerveau, poumons, foie, membres	Cerveau	13q32
<b>GPC6</b>		Quelques tissus	Quelques tissus, ovaire, reins, foie, intestin	13q32

### **Tableau 2.** La famille des glypicans chez l'homme

Localisation des gènes sur les chromosomes et expression tissulaire chez l'adulte et l'embryon humains.

D'après Filmus, 2000

Les glypicans sont exprimés principalement au cours du développement embryonnaire, à des taux variables en fonction du stade de développement et du type de tissu. Chez l'homme, la mutation ponctuelle du gène codant pour le glypican-3 (GPC3) est responsable du syndrome de Simpson-Golabi-Behmel (SGBS). Le tableau clinique de ce syndrome est très lourd, il est

caractérisé par une macromégalie faciale, une surcroissance prénatale et postnatale, une prédisposition tumorale, un défaut cardiaque congénital, une syndactylie, une polydactylie, une dysplasie rénale et d'autres manifestations cliniques selon le phénotype du patient. Ainsi, le phénotype est corrélé avec la localisation de la mutation sur le gène GPC3. Ces mutations produisent des glypicans dont le rôle dans la signalisation des facteurs de croissance au cours du développement est altéré (Duenas *et al.*, 1998).

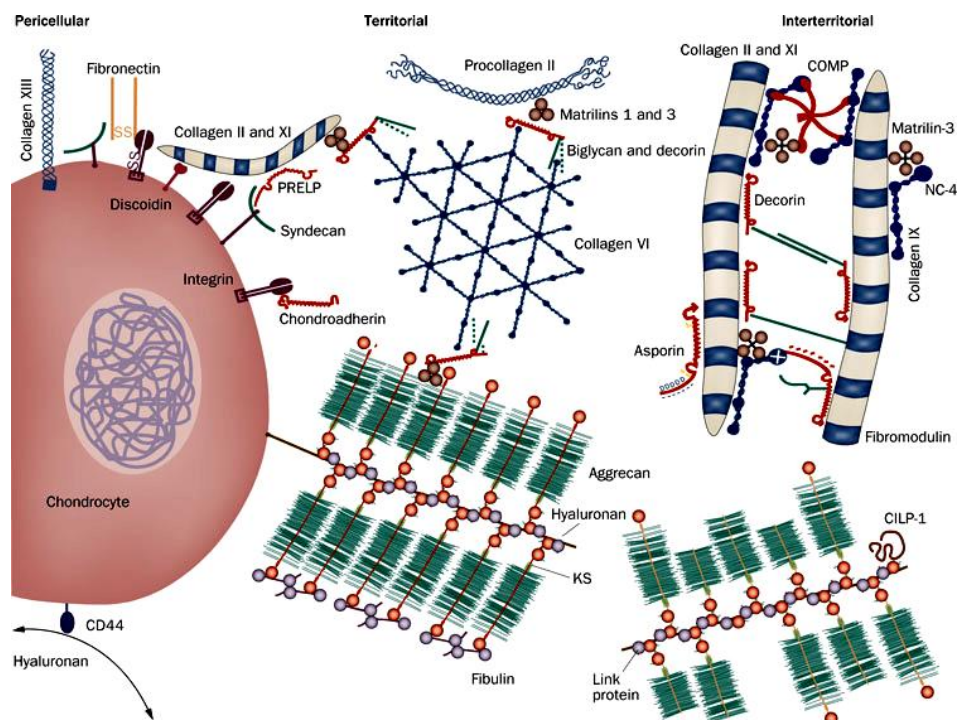
### **III.2. Les PGs intracellulaires**

De nombreuses études ont porté sur l'identification et la caractérisation des fonctions biologiques des PGs extracellulaires et de surface. En revanche, les PGs intracellulaires ont suscité moins d'attention. Le PG intracellulaire le plus connu est la serglycine (SG), c'est la seule protéine "core" qui porte des chaînes d'Hp. Elle peut également porter des chaînes de CS et d'HS. C'est l'un des plus petits PGs connus, avec une masse moléculaire de 17 à 19 kDa. Le nom de serglycine provient des longues séquences répétitives de Ser-Gly (sites de fixation des chaînes de GAGs sur la protéine "core") présentes au niveau de la région centrale de la protéine.

La serglycine est exprimée majoritairement dans les granules de sécrétion des mastocytes, les lymphocytes T cytotoxiques (LT<sub>C</sub>), les cellules NK (Natural Killer), les neutrophiles et les plaquettes. Elle est aussi détectée dans des cellules d'origine non hépatopoïétique, notamment dans les cellules endothéliales (Schick *et al.*, 2001) et l'endoderme pariétal (Grover *et al.*, 1987; Ho *et al.*, 2001). Le nombre et la longueur des chaînes de GAGs dépendent du type cellulaire. La serglycine est indispensable pour le stockage de diverses enzymes dans les vésicules de sécrétion. Ainsi, ce PG joue un rôle crucial dans la formation des granules de sécrétion dans les mastocytes. De plus, il est devenu clair que la serglycine et surtout les chaînes d'Hp sont des acteurs clés dans les processus de thrombose et l'agrégation plaquettaire.

### **III.3. Les PGs extracellulaires**

Les PGs extracellulaires constituent un groupe de molécules hétérogènes très diversifiées qui ont suscité un grand intérêt depuis leur découverte (Figure 11).



**Figure 11. Représentation schématique de la MEC du cartilage artériel**

La MEC du cartilage artériel est constituée principalement de collagènes (essentiellement de type II), de PGs de grande taille: l'agrécane, de PGs de petite taille: les SLRPs (décorine, biglycan et fibromoduline) et de protéines non collagéniques COMP (*Cartilage Oligomeric Matrix Protein*).

D'après Heinegård and Saxne, 2010

### III.3.1. L'agrécane

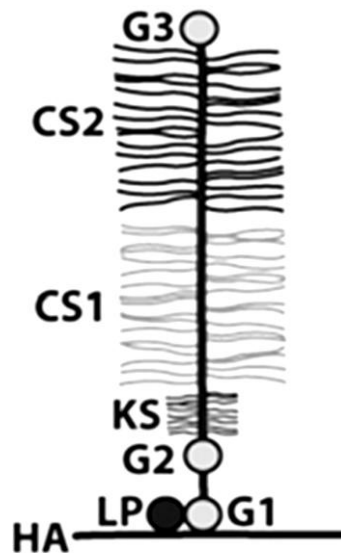
L'agrécane est le PG le plus abondant de la MEC du cartilage artériel. Il fait partie des PGs qui forment un complexe multimoléculaire en se liant à l'AH (Iozzo, 1998). Le versican, le neurocan, le brévican et l'agrécane sont les quatre membres de la famille des hyalectans appelée aussi lecticans. Ils possèdent tous un domaine globulaire N-terminal qui interagit avec l'AH et un domaine globulaire C-terminal qui présente une forte homologie avec la lectine de type C, d'où le nom de cette famille.

L'agrécane est un PG de grande taille très fortement glycosylé. Il contient plus de 100 chaînes de CS et environ 30 chaînes de KS liées à une protéine "core" de très grande taille (220,000 kDa approximativement). Chaque chaîne de CS est 4-sulfatée et/ou 6-sulfatée.

Il interagit avec l'AH pour former une macromolécule qui a l'apparence d'un agrégat (d'où le nom d'agrécane). Cette liaison agrécane-AH est stabilisée par une glycoprotéine appelée protéine de liaison (*link protein*).

### III.3.1.1. L'organisation structurale de l'agrécan

La protéine "core" est constituée de trois régions globulaires appelées G<sub>1</sub>, G<sub>2</sub> et G<sub>3</sub>, chacun de ces domaines contient des résidus Cys qui forment des ponts disulfures. Les régions G<sub>1</sub> et G<sub>2</sub> sont séparées par un court domaine interglobulaire appelé IGD (*Inter Globular Domain*) et le domaine globulaire G<sub>2</sub> est séparé du domaine globulaire C-terminal (G<sub>3</sub>) par une région de fixation des chaînes de KS et deux domaines CS<sub>1</sub> et CS<sub>2</sub> sur lesquels sont greffées les chaînes de CS (Figure 12). Le domaine G<sub>3</sub> est situé à l'extrémité C-terminale de la protéine "core". Il est constitué de domaines structuraux distincts et semble jouer un rôle essentiel dans les processus de modifications post-traductionnelles de la protéine "core" de l'agrécan et sa sécrétion (Zheng *et al.*, 1998).



**Figure 12. Représentation schématique de l'organisation d'un monomère d'agrécan et de ses domaines structuraux**

L'agrécan est lié à l'AH par le domaine globulaire G<sub>1</sub>, cette liaison est stabilisée par la LP (*link protein*). On distingue deux autres domaines globulaires G<sub>2</sub> et G<sub>3</sub> qui délimitent trois régions: KS sur laquelle se fixent les chaînes de KS, puis CS<sub>1</sub> et CS<sub>2</sub> sur lesquelles sont greffées les chaînes de CS.

D'après Roughley, 2006

### III.3.1.2. Les fonctions de l'agrécan

L'agrécan joue un rôle majeur dans les propriétés rhéologiques du cartilage articulaire, en particulier sur la résistance de ce tissu aux forces de compression. C'est un élément

déterminant du maintien de l'intégrité de la MEC du cartilage articulaire. Il joue également un rôle important dans les interactions chondrocyte-chondrocyte et chondrocyte-MEC. L'altération de la qualité de l'agrécane et de ces chaînes de GAGs présent dans cette MEC est associée au processus arthrosique.

De façon intéressante, il a été démontré chez l'homme que le gène codant pour l'agrécane, présente un polymorphisme génétique qui touche le domaine CS<sub>1</sub> présentant une taille variable. Ce domaine est composé de motifs Ser-Gly répétés 13 à 33 (Doege *et al.*, 1997). Les individus qui portent l'allèle le plus court expriment un agrécane caractérisé par un nombre réduit de chaînes de CS. Ils présenteraient un risque de dégénérescence prématurée du cartilage.

La composition de l'agrécane, notamment le nombre, la longueur et le profil de sulfatation des chaînes de GAGs (KS et CS) varient également en fonction de l'âge. Enfin, à cela s'ajoute l'hétérogénéité de l'agrécane générée par l'action des agrécanases et des métalloprotéinases au cours des pathologies dégénératives du cartilage.

Enfin, des phénotypes chondrodystrophiques ont été décrits chez l'homme, en conséquence de la diminution de la sulfatation des CS, elle-même due à la mutation du gène codant pour le transporteur de sulfates (Superti-Furga *et al.*, 1996).

### **III.3.2. Les PGs de la membrane basale**

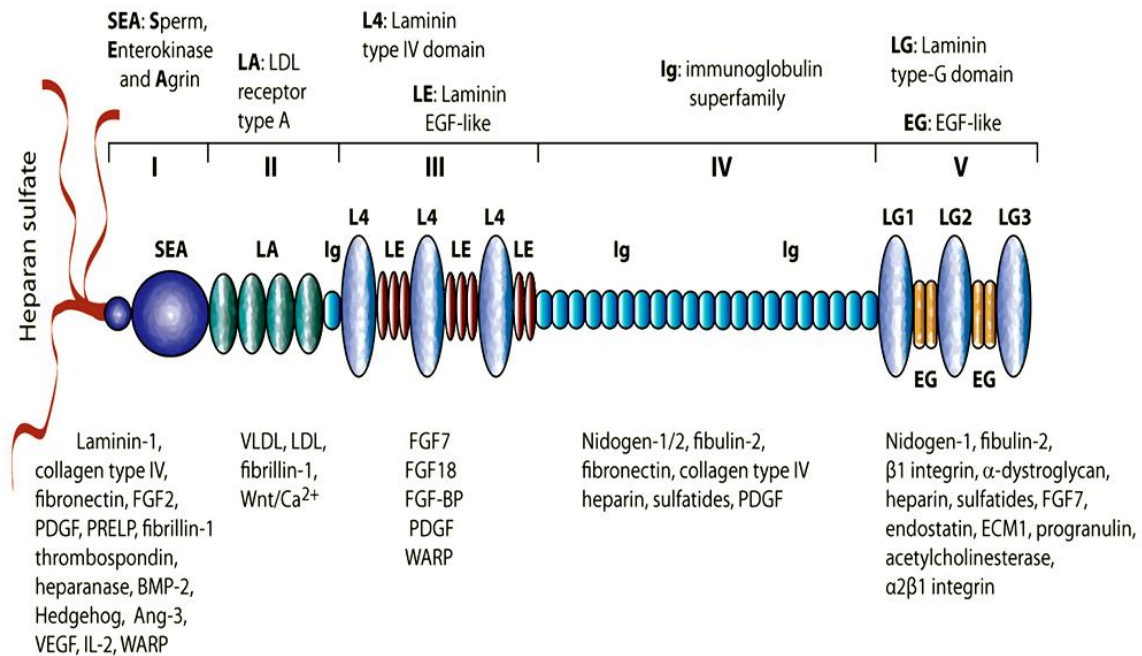
La membrane basale (MB) est considérée comme une spécialisation de la MEC. Elle se présente comme un mince feuillet de 50-800 nm d'épaisseur qui forme une barrière entre les cellules épithéliales, endothéliales, musculaires et le tissu conjonctif. Cette membrane est composée principalement de laminine, collagène-IV, fibronectine, entacine (nidogène), et d'HSPGs (perlecan, bamacan et agrine) (Erickson and Couchman, 2000).

#### **III.3.2.1. Le perlecan**

Le perlecan a été isolé pour la première fois par Hassel *et al.* (1980) à partir du sarcome d'Engelbreth-Holm-Swarm (EHS). C'est le PG de la MB le plus étudié et le plus abondant. Le nom de "perlecan" provient de l'apparence structurale de ce PG qui ressemble à un collier de perles. Cet aspect est dû à la présence de domaines globulaires dispersés le long d'une structure linéaire (Hassell *et al.*, 2003).

### III.3.2.1.a. L'organisation structurale du perlecan

Le perlecan est un HSPG dont la protéine "core" est divisée en cinq domaines distincts (Figure 13). Ces derniers présentent des homologies de séquences avec d'autres protéines connues comme certains facteurs de croissance et des protéines impliquées dans le métabolisme des lipides (Hassel *et al.*, 1980; Noonan *et al.*, 1991; Kallunki and Tryggvason, 1992; Murdoch *et al.*, 1992). Le domaine I N-terminal est spécifique du perlecan, il porte trois sites potentiels pour l'attachement de chaînes d'HS, et un module SEA (*Sperm protein, Enterokinase, Agrin*).



**Figure 13. Présentation schématique de la structure du perlecan**

Organisation de la protéine "core" de ce HSPG en cinq domaines structuraux (indiqués en chiffres romains). Le schéma montre la présence de modules similaires à des protéines de la MEC. La liste située sous le schéma indique les molécules capables d'interagir avec chaque domaine du perlecan.

D'après Whitelock *et al.*, 2008

### III.3.2.1.b. Les fonctions du perlecan

Le perlecan joue un rôle important au niveau de la MB, il interagit avec les molécules présentes dans la MEC, à la fois par sa protéine "core" et par ses chaînes d'HS. Le domaine

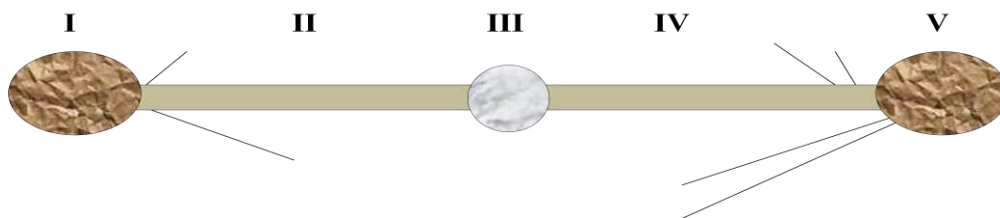
IV, par exemple, peut fixer l'entactine/nidogène-1 et -2, la fibuline-2 et la fibronectine (Hopf *et al.*, 2001). Quand au domaine V, il interagit avec la fibuline-2, le complexe laminine-entactine/nidogène ainsi que d'autres molécules (Friedrich *et al.*, 1999). Le perlecan a une expression ubiquitaire, il joue au niveau de la MB des MEC un rôle crucial dans l'adhésion, la prolifération cellulaire et dans l'angiogénèse.

Chez l'homme, deux mutations du perlecan ont été identifiées. Elles sont responsables de deux maladies génétiques: le syndrome de Schwartz-Jampel de type I (SSJI) et la dysplasie dysségmentale de type Silverman-Handmaker (DDSH). Les nouveau-nés DDSH présentent de très nombreuses anomalies squelettiques, et décèdent rapidement après la naissance. Les patients atteints du SSJI vivent souvent jusqu'à l'âge adulte avec de nombreux désordres liés au degré de l'haploinsuffisance et souffrent de raideurs musculaires sévères (Rodgers *et al.*, 2007).

### III.3.2.2. Le bamacan

C'est un PG à CS abondant dans la MB. Le nom de bamacan provient de la localisation de cette molécule et de sa composition en GAGs (*Basement membrane-associated chondroitin proteoglycan*) (Couchman *et al.*, 1996).

Le bamacan présente une forte homologie de séquence avec les protéines de la famille des SMC (*Structural Maintenance of Chromosome*), et plus particulièrement, avec la sous-classe SMC3. Il possède trois domaines qui ont une conformation globulaire (I, III et V) séparés par deux hélices- $\alpha$  superenroulées (domaines II et IV). Les régions II, IV et V contiennent des séquences consensus Ser-Gly qui constituent des sites potentiels de N-glycosylations et de fixation des chaînes de CS (Figure 14).



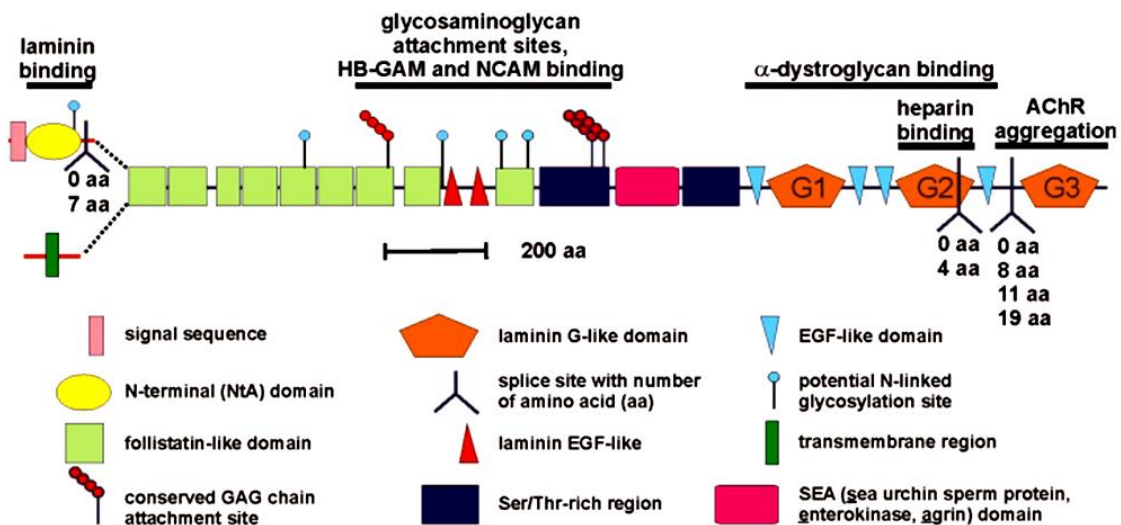
**Figure 14. Présentation schématique du bamacan**

Le bamacan présente une organisation structurale séquentielle, composée de trois domaines globulaires séparés par deux hélices- $\alpha$  super-enroulées. Ce schéma montre la présence de trois sites de N-glycosylation (traits courts) et trois sites de fixation de chaînes de CS (traits longs).

Le bamacan est un CSPG qui joue un rôle de stabilisateur de la MB (Couchman and Woods, 1993), Chez les patients atteints de la forme dystrophique d'épidermolyse bulleuse et dans certains cas de tumeurs de l'épiderme, le taux de chaînes de CS présent dans la MB située entre le derme et l'épiderme est anormalement bas (Fine and Couchman, 1989; Oguro *et al.*, 1991). D'un autre côté, il a été constaté que l'expression du bamacan est augmentée dans les lignées cellulaires et certains tissus tumoraux, ce qui suggère un rôle de ce PG dans la tumorigenèse.

### III.3.2.3. L'agrine

C'est un HSPG hautement glycosylé (Tsen *et al.*, 1995) présent dans la MEC. Il a été identifié pour la première fois par McMahan *et al.*, (1990), comme un polypeptide de 200 kDa. L'agrine exprimée à la fois dans les muscles et les neurones est impliquée dans la formation, le maintien et la régénération de la jonction neuromusculaire. Un épissage alternatif produit deux formes d'agrine: une forme transmembranaire (de type II) et une forme soluble présente dans la MB de la MEC.



**Figure 15. Présentation schématique de l'agrine**

L'agrine présente une organisation structurale modulaire constituée de nombreux domaines distincts. Les sites d'épissage alternatif de l'ARNm avec le nombre possible d'acides aminés sont indiqués sur le schéma ci-dessus. NCAM, neural *cell adhesion molecule*; HB-GAM, *heparin-binding growth-associated-molecule*.

D'après Kröger and Schröder, 2002



L'agrine présente une organisation multimodulaire. Elle comprend neuf domaines "*folliculin-like*", une répétition de deux domaines "*laminin-like-EGF*", deux domaines riches en acides aminés proline (Pro), Ser et Thr, et un module SEA (Figure 15). Du côté de l'extrémité C-terminale, on trouve quatre domaines "*EGF-like*" et trois domaines "*laminin-G-like*".

L'agrine est faiblement exprimée chez le fœtus (Groffen *et al.*, 1998). Chez l'adulte, elle est exprimée majoritairement dans les reins, les poumons, le foie et la glande thyroïdienne. On la trouve également dans les différentes régions du cerveau, notamment dans les neurones et dans la MB microvasculaire.

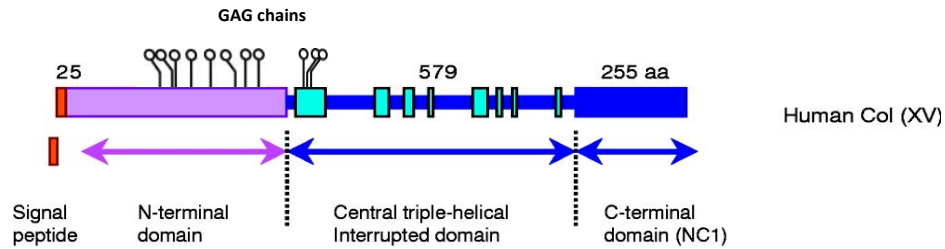
Les chaînes de GAGs de ce PG sont impliquées dans la potentialisation de l'adhésion hémophilique des NCAM (*Neural Cell Adhesion Molecule*) et dans la fixation des facteurs de croissance de type HB-GAM (*Heparin-Binding Growth-Associated-Molecule*). Chez les patients atteints de la maladie d'Alzheimer, une forme soluble d'agrine a été identifiée à de fortes concentrations aussi bien dans les plaques neuritiques, appelées aussi plaques séniles, que dans les plaques diffuses du cerveau. Les cerveaux de ces patients montrent des anomalies au niveau de la MB, qui peuvent être associées à l'emprisonnement de l'agrine par la protéine amyloïde (Donahue *et al.*, 1999).

#### **III.3.2.4. Le collagène XV et le collagène XVIII**

Les collagènes de type XV (Col XV) et de type XVIII (Col XVIII) sont des CSPG et HSPG, respectivement (Halfter *et al.*, 1998; Li *et al.*, 2000). Ils appartiennent à la sous-famille des collagènes non fibrillaires (Abe *et al.*, 1993; Oh *et al.*, 1994; Rehn and Pihlajaniemi, 1994). Pour le Col XVIII, un mécanisme d'épissage alternatif produit deux isoformes chez l'homme (longue et courte) (Saarela *et al.*, 1998).

La comparaison des séquences primaires de Col XV et Col XVIII montre une grande similarité, la structure tridimensionnelle correspond à des chaînes  $\alpha 1$  qui forment des homotrimères (trois chaînes  $\alpha 1$ ).

Chez l'homme Col XV comprend une séquence peptide signal N-terminale, suivie de trois régions. La région collagénique centrale est interrompue par 8 segments non collagéniques de taille variable (Figure 16).



**Figure 16. Présentation schématique de l'organisation du Col XV chez l'homme**

Col XV présente trois domaines: un domaine N-terminal qui porte 11 séquences consensus de fixation des chaînes de GAGs, et un domaine C-terminal séparés par une région collagénique centrale interrompue par 8 segments non collagéniques, dont le premier porte trois sites potentiels de fixation de GAGs.

D'après Ortega and Werb, 2002

Le Col XV est un PG qui porte principalement des chaînes de CS sur le domaine N-terminal de la protéine "core" (Li *et al.*, 2000). Cependant, de nouvelles études ont montré que le Col XV peut porter également des chaînes de CS/DS et des chaînes d'HS ce qui suggère que ce PG joue un rôle structural et serait également impliqué dans les interactions cellules-matrices.

	<b>Induction</b>	<b>Inhibition</b>
<b>Collagène XV monomérique soluble</b>	Réduction de la croissance tumorale (Ramchadran <i>et al.</i> , 1999)	Migration des cellules endothéliales induite par le FGF2
<b>Trimère NC1</b>		Angiogenèse de la membrane chorioallantoïque induit par le VEGF (Sasaki <i>et al.</i> , 2000)  Angiogenèse de la membrane chorioallantoïque induit par le VEGF et par le FGF2 (Sasaki <i>et al.</i> , 2000)

**Tableau 3. Exemples de fonctions du col XV**

NC, domaine non collagénique; VEGF, *vascular endothelial growth factor*; FGF, *fibroblast growth factor*.

Ces deux types de PGs ont une distribution tissulaire très différente, ce qui suggère qu'ils utilisent des mécanismes distincts pour inhiber l'angiogenèse. Col XV est exprimé majoritairement dans le muscle cardiaque et squelettique, tandis que Col XVIII est exprimé principalement dans les cellules qui constituent les parois des vaisseaux et dans les MBs.

A l'heure actuelle, les rôles physiologiques des Col XV et Col XVIII ne sont pas encore bien déterminés. Chez la souris la déficience en Col XV est responsable de la dégénérescence progressive des muscles squelettiques et des cellules endothéliales. En revanche, la mutation de l'isoforme courte de Col XVIII chez l'homme est associée au syndrome Knoblok. C'est une maladie génétique autosomale récessive qui se caractérise par une forte myopie, un détachement de la rétine avec des anomalies qui touchent la macula ainsi que par des défauts occipitaux (Sertie *et al.*, 2000).

### **III.3.3. La famille des *Small Leucine Rich Proteoglycans* (SLRPs)**

Les PGs riches en leucine (SLRP) constituent un sous-groupe de la famille des LRR (*Leucine Rich Repeat*). Ils sont caractérisés principalement par la présence de 6-10 séquences LRR, entourées de motifs riches en résidus Cys aux extrémités N- et C-terminales (Fisher *et al.*, 1989; Oldberg *et al.*, 1989; Blochberger *et al.*, 1992). Ce sont des PGs de petite taille (40 kDa environ). La protéine "core" porte une ou deux chaînes de CS ou DS, ou plusieurs chaînes de KS (Figure 17). La famille des SLRPs ne cesse de s'agrandir. Actuellement, elle contient 13 membres (voir tableau 4) essentiellement présents dans la MEC (Bech-Hansen *et al.*, 2000; Push *et al.*, 2000). La plupart des SLRPs sont des PGs sauf l'asporine, l'opticine/oculoglycan, la chondroadhérine et la nyctalopine qui ne portent pas de chaînes de GAGs.

La classification des SLRPs est basée sur l'organisation génomique, la structure et le nombre de domaines riches en Leu et sur la nature des chaînes de GAGs. Ainsi, on distingue trois classes principales de SLRPs: la classe I contient la décorine (DCN) le biglycan (BGN) et l'asporine. Ces derniers sont codés par 8 exons, et sont caractérisés par la présence de 10 motifs LRR et d'une séquence N-terminale CX<sub>3</sub>CXCX<sub>6</sub>C, (C correspondant à une Cys et X à un acide aminé quelconque). Les SLRPs de classe II sont la fibromoduline (FM), le lumican (LUM), le kératocan, le PRELP et l'ostéo-adhérine/ostéo-moduline. Toutes ces molécules sont des KSPGs, et sont caractérisées par la présence de 10 motifs LRR, d'une séquence de Cys N-terminale CX<sub>3</sub>CXCX<sub>9</sub>C et sont codées par 3 exons seulement. La classe III contient l'épiphican/Pg-Lb/DSPG<sub>3</sub> et l'ostéoglycan/mimican qui portent respectivement des chaînes

de KS et de DS, ces SLRPs se distinguent par la présence de 6 motifs LRR seulement et d'une séquence N-terminale CX<sub>2</sub>CXCX<sub>6</sub>C. Les SLRPs de classe III sont codés par 7 exons.

	Nom de la protéine "core"	Masse moléculaire (kDa)	GAG (nombre de chaînes)	Structure des gènes et des protéines, région N-terminale riche en Cys
<b>Classe I</b>	Biglycan	40	CS/DS (1)	8 exons
	Décorine	40	CS/DS (2)	10 LRR
	Asporine			CX <sub>3</sub> CXCX <sub>6</sub> C
<b>Classe II</b>	Fibromoduline	42	KS (2-3)	3 exons
	Lumican	38	KS (3-4)	10 LRR
	Kératocan	38	KS (3-5)	CX <sub>3</sub> CXCX <sub>9</sub> C
	PRELP	44	KS (2-3)	
	Ostéoadhérine/Ostéomoduline	42	KS (2-3)	
<b>Classe III</b>	Epiphican/PG-Lb/DSPG3	35	DS (2-3)	7 exons
	Mimican/ostéoglycan	35	KS (2-3)	6 LRR CX <sub>2</sub> CXCX <sub>6</sub> C
<b>Autres (classe IV et classe V)</b>	Opticine/Oculoglycan			3 exons
	Chondroadhérine			11 LRR
	Nyctalopine			CX <sub>3</sub> CXCX <sub>8</sub> C ou CX <sub>3</sub> CXCX <sub>6</sub> C

#### **Tableau 4. Classification de la famille des SLRPs**

La classification des SLRPs est basée sur l'organisation génique, le nombre de domaines LRR, la nature de la région C-terminale riche en Cys et le type de chaînes de GAGs.

L'asporine, l'opticine/oculoglycan, la chondroadhérine et la nyctalopine appartiennent à la famille des LRRs mais ne sont pas des PGs. DS, dermatane sulfates; KS, kératane sulfates; LRR, *leucine rich repeat*; C, Cys; X, acide aminé quelconque.

## Small Leucine-Rich Proteoglycans (SLRPs)

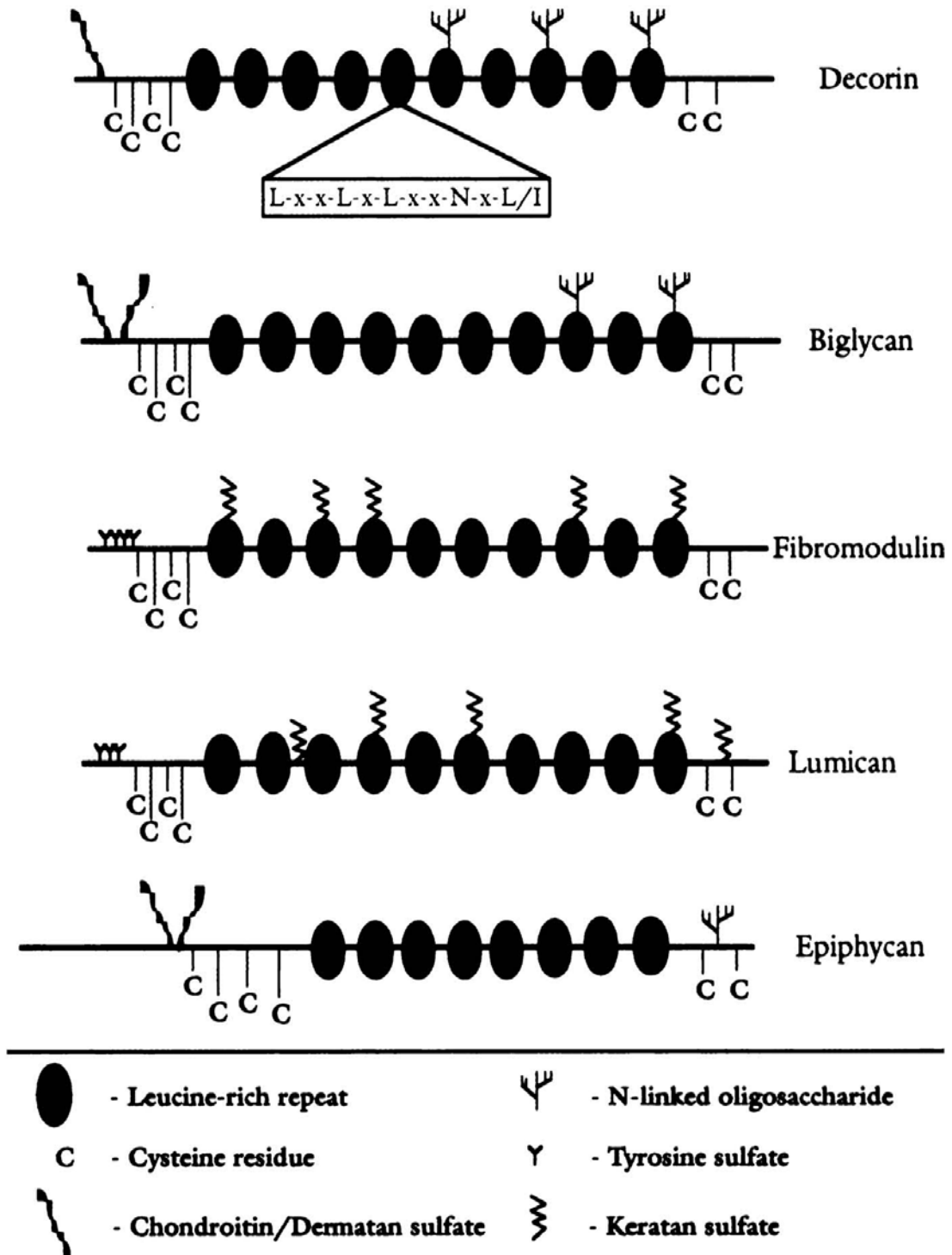


Figure 17. Présentation schématique de la famille des SLRPs

La séquence consensus pour les LRRs est entourée d'un rectangle. L, Leu; X, acide aminé quelconque.

D'après Iozzo and Murdoch, 1996

Les SLRPs jouent un rôle important dans la régulation de la fibrogenèse du collagène et dans le maintien de la structure du stroma. Ils sont également impliqués dans la prolifération cellulaire, la migration et la croissance tumorale. Ce groupe de molécules est capable d'interagir avec différentes protéines de la MEC notamment avec les facteurs de croissance tels que le TGF $\beta$  et l'EGF ainsi qu'avec d'autres molécules impliquées dans la prolifération cellulaire et la morphogenèse des tissus.

### **III.3.3.1. Les principaux SLRPs**

#### **III.3.3.1.1. La fibromoduline**

C'est un PG qui appartient à la classe II de la famille des SLRPs et qui porte 2 à 3 chaînes de KS fixées sur la protéine "core". La fibromoduline présente 9 résidus Tyr proches de l'extrémité N-terminale, qui constituent des sites de phosphorylation (Antonsson *et al.*, 1991; Onnerfjord *et al.*, 2004). Elle contient aussi 5 sites de N-glycosylation dans le domaine LRR (Plaas *et al.*, 1990).

La fonction biologique des chaînes de GAGs n'est pas complètement déterminée. Cependant, il est connu que l'absence des chaînes de KS sur la fibromoduline n'affecte pas son interaction avec le collagène (Font *et al.*, 1998). Des souris KO (knockout) pour la fibromoduline montrent un développement de l'ostéoarthrose, une augmentation (compensatoire) du taux d'expression du lumican et une hyperlaxité des tendons. Un nombre important de fibrilles de collagène à diamètre réduit a été également constaté chez ces souris. La fibromoduline participe à la résistance des tendons (Jepsen *et al.*, 2002). Ainsi, son absence conduit à l'altération de ce tissu (Ezura *et al.*, 2000). Différents travaux suggèrent un rôle de cette molécule dans le processus métastatique (Clark *et al.*, 2000). Elle est aussi impliquée dans la réponse inflammatoire du chondrocyte (Sjoberg *et al.*, 2005; Heathfield *et al.*, 2004) et elle est régulée par le TGF $\beta$  au cours développement embryonnaire (Baffi *et al.*, 2006).

#### **III.3.3.1.2. Le lumican**

Le lumican est un PG qui appartient à la classe II de la famille des SLRPs et qui porte des chaînes de KS. Il présente des homologies structurales avec la fibromoduline. L'analyse des séquences primaires montre que ces deux protéines "core" sont identiques à 47%. Le lumican est exprimé dans différents tissus et plus particulièrement dans les tissus connectifs. Il a la particularité d'être présent sous deux formes. En effet, dans la cornée de l'adulte, le lumican

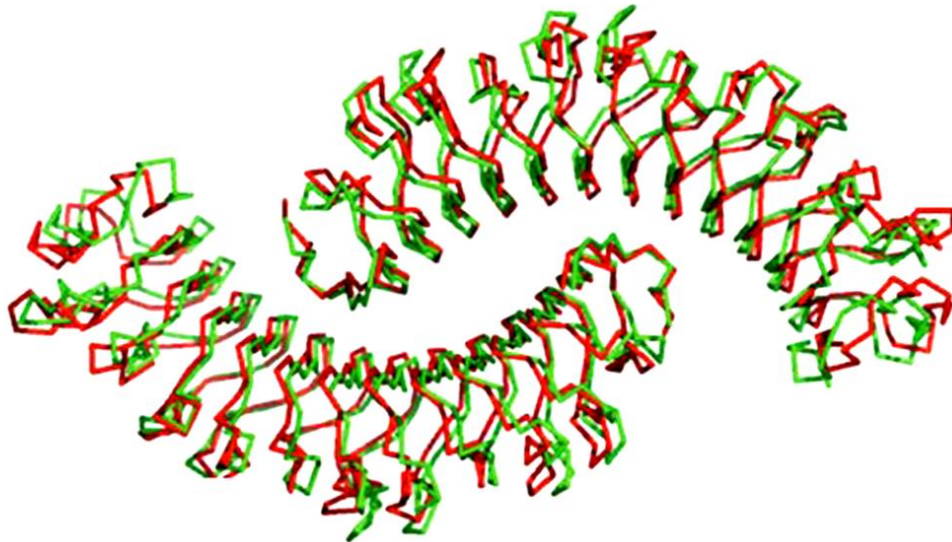
est présent sous forme de PG alors que dans la cornée de l'embryon, il est sous forme de glycoprotéine.

Le lumican constitue la composante majeure de la cornée. Les souris qui n'expriment pas cette molécule montrent une peau fragile et une cornée opaque, principalement dues à la formation de faisceaux de collagène non homogènes. Au cours de l'athérosclérose, il a été démontré que le lumican est surexprimé dans les cellules musculaires lisses des vaisseaux, ces cellules vont alors migrer dans l'intima qui a antérieurement subi un épaissement. Le lumican est impliqué dans différents types de cancers, notamment le cancer du col utérin, et le cancer colorectal. Dans le cas de l'adénocarcinome du sein, un faible taux d'expression du lumican est associé à une progression très rapide du cancer et à une faible chance de survie (Troup *et al.*, 2003).

#### **III.3.3.1.3. Le biglycan**

Comme la décorine, le biglycan est un membre de la classe I de la famille des SLRPs. La protéine "core" est composée de 330 acides aminés et montre une séquence identique à 57% à celle de la décorine. Le biglycan porte deux chaînes de DS proches de l'extrémité N-terminale et deux sites de N-glycosylation (Fisher *et al.*, 1989). Il est exprimé essentiellement dans l'os, le cartilage et le tendon et constitue le SLRP le plus abondant du cartilage articulaire et les fibrocartilages tel que le ménisque du genou (Scott *et al.*, 1997). Il est également présent dans la peau et dans les vaisseaux sanguins.

Comme les autres membres de la famille des SLRPs, la décorine et le biglycan présentent une organisation dite en "fer à cheval" (Figure 18). Ils forment chacun un dimère stable en solution, essentiel pour la stabilité de la protéine. Il a été démontré que ces deux macromolécules ne sont pas en compétition pour la fixation du collagène.



**Figure 18. Représentation schématique de la superposition des dimères de la décorine et du biglycan**

Superposition des chaînes C- $\alpha$  de la décorine (rouge) et du biglycan (vert). Ces deux SLRPs ont une forte homologie de séquence et montrent une superposition quasi-parfaite. Les deux structures ont été exposées dans Xfit.

D'après Scott *et al.*, 2006

Chez les souris, l'inactivation du gène codant pour le biglycan est à l'origine du développement d'un phénotype ostéoporosique sévère. Cette constatation suggère que ce SLRP est impliqué dans l'ostéogenèse (Xu, *et al.*, 1998). Chez les patients atteints de syndrome d'Ullrich-Turner (présence d'un seul chromosome X), la réduction de l'expression du biglycan est associée au nanisme de ces individus. En revanche, la présence d'un chromosome X surnuméraire (syndrome de Klinefelters) est accompagnée d'une augmentation de l'expression du biglycan et d'un développement important de la densité et de la longueur de l'os (Heegaard *et al.*, 1997).

Le biglycan joue un rôle important dans l'assemblage de la MEC, il présente une grande affinité pour le TGF $\beta$  et participe ainsi à la modulation de la prolifération cellulaire.

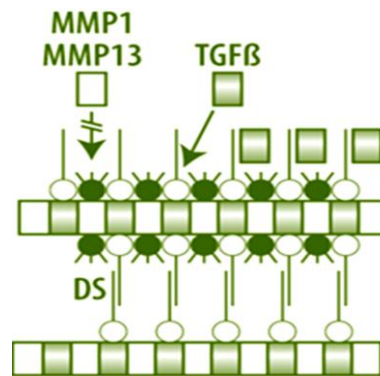
#### **III.3.3.1.4. La décorine**

La décorine est le membre le plus étudié et le mieux caractérisé de la classe I de la famille des SLRPs. Elle porte une séquence pro-peptide fortement conservée entre les espèces qui jouerait le rôle de signal de reconnaissance pour la xylosyltransférase (XT), la première enzyme impliquée dans la synthèse des chaînes de GAGs.



La décorine présente une organisation similaire à celle du biglycan, elle est codée par 8 exons et est caractérisée par une séquence N-terminale riche en Cys: CX<sub>3</sub>CXCX<sub>6</sub>C. La protéine "core" présente 3 sites de N-glycosylation sur le motif LRR et porte une chaîne unique de GAG, greffée sur le domaine N-terminal, et qui peut être un CS ou DS selon le type cellulaire ou tissulaire. La décorine est présente principalement dans la MEC des tissus conjonctifs, la protéine "core" est produite dans différentes cellules, notamment dans les chondrocytes, les fibroblastes, les cellules endothéliales et les cellules musculaires lisses.

La décorine interagit avec de nombreuses molécules telles que la fibronectine et la thrombospondine, et elle est capable de se lier à presque tous les types de collagène, et plus particulièrement le collagène de type I, II, III et VI, et régule ainsi sa fibrogenèse. Il a été démontré qu'elle est impliquée dans la régulation, la formation et/ou le maintien des fibres de collagène (Scott, 1993; Danielson *et al.*, 1997) (Figure 19). Elle interagit aussi avec une grande variété de facteurs de croissance et de cytokines, et présente une forte affinité pour le TGFβ, grâce à sa structure en "fer à cheval" qui permet l'exposition d'un large domaine peptidique aux nombreuses molécules qui se trouvent dans la MEC. Les souris déficientes en décorine (KO), montrent une peau fragile avec une réduction importante de son élasticité et de ses propriétés mécaniques, ainsi que des anomalies qui touchent l'organisation et l'orientation des fibres de collagène.



**Figure 19. Représentation schématique de l'interaction entre les SLRPs et les fibres de collagène.**

La décorine (cercle vide) et la fibromoduline ou lumican (cercle plein) interagissent avec différentes régions du collagène pour couvrir sa surface. Les chaînes de DS de la décorine peuvent s'associer entre elles pour faciliter les interactions entre les fibres de collagène. Les DS de la décorine sont également capables d'interagir avec le TGFβ et le piéger dans la MEC. Le "manteau" formé par les SLRPs empêche les MMPs (MMP1 et MMP13) d'accéder aux fibres de collagène et retarde sa dégradation. MMP, *matrix metallo-proteinase*.

En plus de son rôle dans l'organisation de la MEC, la décorine contrôle la prolifération cellulaire et la fibrogenèse tissulaire (Iozzo and Murdoch, 1996). En effet, ce PG joue un rôle crucial dans l'inhibition de la propagation des cellules cancéreuses, et participe également à la réponse immunitaire.

La décorine inhibe la croissance et la prolifération de nombreuses cellules tumorales en bloquant la voie de signalisation du TGF $\beta$ . En effet, il a été démontré que cette molécule atténue fortement la croissance tumorale dans le cas du carcinome du colon (Reed *et al.*, 2002), du poumon, du foie (Biglari *et al.*, 2004) et dans le cas du gliome (Reed *et al.*, 2005). La décorine est considérée comme un suppresseur de tumeurs. Son expression est réprimée dans de nombreuses cellules cancéreuses à l'exception de l'ostéosarcome. Zafiopoulos *et al.* (2008) ont récemment mis en évidence, pour la première fois, la présence d'une nouvelle voie de signalisation dans les cellules de l'ostéosarcome qui contrecarre l'effet bénéfique de la décorine sur la prolifération cellulaire.

La synthèse défectueuse de l'unique chaîne de GAG de la décorine est associée à une maladie génétique connue sous le nom de syndrome progéroïde d'Ehlers-Danlos (ED). Il a été démontré que ce syndrome génétique est une conséquence directe de mutations ponctuelles du gène codant pour l'enzyme  $\beta$ 1,4-galactosyltransférase 7 ( $\beta$ 4GalT7) humaine. Elle est caractérisée par un tableau clinique lourd avec des altérations qui touchent principalement les tissus conjonctifs. Ainsi, ces patients montrent des anomalies assez proches de celles constatées chez les souris qui n'expriment pas la décorine, avec la présence d'une peau fragile, un retard de développement, une disproportion cranio-faciale, un vieillissement prématuré, des articulations hypermobiles, une ostéopénie, ainsi que des muscles hypotoniques.

#### **III.3.4. Les PGs circulants: exemple de l'endocan**

Les PGs circulants jouent un rôle important dans la médiation de la réaction inflammatoire et dans la signalisation cellulaire. Le PG le plus étudié de ce groupe de molécules est l'endocan.

L'endocan est un DSPG qui porte une seule chaîne de GAG fixée sur la Ser137 de la protéine "core". Il est anciennement connu sous le nom d'ESM-1 (*Endothelial cell-Specific Molecule-1*) (Lassalle *et al.*, 1996; Béchard *et al.*, 2001). C'est un PG qui a la particularité d'être circulant, et a donc été mis en évidence sous forme libre dans la circulation sanguine. Il est sécrété spécifiquement par les cellules endothéliales et est exprimé préférentiellement dans les

tissus rénaux et pulmonaires. L'endocan a été cloné pour la première fois en 1996 par Lassalle *et al.*, à partir d'une banque d'ADNc de cellules HUVEC (*Human Umbilical Vein Endothelial Cell*). Le gène codant est situé sur le chromosome 5 (5q11.2) et il est constitué de 3 exons séparés par deux introns. Le dernier exon code pour la région C-terminale de la protéine "core" de l'endocan qui contient la Ser137 sur laquelle se fixe la chaîne de DS (Depontiou *et al.*, 2008).

Depuis la découverte de ce PG, de nombreuses études ont montré que l'endocan joue un rôle clé dans la régulation de certains processus physiologiques tels que l'adhésion cellulaire et la réponse inflammatoire, ainsi que dans la régulation de processus pathologiques, conduisant à la progression métastatique dans différents types de cancers. Il a été rapporté que l'expression de l'endocan joue un rôle capital dans la pathogenèse du cancer colorectal (Zuo *et al.*, 2008). Ainsi, il a été proposé comme biomarqueur dans le diagnostic du cancer colorectal (Zuo *et al.*, 2008). L'expression de l'endocan augmente de manière remarquable chez les patients atteints de cancer du poumon, ainsi, ce phénomène est associé à l'invasion tumorale (Scherpereel *et al.*, 2003; Grigoriu *et al.*, 2006). Il a été suggéré que cette molécule constituerait un indicateur de mauvais pronostic dans le cas du cancer du sein (Van't Veer *et al.*, 2002), du poumon (Borczuk *et al.*, 2004) et dans le cancer de la prostate (Glinskii *et al.*, 2003).

## ***Chapitre 2 : Les glycosyltransférases (GTs)***

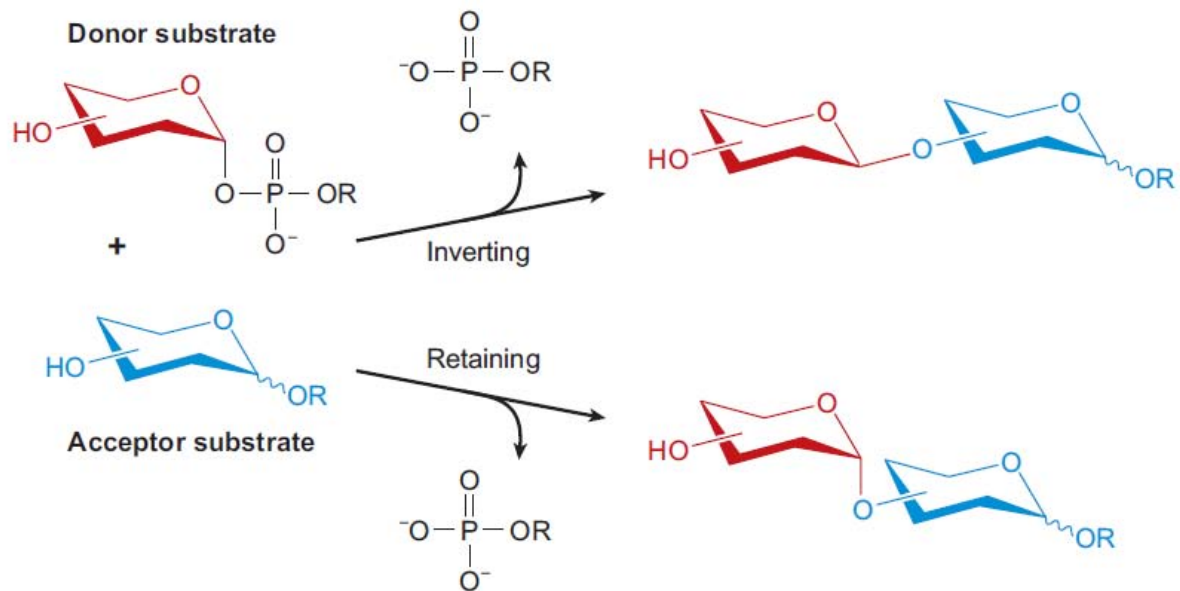
### **Généralités**

La grande complexité des structures oligosaccharidiques rencontrées dans la nature est le résultat d'une parfaite harmonisation de multiples réactions enzymatiques catalysées principalement par les GTs (Rademacher, 1988). Le transfert enzymatique d'un résidu glycosidique est quantitativement l'une des réactions les plus importantes au sein des organismes vivants. Les GTs constituent un vaste groupe d'enzymes impliqué dans la biosynthèse d'oligosaccharides et de polysaccharides. La synthèse des carbohydrates est d'une importance biologique remarquable. Ces molécules contrôlent un large éventail de fonctions cellulaires, notamment le stockage de l'énergie, la glycosylation des protéines, la signalisation cellulaire, la structure de la paroi cellulaire, les interactions cellule-cellule et les interactions hôte/pathogène (Rudd *et al.*, 2001; Wells *et al.*, 2001; Bertozzi *et al.*, 2001).

### **I. CLASSIFICATION DES GTs**

Les GTs constituent un groupe d'enzymes ubiquitaires qui catalysent le transfert d'un résidu saccharidique, principalement à partir d'un sucre nucléotidique donneur, sur différents types de substrats accepteurs, pour former une liaison glycosidique. La classification de ces protéines est basée sur différents critères notamment sur leur homologie de séquence primaire (Campbell *et al.*, 1997; Coutinho *et al.*, 2003). En 2003, la base des données CAZy (Carbohydrate Active enZymes) permettait de distinguer 65 familles de GTs. Ce groupe d'enzymes n'a cessé de s'agrandir depuis le lancement du projet de classification en 1997. En effet, on trouve actuellement près de 90 familles de GTs. L'augmentation du nombre de séquences est le résultat de divers facteurs. En particulier, le séquençage du génome d'organismes eucaryotes supérieurs a permis l'identification de nouvelles GTs. Il a été constaté que certains organismes possèdent des centaines de GTs, démontrant l'impact de la glycosylation dans de multiples fonctions cellulaires. De plus, la reconnaissance de l'importance de la glycobiologie a conduit à l'identification fonctionnelle de nouvelles GTs. La classification de ces protéines prend également en compte le mécanisme réactionnel de ces enzymes et leur structure (Henrissat *et al.*, 1997). Au sein d'une même famille, les GTs catalysent la formation d'une liaison glycosidique de nature stéréochimique identique. Selon la configuration anomérique du substrat donneur, le produit formé peut garder la

configuration anomérique de départ ou subir une inversion de configuration (Coutinho *et al.*, 2003) (Figure 20).



**Figure 20. Mécanisme de la réaction catalysée par les GTs avec rétention ou inversion de la configuration anomérique du sucre donneur**

La réaction enzymatique catalysée par les GTs peut se produire avec inversion ou rétention de la configuration du glycoside donneur. R (couleur noire), nucléoside monophosphate; R (couleur bleue), de nature variable, peut être un sucre, un lipide, un antibiotique, un acide nucléique ou autre.

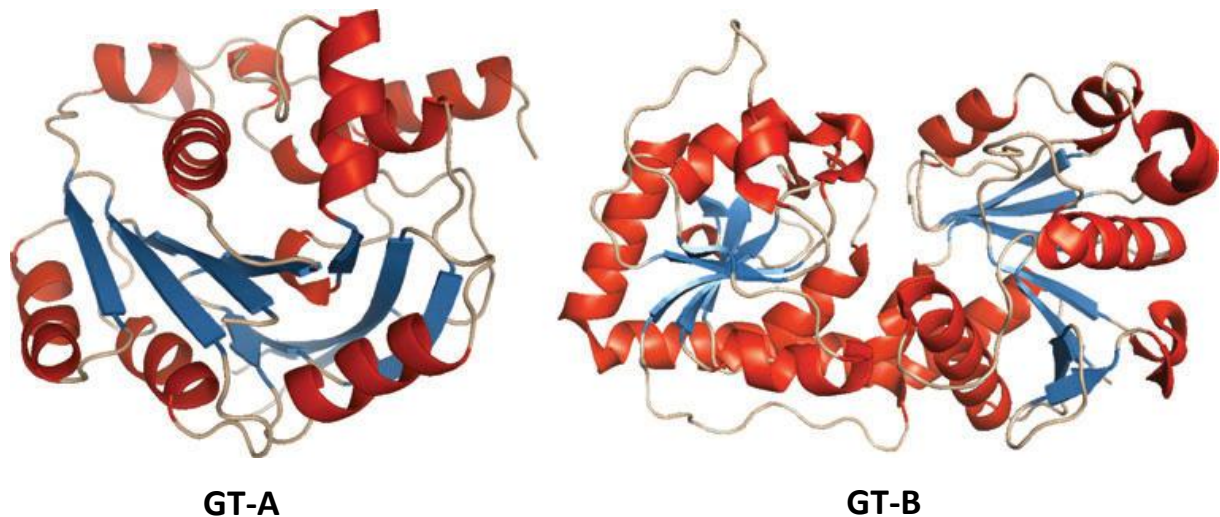
## II. LE REPLIEMENT DES GTs, "FOLDS"

Malgré leur diversité de séquence primaire, la comparaison de la structure des GTs a révélé qu'elles adoptent principalement deux types de repliements (Breton *et al.*, 2006). On parle de GTs qui présentent un "fold" de type GT-A ou GT-B formant deux super-familles (Bourne *et al.*, 200; Breton *et al.*, 2002) (Figure 21).

### II.1. La structure de type GT-A

Cette structure 3D a été décrite pour la première fois pour l'enzyme SpsA de *Bacillus Subtilis*. Le repliement ou organisation GT-A est considéré comme une association de deux domaines de type Rossmann *like* (Rossmann *et al.*, 1974) (structure  $\alpha/\beta/\alpha$ ) associés pour former un feuillet central continu, constitué d'au moins 8 brins  $\beta$ . Pour cette raison, certains auteurs décrivent la structure GT-A comme formant un seul domaine de Rossmann (Figure 21). Ce

dernier est également appelé domaine de liaison des nucléotides car il est typiquement associé à la structure des protéines capables de fixer ces molécules.



**Figure 21. Principales structures 3D des GTS**

Le repliement de type GT-A est représenté par l'enzyme SpsA de *Bacillus Subtilis* et l'organisation GT-B est représentée par la structure de la  $\beta$ -glucosyltransférase du bactériophage T4.

D'après Lairson *et al.*, 2008

Chez les eucaryotes, les protéines de type GT-A possèdent généralement un court domaine cytoplasmique N-terminal suivi d'un segment transmembranaire, lié au domaine catalytique globulaire (région cristallisée) par une région tige ou "stem" (Breton *et al.*, 1999). Il est important de noter que toutes les enzymes qui possèdent une structure de type GT-A ne sont pas forcément des GTs (Brown *et al.*, 1999). La plupart des enzymes GT-A présentent une séquence "signature" ou motif DXD dans lequel l'un ou l'autre des carboxylates forme une liaison de coordination avec un cation divalent (et/ou le ribose du sucre nucléotidique) (Wiggins *et al.*, 1998; Breton *et al.*, 1998). De plus, de nombreuses GT-A utilisent un acide carboxylique, Asp (acide aspartique) ou Glu (acide glutamique), qui joue le rôle de base catalytique (Tarbouriech *et al.*, 2001).

## II.2. La structure de type GT-B

Les enzymes GT-B sont constituées de deux vrais domaines de Rossmann  $\alpha/\beta/\alpha$ . Contrairement aux GT-A, ces deux domaines ne sont pas associés, mais disposés face à face.

La fixation du ligand est accompagnée d'un changement conformationnel de la protéine. Le site actif de ces enzymes est situé à l'intérieur du sillon délimité par les deux domaines de Rossmann (Figure 21). De la même façon que les GT-A, les deux domaines des enzymes GT-B sont associés aux sites de fixation des substrats donneur et accepteur. Hormis les GTs, d'autres types d'enzymes peuvent également adopter un repliement GT-B, telles que l'UDP-GlcNAc-2-épipimérase (Campbell *et al.*, 2000) et la GlcN-1-phosphate-pyrophosphorylase (Brown *et al.*, 1999).

### III. LES GALACTOSYLTRANSFERASES (GalTs)

Les GalTs appartiennent à la super-famille des GTs. Les galactosylations sont des réactions qui contribuent à la grande diversité structurale des oligosaccharides exprimés dans les organismes vivants. Le transfert du Gal à partir d'un sucre nucléotidique sur un substrat accepteur est une réaction enzymatique catalysée par différentes GalTs. Cette activité enzymatique est largement répandue, aussi bien chez les organismes procaryotes que chez les eucaryotes. Chez les procaryotes, les GalTs sont impliquées dans la biosynthèse des polysaccharides capsulaires (Kolkman *et al.*, 1997), des glycanes associés à la piline (Jennings *et al.*, 1998) et dans la formation des lipopolysaccharides attachés à un antigène par une liaison O-glycosidique. Dans les levures *Schizosaccharomyces pombe*, une GalT participe à la formation du galactomannane de la paroi cellulaire (Chappell *et al.*, 1994). Dans différents types de plantes, les GalTs interviennent également dans la synthèse des galactolipides plastidiques (Miege *et al.*, 1999; Awai *et al.*, 2001), des flavonols galactosylés (Miller *et al.*, 1999) et dans la formation des galactomannanes (Edwards *et al.*, 1999). Chez les eucaryotes, la principale fonction des GalTs est d'assurer la biosynthèse des glycoprotéines et des glycolipides.

Chez les eucaryotes, les GalTs sont localisées au niveau de l'appareil de Golgi et sont impliquées dans la synthèse du lactose et de différentes classes de glycoconjugués. Les GalTs présentent entre-elles une faible homologie de séquence. Elles catalysent le transfert d'un Gal sur divers substrats accepteurs en formant différents types de liaisons osidiques. Ainsi, les GalTs catalysent la formation de liaisons glycosidiques de type:  $\alpha$ 1,2,  $\alpha$ 1,3,  $\alpha$ 1,4,  $\alpha$ 1,6,  $\beta$ 1,3 ou  $\beta$ 1,4. En se basant sur la nature de la liaison formée, les GalTs sont classées en familles d'enzymes. Ainsi, on distingue la famille des  $\alpha$ 2GalTs,  $\alpha$ 3GalTs,  $\alpha$ 4GalTs,  $\alpha$ 6GalTs,  $\beta$ 3GalTs et  $\beta$ 4GalTs. Seule cette dernière famille de GalTs sera présentée dans ce manuscrit.

### III.1. La famille des $\beta$ 4GalTs

Ce sont des protéines transmembranaires qui présentent une topologie de type II, avec un court segment cytosolique N-terminal, un seul domaine transmembranaire suivi d'une région en forme de tige, appelée région "stem", et un large domaine catalytique luminal C-terminal. Les membres de cette famille possèdent de nombreuses caractéristiques communes, notamment (1) l'utilisation de l'UDP-Gal comme substrat donneur, (2) la formation d'une liaison glycosidique  $\beta$ 1,4, (3) l'utilisation de substrats accepteurs osidiques similaires: GlcNAc, Glc ou Xyl. De plus, ces enzymes présentent une séquence "signature", DXD, comme la plupart des GTs possédant une structure de type GT-A.

Cette famille d'enzymes est constituée de 7 sous-familles:  $\beta$ 4GalT1 à  $\beta$ 4GalT7. La caractérisation des gènes codant pour les  $\beta$ 4GalTs met en évidence une relation évolutive avec une organisation génomique similaire et de fortes homologies de séquence primaires. A l'exception de la  $\beta$ 4GalT7, les  $\beta$ 4GalTs contiennent quatre résidus Cys au niveau de leur domaine catalytique. Les  $\beta$ 4GalTs subissent des modifications post-traductionnelles de type N-glycosylation. En effet, les  $\beta$ 4GalT5 et T6 portent un grand nombre de sites potentiels de N-glycosylation au niveau de la région "stem". En outre, un site de glycosylation conservé a été identifié sur l'extrémité C-terminale des  $\beta$ 4GalT2, T3, T4, T5 et T6.

La famille des  $\beta$ 4GalTs présente une homologie de séquence avec les polypeptides  $\alpha$ GalNAcTs. Plus précisément, les enzymes  $\beta$ 4GalTs possèdent un motif peptidique (WG(G/R)EDD(E/D)) similaire au motif (WGGENXE) identifié dans différents membres de la famille de GalNAcTs (Clausen and Bennett, 1996; Amado *et al.*, 1999).

Bien que les  $\beta$ 4GalTs présentent des activités enzymatiques similaires, Furukawa *et al.* (1999) ont montré que leurs spécificités vis-à-vis des substrats accepteur et donneur sont différentes et que leur expression est tissu-spécifique. Les  $\beta$ 4GalT1 et T2 participent à la synthèse de la N-acétyl-lactosamine dans les glycolipides et les glycoprotéines et possèdent également une activité lactose synthase qui peut être modulée par l' $\alpha$ -lactalbumine (LA). Les  $\beta$ 4GalT3 et T4 synthétisent la N-acétyl-lactosamine et cette activité n'est pas régulée par la LA (Tableau 5).



	Localisation chromosomique (chez l'homme)	Activité enzymatique	Expression	Produit formé	Voie de biosynthèse
<b><math>\beta</math>4GalT1</b>	9p13	Gal $\beta$ 1 $\rightarrow$ 4GlcNAc $\beta$ -	Ubiquitaire	N-acétyl-lactosamine Lactose (LA)	Glycolipide Glycoprotéine
<b><math>\beta</math>4GalT2</b>	1p34-p33	Gal $\beta$ 1 $\rightarrow$ 4GlcNAc $\beta$ -	Ubiquitaire	N-acétyl-lactosamine Lactose (LA)	Glycolipide Glycoprotéine
<b><math>\beta</math>4GalT3</b>	1q23	Gal $\beta$ 1 $\rightarrow$ 4GlcNAc $\beta$ -	Ubiquitaire	N-acétyl-lactosamine Lactose	Glycolipide Glycoprotéine
<b><math>\beta</math>4GalT4</b>	3q13.3	Gal $\beta$ 1 $\rightarrow$ 4GlcNAc $\beta$ -	Ubiquitaire	N-acétyl-lactosamine Lactose	Glycolipide Glycoprotéine
<b><math>\beta</math>4GalT5</b>	20q13	Gal $\beta$ 1 $\rightarrow$ 4GlcNAc $\beta$ -	Ubiquitaire	Non déterminée	Non déterminée
<b><math>\beta</math>4GalT6</b>	18q11	Gal $\beta$ 1 $\rightarrow$ 4Glc $\beta$ 1-Cer	Cerveau Poumon Pancréas Moelle épinière	Lactosylcéramide	Glycolipide Glycoprotéine
<b><math>\beta</math>4GalT7</b>	5q35.1-35.3	Gal $\beta$ 1 $\rightarrow$ 4Xyl $\beta$ 1	Ubiquitaire	[Xyl $\beta$ 1,4Gal]	PGs

**Tableau 5:** Principales caractéristiques fonctionnelles et distribution tissulaire des  $\beta$ 4GalTs

Cer, céramide; Gal, galactose; Glc, glucose; GlcNAc, N-acétylglucosamine; Xyl, xylose; LA,  $\alpha$ -lactalbumine.

D'après Hennet, 2002

De façon importante, il a été démontré récemment que la  $\beta$ 4GalT4 est impliquée dans la formation des chaînes de KS qui sont indispensables à la structure de la cornée (Kitayama *et al.*, 2007). Les fonctions biologiques des  $\beta$ 4GalT5 et T6 ne sont pas bien déterminées (Sato *et*

*al.*, 1998; Van Die I., *et al.*, 1999). Néanmoins, il est établi que la  $\beta$ 4GalT6 possède une activité lactocéramide synthase, importante pour la biosynthèse des glycolipides (Nomura *et al.*, 1998). La  $\beta$ 4GalT7 est le dernier membre de la famille des  $\beta$ 4GalTs. Elle catalyse le transfert du premier Gal de l'amorce tétrasaccharidique sur le Xyl lié à la protéine "core" des PGs.

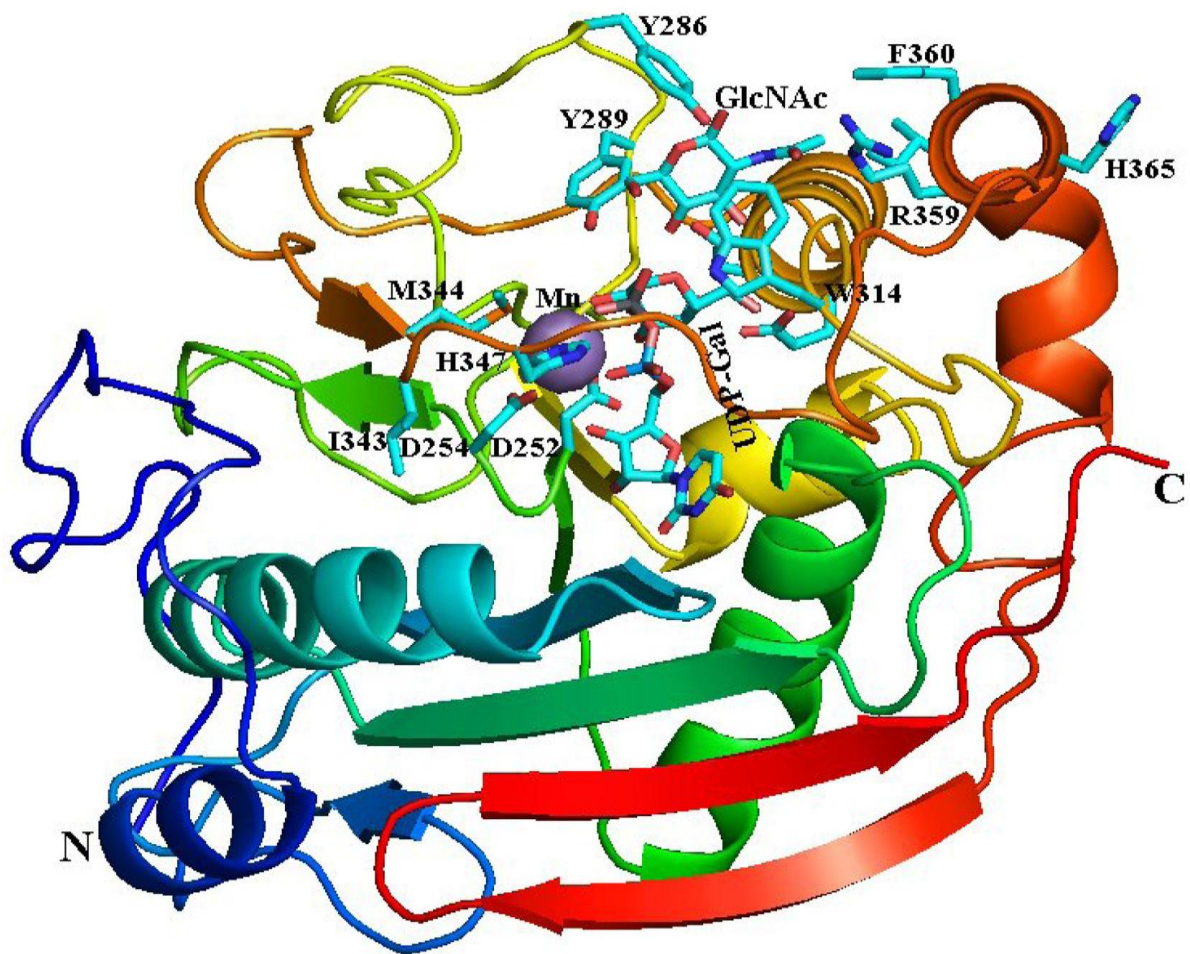
### III.2. La $\beta$ 4GalT1 bovine

La  $\beta$ 4GalT1 catalyse le transfert d'un Gal à partir de l'UDP-Gal sur un résidu GlcNAc, c'est une réaction qui nécessite des ions  $Mn^{2+}$ . En présence de LA, qui est une protéine spécifique des glandes mammaires, l'enzyme forme un complexe lactose synthase qui permet le transfert d'un Gal sur un résidu Glc libre, *via* la formation d'une liaison glycosidique de type  $\beta$ 1,4 et permet ainsi la production du lactose (Gal $\beta$ 1,4Glc) (Morrison et Ebner, 1971). Seules la  $\beta$ 4GalT1 et T2 ont une activité lactose synthase en présence de LA (Almeida *et al.*, 1997).

La  $\beta$ 4GalT1 bovine est la première GalT à avoir été isolée et clonée (Narimatsu *et al.*, 1986) et sa structure 3D a été résolue pour la première fois par Gastinel *et al.* (Gastinel *et al.*, 1999). L'étude de cette structure en absence d'UDP-Gal montre la présence d'un seul domaine catalytique conique, avec une large poche ouverte située à sa base. La fixation de l'UDP-Gal induit un changement de conformation très important, et la poche catalytique vient recouvrir les résidus d'acides aminés qui interagissent avec le sucre nucléotidique et passe ainsi d'une conformation I (poche catalytique ouverte) à une conformation II (poche catalytique fermée) (Ramakrishnan and Qasba, 2001).

Le domaine catalytique de la  $\beta$ 4GalT1 bovine possède une séquence <sup>252</sup>DVD<sup>254</sup> qui présente un site potentiel de fixation des cations divalents métalliques. Ce motif DXD est conservé dans toutes les  $\beta$ 4GalTs et il a été démontré qu'il présente une forte affinité envers les cations  $Mn^{2+}$  et qu'il ne peut pas fixer le  $Co^{2+}$  (Boeggeman and Qasba, 2002).

L'analyse de la structure de la  $\beta$ 4GalT1 bovine cristallisée en présence d'UDP et de LA a montré que les ions  $Mn^{2+}$  forment cinq liaisons de coordination: deux liaisons avec le pyrophosphate du sucre nucléotidique et trois liaisons avec le C $\delta$  de l'aspartate en position 254 (Asp254), le C $\epsilon$ 2 de l'histidine en position 347 (His347) et le S $\delta$  de la méthionine en position 344 (Met344) (Figure 22).



**Figure 22. Structure cristallographique du domaine catalytique de la  $\beta$ 4GalT1 bovine en présence d'UDP-Gal et d'ions  $Mn^{2+}$**

L'analyse de la structure cristallographique montre une organisation globulaire pliée qui forme une poche conique avec deux extrémités dissemblables, N-terminale en bleu et C-terminale en rouge. Le domaine N-terminal qui intervient dans la reconnaissance du sucre nucléotidique contient des feuillets  $\beta$  (présentés par des flèches), qui sont disposés de part et d'autre d'hélices  $\alpha$  présentées par des rubans. Le domaine C-terminal qui est impliqué dans la reconnaissance du substrat accepteur est constitué principalement de brins  $\beta$ . La liaison entre l'UDP-Gal et la  $\beta$ 4GalT1 bovine produit une conformation de type II (fermée) dans laquelle le résidu W314 (équivalent du W224 de la  $\beta$ 4GalT7 humaine) situé à l'intérieur de la poche catalytique, entre en interaction à la fois avec le substrat donneur et le substrat accepteur.

D'après Qasba *et al.*, 2008

La séquence <sup>312</sup>WGWGGEDDD<sup>320</sup> de la  $\beta$ 4GalT1 bovine présente une structure en forme de boucle, dont les résidus aminés Asp318 et Asp319 sont directement impliqués dans la fixation du substrat accepteur GlcNAc (Boeggeman and Qasba, 2002). Plus précisément, l'oxygène du

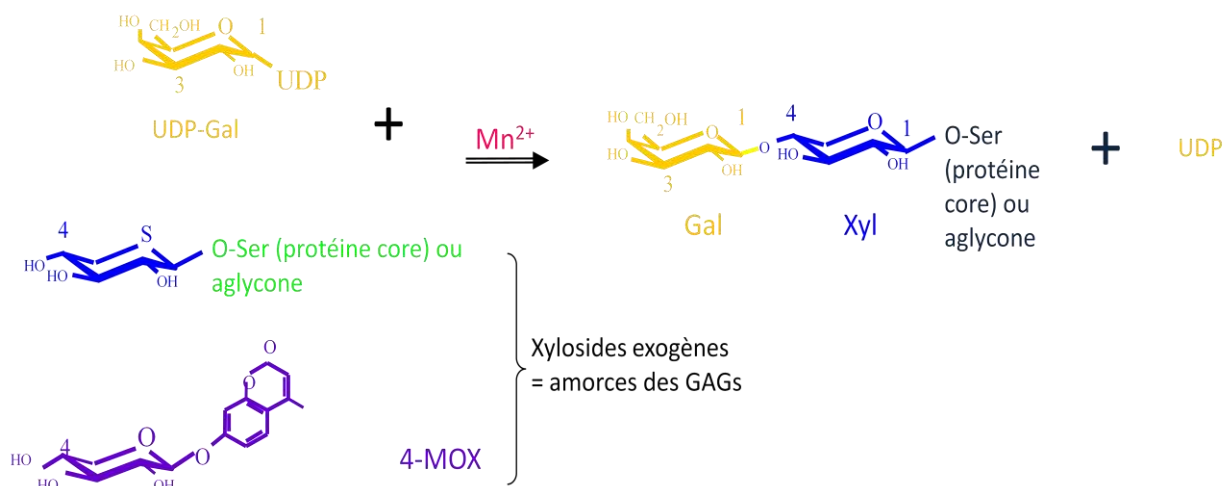
groupement carboxylate de l'Asp228 forme des liaisons hydrogènes avec O3 et O4 du GlcNAc (ou Glc) (Ramakrishnan and Qasba, 2001). Néanmoins, l'étude réalisée par Zhang (Zhang *et al.*, 2003) suggère que l'Asp228 n'affecte pas directement la fixation du substrat donneur. Cela contredit l'hypothèse de Hatanaka *et al.* (2001) qui a identifié, en utilisant un marquage de photoaffinité du GlcNAc, l'Asp228 comme étant le résidu aminé qui fixe le substrat accepteur.

### III.3. Les $\beta$ 4GalT7s

#### III.3.1. Les propriétés enzymatiques

La  $\beta$ 4GalT7 (xylosylprotéine- $\beta$ 1,4-galactosyltransferase; EC 2.4.1.133) catalyse le transfert d'un Gal à partir d'un substrat donneur, l'UDP-Gal, sur l'hydroxyle en position 4 du Xyl lié à une Ser de la protéine "core" des PGs.

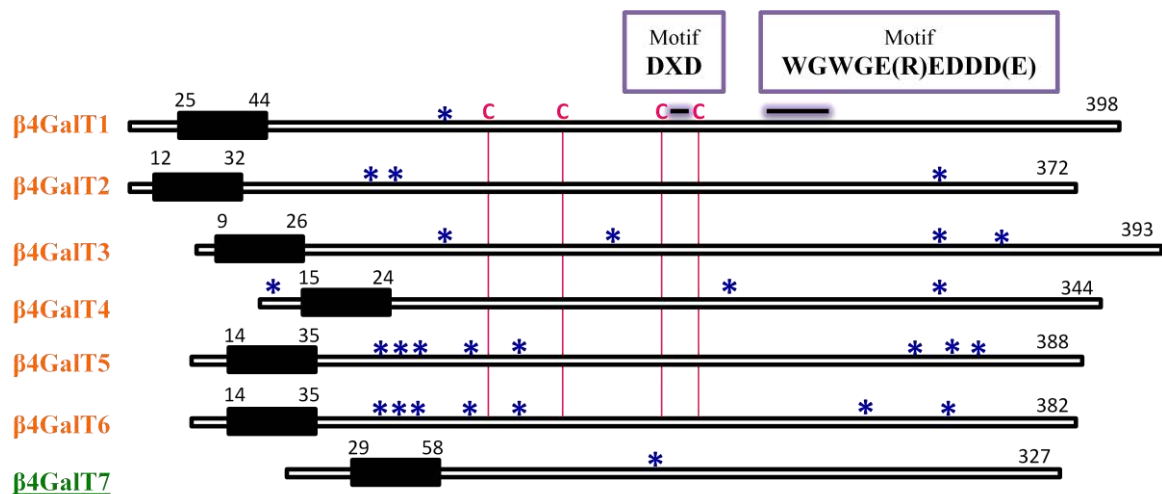
Cette enzyme est capable d'utiliser des substrats accepteurs exogènes de type xylosidique, comprenant un Xyl lié à une aglycone, notamment le 4-méthylombelliferyl-xylopyranoside (4-MOX) et le 4-nitrophényl-xylopyranoside (4-NP-Xyl), pour initier la biosynthèse des GAGs *in vivo* et *in vitro*. La réaction enzymatique catalysée par la  $\beta$ 4GalT7 nécessite la présence obligatoire de cations divalents, principalement le manganèse, comme cofacteur (Figure 23).



**Figure 23. Représentation schématique de la réaction enzymatique catalysée par la  $\beta$ 4GalT7**

En présence d'ions manganèse, la  $\beta$ 4GalT7 catalyse le transfert du Gal sur le Xyl d'un substrat accepteur endogène (Xyl-protéine "core") ou exogène (Xyl-aglycone) pour initier la synthèse des chaînes de GAGs.

Chez l'homme, le gène codant pour cette enzyme a été identifié en 1999 par la recherche d'homologies de séquence dans la base des données EST (Expressed Sequence Tags) (Almeida *et al.*, 1999). L'organisation génomique de la  $\beta$ 4GalT7 est différente de celle des autres  $\beta$ 4GalTs. Son gène ne possède pas les cinq domaines intron/exon conservés dans les quatre premières  $\beta$ 4GalTs mais contient six exons (1-50, 51-413, 414-369, 640-723, 724-828 et 829-984). L'analyse de la séquence primaire de la  $\beta$ 4GalT7 montre également que ce dernier est le membre le plus distant de cette famille d'enzymes. Quatre résidus Cys du site catalytique, conservés au sein des six premières  $\beta$ 4GalTs, sont absents dans la  $\beta$ 4GalT7. De plus, des motifs peptidiques hautement conservés dans les  $\beta$ 4GalTs montrent des variations significatives dans la  $\beta$ 4GalT7, notamment au niveau de la séquence WGWGE(R)EDDD(E) (Figure 24). Enfin, l'analyse du produit de la réaction catalysée par cette enzyme a révélé la formation d'une liaison Gal $\beta$ 1,4Xyl $\beta$ 1-R, impliquée dans la biosynthèse des GAGs (Okajima *et al.*, 1999a; Almeida *et al.*, 1999).



**Figure 24. Comparaison schématique de l'organisation des membres de la famille des  $\beta$ 4GalTs chez l'homme**

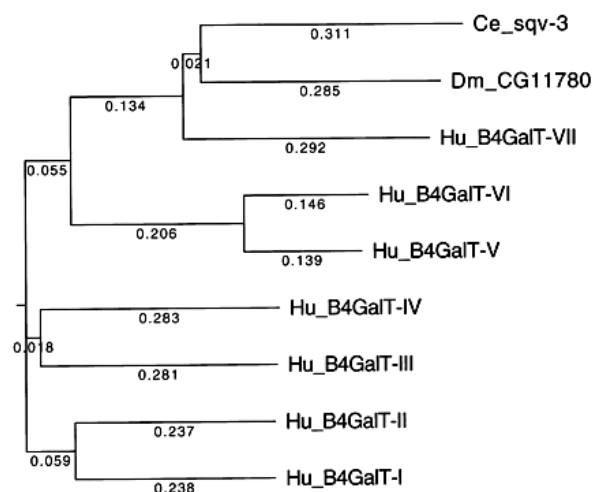
Le domaine catalytique des  $\beta$ 4GalTs contient un premier motif conservé DXD et un deuxième motif conservé qui présente une mutation non conservative E/R et une mutation conservative D/E sur la séquence de la  $\beta$ 4GalT7: WGWGE(R)EDDD(E). Les résidus Cys conservés dans les six premiers membres de cette famille d'enzymes sont indiqués par un trait rose et les sites potentiels de N-glycosylation par des étoiles bleues. Les domaines transmembranaires sont représentés par des rectangles noirs.

D'après Amado *et al.*, 1999

La  $\beta$ 4GalT7 humaine est une protéine membranaire de type II, composée de 327 acides aminés. Le domaine cytoplasmique (28 acides aminés) situé à l'extrémité N-terminale est suivi du domaine transmembranaire composé de 30 acides aminés. Cette enzyme présente un large domaine C-terminal (269 acides aminés) situé dans le lumen de l'appareil de Golgi. Ce domaine contient la région "stem" et le site catalytique de l'enzyme qui comprend des motifs conservés au sein de la famille des  $\beta$ 4GalTs. La  $\beta$ 4GalT7 humaine présente une masse moléculaire de 37 kDa et possède un seul site de N-glycosylation (Figure 24).

Okajima *et al.* (1999a) ont démontré que le gène codant la  $\beta$ 4GalT7 humaine est un orthologue du gène *sqv-3* de *Caenorhabditis elegans* (*C. elegans*): leur séquence primaire présente 38% d'identité. Il a été suggéré dans de nombreuses études qu'il existe une relation évolutive entre les membres de la famille des  $\beta$ 4GalTs au sein de différentes espèces (vertébrés et invertébrés). Les gènes codant sont issus d'événements de duplication à partir d'un gène ancestral commun (Hennet, 2002). La séquence protéique de la  $\beta$ 4GalT7 humaine et de drosophile sont similaires à 58% et l'alignement de séquences de leur domaine catalytique montre une homologie de l'ordre de 73% (Ramakrishnan and Qasba, 2010). De la même façon, l'analyse de la séquence protéique de la  $\beta$ 4GalT7 humaine et de la  $\beta$ 4GalT1 bovine montre une similarité de 33%.

Au cours de ce travail, une étude phylogénétique approfondie de la famille des  $\beta$ 4GalT7s a été réalisée. Elle sera présentée dans la partie « résultats » sous forme de publication (Publication n°1).



**Figure 25. Analyse de la relation phylogénétique entre les  $\beta$ 4GalTs humaines, la  $\beta$ 4GalT7 de drosophile et de *sqv-3***

D'après Vadaie *et al.*, 2002

Il est important de noter que les principales études qui ont été réalisées sur la  $\beta$ 4GalT7 humaine ont été basées sur la structure 3D de la  $\beta$ 4GalT1 bovine, qui était la seule  $\beta$ 4GalT cristallisée. Depuis peu, on dispose également de la structure 3D de la  $\beta$ 4GalT7 de drosophile (Ramakrishnan and Qasba, 2010). La  $\beta$ 4GalT7 de drosophile présente une forte similarité de séquences avec la  $\beta$ 4GalT7 humaine et la  $\beta$ 4GalT1 bovine. Ramakrishnan and Qasba (2010) ont réussi à élucider l'organisation 3D du site catalytique de cette enzyme, en présence d'UDP et d'ions  $Mn^{2+}$ , en produisant une protéine chimérique. En effet, différentes constructions ont été réalisées avant d'arriver à la production d'une forme chimérique, active après solubilisation et cristallisable. L'alignement de séquence des  $\beta$ 4GalT7s humaine et de drosophile montre que ces deux protéines présentent une forte similarité à partir de l'His93 et de l'His75 respectivement. Il a été d'ailleurs suggéré que le domaine catalytique pourrait commencer à partir de ces acides aminés. La cristallisation de la  $\beta$ 4GalT7 de drosophile délétée des 70 premiers acides aminés de l'extrémité N-terminale n'a pas abouti. Il a été proposé que les acides aminés de l'extrémité C-terminale, caractéristiques de la séquence de drosophile, pourraient altérer la cristallisation de l'enzyme. Pour cette raison, ce groupe de chercheurs a réalisé une 2<sup>ème</sup> construction qui a permis de produire une protéine dépourvue des 70 acides aminés N-terminaux et des 11 acides aminés C-terminaux. La cristallisation de cette protéine tronquée n'a également pas pu être réussie. Enfin, la comparaison de la  $\beta$ 4GalT7 de drosophile avec la  $\beta$ 4GalT1 bovine par alignement de séquence a permis d'identifier une forte similarité au niveau de leur domaine catalytique respectif, qui commence à partir du résidu aminé 176 de la protéine bovine. Toutefois, cette approche a permis de constater que le domaine catalytique de la  $\beta$ 4GalT7 de drosophile est dépourvu d'une séquence N-terminale correspondant au peptide situé entre les acides aminés 129 et 175 de l'enzyme bovine. Il a été suggéré que l'absence de cette séquence dans la  $\beta$ 4GalT7 de drosophile expliquerait, vraisemblablement, les difficultés rencontrées lors des premières tentatives de cristallisation de cette protéine. La fusion du peptide 143-175 de la  $\beta$ 4GalT1 bovine à l'extrémité N-terminale de la  $\beta$ 4GalT7 de drosophile sous forme tronquée, dépourvue des 70 acides aminés N-terminaux et des 11 acides aminés de l'extrémité C-terminale (comme décrit précédemment), a rendu la cristallisation de cette protéine possible. La  $\beta$ 4GalT7 de drosophile a été donc cristallisée en présence de l'UDP et du cofacteur métallique  $Mn^{2+}$ .

Quand au résidu Asp320 (équivalent de l'Asp230 de la  $\beta$ 4GalT7 humaine), la structure cristallographique de la  $\beta$ 4GalT1 bovine n'a montré aucun contact entre cet acide aminé et la GlcNAc. Le groupement carboxylate de l'Asp320 forme une liaison hydrogène avec la Tyr311 (équivalent à Phe221 de la  $\beta$ 4GalT7 humaine), et il a été suggéré que cette liaison participe à la stabilisation de la structure de la boucle formée par le peptide  $^{312}\text{WGWGGEDDD}^{320}$ .

Une étude réalisée par Boeggeman et Qasba (Boeggeman and Qasba, 2002) a montré que la mutation du Glu317 (équivalent du Glu227 de la  $\beta$ 4GalT7 humaine qui a fait l'objet de cette étude) ou de l'Asp310 provoque une modification de l'activation de l'enzyme par les cations  $\text{Ca}^{2+}$  et  $\text{Co}^{2+}$ , par rapport à la  $\beta$ 4GalT1 bovine sauvage. Cependant, à l'heure actuelle les données structurales n'indiquent pas la fixation de cations métalliques au niveau de la région  $^{314}\text{WGGEDDDIY}^{322}$  de la  $\beta$ 4GalT1 bovine.

### III.3.2. La $\beta$ 4GalT7 de drosophile (*Drosophila melanogaster*)

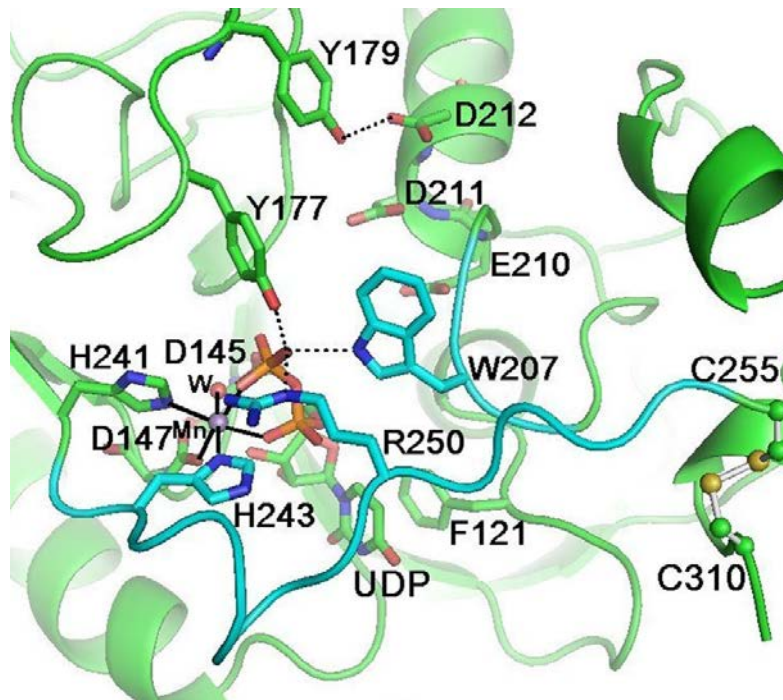
La  $\beta$ 4GalT7 de drosophile est l'orthologue de la  $\beta$ 4GalT7 humaine. C'est une protéine constituée de 322 acides aminés et qui présente une activité enzymatique  $\beta$ 4GalT, comme décrit précédemment.

La structure 3D de la  $\beta$ 4GalT7 de drosophile a été récemment résolue par Ramakrishnan et Qasba (2010). L'organisation cristallographique de cette enzyme montre que, l'Asp147 du motif DXD forme une liaison de coordination avec le cation  $\text{Mn}^{2+}$  (Figure 26).

De façon intéressante, la cristallisation de la  $\beta$ 4GalT7 bovine a permis d'identifier, pour la première fois, un nouveau motif  $^{241}\text{HXH}^{243}$  de fixation du cofacteur métallique, jamais identifié dans la structure 3D des GTs. De plus, l'alignement de séquence des  $\beta$ 4GalT7s montre que ce motif est conservé dans toutes les espèces.

La structure cristallographique de la  $\beta$ 4GalT7 de drosophile montre que le manganèse forme six liaisons de coordination: une avec l'Asp147 (équivalent de l'Asp165 de la  $\beta$ 4GalT7 humaine), deux liaisons avec les atomes d'azote N $\epsilon$ 2 et N $\delta$ 1 de l'His241 et His243, respectivement (équivalents de l'His257 et His259 de la  $\beta$ 4GalT7 humaine), deux liaisons avec les atomes d'oxygène de chacun des deux groupements phosphates de la molécule d'UDP, et une liaison avec une molécule d'eau (Figure 26).





**Figure 26. Organisation cristallographique de la fixation du manganèse et de l'UDP sur la  $\beta$ 4GalT7 de drosophile**

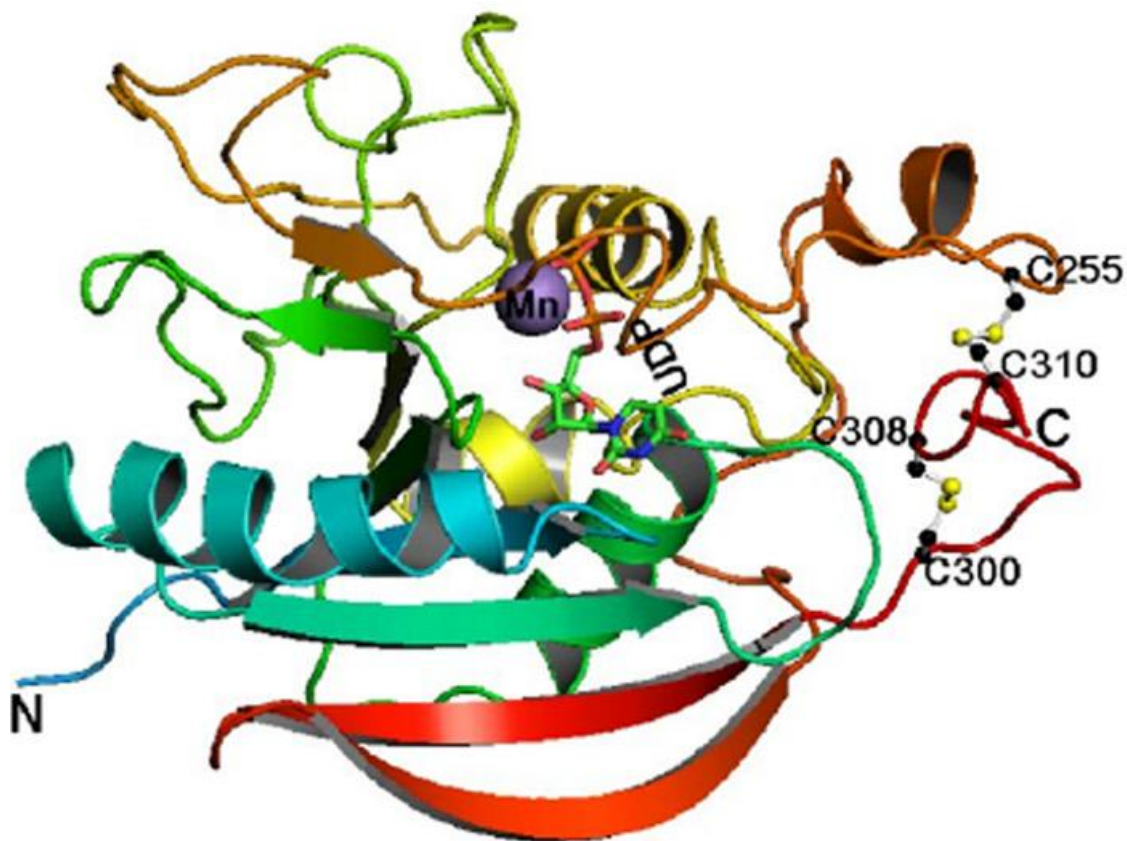
Le manganèse forme six liaisons de coordination avec l'Asp147, l'His241, l'His243, une molécule d'eau et deux liaisons avec les groupements phosphates de l'UDP.

D'après Ramakrishnan and Qasba, 2010

La superposition des structures 3D de la  $\beta$ 4GalT7 de drosophile et de la  $\beta$ 4GalT1 bovine a montré une organisation similaire du site de fixation du substrat donneur et de la poche catalytique. En revanche, la comparaison des sites de fixation du substrat accepteur a montré des différences importantes, probablement dues à la spécificité différente de ces deux enzymes vis-à-vis du substrat accepteur. La deuxième divergence importante de l'organisation 3D a été observée au niveau des extrémités N- et C-terminales. En effet, la  $\beta$ 4GalT7 de drosophile est dépourvue d'un peptide N-terminal identifié dans la  $\beta$ 4GalT1 bovine. Cependant, cette dernière ne possède pas le peptide C-terminal identifié dans l'enzyme de drosophile. Il a été suggéré que ces deux régions peptidiques pourraient structurellement se compenser (Ramakrishnan and Qasba, 2010) (Figure 27).

Enfin la modélisation du site de fixation du Xyl, substrat accepteur de la  $\beta$ 4GalT7, suggère que la chaîne aromatique du résidu Tyr177 interagirait avec l'atome C5 de cette molécule, produisant un encombrement stérique qui empêche d'autres interactions avec cet atome

(Ramakrishnan and Qasba, 2010). Il a été suggéré que la fixation du Xyl sur la  $\beta$ 4GalT7 de drosophile serait similaire à la fixation du substrat GlcNAc ou Glc sur la  $\beta$ 4GalT1 bovine. Ainsi les atomes O2, O3 et O4 des groupements hydroxyles du résidu glycosidique du substrat accepteur formeraient des liaisons hydrogène avec les atomes d'oxygène des carboxylates de l'Asp212 et Asp211 de la  $\beta$ 4GalT7 de drosophile (Asp319 et Asp318 de la  $\beta$ 4GalT1 bovine). Ainsi, l'Asp211 constituerait la base catalytique de cette enzyme (Ramakrishnan and Qasba, 2010).



**Figure 27. Structure cristallographique de la  $\beta$ 4GalT7 de drosophile en présence d'ions  $Mn^{2+}$  et d'une molécule d'UDP**

La couleur du diagramme du ruban de la  $\beta$ 4GalT7 change progressivement de couleur pour passer de la couleur bleue des résidus de l'extrémité N-terminale à la couleur rouge des résidus de l'extrémité C-terminale. Les ponts disulfures formés par les résidus Cys (255, 300, 308 et 310) sont situés du côté C-terminal. La molécule de  $Mn^{2+}$  est représentée par une boule violette. L'organisation 3D de la  $\beta$ 4GalT7 de drosophile est très proche de la structure cristallographique de la  $\beta$ 4GalT1 bovine.

D'après Ramakrishnan and Qasba, 2010

### ***Chapitre 3 : Les voies de biosynthèse des PGs***

Les PGs et leurs chaînes de GAGs sont des molécules diversifiées et très complexes. Ils possèdent un vaste spectre d'action et jouent des rôles majeurs dans de nombreuses pathologies, notamment dans le cas des arthropathies d'origine dégénérative, inflammatoire, et cancéreuse. La modification de l'expression des enzymes impliquées dans la synthèse ou la dégradation des GAGs est observée au cours des différents stades de la progression tumorale et métastatique. De façon intéressante, un effet bénéfique de l'Hp associée aux traitements anti-cancéreux a été récemment mis en évidence. Un essai clinique a été consacré à l'étude de l'effet de la nadroparine (Hp de faible poids moléculaire) chez des patients atteints de cancers à un stade avancé sans thromboembolisme veineux. Cette étude a montré que le traitement à l'Hp influence favorablement la survie des malades (Klerk *et al.*, 2005). Plusieurs études suggèrent un effet favorable de l'association entre la chimiothérapie conventionnelle et l'Hp. Un essai clinique de traitement des patients atteints d'une forme agressive de cancer des poumons a démontré que l'association d'Hp de faible poids moléculaire avec la phosphocyclamide, l'épirubicine et la vincristine permet d'améliorer la réponse à ces traitements comparée aux résultats obtenus avec la chimiothérapie conventionnelle (Altinbas *et al.*, 2004).

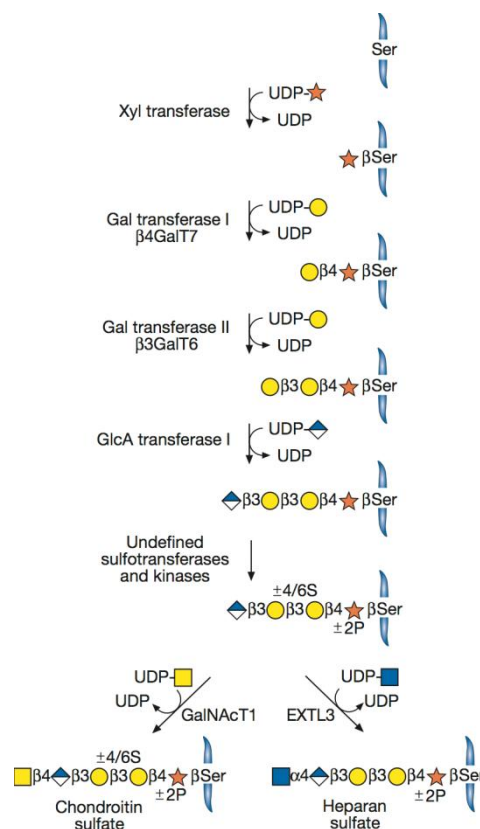
Les GAGs constituent donc des cibles et des outils pharmacologiques dans diverses situations pathologiques, d'où l'intérêt de la connaissance de leur structure, la compréhension de leur mécanisme d'action et de la caractérisation des enzymes impliquées dans leur métabolisme.

#### **I. LA BIOSYNTHESE DES PGs**

A l'exception de l'AH, la biosynthèse des PGs commence toujours par la formation de la protéine "core" qui possède des séquences consensus d'acides aminés constituant les sites d'ancrage des chaînes de GAGs. La formation des GAGs fait ensuite intervenir plusieurs familles d'enzymes qui agissent de façon coordonnée. Il s'agit de GTs, de sulfotransférases (STs) et d'épimérasés qui sont localisées principalement dans le RE et dans les différents compartiments de l'appareil de Golgi.

## II. LA FORMATION DE L'AMORCE TETRASACCHARIDIQUE

A l'exception des chaînes de KS (Baker *et al.*, 1975; Stein *et al.*, 1982), la biosynthèse des GAGs est initiée par la formation d'une amorce tétrasaccharidique [GlcA $\beta$ 1,3Gal $\beta$ 1,3Gal $\beta$ 1,4Xyl $\beta$ 1] (appelée également région de liaison ou "linkage region" entre la protéine "core" et les chaînes de GAGs) commune aux Hp/HS et CS/DS. Les premières enzymes impliquées dans ce processus sont les XTs (XTI/XTII), qui catalysent le transfert d'un Xyl sur un résidu d'acide aminé Ser (ou Thr) de la protéine "core" en formant une liaison O-glycosidique. Puis l'enzyme  $\beta$ 4GalT7 ajoute le premier Gal sur l'extrémité non réductrice du Xyl, suivie de la  $\beta$ 3GalT6 qui transfère le deuxième Gal. Enfin, l'ajout du dernier sucre, GlcA, sur le deuxième Gal de l'amorce tétrasaccharidique est assuré par la  $\beta$ 1,3-glucuronosyltransférase I (GlcAT-I) (Figure 28).



**Figure 28. Les enzymes impliquées dans la biosynthèse de l'amorce tétrasaccharidique commune aux CS/DS et Hp/HS**

Illustration de la synthèse et des modifications de l'amorce tétrasaccharidique commune aux différentes chaînes de GAGs.

D'après Esko *et al.*, 2009, Essentials of Glycobiology

La  $\beta$ 4GalT7 (appelée aussi GalT-I) est une enzyme qui appartient à la famille des  $\beta$ 4GalTs (voir chapitre 2). Elle intervient à un stade précoce de l'initiation de la synthèse des GAGs, en catalysant le transfert du premier Gal de l'amorce tétrasaccharidique sur le Xyl lié à la protéine "core" des PGs. Chez l'homme, des mutations ponctuelles (A186D, L206P et R270C) de la  $\beta$ 4GalT7 sont liées à une maladie d'origine génétique, la forme progéroïde du syndrome d'Ehlers-Danlos (Almeida *et al.*, 1999; Faiyaz *et al.*, 2004). Les patients atteints de ce syndrome présentent un tableau clinique sévère affectant en particulier le système musculo-squelettique.

Les conséquences des mutations de la  $\beta$ 4GalT7 dans le contexte de ce syndrome sur la synthèse des chaînes de GAGs sont partiellement caractérisées. Au niveau des tissus conjonctifs, une substitution aberrante des chaînes de CS/DS de la décorine est observée (voir première partie de l'introduction bibliographique). L'équipe de Götte a également suggéré l'implication d'une altération des chaînes d'HS dans le retard du processus de cicatrisation observé chez les patients atteints du syndrome progéroïde d'Ehlers-Danlos (Götte *et al.*, 2008).

Par ailleurs, Esko *et al.* ont démontré en 1985 que les cellules CHO-pgsB618 (*Chinese Hamster Ovary*), déficientes en  $\beta$ 4GalT7, montrent une absence totale de chaînes de GAGs. Enfin, la  $\beta$ 4GalT7 présente la particularité de pouvoir initier la synthèse des chaînes de GAGs *in vivo* et *in vitro*, à partir de substrats xylosidiques exogènes (Esko *et al.*, 1987). De façon intéressante, ces molécules peuvent avoir des propriétés thérapeutiques importantes en stimulant ou inhibant la synthèse des chaînes de GAGs. Ainsi, il a été suggéré que des inhibiteurs de la biosynthèse des GAGs auraient des propriétés anti-cancéreuses (Mani *et al.*, 1998; Belting *et al.*, 2002; Schuksz *et al.*, 2008). Des xylosides initiant la formation des chaînes de GAGs constituent des médicaments anti-thrombotiques potentiels (Martin *et al.*, 1996) ou encore des agents thérapeutiques proposés dans le traitement de la maladie d'Alzheimer grâce à leurs propriétés anti-amyloïdes (Kisilevsky *et al.*, 2004).

L'ensemble de ces observations indique un rôle clé de la  $\beta$ 4GalT7 dans l'initiation des voies de biosynthèse des GAGs. De ce fait, cette protéine pourrait constituer une cible pharmacologique dans l'objectif de développer des stratégies thérapeutiques visant à pallier la perturbation de la synthèse des GAGs observée au cours de diverses pathologies.

### III. LA POLYMERISATION DES CHAINES DE GAGs

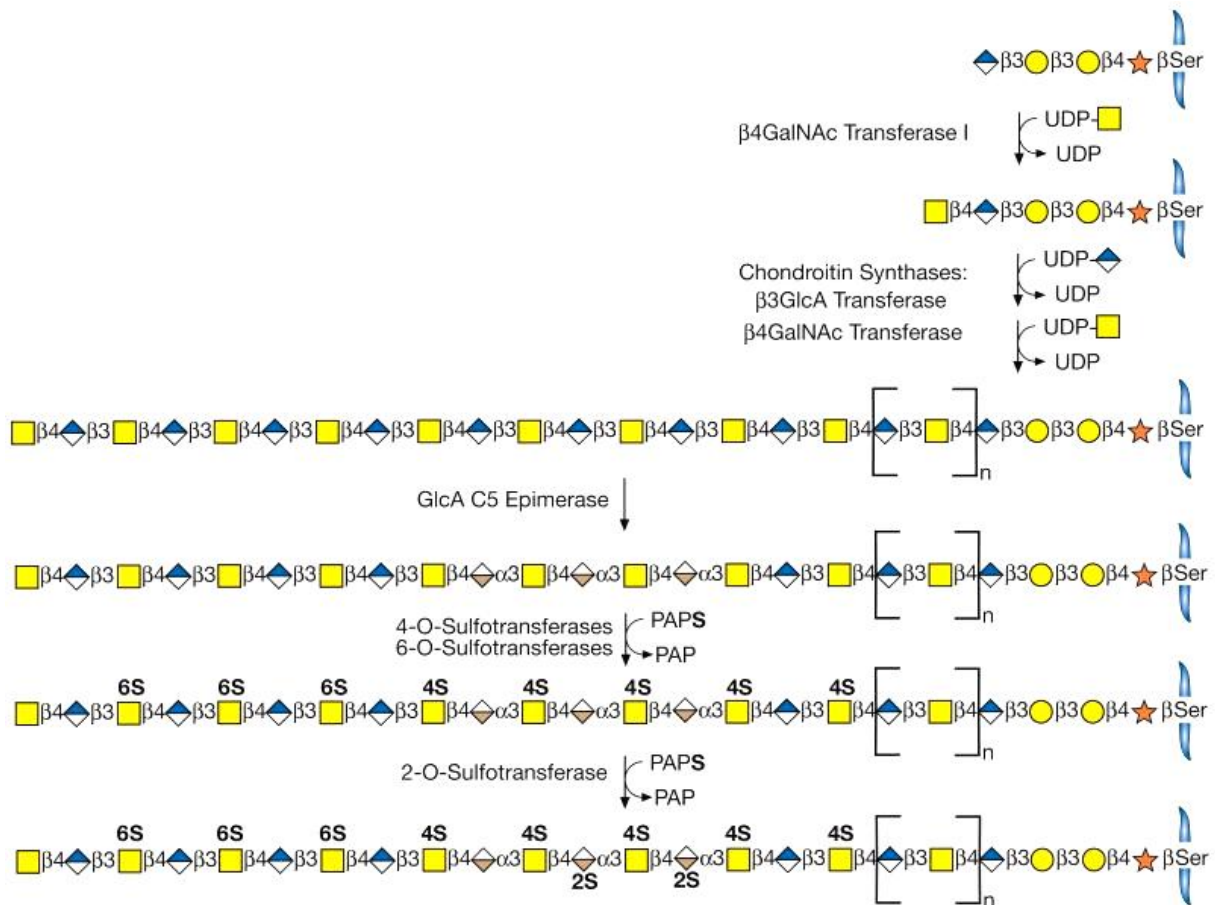
Après la formation de l'amorce tétrasaccharidique [GlcA $\beta$ 1,3Gal $\beta$ 1,3Gal $\beta$ 1,4Xyl $\beta$ 1], commence l'élongation des chaînes de GAGs qui oriente la machinerie cellulaire vers la synthèse des chaînes d'Hp/HS ou de CS/DS. La nature du cinquième résidu saccharidique transféré sur l'amorce tétrasaccharidique détermine le type de chaînes de GAGs synthétisées. Il s'agit d'un résidu GlcNAc pour les Hp/HS et d'un résidu GalNAc pour les CS/DS. D'une façon générale, l'extension des chaînes se fait par l'addition alternative de deux monosaccharides: une hexosamine et un GlcA. Ces réactions sont catalysées respectivement par des HS-synthases (EXTL1/2/3 et EXT1/2) et des CS-synthases.

Dans le cas des KS, les GAGs sont formées comme des oligosaccharides attachés par une liaison glycosidique de type N ou O à une protéine "core", et l'élongation se fait par l'addition de résidus GlcNAc et Gal.

#### III.1. La polymérisation des chaînes de CS/DS

Après l'ajout d'un résidu GalNAc sur l'amorce tétrasaccharidique par la CS-GalNAcT-I (ChGn), la polymérisation des chaînes de CS/DS a lieu par addition alternative d'un résidu GlcA et d'un résidu GalNAc. Chez les vertébrés, la  $\beta$ 3GlcAT et la  $\beta$ 4GalNAcT sont considérées comme les deux enzymes responsables de la polymérisation de l'unité disaccharidique élémentaire des chaînes de CS/DS (Richmond *et al.*, 1973) (Figure 29). La GlcAT qui intervient dans la polymérisation des chaînes de CS/DS est appelée  $\beta$ 3GlcAT-II, elle est distincte de l'enzyme qui catalyse le transfert du GlcA sur l'amorce trisaccharidique (Sugumaran *et al.*, 1998).

Le GlcA des chaînes de CS subit une épimérisation du carbone en position 5 (C5) pour former un IdoA. Cette modification permet de distinguer les chaînes de DS des chaînes de CS, qui sont composées de deux types d'acide uronique (GlcA et IdoA), comme indiqué dans la Figure 29.



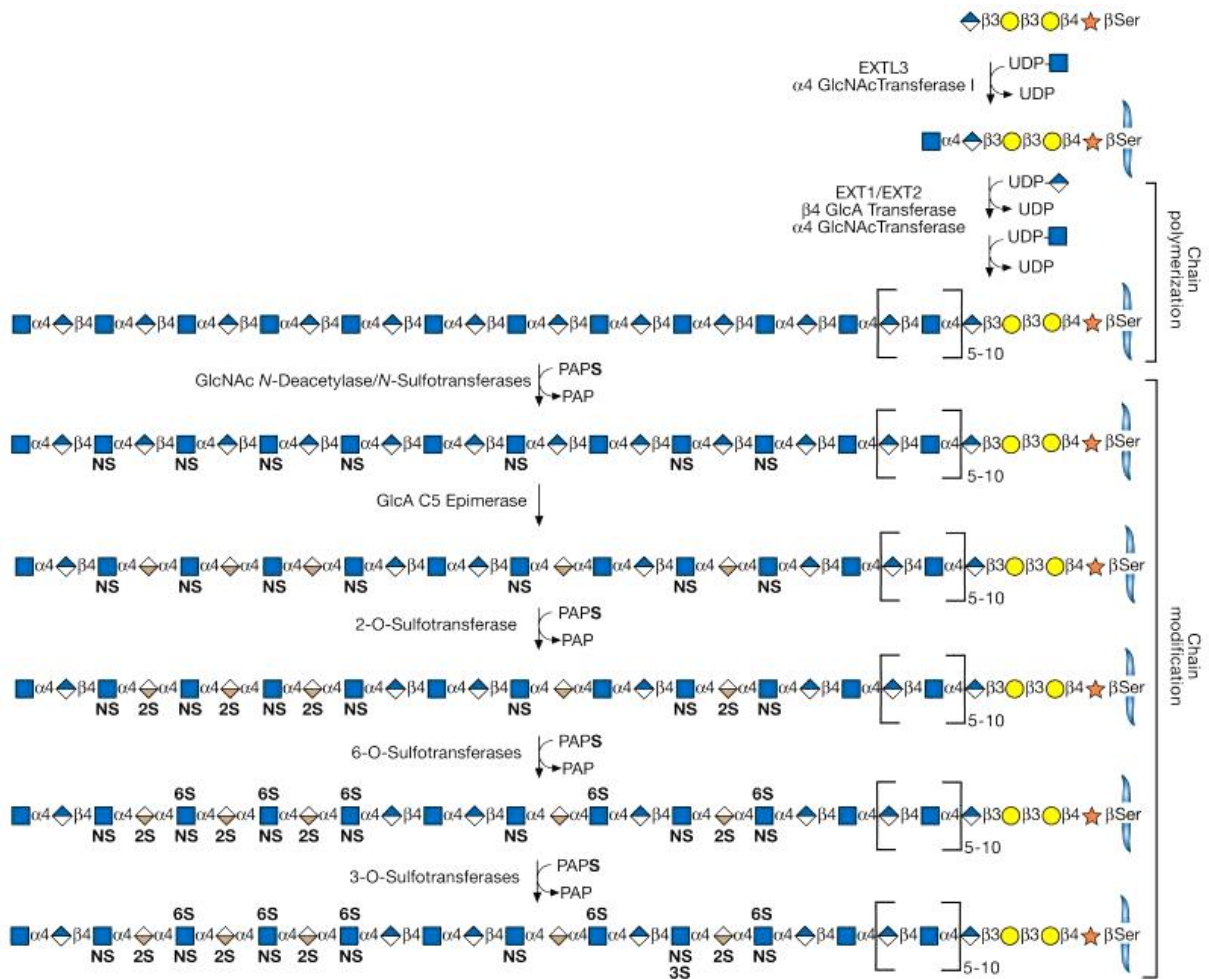
**Figure 29. Représentation schématique des voies d'élongation des chaînes de CS/DS et des enzymes impliquées dans ces réactions**

Le GlcA subit une épimérisation en position C5 pour former un IdoA caractéristique des chaînes de DS. De plus, ces dernières peuvent être sulfatées à différentes positions (C2, 4 ou 6). PAPS, 3'-phosphoadénosyl-5'-phosphosulfate.

D'après Esko *et al.*, 2009, *Essentials of Glycobiology*

### III.2. La polymérisation des chaînes d'Hp/HS

La synthèse des chaînes d'HS est associée aux produits de la famille des gènes *EXT*. Le transfert du résidu GlcNAc sur l'amorce tétrasaccharidique est catalysé par l'enzyme EXTL3 ( $\alpha$ 4GlcNAcT-I). Cette réaction est suivie de l'addition successive d'un résidu GlcA et d'un résidu GlcNAc, catalysée par le complexe EXT1/EXT2 qui possèdent des activités  $\beta$ 4GlcAT- et  $\alpha$ 4GlcNAcT. Les chaînes d'Hp/HS subissent une série de modifications importantes, notamment, des sulfatations en position 3 ou 6 de la GlcNAc, ou une épimérisation du GlcA qui peut à son tour être sulfaté en position 2 (Figure 30).



**Figure 30. Représentation schématique de la polymérisation des chaînes d'Hp/HS et enzymes impliquées dans ces réactions**

Cette figure montre les séries de modifications que les chaînes d'Hp/HS peuvent subir. NS, N-sulfatation; PAPS, 3'phosphoadénosyl-5'-phosphosulfate.

D'après Esko *et al.*, 2009, Essentials of Glycobiology

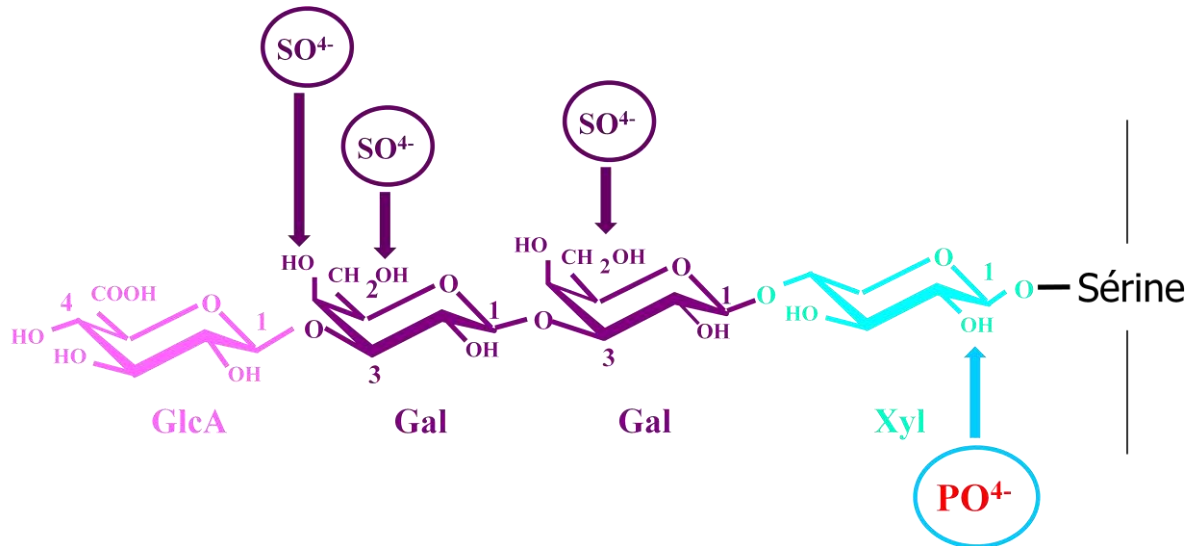
#### IV. LES REACTIONS DE MODIFICATION DES CHAINES DE GAGs DES PGs

##### IV.1. Les réactions de modification de l'amorce tétrasaccharidique

La région de liaison entre la protéine "core" et les chaînes de GAGs (l'amorce tétrasaccharidique) peut subir diverses modifications, notamment, une phosphorylation en C2 du Xyl et une sulfatation des Gal à différentes positions (C4 et/ou C6) comme représenté dans la Figure 31. De façon intéressante, et contrairement au "dogme" selon lequel la structure ou la composition de l'amorce tétrasaccharidique est commune aux CS/DS et Hp/HS, il a été



démontré que dans certains PGs isolés de l'aorte bovine, le GlcA serait épimérisé en C5 pour former l'IdoA dans les chaînes de DS (Sugahara *et al.*, 1995a and 1995b).



**Figure 31. Modification de la région de liaison des chaînes de GAGs**

Durant la synthèse des PGs, l'amorce tétrasaccharidique peut être modifiée à différentes positions. Ainsi, le Xyl peut être phosphorylé en position 2, tandis que chacun des deux Gal peut être sulfaté en position 4 et/ou 6.

#### IV.1.1. La phosphorylation en C2 du xylose de l'amorce tétrasaccharidique

Le C2 du Xyl constitue un site de phosphorylation majeur, à la fois pour les CSPGs (Oegma *et al.*, 1984) et pour les HSPGs (Fransson *et al.*, 1985). Des études ont montré que la phosphorylation est maximale après le transfert des deux résidus Gal de l'amorce tétrasaccharidique (Moses *et al.*, 1999), et que l'addition du GlcA serait suivie d'une déphosphorylation rapide. En outre, les PGs des tissus ayant une MEC abondante, ne portent pas de groupement phosphate sur le Xyl de l'amorce tétrasaccharidique (Sugahara *et al.*, 1995a).

Le rôle biologique de cette phosphorylation n'est pas complètement compris. Il a été proposé que cette modification pourrait constituer un signal pour le transport et la sécrétion des PGs, ou encore pour les modifications ultérieures des chaînes de GAGs (Moese *et al.*, 1997; 1999). Dans les PGs produits par les cellules de mélanome, la déphosphorylation du Xyl est

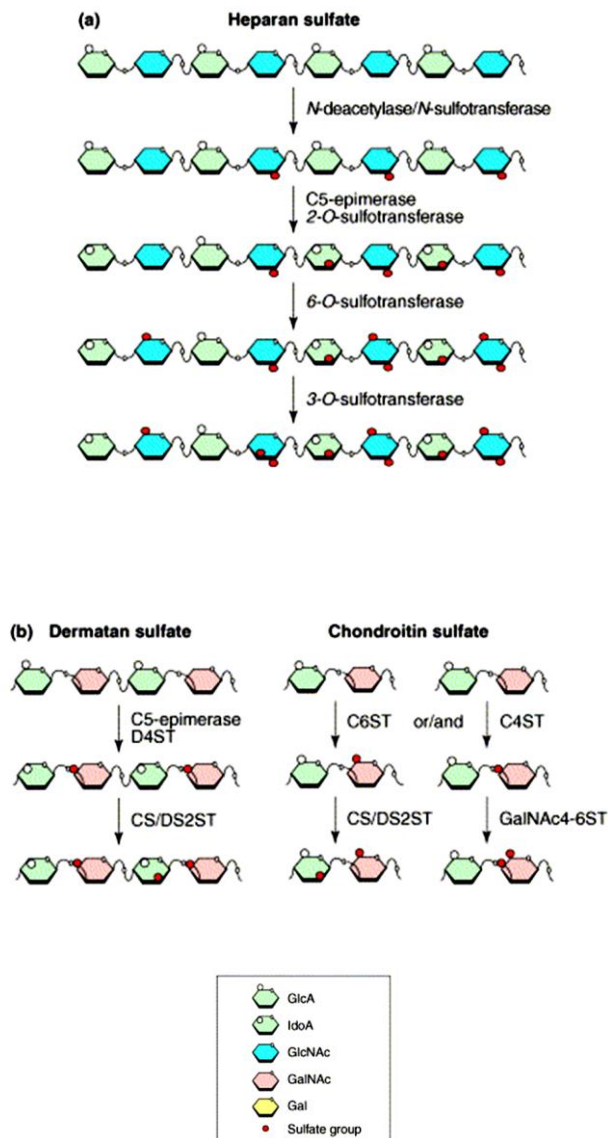
complète (Spiro *et al.*, 1991) tandis que dans les cellules de chondrosarcome, l'agrécan produit porte un groupement phosphate sur le Xyl de l'amorce tétrasaccharidique (Oegma *et al.*, 1984). Il a été suggéré que la déphosphorylation peut être inefficace dans certaines lignées cancéreuses, et que ce phénomène serait associé à l'altération de la biosynthèse des PGs dans ces cellules (Prydz *et al.*, 2000). De plus, des travaux réalisés dans mon équipe ont démontré que l'activité enzymatique de la  $\beta$ 4GalT7 est complètement abolie *in vitro* lorsque le Xyl est phosphorylé en C2 (Gulberti *et al.*, 2005). L'ensemble de ces résultats suggère que cette modification de l'amorce tétrasaccharidique joue un rôle important dans la régulation de la synthèse des GAGs, selon un mécanisme qui reste à déterminer.

#### **IV.1.2. La sulfatation des galactoses de l'amorce tétrasaccharidique**

De façon intéressante, la sulfatation des Gal de l'amorce tétrasaccharidique des PGs a été observée au niveau des chaînes de CS/DS, alors qu'elle est absente dans les Hp/HS. Cette observation suggère que les groupements sulfates de ces résidus pourraient constituer un signal d'orientation des voies de biosynthèse vers l'un ou l'autre des types de GAGs. Il a été également proposé que cette modification pourrait être impliquée dans le recyclage des PGs. De plus, il a été démontré que l'activité enzymatique de la GlcNAcT-I, qui joue un rôle important dans l'initiation de la synthèse des HS, est inhibée par la 4-O-sulfatation du Gal (Kitagawa *et al.*, 1995; Lidholt *et al.*, 1997).

Les CS/DS peuvent être sulfatés en C4 et/ou C6 sur les deux résidus Gal (Sugahara *et al.*, 1988). Une étude réalisée par mon équipe a mis en évidence l'importance de cette modification dans l'initiation de la biosynthèse des GAGs. En effet, la O-sulfatation du C4 ou C6 du premier et/ou du deuxième Gal régule l'efficacité enzymatique de la GlcAT-I (Gulberti *et al.*, 2005). En particulier, cette enzyme présente une efficacité 7 fois supérieure vis-à-vis d'un oligosaccharide accepteur sulfaté en position 6 sur le premier résidu Gal comparé à son analogue non sulfaté. Au contraire, toutes les autres positions de sulfatation abolissent l'activité de la GlcAT-I.

De façon intéressante, la phosphorylation du Xyl n'a jamais été identifiée en présence de groupements sulfates sur les résidus Gal de l'amorce tétrasaccharidique des chaînes glycosaminoglycaniques de CS/DS (Prydz *et al.*, 2000).



**Figure 32. Réactions de modifications des chaînes d’HS et de CS/DS**

**a)** Les réactions de modifications des HS sont initiées par la N-acétylation/N-sulfatation de la GlcNAc suivie d’une sulfatation rapide au niveau des groupements amines libres pour former des GlcNS. **b)** Dans le cas de la modification des chaînes de DS, les disaccharides subissent une 4-O-sulfatation directement après épimerisation, pour laisser place à l’action de la DS2ST qui catalyse la sulfatation de l’IdoA en C2. Enfin, dans les chaînes de CS, la réaction de 6-O-sulfatation des GalNAc et GalNAc(4S) est catalysée par deux enzymes distinctes.

D’après Kusche-Gulberg et Kjellen, 2003

## IV.2. Les modifications des différentes chaînes de GAGs

Au cours du processus de polymérisation des chaînes de GAGs, les unités glycosidiques sont modifiées à différentes positions. Ainsi, on distingue : 1) les réactions de N-déacétylation/N-

sulfatation de la GlcNAc de l'unité disaccharidique de l'Hp et des HS; 2) l'épimérisation du GlcA en IdoA dans l'Hp et les HS (et DS); 3) les O-sulfatations des disaccharides répétitifs de l'Hp, des HS, CS et DS à des positions variables (C2, 3, 4 ou 6).

La réaction de N-sulfatation est une modification spécifique des chaînes d'HS et elle n'a jamais été observée dans les CS/DS ou KS. Une famille d'enzymes bifonctionnelles appelée NDSTs (GlcNAc-N-déacétylase/sulfotransférase) catalysent la N-déacétylation du GlcNAc pour générer une GlcN, sur laquelle elles transfèrent un groupement N-sulfate (Wei *et al.*, 1993) (Figure 32).

Au cours de la polymérisation des chaînes de GAGs, les HS subissent une série de réactions de modifications (Figures 30 et 32). A la suite de l'action des NDSTs, les STs et les épimérasés peuvent intervenir pour assurer la modification des chaînes de GAGs et augmenter la diversité structurale de ces molécules.

#### **IV.2.1. La sulfatation des Hp/HS**

##### **IV.2.1.1. La N-déacétylation/N-sulfatation des HS**

Chez les mammifères, la famille des NDSTs est composée de quatre membres, NDST1, NDST2, NDST3 et NDST4, de spécificité de substrat différente (Aikawa *et al.*, 2001). Ces enzymes se distinguent également par leur profil d'expression, ainsi, les NDST1 et 2 sont exprimées ubiquitairement au cours du développement embryonnaire ainsi que chez l'adulte, tandis que la NDST3 et 4 sont exprimées uniquement au stade embryonnaire (Grobe *et al.*, 2002).

La modification de la structure des chaînes d'HS confère à ces GAGs une organisation particulière. En effet, contrairement aux chaînes de CS, la N-sulfatation des HS engendre la formation successive de domaines sulfatés et de domaines dépourvus de groupements sulfates. Cette organisation permet de différencier des domaines N-acétylés (NA), N-sulfatés (NS), et des domaines mixtes (NA/NS) (Figure 30). Ces régions sont le plus souvent présentes dans cet ordre et elles constitueraient des sites de fixation de divers ligands, notamment les FGFs (*Fibroblast Growth Factors*) et l'anti-thrombine.

##### **IV.2.1.2. La O-Sulfatation des Hp/HS**

La O-sulfatation intervient après la N-sulfatation et l'épimérisation en C5 du GlcA en IdoA. La 2-O-sulfatation se produit préférentiellement en C2 de l'IdoA, mais également en C2 du

GlcA (Rong *et al.*, 2001; Merry and Wilson, 2002). Les acides hexuroniques 2-O-sulfatés se trouvent, à quelques exceptions près, exclusivement à côté des domaines NS du polysaccharide. L'enzyme qui catalyse cette réaction est la 2-OST (ou HS2ST), il a été démontré que cette protéine forme un complexe avec la C5 épimérase (Esko and Selleck, 2002).

Comme il est mentionné précédemment, les Hp/HS peuvent subir également des modifications par 6-O-sulfatation du résidu GlcNAc. A ce jour, trois enzymes GlcNAc-6-OSTs ont été identifiées (Smeds *et al.*, 2003): 6-OST1, 6-OST2 et 6-OST3 (HS6ST-1, HS6ST-2, HS6ST-3). Il est intéressant de noter que l'expression de ces protéines est tissu-spécifique. Chez la souris adulte, la 6-OST1 est préférentiellement exprimée dans le foie, tandis que la 6-OST2 est présente dans le cerveau et la 6-OST3 possède une expression ubiquitaire. Chez l'homme, la 6-OST2 existe sous deux isoformes (voir Tableau 6).

Enfin, au cours de la biosynthèse des Hp/HS les résidus GlcNAc peuvent être modifiés par 3-O-sulfatation. C'est une réaction rare, spécifique de ce type de chaînes. Elle est catalysée par une famille de 3-OSTs composée de sept membres: 3-OST1, 3-OST2, 3-OST3A, 3-OST3B, 3-OST4, 3-OST5 et 3-OST6 (Showrak *et al.*, 1999; Xia *et al.*, 2002; Mochizuki *et al.*, 2003). De nombreuses études ont suggéré que la 3-O-sulfatation régule sélectivement la fonction biologique de l'Hp et des HS. En particulier, la présence de la 3-O-sulfatation sur les HS est indispensable pour la reconnaissance de l'anti-thrombine (Petitou *et al.*, 2003).

Le virus de l'herpès de type 1 (HSV-1: *Herpes Simplex virus*) est un des nombreux agents pathogènes qui utilisent les HS des PGs de la surface cellulaire comme récepteur. Ainsi, les chaînes d'HS interagissent avec les glycoprotéines gB et gC de l'enveloppe virale pendant la fixation du virus sur la cellule. Cela constitue la première étape du processus de l'entrée du HSV-1 dans la cellule hôte. De façon importante, il a été démontré que c'est la 3-O-sulfatation des HS qui permet l'interaction entre ces chaînes de GAGs cellulaires et la glycoprotéine virale gD, induisant ainsi la fusion entre l'enveloppe virale du HSV-1 et la membrane de la cellule infectée (O'Donnell *et al.*, 2005; Shukla *et al.*, 1999).

#### **IV.2.2. La sulfatation des CS/DS**

La modification des chaînes de CS/DS par des réactions de sulfatation est catalysée par trois familles de STs. Les acides hexuroniques sont 2-O-sulfatés par l'enzyme CS/DS2ST, la seule ST impliquée dans la 2-O-sulfatation de ces chaînes de GAGs identifiée à ce jour (Kobayashi

*et al.*, 1999). Elle catalyse le transfert d'un groupement sulfate sur le C2 du GlcA et de l'IdoA.

Quant aux résidus GalNAc, ils peuvent être O-sulfatés en C4 et C6. Ces réactions sont catalysées respectivement par la famille des 4-O-sulfotransférases: C4ST/D4ST (Yamauchi *et al.*, 1999) et la famille des 6-O-sulfotransférases C6STs (Uchimura *et al.*, 1998) (voir Tableau 6).

## V. LA DEGRADATION ET LE RECYCLAGE OU "*TURNOVER*" DES PGs

De nombreuses études ont montré la présence d'enzymes impliquées dans la dégradation des chaînes de GAGs, permettant ainsi le recyclage de ces molécules par l'organisme. En premier lieu, les PGs sont internalisés par les cellules. Au cours de l'endocytose, ils subissent un clivage protéolytique qui permet la libération de la protéine "core", tandis que les chaînes de GAGs sont partiellement clivées par des enzymes telles que l'endo- $\beta$ -glucuronidase ou l'endohexosaminidase. Ces enzymes reconnaissent des séquences consensus spécifiques, qui dépendent de la composition des chaînes de GAGs. Elles produisent des fragments glycosidiques de taille variable qui subissent une dégradation complète par l'action d'une série d'exoglycosidases et de sulfatases.

Une famille d'enzymes appelée endosulfatases a été récemment découverte. Chez l'homme, deux ADNc ont été identifiés: SULF1 et SULF2 (Shridhar *et al.*, 2002). Ces deux gènes codent pour des protéines membranaires capables de désulfater les chaînes de GAGs. La découverte de cette réaction ainsi que de l'action des héparanases extracellulaires a suggéré la présence d'autres processus enzymatiques capables de modifier la structure des GAGs et de les dégrader, notamment celle des HS après leur sécrétion.

De façon intéressante, l'altération de l'expression de SULF1 a été associée à de multiples cancers, en effet, la SULF1 humaine est considérée comme un suppresseur de tumeurs dans la carcinogénèse (Lai *et al.*, 2008). Cette enzyme est responsable de la désulfatation des HSPGs cellulaires. Elle est impliquée dans la croissance cellulaire, la prolifération tumorale, la migration, l'invasion et l'angiogénèse. En revanche, la SULF2 a été décrite comme une enzyme qui présenterait des propriétés pro-oncogéniques (Lai *et al.*, 2010).

Enzyme	Abréviation	Substrat prédominant	Expression chez l'homme
<u>Chondroïtine/dermatane sulfates</u>			
<b>Chondroïtine 4-O-sulfotransférase-1</b>	C4ST1	-GlcA-GalNAc-	Tissus hématopoïétiques
<b>Chondroïtine 4-O-sulfotransférase-2</b>	C4ST2	-GlcA-GalNAc-	Ubiquitaire
<b>Chondroïtine 4-O-sulfotransférase-3</b>	C4ST3	-GlcA-GalNAc-	Foie
<b>Dermatane 4-O-sulfotransférase-1</b>	D4ST1	-GlcA-GalNAc-	Placenta
<b>Chondroïtine 6-O-sulfotransférase</b>	CS6ST	-GlcA-GalNAc-	Aorte, testicules
<b>Chondroïtine 6-O-sulfotransférase-2</b>	CS6ST2	-GlcA-GalNAc-	Cœur, ovaires
<b>Chondroïtine 4-sulfate 6-O-sulfotransférase Galactosaminyl uronyl</b>	GalNAc4S-6ST	-GlcA-GalNAc4S-	Ubiquitaire
<b>2-O-sulfotransférase</b>	CS/DS2ST	-GlcA-GalNAc6S- -IdoA-GalNAc4S-	
<u>Héparine/héparane sulfates</u>			
<b>N-déacétylase/N-sulfotransférase 1</b>	NDST1	-GlcA-GlcNAc-	Ubiquitaire
<b>N-déacétylase/N-sulfotransférase 2</b>	NDST2	-GlcA-GlcNAc-	Ubiquitaire
<b>N-déacétylase/N-sulfotransférase 3</b>	NDST3	-GlcA-GlcNAc-	Ubiquitaire
<b>N-déacétylase/N-sulfotransférase 4</b>	NDST4	-GlcA-GlcNAc-	Ubiquitaire
<b>Héparane sulfate 2-O-sulfotransférase</b>	2-OST	-HexA-GlcNS-	Ubiquitaire
<b>Héparane sulfate 6-O-sulfotransférase 1</b>	6-OST1	-HexA±2-GlcNS/Ac-	Ubiquitaire
<b>Héparane sulfate 6-O-sulfotransférase 2</b>	6-OST2	-HexA±2S-GlcNS/Ac-	Cerveau
<b>Héparane sulfate 6-O-sulfotransférase 2S</b>	6-OST2S	-HexA±2S-GlcNS/Ac-	Ovaires, placenta, reins du fœtus
<b>Héparane sulfate 6-O-sulfotransférase 3</b>	6-OST3	-HexA±2S-GlcNS/Ac-	-
<b>Héparane sulfate 3-O-sulfotransférase 1</b>	3-OST1	-GlcA-GlcNS±6S-	Cerveau, reins
<b>Héparane sulfate 3-O-sulfotransférase 2</b>	3-OST2	-HexA2S-GlcNS-	Cerveau
<b>Héparane sulfate 3-O-sulfotransférase 3A</b>	3-OST3A	-IdoA2S-GlcNH2±6S-	Cœur, placenta
<b>Héparane sulfate 3-O-sulfotransférase 3B</b>	3-OST3B	-IdoA2S-GlcNH2±6S-	Foie, placenta
<b>Héparane sulfate 3-O-sulfotransférase 4</b>	3-OST4	?	Cerveau
<b>Héparane sulfate 3-O-sulfotransférase 5</b>	3-OST5	-GlcA-GlcNS±6S-, -IdoA2S-GlcNS±6S-	Cerveau, corde spinale

**Tableau 6:** Les différentes sulfotransférases impliquées dans les réactions de modifications de biosynthèse des GAGs

Ce tableau résume les principaux substrats reconnus par les STs et le profil d'expression de ces enzymes.

*Présentation  
de l'étude*



## Présentation de l'étude

Les récents progrès de la glycobiologie ont révélé l'importance majeure des carbohydrates, et en particulier celle des PGs. Ces macromolécules jouent un rôle crucial dans de nombreux processus biologiques (physiologiques et pathologiques), notamment *via* leurs chaînes de GAGs, impliquées dans de multiples interactions cellulaires associées au développement embryonnaire, à la coagulation, l'inflammation, la cancérogenèse, ou encore aux infections virales. En raison de leur implication dans de nombreuses maladies, ces molécules constituent des cibles pharmacologiques potentielles. Elles peuvent également être utilisées comme agents thérapeutiques et représenter des marqueurs diagnostiques.

La grande complexité du système de biosynthèse, de modification et de dégradation des GAGs est à l'origine de la variété structurale considérable de ces polysaccharides. Bien que la majorité des enzymes impliquées dans la biosynthèse des GAGs soit identifiée, la caractérisation structurale et fonctionnelle de chacune d'entre-elles est très incomplète. Une meilleure compréhension des voies de biosynthèse des GAGs ouvrirait des perspectives thérapeutiques au regard de nombreuses pathologies. L'énorme potentiel de ces molécules a été longtemps négligé. Néanmoins, au cours des dernières décennies, elles attirent de plus en plus l'attention de la communauté scientifique et médicale. Ces travaux ont déjà permis le développement et l'utilisation de médicaments dans le traitement de plusieurs affections associées à des désordres de la biosynthèse des PGs. L'injection intra-articulaire d'AH est utilisée dans le traitement de l'arthrose (Migliore and Granata, 2008). Des approches thérapeutiques ont conduit à l'utilisation des GAGs afin de cibler spécifiquement les cellules tumorales, tout en véhiculant des molécules anti-cancéreuses (Yip *et al.*, 2006). L'Hp est un puissant anti-coagulant, utilisé depuis plusieurs décennies dans la prévention et le traitement de la thrombose veineuse. Dans le cas du cancer, l'activité de l'Hp atténue la progression tumorale, notamment en réduisant la production de thrombine et la formation de fibrine (Prandoni *et al.*, 2005). De façon intéressante, plusieurs essais cliniques suggèrent que l'association de la chimiothérapie conventionnelle au traitement par l'Hp serait favorable en termes de survie des patients (Altinbas *et al.*, 2004; Klerk *et al.*, 2005).

Par conséquent, l'élucidation des mécanismes d'action des enzymes impliquées dans la biosynthèse des PGs et leurs chaînes de GAGs constitue un véritable enjeu médical. Leurs voies de biosynthèse nécessitent la participation de plusieurs familles d'enzymes, notamment des GTs, des STs, des épimérasés, des sulfatases et des phosphorylases assurant la diversité et

l'hétérogénéité des GAGs. La perturbation de l'activité de ces enzymes, en particulier des GTs, est associée à plusieurs pathologies. Le syndrome progéroïde d'Ehlers-Danlos est une maladie génétique résultant de mutations de la  $\beta$ 4GalT7, enzyme qui intervient à un stade précoce de l'initiation de la synthèse des chaînes de GAGs des PGs (Quentin *et al.*, 1990). Trois mutations ont été identifiées chez les patients atteints de ce syndrome héréditaire. Il s'agit des mutations A186D, L206P et R270C, qui diminuent ou abolissent l'activité de la  $\beta$ 4GalT7 affectant la substitution des chaînes de GAGs des PGs, en particulier la décorine. Ces altérations se traduisent par des anomalies fonctionnelles, notamment au niveau des tissus conjonctifs. La forme progéroïde du syndrome d'Ehlers-Danlos est en effet caractérisée par un tableau clinique sévère. Les patients atteints présentent un aspect progéroïde, un retard de développement, une disproportion cranio-faciale, une ostéopénie généralisée, des défauts de cicatrisation, des articulations hypermobiles, des muscles hypotoniques, une peau fragile et hyperextensible, ainsi que des anomalies du système musculo-squelettique.

A la lumière des récentes découvertes démontrant l'importance des chaînes de GAGs et leur potentiel thérapeutique, l'enjeu de ma thèse est de mieux comprendre les mécanismes d'action des enzymes intervenant dans leur biosynthèse. Mes travaux portent plus spécifiquement sur la caractérisation structurale et fonctionnelle de la  $\beta$ 4GalT7 humaine. Dans ce contexte, j'ai développé deux systèmes d'expression hétérologue de la  $\beta$ 4GalT7 humaine recombinante, permettant la production de la protéine sauvage et mutée dans les cellules eucaryotes et dans la bactérie. Ainsi, j'ai réalisé une étude comparative des effets *in vitro/ex vivo* de la surexpression de la  $\beta$ 4GalT7 sur l'activité GalT et sur la capacité de l'enzyme à stimuler l'initiation de la synthèse des GAGs dans des cellules déficientes. Nous avons également étudié les conséquences de la mutation des acides aminés situés dans des motifs hautement conservés au sein des  $\beta$ 4GalTs, potentiellement impliqués dans la fixation des substrats et/ou l'organisation structurale de la  $\beta$ 4GalT7 humaine, sur les constantes cinétiques et sur l'activité enzymatique évaluée *in vitro* et dans les cellules CHO-pgsB618 en culture.

Une première démarche a consisté à mener une étude phylogénétique afin d'identifier les membres de la famille des  $\beta$ 4GalTs et la sous-famille des  $\beta$ 4GalT7s de différentes espèces. Nous avons ensuite réalisé un alignement de séquences de cette famille d'enzymes qui a permis de déterminer des domaines peptidiques conservés au sein des 7  $\beta$ 4GalTs humaines et de 48  $\beta$ 4GalT7s de différentes espèces (cette partie du projet a été réalisée en collaboration avec le Dr. Rafael Oriol, INSERM U1004, Villejuif). Parallèlement, nous avons développé avec le Dr. Guillermo Mulliert (UMR 7036 CNRS-UHP Nancy 1) un modèle moléculaire de

la  $\beta$ 4GalT7 humaine basé sur la structure cristallographique de la  $\beta$ 4GalT1 bovine qui était la seule  $\beta$ 4GalT dont la structure 3D était élucidée. Récemment, Ramakrishnan et Qasba (2010) ont élucidé la structure 3D de la  $\beta$ 4GalT7 de drosophile, ce qui nous a permis de développer un nouveau modèle moléculaire de la  $\beta$ 4GalT7 humaine basé sur cette structure. La superposition des deux modèles a confirmé la pertinence de l'organisation structurale du site actif de l'enzyme humaine fondée sur le premier modèle moléculaire. L'ensemble de ces stratégies a permis d'identifier la présence de motifs peptidiques conservés potentiellement impliqués dans la reconnaissance de substrat, la catalyse enzymatique et/ou l'organisation structurale du site actif de la  $\beta$ 4GalT7 humaine. Il s'agit des domaines  $^{163}\text{DVD}^{165}$ ,  $^{221}\text{FWG}\text{WGREDDE}^{230}$  et  $^{257}\text{HLH}^{259}$ . Les données bibliographiques et les modèles moléculaires développés proposent l'implication du motif  $^{163}\text{DVD}^{165}$  dans la reconnaissance du substrat donneur et la formation de liaisons de coordination avec les cations divalents ( $\text{Mn}^{2+}$ ). Le deuxième motif contient les résidus carboxyliques Glu227 et Asp228, potentiellement impliqués dans la catalyse enzymatique. Il contient également des acides aminés (résidus Gly223 et Gly225) qui auraient une fonction structurale et seraient indispensables aux changements conformationnels de la poche catalytique, permettant ainsi les interactions entre le substrat donneur et le substrat accepteur. Quant au dernier motif, il serait essentiellement impliqué dans des interactions avec les ions  $\text{Mn}^{2+}$ . Pour vérifier ces postulats, nous avons procédé à la mutagenèse dirigée systématique par PCR de ces acides aminés, afin d'analyser leur importance fonctionnelle. Dans l'objectif de déterminer le rôle de chacun de ces acides aminés, nous avons analysé et comparé les conséquences de chaque mutation sur les propriétés enzymatiques de la  $\beta$ 4GalT7 *in vitro*. Pour ce faire, j'ai notamment développé un système d'expression bactérien permettant de produire et de purifier cette enzyme en quantité compatible avec des études cinétiques approfondies.

Nous avons également étudié l'effet de la surexpression de l'enzyme sauvage ou mutée dans les cellules en culture, sur l'initiation de la biosynthèse des GAGs à partir d'un substrat xylosidique exogène ou en surexprimant la protéine "core" de la décorine servant d'ancrage aux chaînes de GAGs. L'ensemble de ce travail a permis l'identification d'acides aminés essentiels à la fonction de la  $\beta$ 4GalT7 humaine. J'ai en particulier démontré le rôle clé du résidu Trp224 qui semble être impliqué à la fois dans l'interaction avec les substrats donneur et accepteur, et dont la mutation en His module la spécificité vis-à-vis de ce dernier (**Publication n°1, Talhaoui *et al.*, 2010**).

Dans le cadre de l'étude des bases moléculaires de la forme progéroïde du syndrome d'Ehlers-Danlos, nous avons surexprimé la  $\beta$ 4GalT7 recombinante humaine portant chacune des mutations associées à cette pathologie dans les cellules eucaryotes en culture et dans la bactérie en fusion avec la GST, en vue de déterminer les conséquences de ce désordre génétique sur les propriétés catalytiques de cette enzyme. Enfin, j'ai analysé l'effet de l'apport de substrat xylosidique sur la capacité de l'enzyme  $\beta$ 4GalT7 sauvage et mutée, à restaurer l'initiation de la biosynthèse des GAGs (**Publication n°2, Talhaoui *et al.*, 2010**).

Ces deux parties de mon projet de thèse ont pour objectif d'établir des bases moléculaires de la synthèse des chaînes de GAGs des PGs vers le développement de molécules thérapeutiques destinées au traitement des pathologies associées à la perturbation de ce processus. A l'issue de ce travail, une étude structurale de la  $\beta$ 4GalT7 humaine purifiée, exprimée dans la bactérie, en fusion avec la GST, par cristallographie a été engagée. Cependant, cette étude est encore à un stade préliminaire, et un travail important reste à fournir afin d'obtenir les cristaux de cette protéine et d'élucider sa structure 3D.

De façon complémentaire, mon équipe s'intéresse aux mécanismes régissant l'assemblage des GAGs *in vivo* et à leurs anomalies au cours des pathologies articulaires. Dans ce cadre, j'ai participé à une étude originale fondée sur l'exploration de l'implication d'une modification épigénétique dans la régulation de la synthèse des PGs.

Cette étude est menée sur un modèle cellulaire humain, la lignée de chondrosarcome H-EMC-SS (HEMC). Dans un premier temps, nous avons analysé la séquence de la région promotrice proximale et du premier exon des gènes codant pour les enzymes impliquées dans la synthèse et la modification des chaînes de GAGs. Cette approche a permis de mettre en évidence la présence d'îlots CpG, susceptibles de subir des méthylations aberrantes de l'ADN. L'utilisation de la méthode de conversion au bisulfite de l'ADN génomique a ensuite permis de cartographier la méthylation des gènes des GTs et des STs dans les cellules HEMC.

De façon intéressante, après analyse du profil de méthylation des gènes codant pour les quatre GTs intervenant dans la formation de l'amorce tétrasaccharidique et les gènes impliqués dans l'élongation et la sulfatation des chaînes d'HS, nous avons démontré que seuls les gènes des STs, NDSTs et OSTs, et plus particulièrement la famille des 3-OSTs, présentent une hyperméthylation, et sont donc potentiellement soumis à un mécanisme de régulation épigénétique *via* la méthylation de l'ADN. Par contre, les gènes de la  $\beta$ 4GalT7 et des autres

GTs impliquées dans la formation des chaînes d'HS ne présentent pas de modification de leur profil de méthylation comparé à l'ADN de cellules non cancéreuses.

Ces résultats nous ont conduit à analyser par RT-PCR semi-quantitative, le taux d'ARNm des gènes hyperméthylés, identifiés comme potentiellement soumis à un mécanisme de régulation épigénétique, dans des cellules HEMC traitées par un inhibiteur de méthylation de l'ADN, la 5-aza-2'-déoxycytidine (5-Aza-dc). Nous avons notamment analysé l'expression des gènes de la famille 3-OSTs, afin de vérifier l'hypothèse selon laquelle l'hyperméthylation de leur promoteur serait responsable d'une répression de leur expression dans les cellules HEMC de chondrosarcome humain.

Cette investigation a été poursuivie par la caractérisation de l'effet de la surexpression, par transfection, des gènes hyperméthylés 3-OST1 et 3-OST3A dans les cellules HEMC sur le degré de sulfatation et le taux de biosynthèse des HSPGs cellulaires. Le résultat obtenu a été comparé à la caractérisation de l'impact de l'induction de l'expression de ces gènes par inhibition de la méthylation, dans les cellules traitées par la 5-Aza-dc. Enfin, nous avons finalisé cette étude par la détermination des conséquences de la surexpression de la 3-OST1 et la 3-OST3A dans les cellules HEMC, par transfert de gènes ou par traitement à la 5-Aza-dc sur l'adhésion, la prolifération et la migration cellulaires. Ce travail a fait l'objet d'une publication qui démontre l'implication de la répression de l'expression des 3-OSTs par un processus de régulation épigénétique dans le développement du phénotype invasif des cellules de chondrosarcome HEMC (**Publication n°3, Bui *et al.*, 2010**).

# *Matériels et méthodes*

# Matériels et méthodes

## *Matériels*

### 1. Lignées cellulaires et réactifs de culture cellulaire

Les cellules H-EMC-SS (HEMC) de chondrosarcome humain et les cellules HeLa proviennent de la banque « European Collection of Cell Culture » (ECACC, Sigma-Aldrich, Saint-Quentin Fallavier, France). Les cellules CHO-pgsB618 (cellules d'ovaire de hamster chinois, déficientes en  $\beta$ 4GalT7, CRL-2241) ont été obtenues auprès de la banque « American Type Culture Collection » (ATCC, LGC Standards, Molsheim, France).

Les milieux de culture DMEM, DMEM/F12 et Fischer sont fournis par GIBCO-Life Technologies (Invitrogen, Cergy-Pontoise, France).

### 2. Les réactifs de biologie moléculaire

Les enzymes de restriction proviennent de New England Biolabs (Hitchin, Grande-Bretagne). Le vecteur d'expression eucaryote pcDNA-3.1a(+) et les bactéries compétentes TOP 10 sont fournis par Invitrogen-Fischer-Scientific. Le vecteur d'expression bactérienne pET-41a(+) et les bactéries *Escherichia coli* (*E. coli*) BL21 (DE3) proviennent de Novagen EMD4Biosciences (Nottingham, UK). Le kit de mutagenèse dirigée « QuickChange site-directed mutagenesis kit » et la *Pfu* Turbo DNA Polymerase sont fournis par Stratagene (La Jolla, CA). Le kit d'extraction d'ADN plasmidique provient de Qiagen (Courtabœuf, France)

### 3. Les réactifs chimiques

Le 4-Méthylombelliféryl- $\beta$ -D-xylopyranoside (4-MOX), le 4-nitrophénol- $\beta$ -D-xylopyranoside (4-NP-Xyl), l'UDP-Gal, l'UDP- $\alpha$ -D-glucose (UDP-Glc), l'UDP- $\alpha$ -D-N-acétylglucosamine (UDP-GlcNAc), l'UDP- $\alpha$ -D-N-acétylgalactosamine (UDP-GalNAc), l'UDP- $\alpha$ -D-acide glucuronique (UDP-GlcA) et l'UDP- $\alpha$ -D-xylose (UDP-Xyl) sont fournis par Sigma-Aldrich. Les substrats accepteurs 7-méthoxy-naphtyl- $\beta$ -D-xylopyranoside et le 7-méthoxy-naphtyl- $\beta$ -D-xylopyranoside-2-phosphate ont été synthétisés par l'équipe du Dr. Jean-Claude Jacquinet (UMR 6005 CNRS, Orléans).

#### 4. Les réactifs radiomarqués

Le sulfate de sodium  $\text{Na}_2[^{35}\text{S}]\text{SO}_4$ , l'UDP[ $^{14}\text{C}$ ]-Gal, l'UDP[ $^{14}\text{C}$ ]-GlcA, l'UDP[ $^{14}\text{C}$ ]-Xyl, l'UDP[ $^{14}\text{C}$ ]-GalNAc, l'UDP[ $^{14}\text{C}$ ]-GlcNAc et le mélange cystéine-méthionine « EasyTag express protein labeling, [ $^{35}\text{S}$ ] » proviennent de Perkin-Elmer Life Science (Courtaboeuf, France). L'UDP[ $^{14}\text{C}$ ]-Glc et le marqueur de masse moléculaire « Rainbow<sup>TM</sup> [ $^{14}\text{C}$ ] methylated protein molecular weight » sont fournis par Amersham Biosciences (GE Healthcare, Saclay, France).

#### 5. Les réactifs biochimiques

La papaïne, la généticine G418, l'isopropyl- $\beta$ -D-galactopyranoside (IPTG), l'anticorps polyclonal anti-IgG de souris, l'anticorps polyclonal anti-IgG de chèvre, proviennent de Sigma-Aldrich. Le kit « MagneHis Protein Purification System » est fourni par Promega. La résine Sephadex G50, la résine Sepharose Glutathion 4B et les films Hyperfilm MP proviennent d'Amersham Biosciences (GE Healthcare). L'anticorps monoclonal anti-décorine provient de Calbiochem. L'anticorps monoclonal anti-GST est fourni par Novagen.

Les gels à gradient (4-12% Bis-Tris) « Gel Criterion XT », le tampon XT MOPS (acide *N*-morpholino-3-propane sulfonique) et le gel de chromatographie « Macro-Prep DEAE Support » sont fournis par Bio-Rad (Marnes-la-Coquette, France).

## *Méthodes expérimentales*

### I. CLONAGE DE LA $\beta$ 4GALT7 HUMAINE ET CONSTRUCTION DES VECTEURS D'EXPRESSION

La séquence codant pour la  $\beta$ 4GalT7 a été amplifiée par une procédure standard d'amplification par PCR « Polymerisation Chain Reaction » à partir d'une banque d'ADNc de placenta (Clontech, Palo Alto, CA) en utilisant des amorces 3' et correspondant à la séquence publiée par Almeida *et al.* (1999).

Amorce sens : 5'-CAGGTACCATGTTCCCCTCG-3'

Amorce antisens : 5'-CTCTAGACTACAGATCCTCTTCAGAGATGAGTTTCTGCTCGCTGAATGTG-3'

L'ADNc a été ensuite sous-cloné dans le vecteur PCR2.1 TOPO (Invitrogen, Cergy Pontoise, France) comme précédemment décrit (Gulberti *et al.*, 2005). La construction PCR2.1-



$\beta$ 4GalT7 obtenue a été utilisée pour construire les vecteurs d'expression nécessaires à la réalisation de la suite de ce travail.

### 1. Construction du vecteur d'expression pcDNA- $\beta$ 4GalT7

Pour l'expression de l'enzyme  $\beta$ 4GalT7 dans des cellules eucaryotes, la séquence précédente a été modifiée par PCR en introduisant dans l'amorce 5' un site *KpnI* et une séquence Kozak d'optimisation de la traduction; et dans l'amorce 3' un site *XbaI* et une séquence codant pour l'étiquette "Myc" (EQKLISEEDL). Le produit de PCR est ainsi cloné dans les sites *ad hoc* du vecteur pcDNA-3.1a(+) qui porte le gène de résistance à l'ampicilline, ce qui a permis la construction du vecteur d'expression pcDNA- $\beta$ 4GalT7. Le plasmide portant le gène codant pour la  $\beta$ 4GalT7 sauvage ou mutée a été utilisé pour transférer des cellules HeLa (cellules épithéliales humaines de cancer de col utérin) et des cellules CHO-pgsB618 (*Chinese Hamster Ovary*) déficientes en  $\beta$ 4GalT7.

Les cellules HeLa ont été utilisées pour la réalisation des cinétiques enzymatiques et la détermination des constantes enzymatiques apparentes. Les cellules CHO-pgsB618, qui sont déficientes en  $\beta$ 4GalT7 et ne produisent pas de GAGs endogènes, ont été utilisées pour la réalisation des études *ex vivo*.

### 2. Construction du vecteur d'expression pET- $\beta$ 4GalT7

Pour la construction du vecteur d'expression bactérien, l'ADNc codant une forme tronquée de  $\beta$ 4GalT7, déléetée des 60 acides aminés N-terminaux, a été amplifié par PCR au moyen d'amorces contenant un site *NcoI* en 5' et un site *XhoI* en 3' et sous-cloné dans le vecteur pET-41a(+) en fusion avec le gène de la GST (Glutathion-S-Transférase), pour générer la construction pET- $\beta$ 4GalT7. Ce plasmide a été ensuite utilisé pour transformer des cellules *E. coli* BL21 (DE3) et pour exprimer la  $\beta$ 4GalT7 sauvage ou mutée en fusion avec la GST en N-terminal, avec un site de coupure à l'enterokinase (Invitrogen).



**Figure 33. Schéma de la construction GST- $\beta$ 4GalT7 recombinante humaine exprimée dans des cellules BL21**

## **II. CLONAGE DE LA DECORINE HUMAINE ET CONSTRUCTION DU VECTEUR D'EXPRESSION pcDNA-DECORINE**

L'ADNc obtenu à partir d'une banque d'ADNc de placenta a été amplifié par PCR en utilisant l'amorce sens (5'-ATGAAGGCCACTATCATCCTCCTTCTG-3') et l'amorce antisens (5'-CTACTTATAGTTTCCGAGTTGAATGGC-3'). Un site *Afl* et une séquence Kozak du côté 5' et un site *Xho* et une séquence codant 5 His (pentaHis) du côté 3' ont ensuite été introduits par PCR. L'ADNc modifié a été cloné dans le vecteur d'expression pcDNA3.1a(+) pour obtenir la construction pcDNA-décorine qui a été transfectée dans des cellules CHO-pgsB618.

## **III. MUTAGENESE DIRIGEE PAR PCR**

Dans ce travail, les mutations réalisées sont des mutations ponctuelles (mutations conservatives et mutations non conservatives) réalisées par la technique de mutagenèse dirigée par PCR en utilisant le kit « QuickChange site-directed mutagenesis kit ». Le plasmide pcDNA-β4GalT7 ou pET-β4GalT7 a été utilisé comme matrice avec les amorces sens et antisens qui portent le codon de la mutation d'intérêt (en gras souligné). Les séquences des amorces utilisées sont fournies dans le tableau situé en annexes.

Les mélanges réactionnels de PCR contiennent 50 ng d'ADN parental, 125 ng de chaque amorce nucléotidique, 10 mM dNTP, le tampon de la réaction et 2 U de l'enzyme *Pfu* Turbo dans un volume final de 50 µl. Le programme de PCR utilisé est le suivant: 95°C pendant 30 secondes, 95°C pendant 1 minute, 55°C pendant 1 minute et 65°C pendant 1 minute x le nombre de kb du plasmide. Ce programme a été répété 16 fois. Le produit de PCR est digéré par l'enzyme de restriction *DpnI* pendant 1 heure à 37°C. Cette enzyme clive de façon spécifique l'ADN méthylé parental permettant ainsi de sélectionner l'ADN muté non méthylé et d'éliminer l'ADN non muté parental.

Après transformation des cellules compétentes *E. coli* (TOP 10) par choc thermique, les clones recombinants ont été sélectionnés grâce à la présence d'une résistance à l'ampicilline (Amp) pour la construction pcDNA-β4GalT7 et d'une résistance à la kanamycine (Kana) pour la construction pET-β4GalT7.

Les clones qui présentent une résistance à l'antibiotique ont été mis en culture dans du milieu Luria-Bertani (LB) supplémenté en kanamycine ou ampicilline afin d'extraire et de purifier les plasmides d'intérêt. Les mutations sont systématiquement vérifiées par séquençage.

## **IV. EXPRESSION HETEROLOGUE DE LA $\beta$ 4GalT7 RECOMBINANTE HUMAINE**

### **1. Expression de la $\beta$ 4GalT7 dans des cellules eucaryotes**

Pour exprimer la  $\beta$ 4GalT7 dans des cellules eucaryotes, les cellules HeLa ou les cellules CHO-pgsB618 ont été transfectées avec le plasmide pcDNA- $\beta$ 4GalT7 codant la protéine sauvage ou mutée. Les cellules ont été maintenues en culture monocouche à 37°C et à 5% de CO<sub>2</sub> dans du milieu DMEM pour les HeLa et F12/DMEM (1:1) pour les CHO-pgsB618, contenant chacun 10% (v/v) de sérum de veau fœtal (SVF, GIBCO-Invitrogen), 1 mM de L-Glutamine et pénicilline (100 U/ml)/streptomycine (100 mg/ml).

#### **1.1. Transfection transitoire des cellules eucaryotes (HeLa et CHO-pgsB618) avec la $\beta$ 4GalT7**

Les cellules sontensemencées la veille de la transfection dans des plaques de culture de 6 puits afin d'obtenir une confluence de 70%. Elles sont ensuite transfectées transitoirement avec le plasmide pcDNA- $\beta$ 4GalT7 sauvage ou muté au moyen d'un agent de transfection à base de polyéthylénimine (PEI), commercialisé sous le nom d'ExGen 500 (Euromedex, Souffelweyersheim, France). Un mélange de 2  $\mu$ g d'ADN et de 15  $\mu$ l d'ExGen 500 dilués dans une solution de NaCl (0.9%) est ajouté à chaque puit, selon les recommandations du fournisseur. Les cellules sont collectées 48 h après transfection, puis centrifugées dans le tampon PBS (*Phosphate Buffered Saline* pH 7.8). Les culots cellulaires sont repris dans du tampon 5 mM HEPES, 0.25 M saccharose (pH 7.4) puis soniqués pendant 10 secondes sur la glace (2 x 5 secondes avec un intervalle de 10 secondes). Les homogénats cellulaires obtenus sont utilisés pour la réalisation des tests d'activité enzymatique et pour l'analyse de l'expression de la protéine par immunoblot.

#### **2. Expression stable de la décorine dans des cellules CHO-pgsB618**

Le vecteur pcDNA-décorine a été transfecté, comme décrit précédemment, dans des cellules CHO-pgsB618. La G418 ou généticine (4 g/l) a été utilisée pour sélectionner les cellules qui ont intégré le gène d'intérêt dans leur génome. Les colonies résistantes ont été ensuite isolées et mises en culture. La décorine recombinante excrétée dans le milieu a été purifiée par chromatographie d'affinité sur des billes magnétiques de nickel. L'expression de la protéine est analysée par SDS-PAGE (*Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis*) et immunoblot.

### **a. Purification de la décorine par chromatographie d'affinité (*Immobilized Metal Affinity Chromatography*, IMAC)**

**Principe :** La technique que nous avons utilisée pour purifier la décorine recombinante présente dans le milieu de culture est une chromatographie d'affinité de type IMAC. Elle nécessite l'immobilisation d'un atome de nickel sur une résine, ce métal va établir des liens de coordination avec les chaînes latérales des résidus His (Tag-His) présents sur la décorine.

**Protocole de purification :** La décorine recombinante humaine exprimée de façon stable dans les cellules CHO-pgsB618, été donc purifiée, à partir du milieu de culture, par le kit « MagneHis Protein Purification System » selon les recommandations du fournisseur. La décorine purifiée a été analysée par SDS-PAGE puis immunodétection (comme décrit ci-dessous), en utilisant un anticorps polyclonal anti-décorine, et un anticorps secondaire anti-IgG de chèvre.

### **1.3 Analyse de l'expression protéique par immunoblot**

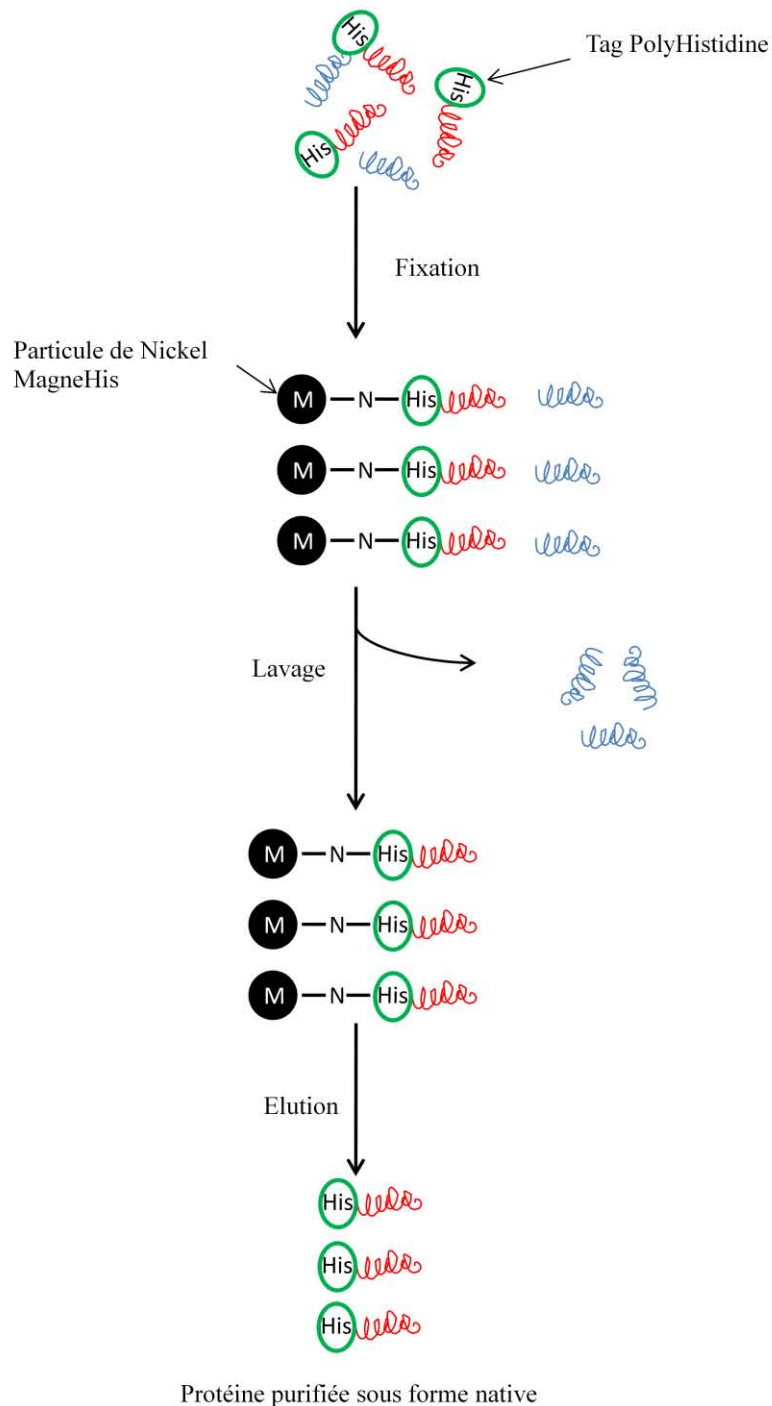
**Séparation des protéines par électrophorèse :** L'expression de la protéine recombinante a été analysée par SDS-PAGE en conditions dénaturantes. Les échantillons ont été dilués (v/v) dans le tampon de Laemmli (1970) composé de 125 mM Tris-HCl (pH 6.8), 4% (p/v) SDS, 20% (v/v) glycérol, 10% (p/v)  $\beta$ -mercaptoéthanol, 0.004% (p/v) de bromophénol, et la dénaturation des protéines est complétée par ébullition pendant 5 min avant dépôt sur gel.

La séparation des protéines se fait parallèlement à un marqueur de masse moléculaire (10-250 kDa, Bio-Rad), dans un gel de migration 10% (p/v) acrylamide-bis acrylamide (pH 8.8).

La migration est effectuée dans une mini-cuve (Novex, Paris, France) contenant un tampon de migration constitué de 200 mM de glycine, 25 mM de Tris-base et de 10 g de SDS, à ampérage constant de 20 à 30 mA (4 mA/cm<sup>2</sup> gel).

**Transfert des protéines sur membrane de polyvinylidène difluorure (PVDF) et immunodétection :** Après migration, les protéines sont transférées sur une membrane de PVDF en milieu semi-sec, dans un appareil d'électrotransfert Trans-Blot (Bio-Rad), sous un voltage constant de 12 mV pour un gel (2 mV/cm<sup>2</sup> gel). Après le transfert, la membrane est saturée pendant 1 h sous agitation dans une solution d'albumine sérique de bovin (ASB) à 3% (p/v) préparée dans le tampon TBST (*Tris Buffer Saline Tween*, 150 mM Tris-HCl (pH 8.5), 10 mM NaCl, 0.5% (v/v) Tween 20, 0.02% (p/v) azide de sodium), puis la membrane est hybridée avec un premier anticorps pendant toute la nuit avant d'être lavée et hybridée avec

un anticorps secondaire, anti-IgG couplé à la phosphatase alcaline (Sigma). La protéine recombinante est révélée par l'ajout de BCIP/NBT (5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium).



**Figure 34. Schéma du protocole de purification de la décorine par le kit MagneHis Protein Purification System de Promega.**

## 2. Expression de la $\beta$ 4GalT7 sous forme soluble dans les bactéries BL21

Les cellules BL21 (DE3) transformées par choc thermique avec le plasmide pET- $\beta$ 4GalT7 sauvage ou muté sont mises en culture dans du milieu LB supplémenté de 100  $\mu$ g/ml de kanamycine à 37°C sous agitation (160 rpm) pendant 12 h. La préculture est inoculée à 1/100<sup>ème</sup> dans 200 ml de milieu LB-kanamycine puis incubée à 37°C sous agitation (160 rpm) jusqu'à atteindre une densité optique (DO) à 600 nm comprise entre 0.6 et 0.8 unités. L'expression de la protéine est alors induite par l'ajout de 1 mM d'IPTG et les cellules sont laissées sous agitation (160 rpm) 1é h à 20°C. Les bactéries sont centrifugées à 5000 x g pendant 15 min puis elles sont lavées deux fois à 4°C avec le tampon de lyse (50 mM de phosphate de sodium (pH 7.4), 1 mM phénylméthylsulfonyfluoride (PMSF), 1 mM d'acide éthylène diamine tétra acétique (EDTA), 5% (v/v) glycérol. Les culots de bactéries obtenus sont repris dans le tampon de lyse supplémenté d'un cocktail complet d'inhibiteur de protéases sans EDTA (Roche Diagnostics GmbH, Mannheim, Germany) et lysés par sonication (4 min de sonication à 4°C par pulses de 30 secondes espacées de 30 secondes). Le lysat de bactéries est centrifugé à 4°C pendant 5 min à 3000 x g pour éliminer les débris cellulaires puis à 12000 x g pendant 15 min à 4°C afin de séparer les protéines solubles du matériel insoluble. Le surnageant contenant les protéines solubles a ensuite été utilisé pour purifier la  $\beta$ 4GalT7.

### 2.1. Purification de la protéine par chromatographie d'affinité

**Principe** : La  $\beta$ 4GalT7 recombinante humaine a été exprimée dans les bactéries BL21 en fusion avec la GST ce qui nous a permis de purifier la protéine GST- $\beta$ 4GalT7 par chromatographie d'affinité sur une colonne contenant une résine de type Sepharose 4B glutathion sur laquelle se fixe spécifiquement la GST.

**Protocole de purification** : Les protéines solubles sont chargées sur la colonne préalablement équilibrée avec le tampon PBS 1x, et les protéines non fixées sur le glutathion ont été éliminées par trois lavages équivalents à dix fois le volume de la matrice. La protéine GST- $\beta$ 4GalT7 sauvage ou mutée a été éluée dans le tampon d'éluion (50 mM Tris-HCl pH 8, 10 mM glutathion réduit) et les fractions protéiques obtenues ont été analysées par SDS-PAGE puis coloration du gel au bleu de Coomassie. Les fractions contenant la  $\beta$ 4GalT7 sont ensuite regroupées, puis concentrées par Spectra/Gel<sup>®</sup> Absorbent (Spectrum Labs). La protéine

purifiée ainsi obtenue a été alors aliquotée et conservée à -80°C pour une utilisation ultérieure.

### **3. Dosage de protéines**

Le dosage de protéines totales dans les homogénats de cellules eucaryotes a été réalisé par dosage colorimétrique selon la méthode de Bradford (1976).

Pour les protéines exprimées dans les bactéries, la concentration en protéines totales a été déterminée par le kit de dosage de protéines Qubit (Invitrogen), et par la puce à protéine Agilent (Agilent 2100 bioanalyser, Protein 230 kit Agilent Technologies) pour les protéines purifiées, selon les recommandations des fournisseurs. Cette technique permet de déterminer le degré de pureté de la protéine purifiée et sa concentration.

### **4. Mesure de l'activité enzymatique**

Nous avons utilisé deux méthodes pour déterminer l'activité enzymatique GalT de la protéine recombinant humaine. La chromatographie liquide haute performance en phase inverse (CLHP) a été utilisée pour mesurer l'activité de l'enzyme sauvage ou mutée, exprimée dans des cellules eucaryotes ou dans des bactéries, en présence d'UDP-Gal comme substrat donneur et de 4-MOX comme substrat accepteur. Les constantes cinétiques apparentes et réelles ont été déterminées au moyen de la même technique. La chromatographie sur couche mince (CCM) a été utilisée pour l'étude de la spécificité de l'enzyme vis-à-vis des différents substrats donneurs et des différents substrats accepteurs.

#### **4.1. Mesure de l'activité enzymatique par CLHP en phase inverse**

La réaction enzymatique a été réalisée dans un volume de 50 µl en présence de 40 µg d'homogénat cellulaire (protéines exprimées dans les cellules HeLa), 1 mM d'UDP-Gal, 0.05 µCi UDP[U<sup>14</sup>C]-Gal, 5 mM de 4-MOX, 100 mM de cacodylate de sodium (pH 7.0) contenant 10 mM d'ions Mn<sup>2+</sup>. Le mélange a été incubé 60 min à 37°C, et la réaction a été arrêtée par l'ajout de 5 µl d'une solution de HCl 6 N. Les essais enzymatiques sont centrifugés à 10,000 × g pendant 10 min à 4°C et les surnageants sont analysés par CLHP en phase inverse grâce à une colonne C18 (4.6 × 150 mm, 4 µm, Waters, Saint-Quentin-en-Yvelines, France) et en utilisant une chaîne de CLHP Waters équipée d'un radiodétecteur Berthold HPLC. La phase mobile qui nous a permis de séparer le produit de la réaction est composée de 14% (v/v)

d'acétonitrile (ACN), 0.02% (v/v) de l'acide trifluoroacétique (TFA) et 86% (v/v) d'eau distillée. L'activité enzymatique a été déterminée par la mesure de la radioactivité associée au produit de la réaction par comptage simultané en scintillation liquide (Ultima-Flo, Perkin Elmer).

L'activité des enzymes purifiées est évaluée comme décrit précédemment, mais en présence de 200 ng de protéine purifiée et une incubation de 30 min à 37°C. L'activité enzymatique est déterminée de la même façon par CLHP.

#### **4.2. Mesure de l'activité enzymatique par CCM selon la méthode de Bansal et Gessner (1980)**

Afin d'étudier la spécificité de l'enzyme vis-à-vis de ses deux substrats, nous avons incubé la protéine recombinante à 37°C pendant 60 min dans un volume réactionnel final de 40 µl, en présence de 5 mM de substrat accepteur 4-MOX, 0.04 mM de substrat donneur UDP-Gal, 0.1 µCi UDP[U<sup>14</sup>C]-Gal, 10 mM MnCl<sub>2</sub>, 100 mM de tampon cacodylate de sodium. La réaction est arrêtée par l'ajout de 40 µl d'éthanol absolu et le surnageant obtenu après 10 min de centrifugation à 10.000 x g a été déposé sur des plaques de gel de silice. Les produits de la réaction enzymatique ont été séparés par la migration par capillarité dans une cuve fermée contenant une phase mobile constituée de *n*-butanol, acétone, acide acétique, ammoniac, et d'eau, 35/25/8/1/30 (v/v). En fin de la migration, les plaques sont séchées avant d'être mises au contact du film dans une cassette d'exposition et gardées à -80°C, 3 jours environ. Finalement, les films sont développés et les plaques de silices sont grattées, afin de réaliser un comptage dans le liquide scintillant Ultima-Gold et de déterminer la radioactivité associée au produit de la réaction.

Nous avons testé les substrats donneurs de type UDP-glycose : UDP-Gal, UDP-Glc, UDP-GalNAc, UDP-GlcNAc, UDP-GlcA et UDP-Xyl; et différents substrats accepteurs de type : 4-MO-Xyl (4-MOX), 4-MO-Gal, 4-MO-Glc, 4-MO-GalNAc, 4-MO-GlcNAc, 4-NP-Xyl, MN-Xyl et MN-Xyl-2P.

#### **5. Détermination des constantes cinétiques $V_{max}$ , $K_m$ et $k_{cat}$**

Comme nous l'avons précisé auparavant, la β4GalT7 est une enzyme à deux substrats; Un substrat donneur, l'UDP-Gal, et un substrat accepteur, le 4MOX. Dans le but de déterminer les paramètres enzymatiques de la protéine vis-à-vis des deux substrats, la cinétique enzymatique a été réalisée dans les conditions standards (comme décrit précédemment) en



fixant la concentration de l'un des substrats et en faisant varier celle de l'autre. Ainsi pour déterminer les constantes cinétiques de l'enzyme vis-à-vis du substrat accepteur, la concentration du substrat donneur a été fixée à 1 mM et celle du substrat accepteur a été variée de 0 mM-10 mM. De la même façon, les constantes cinétiques de l'enzyme vis-à-vis du substrat donneur ont été déterminées à une concentration fixe de 4-MOX de 5 mM à une concentration croissante d'UDP-Gal (0 mM-10 mM).

## 6. Exploitation des données

Nous avons utilisé le logiciel de régression non linéaire Sigma Plot 9.0 (Erkraft, Allemagne) pour calculer les constantes enzymatiques  $V_{max}$ ,  $K_m$  et  $K_{cat}$  selon l'équation de Michaelis-Menten :

$$v = (V_{max} \times [S]) / (K_m + [S])$$

Où

$v$	La vitesse initiale de la réaction
$K_m$	La constante de Michaelis-Menten
$S$	La concentration en substrat
$V_{max}$	La vitesse maximale de la réaction enzymatique

La constante catalytique ( $K_{cat}$ ) de l'enzyme purifiée a été calculée selon l'équation :

$$K_{cat} = V_{max}/[P]$$

Où  $[P]$  correspond à la concentration en protéines.

Pour déterminer la masse molaire de la protéine en g/mol nous avons utilisé le serveur ExPASy Proteomics Server, Suisse Institute of Bioinformatics.

Nous avons également utilisé l'équation ci-dessous correspond à un mécanisme de réaction de type séquentiel afin de déterminer les constantes cinétiques réelles:

$$V = \frac{V_{max} [A][B]}{K_a[B] + K_b[A] + K_{ia} K_b + [A][B]}$$

Où:

$v$	La vitesse initiale de la réaction
$V_{max}$	La vitesse maximale de la réaction enzymatique
$K_a$	$K_m$ du substrat A

$K_b$	$K_m$ du substrat B
$K_{ia}$	La constante d'inhibition du substrat A
$K_{ib}$	La constante d'inhibition du substrat B

Dans cette équation [A] et [B] correspondent respectivement aux concentrations des substrats donneur et accepteur de la  $\beta$ GalT7.

Pour l'ensemble des expérimentations, chaque valeur correspond à la moyenne  $\pm$  SD d'au moins trois expériences indépendantes réalisées en triple. L'étude statistique a été réalisée par le logiciel GraphPad Prism 5 en appliquant le test d'ANOVA avec la correction de Bonferonni.

## V. EXTRACTION ET ANALYSE DES GAGs PAR INCORPORATION DU $[^{35}\text{S}]\text{SO}_4$

### 1. Analyse du taux de biosynthèse des GAGs endogènes

Les cellules CHO-pgsB618 préalablement et indépendamment transfectées avec le plasmide vide (contrôle négatif) et le plasmide portant le gène codant la protéine sauvage ou mutée ont été incubées pendant 8 h dans le milieu Fischer supplémenté avec 10  $\mu\text{Ci/ml}$  de  $[^{35}\text{S}]\text{Na}_2\text{SO}_4$  (Perkin-Elmer Life Sciences) 48 h après transfection. Les PGs ont été extraits et purifiés à partir du milieu selon le protocole adapté de Ha *et al.* (2005). La protéine "core" des PGs a été digérée pendant 3 h à 60°C par l'enzyme papaïne fraîchement préparée dans un tampon qui contient 20 mM  $\text{Na}_2\text{HPO}_4$ , 1 mM EDTA et 5 mM L-Cys (pH 6.8). La digestion a été arrêtée par ébullition pendant 10 min, et la protéine a été précipitée par centrifugation à 8,000 x g pendant 15 min. Trois volumes d'acétate de potassium 5% (p/v) préparé dans l'éthanol absolu ont été ajoutés à un volume de surnageant afin de permettre la précipitation des GAGs toute la nuit à 4°C. Le précipité a été collecté par centrifugation pendant 30 min à 8,000 x g, puis resuspendu dans 500  $\mu\text{l}$  de NaCl 0.2 M et centrifugé de nouveau pendant 30 min à 8,000 x g. Les GAGs se trouvent alors en suspension dans le surnageant; ils ont été précipités à 37°C pendant deux à trois heures en présence de 100  $\mu\text{g}$  de CS de requin et de 100  $\mu\text{l}$  de chlorure de cétypyridinium 5% (CPC), puis collectés par centrifugation à 8,000 x g pendant 30 min. Le surnageant a été ensuite repris dans 1 volume de NaCl 2.5 M et 5 volumes d'éthanol absolu et centrifugé pendant 30 min à 8000 x g. Le culot de GAGs purifiés obtenu a enfin été resuspendu dans le tampon Tris-HCl 0.2 M (pH 8.0) et la radioactivité associée à cet extrait a

été déterminée par comptage dans 4.5 ml du cocktail de scintillation Ultima-Gold (Perkin Elmer).

## **2. Analyse des chaînes de GAGs synthétisées sur la protéine "core" de la décorine par SDS-PAGE et autoradiographie**

Les cellules CHO-pgsB618 qui expriment la décorine recombinante humaine de façon stable ont été transfectées transitoirement et séparément avec le plasmide pcDNA qui porte le gène codant pour la  $\beta$ 4GalT7 humaine sauvage ou mutée ou avec le plasmide vide. 48 heures après transfection, les cellules ont été radiomarquées au  $[^{35}\text{S}]\text{Na}_2\text{SO}_4$  pendant 12 h comme décrit précédemment. L'analyse des chaînes de GAGs néosynthétisées sur la protéine "core" de la décorine purifiée, à partir du milieu, par chromatographie d'affinité en utilisant des billes magnétiques de nickel, comme décrit précédemment.

Les PGs radiomarqués purifiés ont été séparés par SDS-PAGE sur gel à gradient de polyacrylamide 4-12% (des gels Criterion précoulés, BioRad) parallèlement à un marqueur de masse moléculaire radioactif « Rainbow<sup>TM</sup>[ $^{14}\text{C}$ ] methylated protein molecular weight ». Les échantillons sont portés à ébullition pendant 5 min, en présence d'un agent réducteur, avant d'être déposés sur gel. Après électrophorèse, le gel a été fixé dans un mélange composé d'eau distillée, de méthanol et d'acide acétique: 5/1/1 (v/v), puis incubé à température ambiante, dans une solution amplificatrice du signal radioactif (Amplify Reagent, Amersham Biosciences), avant d'être séché et exposé sur film photographique (Amersham Hyperfilm<sup>TM</sup>MP, GE Healthcare). Après une semaine d'exposition à  $-20^\circ\text{C}$  à l'abri de la lumière, le film est développé par autoradiographie puis scanné.

## **3. Analyse de l'initiation de la biosynthèse des chaînes de GAGs sur les xylosides exogènes par SDS-PAGE**

Les cellules CHO-pgsB618 transfectées avec la  $\beta$ 4GalT7 sauvage ou mutée ou avec le plasmide vide ont été incubées pendant 12 h dans le milieu Fischer contenant  $10 \mu\text{Ci/ml}$  de  $[^{35}\text{S}]\text{Na}_2\text{SO}_4$  en présence ou en absence de 4-MOX ( $5 \mu\text{M}$ ). Le milieu collecté a été déposé sur une colonne de Sephadex G-50 afin d'éliminer la radioactivité non incorporée, puis les PGs ont été élués dans deux fractions de 1.5 ml d'eau ultra pure. Les échantillons sont concentrés au *Speed Vac* (Avantec), puis portés à ébullition pendant 5 min en présence d'un agent réducteur, avant d'être déposés sur gel à gradient 4-12% Bis-Tris. La séparation des GAGs se fait parallèlement au marqueur de masse moléculaire. Le gel ainsi obtenu a été fixé,

séché, exposé au film photographique et développé par autoradiographie, comme décrit précédemment.

#### **4. Marquage des cellules CHO-pgsB618 par immunofluorescence indirecte**

##### **4.1. Analyse de l'expression des chaînes d'HS au niveau de la membrane cellulaire**

Les cellules CHO-pgsB618 transfectées avec le plasmide vide, la  $\beta$ 4GalT7 sauvage ou mutée ont été lavées trois fois avec le tampon PBS 1x puis fixées dans une solution de paraformaldéhyde/PBS 3% (v/v) pendant 20 min à température ambiante. Les cellules sont à nouveau lavées avec le tampon PBS, avant d'être incubées en présence de sérum de chèvre/PBS 10% (v/v) pendant 30 min, puis en présence de l'anticorps primaire anti-HS (dilution 1:50) exprimé dans la souris et dirigé contre les HS humaines (épitope 10E4) (Seikagaku Corp., Tokyo, Japan). Après trois lavages dans le PBS, les cellules sont incubées avec l'anticorps secondaire anti-IgG de souris couplé à l'AlexaFluor 488 caractérisé par un spectre d'émission vert. Après une dernière série de lavage avec le PBS, nous avons réalisé un marquage des noyaux cellulaires, au moyen d'une solution de Hoechst. Enfin nous avons réalisé le montage des cellules ainsi obtenues, entre lame et lamelle en utilisant le Moviol. Après séchage à 4°C à l'abri de la lumière, nous avons pu observer et analyser les préparations au microscope à fluorescence (Leica Microsystems, Leica DFC420 C).

##### **4.2. Analyse de l'expression et de la localisation cellulaire de la $\beta$ 4GalT7 par immunofluorescence indirecte**

Afin d'analyser simultanément l'expression de la  $\beta$ 4GalT7 recombinante dans les cellules CHO-pgsB618 transfectées avec l'enzyme sauvage, mutée ou avec le plasmide vide, nous avons procédé également à un marquage en immunofluorescence indirecte. Les cellules sont perméabilisées avec le Triton-X100, après fixation par le paraformaldéhyde comme décrit précédemment. Nous avons utilisé un anticorps monoclonal primaire anti-Myc qui se fixe spécifiquement le Tag Myc de la  $\beta$ 4GalT7 recombinante. Ensuite les cellules ont été incubées avec l'anticorps secondaire anti-IgG de souris, conjugué au fluorochrome AlexaFluor 594 selon le protocole expérimental décrit ci-dessus.

# *Résultats*

*Publication n°1*

# Identification of Key Functional Residues in the Active Site of Human $\beta$ 1,4-Galactosyltransferase 7

## A MAJOR ENZYME IN THE GLYCOSAMINOGLYCAN SYNTHESIS PATHWAY<sup>\*,§</sup>

Received for publication, June 8, 2010, and in revised form, September 14, 2010. Published, JBC Papers in Press, September 15, 2010, DOI 10.1074/jbc.M110.151951

Ibtissam Talhaoui<sup>‡</sup>, Catherine Bui<sup>†1</sup>, Rafael Oriol<sup>§</sup>, Guillermo Mulliert<sup>¶</sup>, Sandrine Gulberti<sup>‡</sup>, Patrick Netter<sup>‡</sup>, Michael W. H. Coughtrie<sup>||</sup>, Mohamed Ouzzine<sup>‡</sup>, and Sylvie Fournel-Gigleux<sup>‡2</sup>

From the <sup>‡</sup>Faculté de Médecine, UMR 7561 CNRS-Université de Nancy I, BP 184, 54505 Vandoeuvre-lès-Nancy, France, the <sup>§</sup>Institut André Lwoff, U1004 INSERM, Université Paris Sud, 94807 Villejuif, France, the <sup>¶</sup>Faculté des Sciences et Techniques, UMR 7036 CNRS-Université de Nancy I, BP 184, 54505 Vandoeuvre-lès-Nancy, France, and the <sup>||</sup>Division of Medical Sciences, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY, Scotland, United Kingdom

Glycosaminoglycans (GAGs) play a central role in many pathophysiological events, and exogenous xyloside substrates of  $\beta$ 1,4-galactosyltransferase 7 ( $\beta$ 4GalT7), a major enzyme of GAG biosynthesis, have interesting biomedical applications. To predict functional peptide regions important for substrate binding and activity of human  $\beta$ 4GalT7, we conducted a phylogenetic analysis of the  $\beta$ 1,4-galactosyltransferase family and generated a molecular model using the x-ray structure of *Drosophila*  $\beta$ 4GalT7-UDP as template. Two evolutionary conserved motifs, <sup>163</sup>DVD<sup>165</sup> and <sup>221</sup>FWGWGREDD<sup>230</sup>, are central in the organization of the enzyme active site. This model was challenged by systematic engineering of point mutations, combined with *in vitro* and *ex vivo* functional assays. Investigation of the kinetic properties of purified recombinant wild-type  $\beta$ 4GalT7 and selected mutants identified Trp<sup>224</sup> as a key residue governing both donor and acceptor substrate binding. Our results also suggested the involvement of the canonical carboxylate residue Asp<sup>228</sup> acting as general base in the reaction catalyzed by human  $\beta$ 4GalT7. Importantly, *ex vivo* functional tests demonstrated that regulation of GAG synthesis is highly responsive to modification of these key active site amino acids. Interestingly, engineering mutants at position 224 allowed us to modify the affinity and to modulate the specificity of human  $\beta$ 4GalT7 toward UDP-sugars and xyloside acceptors. Furthermore, the W224H mutant was able to sustain decorin GAG chain substitution but not GAG synthesis from exogenously added xyloside. Altogether, this study provides novel insight into human  $\beta$ 4GalT7 active site functional domains, allowing manipulation of this enzyme critical for the regulation of GAG synthesis. A better understanding of the mechanism underlying GAG assembly paves the way toward GAG-based therapeutics.

Glycosaminoglycans (GAGs),<sup>3</sup> the polysaccharide chains of proteoglycans, are distributed at the surface of most cells and in extracellular matrices of virtually every tissue. GAGs are implicated as regulators in many biological events related to intracellular signaling, cell-cell interactions, and tissue morphogenesis. Thus, the pathophysiological roles of GAGs are highly diversified, ranging from mechanical support to intricate effects on various cellular processes such as cell adhesion, differentiation, proliferation, and motility (1). These vital functions depend on the interactions of the GAG chains with a variety of molecules, including growth factors, cytokines, and their receptors, enzymes, including matrix proteases and coagulation factors as well as extracellular matrix proteins (2).

The biosynthesis of GAG chains, which governs the expression of these functions, involves the ordered stepwise addition of a series of monosaccharides derived from UDP-sugars onto the polypeptide backbone of a proteoglycan through the action of specific glycosyltransferases (3). This multistep process starts by the addition of Xyl to a serine residue of the core protein followed by the addition of two Gal and one GlcUA to form the tetrasaccharide linkage region (GlcUA $\beta$ 1,3Gal $\beta$ 1,3Gal $\beta$ 1,4Xyl $\beta$ 1) common to most GAGs. This tetrasaccharide structure serves as a primer for the formation of heparin/heparan sulfate and chondroitin/dermatan sulfate chains, which is initiated by the attachment of GlcNAc or GalNAc, respectively. The glucosaminoglycan (heparin/heparan sulfate) and the galactosaminoglycan (chondroitin/dermatan sulfate) chains then assemble by the alternating addition of GlcUA and GlcNAc or GlcUA and GalNAc, respectively. GAG chains are further modified by the cooperative action of epimerases and sulfotransferases that add considerable complexity and functionality to the polysaccharide backbone (4, 5).

The enzyme  $\beta$ 4GalT7 (xylosylprotein  $\beta$ 1,4-galactosyltransferase, EC 2.4.1.133) catalyzes the transfer of a Gal residue provided by UDP- $\alpha$ -D-Gal (UDP-Gal) onto Xyl, a key step in the synthesis of the linkage region of GAG chains (6). Accordingly, we found that phosphorylation of Xyl, preventing  $\beta$ 4GalT7

\* This work was supported by Agence Nationale de la Recherche, GAG Network Grant ANR-08-PCVI-0023-01, INSERM-University of Dundee International Collaboration Contract C21, Royal Society International Joint Grant (to M. W. H. C. and S. F.-G.), and the Région Lorraine. Part of this work was performed under the auspices of a European-associated laboratory between CNRS-UHP Nancy I and the University of Dundee.

§ The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Figs. S1–S2 and Tables S1–S4.

<sup>1</sup> Present address: Catherine Bui's, Institute of Cellular Medicine, Newcastle University, Medical School, Newcastle upon Tyne NE2 4HH, United Kingdom.

<sup>2</sup> To whom correspondence should be addressed. Tel.: 33-3-83-68-39-72; Fax: 33-3-83-68-39-59; E-mail: [sfg@medecine.uhp-nancy.fr](mailto:sfg@medecine.uhp-nancy.fr).

<sup>3</sup> The abbreviations used are: GAG, glycosaminoglycan;  $\beta$ 4GalT7,  $\beta$ 1,4-galactosyltransferase 7; MN-Xyl, 7-methoxy-2-naphthyl- $\beta$ -D-xylopyranoside; MN-Xyl-2P, 7-methoxy-2-naphthyl- $\beta$ -D-xylopyranoside-2-phosphate; 4-MUX, 4-methylumbelliferone- $\beta$ -D-xylopyranoside; 4-NP-Xyl, 4-nitrophenol- $\beta$ -D-xylopyranoside; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; PDB, Protein Data Bank.

activity, may act as a key regulatory event in the initiation of GAG synthesis (7). Furthermore, both in vertebrates and invertebrates, defects in this enzyme, which disrupt GAG synthesis, result in severe biological consequences. In humans, point mutations in the *B4GALT7* gene are associated with the progeroid form of Ehlers-Danlos syndrome, characterized by aged appearance, developmental delay, dwarfism, and several connective tissue disorders (8). In the invertebrate *Drosophila melanogaster*, a reverse genetic approach proved  $\beta$ 4GalT7 to be essential for viability (9). The structure of the fly protein has recently been solved by x-ray crystallography (10). In *Caenorhabditis elegans*, defective  $\beta$ 4GalT7 provokes serious morphological abnormalities during embryonic development (11).

Interestingly, the requirement for a Xyl polypeptide as a starting block to build a GAG chain can be bypassed *in vitro* and *in vivo* by xyloside derivatives consisting of Xyl linked to a hydrophobic aglycone (12). These xylosides serve as substrates for  $\beta$ 4GalT7, thus promoting initiation and polymerization of GAG chains. Such GAG precursors hold promise as anti-thrombotic drugs (13) and anti-amyloid agents (14). On the other hand, inhibitors of GAG synthesis have been proposed as chemotherapeutic agents (15–17). Thus, an understanding of the structure and function of  $\beta$ 4GalT7, a pivotal enzyme in GAG biosynthesis, would represent a major step toward the development of GAG-based therapeutics.

The  $\beta$ 4GalT7 enzyme belongs to the human  $\beta$ 1,4-galactosyltransferase ( $\beta$ 4GalT) family, containing seven members that are involved in the formation of Gal $\beta$ 1,4GlcNAc or Gal $\beta$ 1,4Glc linkages in different glycoconjugates and free saccharides (18). Analysis of their substrate specificity revealed major physiological functions and substantial impact in the pathogenesis of diseases.  $\beta$ 4GalT1 (lactose synthase), one of the first glycosyltransferases to be cloned and characterized, acts on a non-reducing terminal GlcNAc residue as acceptor substrate, whereas in the presence of  $\alpha$ -lactalbumin, the enzyme prefers Glc to GlcNAc (19).  $\beta$ 4GalT2 and  $\beta$ 4GalT3 catalyze the formation of Gal $\beta$ 1,4GlcNAc bonds in several glycoproteins and specific glycolipids.  $\beta$ 4GalT5 and  $\beta$ 4GalT6 are involved in the synthesis of lactosylceramide, which plays a major role in the regulation of cell proliferation, adhesion, migration, and angiogenesis. Recently, the importance of  $\beta$ 4GalT4 in the construction of keratan sulfate chains, essential in the maintenance of corneal matrix structure, has been discovered (20). Interestingly, the  $\beta$ 4GalT genes show sequence homology with those coding for  $\beta$ 1,4-*N*-acetylgalactosaminyltransferases ( $\beta$ 1,4GalNAcTs), among them several enzymes are involved in chondroitin sulfate GAG chains synthesis.

Several short conserved peptide motifs, possibly involved in the donor- and acceptor-binding sites, have been identified in the human enzymes and among  $\beta$ 4GalT orthologs in various animal species (21). Interestingly, these motifs are also conserved between members of the  $\beta$ 1,4-glycosyltransferase family, including seven  $\beta$ 1,4-galactosyltransferases ( $\beta$ 1,4GalTs), two  $\beta$ 1,4GalNAcTs, and four chondroitin synthases (19). Human  $\beta$ 4GalT7 presents 73% protein sequence similarity with its *Drosophila* ortholog (10), and the seven human  $\beta$ 4GalT proteins exhibit 25–55% sequence homology (19).

The  $\beta$ 4GalT7 is the most distant of the seven members of the  $\beta$ 4GalT family (22).

Unlike  $\beta$ 4GalT1, which has been subjected to extensive investigation for more than 20 years, our understanding of the mechanism of action of human  $\beta$ 4GalT7 is limited. To gain further insight into the structure and function of this enzyme, we conducted a phylogenetic analysis of  $\beta$ 1,4GalTs, which identified 48 related sequences containing 8 highly conserved peptide motifs. Using the structural coordinates of *Drosophila*  $\beta$ 4GalT7, we modeled the human  $\beta$ 4GalT7 structure, predicting that the peptide motifs <sup>163</sup>DVD<sup>165</sup> and <sup>221</sup>FWG-WRGEDDE<sup>230</sup> are likely to be important for the organization of the active site. Finally, using site-directed mutagenesis and a combination of *in vitro* and *ex vivo* activity assays, we analyzed the functional implications of these  $\beta$ 4GalT hallmarks. Our results suggest the involvement of a carboxylate residue in the reaction catalyzed by human  $\beta$ 4GalT7 and highlight the central function of Trp<sup>224</sup> in donor and acceptor substrate binding.

## EXPERIMENTAL PROCEDURES

**Chemicals and Reagents**—4-Methylumbelliferyl- $\beta$ -D-xylopyranoside (4-MUX), 4-nitrophenol- $\beta$ -D-xylopyranoside (4-NP-Xyl), UDP-Gal, UDP- $\alpha$ -D-glucose (UDP-Glc), UDP- $\alpha$ -D-N-acetylglucosamine (UDP-GlcNAc), UDP- $\alpha$ -D-N-acetylgalactosamine (UDP-GalNAc), UDP- $\alpha$ -D-glucuronic acid (UDP-GlcUA), UDP- $\alpha$ -D-xylose (UDP-Xyl), and anti-mouse IgG-alkaline phosphatase antibodies were provided from Sigma. UDP[<sup>14</sup>C]Gal, UDP[<sup>14</sup>C]GlcUA, UDP[<sup>14</sup>C]Xyl, UDP[<sup>14</sup>C]GalNAc, UDP[<sup>14</sup>C]GlcNAc, and Na<sub>2</sub>[S<sup>35</sup>]SO<sub>4</sub> were purchased from PerkinElmer Life Sciences, and UDP[<sup>14</sup>C]Glc was from Amersham Biosciences. 7-Methoxy-2-naphthyl- $\beta$ -D-xylopyranoside (MN-Xyl) and 7-methoxy-2-naphthyl- $\beta$ -D-xylopyranoside-2-phosphate (MN-Xyl-2P) were synthesized as described previously (7). Anti-Myc antibodies were from Invitrogen. Cell culture media were provided by Invitrogen. Restriction enzymes and T4 DNA ligase were from New England Biolabs. The eukaryotic expression vector pcDNA3.1(+) and competent One Shot<sup>®</sup> Top10 *Escherichia coli* cells were from Invitrogen. The bacterial expression vector pET-41a(+) and *E. coli* BL21(DE3) cells were from Novagen-EMD Biosciences, and the QuikChange site-directed mutagenesis kit was from Stratagene.

**In Silico Retrieval of  $\beta$ 4GalT7 Sequences**—Only eukaryote sequences were considered in this study. Twenty nine  $\beta$ 4GalT7 homologous sequences from different animal species were retrieved from the NCBI Data Bank by PSI-BLAST (24) with default parameters using B4GT7\_HUMAN as seeding sequence (25). Other  $\beta$ 4GalT7 homologous sequences were searched (26) through exploration of all genomic and expression sequence tags available and sequences from general (NCBI, DDBJ, and ENSEMBL) or specialized databases like the Joint Genome Institute (Department of Energy) for *Branchiostoma floridae*, using BLASTN and TBLASTN with default parameters (e-value cut off at 0.01). Within each animal species, the expression sequence tag contigs for each new  $\beta$ 4GalT7-like sequence were compiled with CAP3 (23). Splice site prediction analysis was achieved at the Berkeley *Drosophila* genome pro-



## $\beta$ 4GalT7 Key Functional Amino Acids

ject. The structure of each gene, in terms of exon/intron boundaries, was deduced from several nonexclusive strategies as follows: (i) comparison of the boundaries proposed by Genscan (MIT Server), (ii) comparison of EST sequences to genomic assemblages (scaffolds or contigs), and (iii) comparison of predicted boundaries to those present in known genes. All nucleotide sequences allowing generation of a complete protein were considered, and the assignment of the new sequences to the  $\beta$ 4GalT7 family was defined by the presence of the eight peptide motifs (Fig. 1). The 19 new sequences of  $\beta$ 4GalT7 reconstructed by these approaches were submitted to EMBL with the following accession numbers: AM231255 *Ciona savigny*; AM231256 *Drosophila ananassae*; AM231257 *Drosophila erecta*; AM231258 *Drosophila mojavensis*; AM231259 *Drosophila sechellia*; AM231260 *Drosophila simulans*; AM231261 *Drosophila yakuba*; AM231262 *Gallus gallus*; AM231263 *Oryzias latipes*; AM231264 *Pan troglodytes*; AM231265 *Schistosoma mansoni*; AM231266 *Tetraodon nigroviridis*; AM231267 *Takifugu rubripes*; FN568101 *Bos taurus*; FN568102 *Glossina morsitans*; FN568103 *Ovis aries*; FN568104 *Rana catesbeiana*; FN568105 *Sus scrofa*; and FN568106 *Squalus acanthias*.

**Phylogenetic Analysis and Sequence Alignment**—Multiple alignments were performed using ClustalW (27) and saved in Pir format. The Pir alignments were used for the selection of informative positions by G-Blocks (28). By this computerized method, 220 informative amino acid positions in eight G-blocks were selected for the phylogeny analyses. Phylogeny was carried out with Phylowin<sup>®</sup> software (29) using pairwise gap removal, BIONJ, Poisson correction, and 500 bootstrap replicates.

**Molecular Modeling**—A molecular model of the catalytic domain of human  $\beta$ 4GalT7 was initially built based on the structure of monomer A of bovine  $\beta$ 4GalT1 complexed with UDP-Gal (PDB code entry 1O0R) (31) using Modeler version 8v2 (30). The resulting model was then minimized using Amber simulation program version 8 (32) with 500 iterations of steepest descent and 150 of conjugate gradient. Docking of UDP-Gal-Mn<sup>2+</sup> molecules in the human  $\beta$ 4GalT7 structure was performed manually based on the position of the donor substrate in the structure of bovine  $\beta$ 4GalT1. Charges of the atoms of UDP-Gal were calculated using the GAUSSIAN94 package (33) and the HF/6-31 G\* basis set. Atom-centered charges were fitted with the antechamber of Amber 8 software package (32). Mn<sup>2+</sup> parameters available from the Amber website were used to build the model of  $\beta$ 4GalT7-UDP-Gal-Mn<sup>2+</sup> complex by energy minimization (500 iterations of steepest descent and 1500 of conjugate gradient). This model structure was next solvated using an octahedron box. The energy of the water molecules was minimized keeping those of the protein, UDP-Gal and Mn<sup>2+</sup>, fixed, with 1000 iterations of steepest descent and 1500 of conjugate gradients. Subsequently, the whole system was submitted to molecular dynamics, by the mean of an equilibration phase of 20 ps allowing the temperature rising from 0 to 300 K followed by a production phase of 600 ps at 300 K. A molecular model of the catalytic domain of human  $\beta$ 4GalT7 was also generated based on the structure of *Drosophila*  $\beta$ 4GalT7 linked to an N-terminal peptide of bovine  $\beta$ 4GalT1 (PDB code entry 3LW6) using Modeler version 8v2 (30). The

resulting model was then minimized using Amber simulation program version 11 (34), as described above. Protein sequence alignments used to generate the models are provided as [supplemental material F1](#).

**Expression Vector Construction**—The human  $\beta$ 4GalT7 sequence (GenBank<sup>®</sup> nucleotide sequence accession number NM\_007255) was cloned by PCR amplification from a placenta cDNA library (Clontech), as described previously (7). For the heterologous expression of the human  $\beta$ 4GalT7 in mammalian cell lines, the full-length cDNA sequence was modified by PCR to include a KpnI site and a Kozak consensus sequence at the 5' end (5'-CGGTACCACCATGTTCCCTCGCGGAGGAAAGCGGGCGC-3'), and a sequence encoding a Myc tag (EQKLIIEDL) and an XbaI site at the 3' end (5'-GTCTAGATCACAGATCCTCTTCAGAGATGAGTTTCTGCTCGCTGAATGTGCACCAGGGTTGGC-3'). The modified cDNA sequence was then subcloned into the KpnI-XbaI sites of the eukaryotic expression vector pcDNA3.1(+) to produce pcDNA- $\beta$ 4GalT7.

For bacterial expression, a truncated form of  $\beta$ 4GalT7 was expressed as a fusion protein with glutathione S-transferase (GST). The sequence lacking the codons of the first 60 N-terminal amino acids was amplified from the full-length cDNA using the oligonucleotide sense (5'-CCATGGTCAGGGGACAAAGGGCAG-3') and antisense primers (5'-GCGGCCGCTCACCTGAATGTGCACCA-3'), including NcoI and NotI sites, respectively, and subcloned into the corresponding sites of pET-41a(+) to produce the plasmid pET- $\beta$ 4GalT7. Final vector constructs were sequenced to verify that no errors had been introduced.

Mutations were constructed using the QuikChange site-directed mutagenesis kit, employing pcDNA- $\beta$ 4GalT7 or pET- $\beta$ 4GalT7 as template, and sense and antisense primers are listed in [supplemental Table 1](#). Mutants were systematically checked by sequencing.

The human decorin cDNA sequence (GenBank<sup>®</sup> accession number NM\_001920.3) was cloned by PCR amplification from a placenta cDNA library (Clontech) using sense (5'-ATGAAGGCCACTATCATCCTCCTCTTG-3') and antisense primers (5'-CTACTTATAGTTTCCGAGTTGAATGGC-3'). For the heterologous expression of the recombinant decorin in eukaryotic cells, the full-length cDNA sequence was modified by PCR to include an AflIII site, a Kozak consensus sequence at the 5' end, a sequence encoding a His<sub>5</sub> tag, and an XhoI site at the 3' end. This sequence was subcloned into pcDNA3.1(+) to produce pcDNA-decorinHis.

**Heterologous Expression of Wild-type and Mutant  $\beta$ 4GalT7 in Eukaryotic Cells**—HeLa cells and CHO pgsB-618 cells purchased from the American Type Culture Collection (ATCC) were cultured in Dulbecco's modified Eagle's medium (DMEM) and F-12/DMEM (1:1), respectively, supplemented with 10% (v/v) fetal bovine serum (FBS, Invitrogen), penicillin (100 units/ml)/streptomycin (100 mg/ml), and 1 mM glutamine.

Wild-type and mutant pcDNA- $\beta$ 4GalT7 plasmids were individually transfected into HeLa or CHO pgsB-618 cells grown to 80% confluency, using ExGen 500 reagent (Euromedex, Soufflweyersheim, France) according to the manufacturer's recommendations. The cells were harvested in phosphate-buffered

saline (PBS) 48 h after transfection and centrifuged for 10 min at  $5,000 \times g$  at 4 °C. The pellet was resuspended in 0.25 M sucrose, 5 mM HEPES buffer (pH 7.4) and sonicated for 5 s. Protein concentration was determined by the method of Bradford (35) prior to SDS-PAGE or activity analyses. Alternatively, transfected cells were transferred to labeling medium prior to GAG extraction, as described below. CHO pgsB-618 cells were transfected with human pcDNA-decorinHis using geneticin G-418 (Sigma) for selection of stable transfectants, as described previously (36), and then used for transient transfection with wild-type and mutant pcDNA- $\beta$ 4GalT7 plasmids.

**Expression and Purification of a Soluble Form of  $\beta$ 4GalT7**—To express wild-type and mutant GST- $\beta$ 4GalT7, *E. coli* BL21(DE3) cells were transformed with pET- $\beta$ 4GalT7 plasmids and selected by kanamycin resistance. Recombinants were inoculated in fresh Luria-Bertani (LB) medium and incubated at 37 °C on an orbital shaker (160 rpm) overnight. Each sample was then inoculated into 200 ml of fresh medium at a ratio of 1:100 and incubated in a shaking incubator at 37 °C until an absorbance (*A*) at 600 nm of 0.6–0.8 was reached. Isopropyl  $\beta$ -D-thiogalactopyranoside (Sigma) was then added to a final concentration of 1 mM, and cells were further cultured at 20 °C for 16 h. Bacterial cells were collected by centrifugation and lysed by sonication in buffer containing 50 mM sodium phosphate (pH 7.4), 1 mM phenylmethylsulfonyl fluoride, 1 mM EDTA, 5% (v/v) glycerol, and supplemented with Protease Inhibitor Mixture tablets (Roche Diagnostics). Cell lysates were centrifuged for 5 min at  $3,000 \times g$  to eliminate cell debris, and the resulting supernatant was centrifuged 15 min at  $12,000 \times g$ . The supernatant and pellet from this centrifugation were then subjected to 10% (w/v) SDS-PAGE, and gels were stained with Coomassie Brilliant Blue. The supernatant from the last centrifugation containing soluble proteins was used for affinity chromatography purification of GST- $\beta$ 4GalT7 using glutathione-Sepharose<sup>TM</sup> 4B packed columns (GE Healthcare), following recommendations of the manufacturer, and the amount of purified protein was quantified using Quant-iT<sup>TM</sup> assay kit and Qubit<sup>TM</sup> spectrofluorimeter.

**Western Blot Analysis**—Protein samples from HeLa or CHO pgsB-618 cells transfected with wild-type or mutant pcDNA- $\beta$ 4GalT7 plasmids were separated on SDS-10% (w/v) polyacrylamide gels and electrophoretically transferred to Immobilon-P<sup>®</sup> polyvinylidene difluoride membranes (Millipore). The blot was developed using monoclonal anti-Myc antibodies and alkaline phosphatase-conjugated anti-mouse IgG as secondary antibodies, as described previously (37). The amount of recombinant wild-type and mutant protein expressed in cells was evaluated from immunoblot analysis using a calibration curve established with 15–50 ng of recombinant GST-Myc protein (Pierce) run on the same gel, as described previously (7). Scanning densitometry was performed using Scion 1.63 Image software.

**Determination of Galactosyltransferase Activity**—The *in vitro* assay of  $\beta$ 4GalT7 activity was performed as described previously (7). Briefly, reactions were performed in 100 mM sodium cacodylate buffer (pH 7.0) containing 10 mM MnCl<sub>2</sub>, 1 mM UDP-Gal, 0.05  $\mu$ Ci of UDP-[U-<sup>14</sup>C]Gal, 5 mM 4-MUX, and 40  $\mu$ g of total cell protein or 0.2  $\mu$ g of purified protein. Incubations

were carried out at 37 °C for 1 h (or 30 min in the case of the purified protein) in a total volume of 50  $\mu$ l and terminated by the addition of 5  $\mu$ l of 6 N HCl. A reaction without 4-MUX was used as a control. The incubation mixture was centrifuged at  $10,000 \times g$  for 10 min at 4 °C, and the supernatant was analyzed by HPLC with a reverse phase C<sub>18</sub> column (Apollo, 4.6  $\times$  150 mm, 4  $\mu$ m, Grace Davison Discovery Sciences, Belgium) using a Waters equipment (Alliance Waters e2695) equipped with a Berthold FLOWStar radioactivity monitor. The mobile phase was composed of 14% (v/v) acetonitrile and 0.02% (v/v) trifluoroacetic acid in water. Radioactivity associated with the reaction product was quantified by scintillation counting using Ultima Flo<sup>TM</sup> scintillant mixture (PerkinElmer Life Sciences).

For the determination of glycosyltransferase activity toward various donor (UDP-Glc, UDP-Xyl, UDP-GlcUA, UDP-GalNAc, and UDP-GlcNAc) or acceptor substrates (4MU-Glc, 4-MU-GlcNAc, 4-MU-GalNAc, 4-MU-Gal, 4-NP-Xyl, MN-Xyl, and MN-Xyl-2P), incubations were performed in similar conditions, and reaction products were separated by thin layer chromatography as described previously (38, 39). Briefly, following protein precipitation, the supernatant was loaded onto plates (LK6DF silica gel, 250 mm, Whatman). The plates were developed with *n*-butanol, acetone, acetic acid, aqueous ammonia (28% v/v), and water (70:50:18:1.5:60, v/v). The radioactivity associated with the reaction product was visualized by autoradiography with X-Omat Kodak film (Sigma) for 3 days. The silica gel areas of the reaction product were scraped off and quantified by scintillation counting using Ultima Gold<sup>TM</sup> scintillant mixture on an LKB spectrophotometer (Packard Instrument Co.).

**Determination of Kinetic Parameters**—Kinetic constants  $K_m$  and  $V_{max}$  or  $k_{cat}$  were determined toward 4-MUX as acceptor substrate by incubating total HeLa cell proteins (40  $\mu$ g) or purified GST- $\beta$ 4GalT7 (0.2  $\mu$ g) with increasing concentrations of 4-MUX (0–10 mM), in the presence of a fixed concentration of UDP-Gal (1 mM). The kinetic constants  $K_m$  and  $V_{max}$  or  $k_{cat}$  toward UDP-Gal donor substrate were determined by incubating cell or purified proteins with increasing concentrations of UDP-Gal (0–10 mM), in the presence of a fixed concentration of 4-MUX (5 mM). Kinetic parameters were determined by non-linear least squares regression analysis of the data fitted to Michaelis-Menten rate equation using the curve-fitter program of Sigmaplot 9.0 (Erkraft, Germany).

**Analysis of GAG Synthesis by Na<sub>2</sub><sup>35</sup>S]SO<sub>4</sub> Incorporation**—To analyze GAG chain synthesis initiated from exogenously added xyloside, CHO pgsB-618 cells transfected with wild-type, mutant pcDNA- $\beta$ 4GalT7, or empty vector were incubated for 12 h in Fischer's medium containing 10  $\mu$ Ci/ml Na<sub>2</sub>[<sup>35</sup>S]SO<sub>4</sub> in the presence or absence of 4-MUX (5  $\mu$ M). 1 ml of medium was applied to a G-50 column to separate nonincorporated radiolabeled sulfate, and radiolabeled GAGs present in the eluates were quantified by scintillation counting or resolved by SDS-PAGE using Criterion Precast gels (4–12% BisTris, Bio-Rad) and visualized by autoradiography on Amersham Biosciences Hyperfilm MP. For analysis of GAG chain synthesis on decorin core protein, CHO pgsB-618 cells stably expressing human recombinant decorin were transiently transfected with pcDNA- $\beta$ 4GalT7 or mutant plasmids and incubated for 12 h in Fischer's

## $\beta$ 4GalT7 Key Functional Amino Acids

medium containing 10  $\mu$ Ci/ml [ $^{35}$ S]SO<sub>4</sub>, as described above. The recombinant His-tagged decorin was purified from 1 ml of culture medium using the MagneHis Protein Purification System according to the recommendations of the manufacturer (Promega). The radiolabeled purified decorin was analyzed by SDS-PAGE using Criterion Precast gels (4–12% BisTris). After incubation in H<sub>2</sub>O/methanol/acetic acid (5:1:1) solution, then in Amplify reagent (Amersham Biosciences), the gel was dried and subjected to autoradiography.

For analysis of total GAG synthesis, CHO pgsB-618 cells, 48 h after transfection, were incubated for 8 h in Fischer's medium supplemented with 10% (v/v) fetal bovine serum without sulfate containing 10  $\mu$ Ci/ml Na<sub>2</sub>[ $^{35}$ S]SO<sub>4</sub>, and proteoglycan extraction was performed as described previously (37). Briefly, 1 ml of culture medium was collected and treated with papain (1 mg/ml) for 3 h at 60 °C. The digestion was stopped by heating at 100 °C for 10 min, and the protein was precipitated by centrifugation at 8,000  $\times$  g for 15 min. Three volumes of 5% potassium acetate (w/v in absolute ethanol) were added to 1 volume of supernatant, and after mixing, the suspension was stored overnight at 4 °C and then centrifuged for 30 min at 8,000  $\times$  g. The precipitate was resuspended in 500  $\mu$ l of 0.2 M NaCl and centrifuged for 30 min at 8,000  $\times$  g. The supernatant was incubated for 2–3 h at 37 °C with 100  $\mu$ l of cetylpyridinium chloride (5%, w/v) in the presence of 100  $\mu$ g of shark chondroitin sulfate to facilitate GAG precipitation. The precipitate was collected by centrifugation for 30 min at 8,000  $\times$  g and washed three times in the cetylpyridinium chloride (5%, w/v), 0.2 M NaCl solution and collected by centrifugation for 10 min at 8,000  $\times$  g. The precipitate was dissolved in 200  $\mu$ l of 2.5 M NaCl, 5 volumes of ethanol, and centrifuged for 30 min at 10,000  $\times$  g, and the final precipitate was dissolved in 200  $\mu$ l of Tris-HCl (0.2 M, pH 8.0). The radioactivity associated to GAGs was determined by scintillation counting.

**Statistical Analyses**—All data are presented as means  $\pm$  S.D. of three determinations of at least three sets of experiments. For comparison among multiple groups, one-way analysis of variance followed by Bonferroni post hoc correction test was used. Statistical analyses were performed using GraphPad Prism software (version 4.0).

## RESULTS

To gain insight into potential functional domains of  $\beta$ 4GalT7 proteins, we performed a multiple alignment of 54 selected sequences, including the 48 animal  $\beta$ 4GalT7 sequences plus six human  $\beta$ 4GalTs (B4GT1 to B4GT6), taken as an outgroup (Fig. 1). The sequence comparison highlighted a high degree of similarity, with eight conserved peptide motifs distributed all along the alignment and separated by intermotif amino acid stretches of similar length (Fig. 1). The size of the N-terminal domain ahead of the first conserved motif was longer in the  $\beta$ 4GalT1–6 outgroup compared with the 48  $\beta$ 4GalT7 proteins. Furthermore, 17 positions were identical only in the  $\beta$ 4GalT7 sequences (Fig. 1, indicated by #) and less well conserved in the outgroup, whereas 32 amino acid positions (Fig. 1, indicated by \*) were identical, both in the outgroup and in the  $\beta$ 4GalT7 family.

Highly conserved  $\beta$ 4GalT amino acids were present in the first seven motifs (Fig. 1) but were particularly abundant in “motif 3” that contains a typical DXD signature sequence ( $^{163}$ DVD $^{165}$  in human  $\beta$ 4GalT7) and in “motif 5” containing a second predicted functional motif  $^{221}$ FWGWGREDDE $^{230}$  (numbered in human  $\beta$ 4GalT7). In the crystal structure of bovine  $\beta$ 4GalT1 and *Drosophila*  $\beta$ 4GalT7, these two polypeptides have been found to be located in the vicinity of the donor substrate UDP-Gal that is common to the  $\beta$ 4GalT enzymes (19). Amino acids specific to  $\beta$ 4GalT7 sequences are found in “motif 6” that contains a new Mn<sup>2+</sup>-binding motif, HXH (10). “Motif 8” exhibits 5 out of the 9 positions highly conserved in the  $\beta$ 4GalT7 family, including two Cys residues forming a disulfide bond, as shown in *Drosophila* (Cys<sup>300</sup>–Cys<sup>308</sup>, corresponding to Cys<sup>316</sup>–Cys<sup>324</sup> in human sequence). In *Drosophila* x-ray structure, this motif following an extended  $\beta$ -sheet has been suggested to compensate for the absence of the N-terminal region present in bovine  $\beta$ 4GalT1, which plays an important role in the catalytic activity. Further functional analysis of human  $\beta$ 4GalT7 is required to test this assumption.

**Phylogeny of  $\beta$ 4GalT Enzymes**—To further define the conserved features of the  $\beta$ 4GalT7 family, we performed a phylogenetic analysis based on the data available from the various genome sequencing programs. This approach led to important insight into the organization of this large glycosyltransferase family. Fig. 2 shows the phylogenetic tree built with the 54 sequences identified. The  $\beta$ 4GalT sequences could be clearly separated into two main groups, one containing the 48 members of the  $\beta$ 4GalT7 family and the other containing the six human  $\beta$ 4GalT1–6 enzymes, which were considered as an outgroup. The vertebrate  $\beta$ 4GalT7 branch is robust with high bootstrap values and constitutes a very homogeneous group in all vertebrate species from fish to man. As expected, the sequences of invertebrate  $\beta$ 4GalT7 have longer branch lengths with low bootstrap values. Thus, within the invertebrate branch, some species do not appear in the expected order, as seen in the case of *H. magnipapillata* and *B. floridae* and urochordata and arthropods branches that are inverted. However, the topology of the phylogenetic tree shows a good fit with the expected evolution order. The  $\beta$ 4GalT7 protein is probably present in all metazoa, because it was found in cnidaria, trematods, nematods, arthropods, urochordata, and cephalochordata. The enzyme activity of these proteins (EC 2.4.1.133) has only been experimentally tested in some vertebrates, and in *D. melanogaster* and *C. elegans*, but the presence of the eight highly conserved peptide motifs distributed at well preserved distances all along the protein suggests that the other proteins may also have similar enzyme activities.

**Structural Model of Human  $\beta$ 4GalT7 Active Site**—Although the structure for human  $\beta$ 4GalT7 has not yet been determined, protein sequence alignment identifies several domains of homology with bovine  $\beta$ 4GalT1 (supplemental Fig. 1), for which a number of x-ray structures has been established (31, 40). Thus, we initially created a molecular model of human  $\beta$ 4GalT7 using bovine  $\beta$ 4GalT1-UDP-Gal-Mn<sup>2+</sup> (PDB entry code 1O0R) as template. The modeled active site revealed the presence of the  $^{163}$ DVD $^{165}$  and the  $^{221}$ FWGWGREDDE $^{230}$  functional domains (numbered in human  $\beta$ 4GalT7) in the

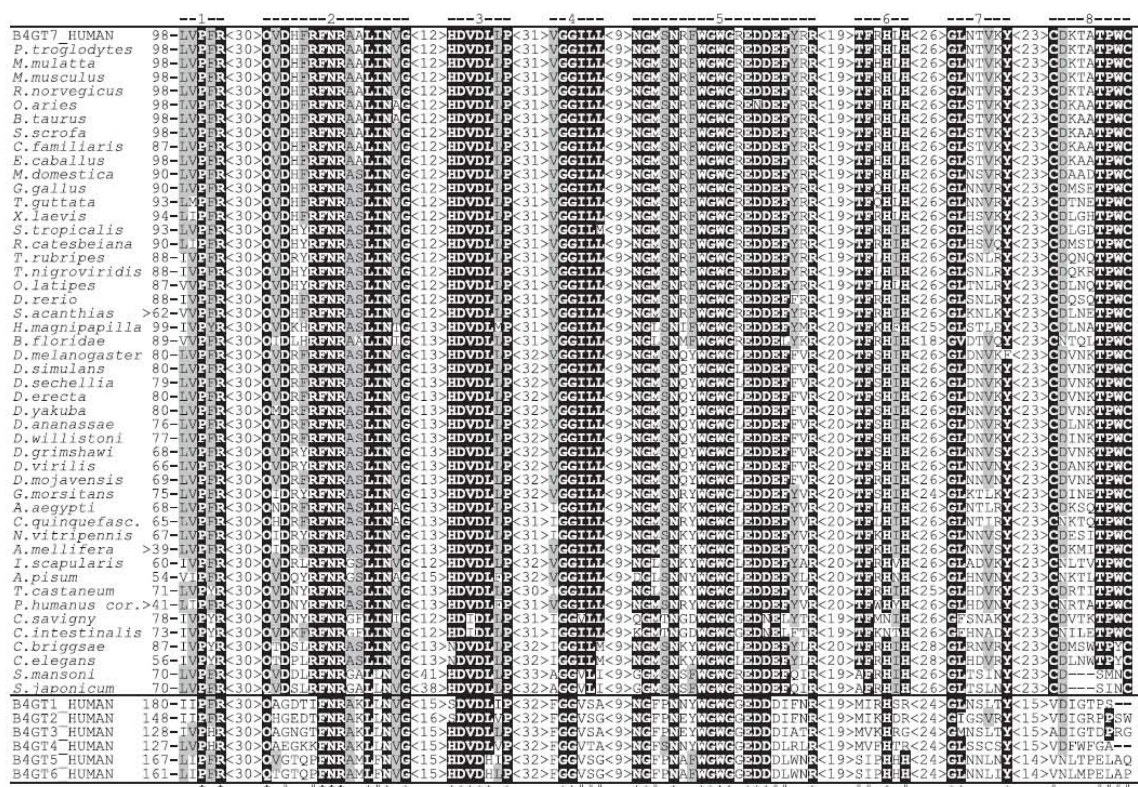


FIGURE 1. Conserved peptide motifs found in the multiple alignment of the 48 animal β4GalT7 proteins and the human β4GalT1-6 sequences. Protein and DNA alignments were performed by ClustalW and saved in Pir format. The Pir alignment was used for the selection of 220 informative positions by G-blocks. The eight highly conserved peptide motifs were visually defined within the ClustalW alignment of the 220 selected positions. Positions identical at more than 90% of the sequences are shown as **bold white letters on a solid background**. Positions identical at more than 50% are shown as **black letters on a gray background**. In the *bottom line*, amino acid positions identical in >90% of all the β4GalT sequences are represented by an *asterisk*, and positions also identical at >90%, but only in the β4GalT7 family are represented by a *number sign*.

vicinity of the donor substrate, UDP-Gal (supplemental Fig. 2). In supplemental Fig. 2., panel A shows predicted interactions between the <sup>163</sup>DVD<sup>165</sup> motif and UDP-Gal-Mn<sup>2+</sup>, including a hydrogen bond between Asp<sup>163</sup> and O3 Gal molecule, and a coordination bond between the side chain carboxyl group of Asp<sup>165</sup> and the Mn<sup>2+</sup> divalent cations. Hydrogen bond formation between Ne1 of Trp<sup>224</sup> belonging to the <sup>221</sup>FWG-WGREDDE<sup>230</sup> conserved motif and a β-phosphate oxygen atom of UDP is highlighted (panel B). We also identified several amino acid residues of human β4GalT7, other than those belonging to the functional peptide regions of interest (listed in supplemental Table 3), which form possible contacts with UDP-Gal-Mn<sup>2+</sup>, including His<sup>257</sup> in coordination bonding with the divalent cations (supplemental Fig. 2).

Because the crystal of *Drosophila* β4GalT7, which shares 73% protein sequence homology with human β4GalT7 (supplemental Fig. 1), was recently solved (10), we next created a molecular model of the human protein based on this structure. For this purpose, we used the structural coordinates of the catalytic domain of *Drosophila* β4GalT7 in complex with UDP (PDB entry code 3LW6) and proceeded to the docking of UDP-Gal into the active site of the modeled human protein. As

shown Fig. 3, the <sup>163</sup>DVD<sup>165</sup> and the <sup>221</sup>FWG-WGREDDE<sup>230</sup> functional motifs were found in a similar position with regard to the donor substrate UDP-Gal, when compared with the model of human β4GalT7 based on bovine β4GalT1. Asp<sup>163</sup> was in hydrogen bond distance to O3 of the Gal molecule, and a coordination bonding was predicted between the side chain carboxyl group of Asp<sup>165</sup> and the Mn<sup>2+</sup> divalent cation (Fig. 3A). Moreover, in addition to the coordination bond between Ne2 of His<sup>257</sup> and Mn<sup>2+</sup>, a coordination bond between Nδ1 of His<sup>259</sup> and the divalent cation was identified. Indeed, <sup>257</sup>VHV<sup>259</sup> corresponds to a new metal-binding motif (10), which is strictly conserved among β4GalT7 proteins, as shown by multiple sequence alignment (see Fig. 1). Furthermore, modeling the UDP-Gal-binding site of human β4GalT7 containing the <sup>221</sup>FWG-WGREDDE<sup>230</sup> motif predicted hydrogen bond formation between the Ne1 of Trp<sup>224</sup> and β-phosphate oxygen atom of UDP of the donor substrate molecule and between Glu<sup>227</sup> and the O4 of the Gal moiety (Fig. 3B). Additional bonds between human β4GalT7 and UDP-Gal were also identified (listed in supplemental Table 4). As found in *Drosophila*, Tyr<sup>194</sup> (corresponding to Tyr<sup>277</sup>) and Arg<sup>250</sup> (corresponding to Arg<sup>266</sup>) were predicted to form a hydrogen bond with an oxy-

Downloaded from www.jbc.org at CNRS, on February 17, 2011

## $\beta$ 4GalT7 Key Functional Amino Acids

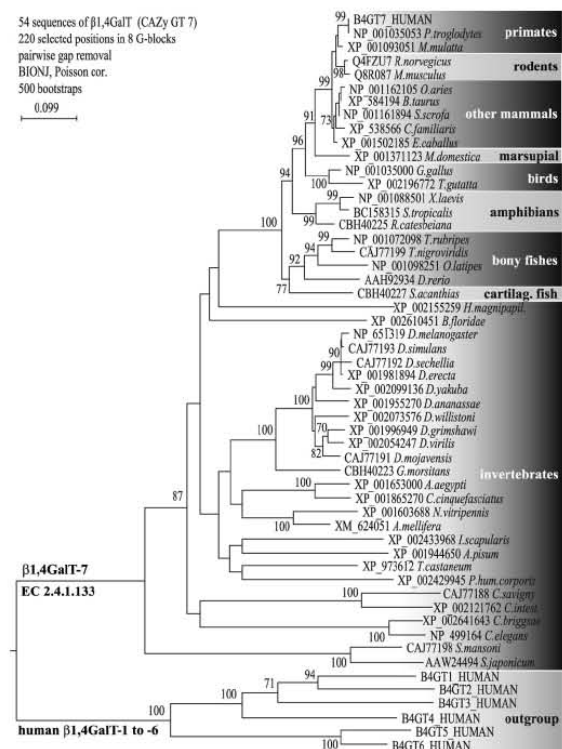


FIGURE 2. Neighbor-joining phylogenetic tree of the 54  $\beta$ 4GalT7 sequences. The robustness of the tree branches was tested by bootstrap calculations from 500 sets of data. Bootstrap values higher than 70% are considered significant and are reported at the left of each node. The scale bar represents the number of substitutions per site for a unit branch length.

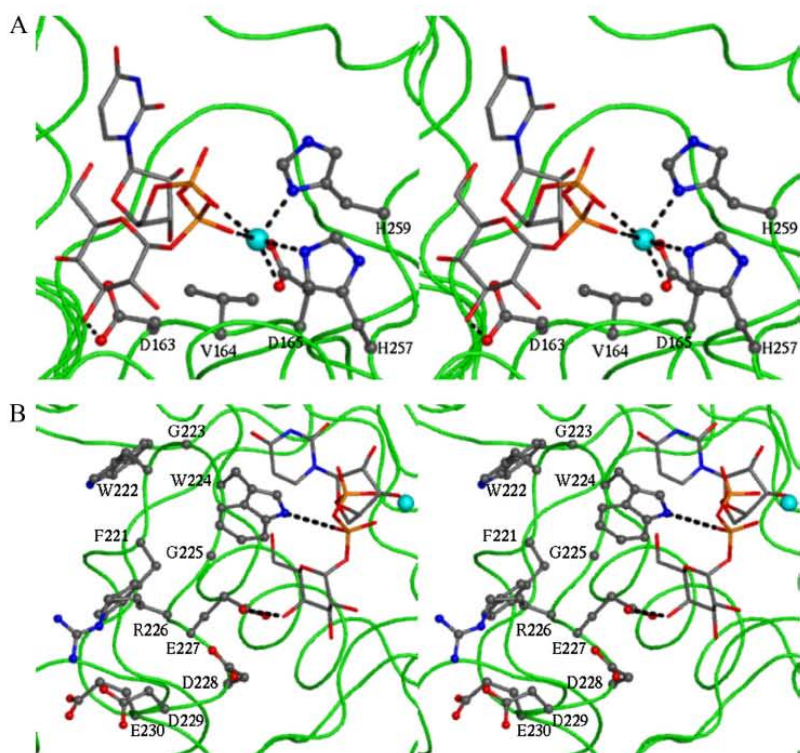
gen atom of the  $\beta$ -phosphate oxygen atom of UDP. In addition, we identified potential interactions between Tyr<sup>199</sup> and O2 and O3 of Gal that are not predicted in the case of its counterpart (Phe<sup>182</sup>) in *Drosophila*.

**In Vitro Characterization of  $\beta$ 4GalT7 Mutants Expressed in Eukaryotic Cells**—To address the functional contribution of two main active site conserved regions of  $\beta$ 4GalT7 located in motif 3 and motif 5, we conducted a systematic site-directed mutagenesis approach. Individual residues of these two peptide regions were subjected to conservative and nonconservative substitutions to generate a library of 21 mutants. In a first set of experiments, wild-type and mutant  $\beta$ 4GalT7 enzymes were expressed in HeLa cells and tested for *in vitro* galactosyltransferase activity using 4-MUX as acceptor substrate. The recombinant enzymes were expressed as Myc-tagged proteins, to quantify the level of expression using recombinant GST-Myc protein as standard, as described under “Experimental Procedures”.

In human  $\beta$ 4GalT7, the canonical DXD motif (<sup>163</sup>DVD<sup>165</sup>) corresponds to a signature characteristic for many glycosyltransferases and other enzymes that use nucleotide substrates and require a divalent metal ion for activity. To gain insight into the functional importance of the <sup>163</sup>DVD<sup>165</sup> motif in this enzyme, mutants were engineered in which each residue was

replaced by Ala and by Glu in the case of Asp<sup>163</sup> and Asp<sup>165</sup>. The consequences of these mutations were evaluated on the *in vitro* galactosyltransferase activity of  $\beta$ 4GalT7, and the apparent  $K_m$  value of wild-type and mutant enzymes was determined toward UDP-Gal and 4-MUX. Immunoblot analysis of HeLa cells expressing the mutants showed that they were produced at a similar level to that of the wild-type protein (Fig. 4A). The activity of the recombinant enzymes was normalized to the amount of expressed protein, and values are presented in Fig. 4B. The full-length form of human  $\beta$ 4GalT7 exhibited high activity up to 4  $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}$  protein, although it was undetectable in mock-transfected cells (Fig. 4B). On the other hand, replacement of each aspartate of the <sup>163</sup>DVD<sup>165</sup> motif to Ala completely abolished the *in vitro* enzyme activity, emphasizing the prevalent function of these residues. Interestingly, although substitution of Asp<sup>163</sup> to Glu did not overcome the loss of activity observed for the Ala mutant, D165E exhibited up to 47% activity compared with wild-type enzyme, allowing the determination of kinetic parameters. The apparent  $K_m$  values of the D165E mutant toward UDP-Gal and 4-MUX were about 3.8- and 5.5-fold, respectively, higher than the wild-type enzyme (Table 1). These results indicate that the carboxyl side chain of Glu could replace the negative charge of Asp in position 165, to some extent, although this modification reduced the affinity of the enzyme toward both donor and acceptor substrates. Furthermore, substitution of Val<sup>164</sup> by an Ala residue did not produce significant changes either in activity or in  $K_m$  values. These results are in agreement with the structure of *Drosophila*  $\beta$ 4GalT7 and the modeled human  $\beta$ 4GalT7 active sites, indicating that the Val residue does not directly or indirectly interact with UDP-Gal, in contrast to the neighboring Asp of the <sup>163</sup>DVD<sup>165</sup> motif (Fig. 3A).

In the next series of experiments, we explored the function of the <sup>221</sup>FWGWGREDD<sup>230</sup> region located in motif 5, which exhibits a high degree of conservation among the human  $\beta$ 4GalT7 family members, as well as among  $\beta$ 4GalT7-related sequences sampled from distant animal species (see Fig. 1). A series of mutants exhibiting conservative and nonconservative substitutions at each position of the peptide motif were engineered, and immunoblot analysis indicated that the mutants were produced at similar level to that of the wild-type protein in HeLa cells (Fig. 4A). Replacement of the first residue of the motif, Phe<sup>221</sup>, by Ala did not affect the *in vitro* galactosyltransferase activity of  $\beta$ 4GalT7 (Fig. 4C). However, although this mutant presented a similar  $K_m$  value to that of the wild-type enzyme toward UDP-Gal, the affinity toward 4-MUX was strongly reduced, as indicated by a 13-fold increase in  $K_m$  (Table 1). This result suggests that introduction of Ala at position 221 may affect binding of 4-MUX and/or the organization of the acceptor site but not of the donor substrate-binding site. On the other hand, the conservative substitution Phe<sup>221</sup> to Tyr led to a mutant that exhibited a similar activity to that of the wild-type, with  $K_m$  values toward acceptor and donor substrates in the same range as  $\beta$ 4GalT7 (Table 1), indicating that Phe and Tyr residues are interchangeable at this position. Concordantly, either Phe or Tyr was found at a position equivalent to 221 in related  $\beta$ 4GalT7 sequences (see Fig. 1). The neighboring Trp<sup>222</sup> residue is invariant among all  $\beta$ 4GalT7 sequences



**FIGURE 3.** Stereo view of the molecular model of the human  $\beta$ 4GalT7 structure in complex with the donor substrate. The UDP-Gal binding domain was modeled by energy minimization and molecular dynamics calculations based on the structure of *Drosophila*  $\beta$ 4GalT7 complexed with UDP (PDB code 3LW6). Representation of the interactions of  $^{163}\text{DVD}^{165}$  and  $^{257}\text{HLH}^{259}$  motifs (A) and of  $^{221}\text{FWGWGREDD}^{230}$  (B) motifs with UDP-Gal molecule is shown. The  $\alpha$ -carbon trace of the protein is shown in green; UDP-Gal is presented in colored "sticks," and  $\text{Mn}^{2+}$  divalent cation is colored in cyan. The residues of the  $^{163}\text{DVD}^{165}$ ,  $^{257}\text{HLH}^{259}$ , and  $^{221}\text{FWGWGREDD}^{230}$  motifs are shown as "ball and sticks." For simplicity, hydrogen atoms and atoms of the protein main chain other than  $\alpha$ -carbons are not shown. Dashed lines indicate the predicted interactions of UDP-Gal- $\text{Mn}^{2+}$  with the side chain of Asp $^{163}$ , Asp $^{165}$ , His $^{257}$ , and His $^{259}$  residues in A and between the  $\beta$ -phosphate group of UDP-Gal and Trp $^{224}$  and between galactose O4 of UDP-Gal and Asp $^{227}$  in B. The atom coordinates are those of the average structure obtained during the production phase of molecular dynamics calculation, as indicated under "Experimental Procedures." Figures are drawn with PyMOL molecular visualization system.

analyzed in this study. As shown in Fig. 4C, a sharp drop in activity was observed upon replacement of this residue by Ala, whereas the conservative mutation to Phe did not affect the enzyme activity. Furthermore, the W222F mutation had no significant effect on the  $K_m$  value toward donor or acceptor substrate. Concordantly, the structural model indicates that the Trp $^{222}$  residue points out of the catalytic center of  $\beta$ 4GalT7. Substitution of Gly residues with Ala at positions 223 and 225 produced different effects on the *in vitro* galactosyltransferase activity of  $\beta$ 4GalT7. The G223A mutant was about 40% less active than the wild-type enzyme (Fig. 4C), and it exhibited over 2-fold increase in apparent  $K_m$  values toward 4-MUX (Table 1), whereas the G225A mutant did not display any *in vitro* activity.

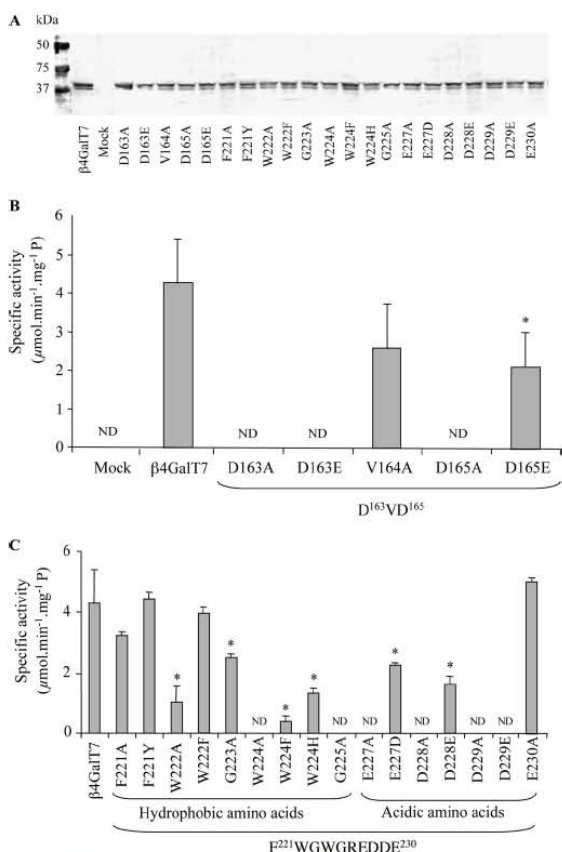
Examination of *Drosophila* and modeled human  $\beta$ 4GalT7 structures highlighted the role of Trp $^{207}$  and Trp $^{224}$ , respectively, as a key active site residue, predicted to interact with the  $\beta$ -phosphate group of the donor substrate (Fig. 3B). Thus, we generated a series of substitutions at this position and analyzed the *in vitro* galactosyltransferase activity of the different recom-

binant human  $\beta$ 4GalT7 proteins. Replacement of Trp $^{224}$  by Ala led to a completely inactive protein, suggesting that the presence of a Trp is indeed critical at this position (Fig. 4C). Trp $^{224}$  was also converted to Phe or His to investigate the influence of the phenyl or imidazole group of these amino acids, in place of the indole side chain of Trp. Replacement of Trp $^{224}$  by Phe was able to recover about 9% of the wild-type enzyme activity, and W224H activity was about 33% that of  $\beta$ 4GalT7. The very low activity observed with the W224F mutant did not allow determination of kinetic parameters. Characterization of the W224H mutant revealed a 4.4-fold decrease in apparent  $K_m$  values toward UDP-Gal together with a large increase in  $K_m$  values toward the acceptor substrate 4-MUX (Table 1), suggesting that Trp $^{224}$  may be involved not only in donor but also acceptor substrate interactions. Altogether, mutations of the N-terminal hydrophobic amino acids of the  $^{221}\text{FWGWGREDD}^{230}$  motif revealed a moderate effect of Phe $^{221}$ , Trp $^{222}$ , and Gly $^{223}$  on enzyme properties, whereas Trp $^{224}$  appears to play an essential role in the organization and function of the catalytic center.

The  $^{221}\text{FWGWGREDD}^{230}$  motif contains a stretch of four conserved acidic amino acid residues ( $^{227}\text{EDDE}^{230}$ ). We also investigated

the role of this motif by site-directed mutagenesis and kinetic analysis of the resulting recombinant enzymes. Substitution of Glu with Ala at position 227 produced an inactive enzyme, supporting the functional importance of this residue, whereas replacement by an Asp residue resulted in retention of ~50% galactosyltransferase activity compared with the wild-type enzyme (Fig. 4C). Furthermore, evaluation of the kinetic properties of E227D showed that this mutant displayed similar apparent  $K_m$  values toward both donor and acceptor substrates, compared with the wild-type enzyme (Table 1). Similarly, replacement of Asp $^{228}$  with Ala abolished activity, whereas the D228E mutant expressed in HeLa cells retained about 40% activity of the wild-type enzyme. This mutant exhibited a  $K_m$  value toward UDP-Gal similar to the wild-type, whereas an 8.0-fold increase in apparent  $K_m$  was observed toward the acceptor substrate (Table 1). In the x-ray structure of bovine  $\beta$ 4GalT1 and of *Drosophila*  $\beta$ 4GalT7, Asp $^{318}$  and Asp $^{211}$  (equivalent to Asp $^{228}$  in human  $\beta$ 4GalT7) are located in the vicinity of the O4 atom of the xylose acceptor substrate and are suggested to act as

## $\beta$ 4GalT7 Key Functional Amino Acids



**FIGURE 4. Expression and activity of recombinant wild-type and mutated  $\beta$ 4GalT7 expressed in HeLa cells.** A, SDS-PAGE and immunoblot analysis of wild-type and mutant  $\beta$ 4GalT7 recombinant proteins probed with anti-Myc antibodies. 15  $\mu$ g of total protein were loaded in each lane. B and C, galactosyltransferase activity of wild-type and mutant  $\beta$ 4GalT7 recombinant enzymes. The reaction was performed using 40  $\mu$ g of total protein in the presence of 1 mM UDP-Gal and 0.05  $\mu$ Ci of UDP[U- $^{14}$ C]Gal as donor substrate and 5 mM 4-MUX as acceptor substrate, as indicated under "Experimental Procedures." The amount of expressed  $\beta$ 4GalT7 protein was quantified using a calibration curve established with GST-Myc protein immunostained under the same conditions. Values are expressed as  $\mu$ mol per min per mg of  $\beta$ 4GalT7, and results are the mean  $\pm$  S.D. of three independent experiments (\*,  $p < 0.05$ , versus transfected cells with wild-type  $\beta$ 4GalT7 cDNA).

the catalytic base (41). Accordingly, the bulkiness of Glu introduced in place of Asp at position 228 in  $\beta$ 4GalT7 may produce steric hindrance, contributing to the observed increase in  $K_m$  value toward 4-MUX. Replacement of the next Asp residue located at position 229 to either Ala or Glu led to completely inactive enzymes *in vitro*, supporting the critical importance of this residue, although the lack of activity did not allow kinetic studies to be carried out (Fig. 4C). However, because this residue was not predicted to be located in the vicinity of the donor substrate, it is not likely to be implicated in direct interactions with UDP-Gal. By contrast, mutation of Glu<sup>230</sup> to Ala did not impair enzyme activity and had little impact on  $K_m$  values (Table 1), indicating a less important function for the last residue of the <sup>221</sup>FWGWGREDD<sup>230</sup> peptide region. Although it

**TABLE 1**

Apparent dissociation constants of wild-type and mutant  $\beta$ 4GalT7 expressed in HeLa cells toward UDP-Gal and 4-MUX

HeLa cells were transiently transfected with wild-type and mutant pcDNA- $\beta$ 4GalT7 plasmids, and enzymatic assays were performed on cell lysates using 40  $\mu$ g of total protein incubated with increasing concentrations of 4-MUX (0–10 mM) at a constant concentration of UDP-Gal (1 mM) or with increasing concentrations of UDP-Gal (0–10 mM) at a constant concentration of 4-MUX (5 mM). Kinetic parameters were determined by nonlinear least squares regression analysis of the data fitted to the Michaelis-Menten rate equation using Sigmaplot 9.0. The results are the means  $\pm$  S.D. of three independent assays. ND indicates no kinetic parameters could be determined using excess acceptor or donor substrate.

Enzyme	$K_m$	
	UDP-Gal	4-MUX
Wild type	0.31 $\pm$ 0.07	0.32 $\pm$ 0.08
V164A	0.48 $\pm$ 0.10	0.53 $\pm$ 0.14
D165E	1.17 $\pm$ 0.13*	1.76 $\pm$ 0.25*
F221A	0.36 $\pm$ 0.04	4.22 $\pm$ 0.40*
F221Y	0.33 $\pm$ 0.03	0.38 $\pm$ 0.05
W222A	0.21 $\pm$ 0.01	ND
W222F	0.45 $\pm$ 0.05	0.16 $\pm$ 0.05
G223A	0.45 $\pm$ 0.06	0.73 $\pm$ 0.10*
W224H	0.07 $\pm$ 0.01*	>4
E227D	0.47 $\pm$ 0.09	0.52 $\pm$ 0.17
D228E	0.47 $\pm$ 0.08	2.57 $\pm$ 0.11*
D230A	0.43 $\pm$ 0.03	0.57 $\pm$ 0.08

\*  $p < 0.05$ .

was found to be conserved among all  $\beta$ 4GalT7 family members, this Glu residue was replaced by Asp in the  $\beta$ 4GalT1–6 enzymes.

**Expression, Purification, and Kinetic Properties of  $\beta$ 4GalT7 Mutants Expressed in *E. coli***—In the next series of experiments, we developed an *E. coli* expression system devoted to the characterization of purified  $\beta$ 4GalT7 mutants. Because the conserved Trp<sup>224</sup>, Glu<sup>227</sup>, and Asp<sup>228</sup> were thought to be important for substrate binding and/or catalysis, based on structural predictions and site-directed mutagenesis studies, we conducted detailed kinetic analyses of point mutants at these positions. The wild-type  $\beta$ 4GalT7 and selected mutants were produced as truncated fusion proteins lacking the 60 N-terminal amino acids (including the transmembrane domain and part of the stem region) linked to GST. The recombinant enzymes were all purified by affinity chromatography using the same conditions, and expression levels (about 2.5 mg of protein/liter of culture) of the mutants were similar to those of the wild-type (Fig. 5). Kinetic analyses indicated that the wild-type  $\beta$ 4GalT7 was highly active with a  $k_{cat}$  up to 1.5 s<sup>-1</sup> and an efficiency of about 5 mM<sup>-1</sup> s<sup>-1</sup> (Table 2). Furthermore, the  $\beta$ 4GalT7 enzyme purified from *E. coli* cells displayed a  $K_m$  value toward either donor or acceptor substrate in the same range as the membrane-bound enzyme expressed in HeLa cells (see Table 1). This indicates that expression of the catalytic domain of  $\beta$ 4GalT7 with the GST fusion partner did not substantially affect substrate binding and catalytic properties of this enzyme.

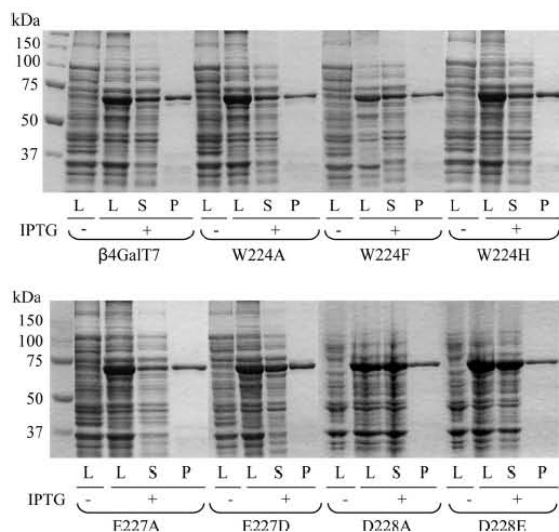
Structural data combined with our systematic site-directed mutagenesis study of the <sup>221</sup>FWGWGREDD<sup>230</sup> motif suggested a key role of Trp<sup>224</sup> in substrate binding, leading us to determine the kinetic properties of purified  $\beta$ 4GalT7 enzyme modified at this position and to evaluate possible changes in specificity upon mutation. The W224A and W224F mutants produced and purified from *E. coli* cells were inactive, confirming the functional importance of the conserved Trp<sup>224</sup> residue. It is noteworthy that the purified W224H mutant displayed a

dramatic change in kinetic behavior compared with the wild-type  $\beta$ 4GalT7 enzyme, characterized by a large increase in affinity toward the donor substrate UDP-Gal (15 times lower  $K_m$  value) as indicated in Table 2. Because the  $k_{cat}$  value for this mutant was reduced, the efficiency of the wild-type enzyme and of the W224H mutant were in the same range (about  $5 \text{ mM}^{-1} \text{ s}^{-1}$ ). On the other hand, the affinity of the purified W224H mutant toward 4-MUX was reduced to such an extent that kinetic parameters toward the acceptor substrate could not be determined.

We next examined whether substitution of Trp at position 224 with His may affect the specificity of  $\beta$ 4GalT7 toward either donor or acceptor substrates. To that aim, the activity of

purified enzymes was first determined toward a series of UDP-sugar analogs, *i.e.* UDP-Gal, UDP-Glc, UDP-Xyl, UDP-GlcUA, UDP-GalNAc, and UDP-GlcNAc. The wild-type  $\beta$ 4GalT7 was able to catalyze the transfer of Glc and Xyl from the corresponding donor substrate onto 4-MUX, although with lower activity (11- and 27-fold, respectively) compared with its galactosyltransferase activity (Fig. 6A). As expected from kinetic studies, the W224H mutant exhibited a large decrease in galactosyltransferase activity compared with wild-type (about 10 times). Notably, activity of this mutant was undetectable toward UDP-Glc and UDP-Xyl. The bulkier UDP-sugars, *i.e.* UDP-GlcUA, UDP-GalNAc, and UDP-GlcNAc did not serve as donor substrate for the wild-type enzyme or the W224H mutant (data not shown). These findings indicate that we engineered a W224H variant of  $\beta$ 4GalT7 that displays a large increase in affinity toward UDP-Gal compared with the wild-type enzyme and shows no detectable activity toward other UDP-sugars.

Furthermore, we analyzed the specificity of wild-type and W224H mutant toward a series of potential acceptor molecules. The wild-type  $\beta$ 4GalT7 was inactive toward glycoside derivatives of 4-MU (*i.e.* 4-MU-GalNAc, 4-MU-GlcNAc, 4-MU-Glc, 4-MU-GlcUA, and 4-MU-Gal) tested as acceptor substrates, indicating a strict selectivity toward the xyloside 4-MUX, in agreement with the critical role of this enzyme in GAG chain synthesis (data not shown). As wild-type  $\beta$ 4GalT7, the W224H mutant was only active toward 4-MUX. This prompted us to determine the activity of  $\beta$ 4GalT7 and the W224H mutant toward xyloside analogs linked to various aglycone moieties. As shown in Fig. 6B, wild-type  $\beta$ 4GalT7 exhibited higher activity toward MN-Xyl compared with 4-MUX and 4-NP-Xyl and was inactive toward MN-Xyl-2P. The use of a purified form of  $\beta$ 4GalT7 showed that the presence of a phosphate group at the position 2 of a xyloside substrate (MN-Xyl-2P) precludes galactosyltransferase activity, in agreement with our previous data (7). Remarkably, the W224H mutant displayed a different substrate profile compared with the wild-type enzyme. Indeed, this mutant exhibited a higher activity toward 4-MUX than for MN-Xyl (Fig. 6B). These results suggest that the side chain of Trp<sup>224</sup> interacts with the aglycone component of the acceptor substrate and that its mutation to His modifies this interaction. These results support the idea that Trp<sup>224</sup> is a



**FIGURE 5. Expression and purification of wild-type and mutated  $\beta$ 4GalT7 expressed as GST fusion proteins in *E. coli*.** The wild-type  $\beta$ 4GalT7, Trp<sup>224</sup>, Glu<sup>227</sup>, and Asp<sup>228</sup> mutants were expressed in *E. coli* and purified by glutathione-Sepharose™ 4B affinity chromatography, as described under "Experimental Procedures." Proteins were analyzed by SDS-PAGE on 10% acrylamide gels and stained with Coomassie Brilliant Blue. For each mutant, shown from left to right, cell lysate from uninduced and isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG)-induced bacteria (L, 25  $\mu$ g), soluble protein from isopropyl  $\beta$ -D-thiogalactopyranoside-induced bacteria (S, 12  $\mu$ g), and purified fusion protein (P, 3  $\mu$ g) are shown. The molecular mass markers are shown in the 1st lane of each gel.

**TABLE 2**  
Steady-state kinetic parameters of wild-type and  $\beta$ 4GalT7 mutants purified from *E. coli*

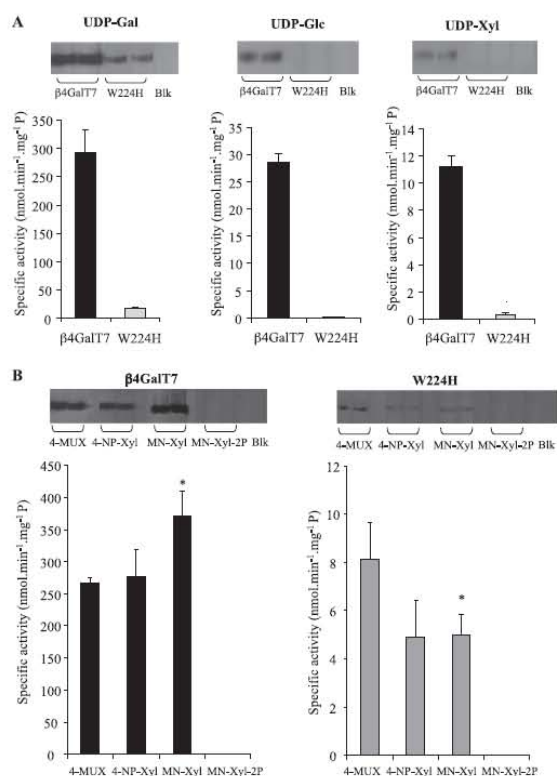
Kinetic parameters toward donor and acceptor substrates were evaluated from initial velocity values of the reaction using 0.2  $\mu$ g of purified GST- $\beta$ 4GalT7 protein at a constant amount of 4-MUX (5 mM) in the presence of increasing concentrations of UDP-Gal (0–10 mM) or increasing concentrations of 4-MUX (0–10 mM) in the presence of a fixed concentration of UDP-Gal (1 mM). Kinetic parameters were determined by nonlinear least squares regression analysis of the data fitted to Michaelis-Menten rate equation using Sigmaplot 9.0. The results are the mean values of three independent determinations  $\pm$  S.D. on assays performed in triplicate. ND indicates no kinetic parameters could be determined using excess acceptor or donor substrate and greater than 0.5  $\mu$ g of enzyme.

Enzyme	UDP-Gal			4-MUX		
	$K_m$ mM	$k_{cat}$ $s^{-1}$	$k_{cat}/K_m$ $\text{mM}^{-1}\text{s}^{-1}$	$K_m$ mM	$k_{cat}$ $s^{-1}$	$k_{cat}/K_m$ $\text{mM}^{-1}\text{s}^{-1}$
$\beta$ 4GalT7	$0.28 \pm 0.13$	$1.52 \pm 0.52$	5.43	$0.33 \pm 0.03$	$1.79 \pm 0.10$	5.42
W224A	ND	ND		ND	ND	
W224F	ND	ND		ND	ND	
W224H	$0.02 \pm 0.01^*$	$0.10 \pm 0.08^*$	5.00	ND	ND	
E227A	ND	ND		ND	ND	
E227D	$0.10 \pm 0.01$	$0.46 \pm 0.06^*$	4.6	$0.74 \pm 0.31$	$0.13 \pm 0.03^*$	0.17
D228A	ND	ND		ND	ND	
D228E	ND	ND		ND	ND	

\* $p < 0.05$ .



## $\beta$ 4GalT7 Key Functional Amino Acids



**FIGURE 6. Specificity of wild-type  $\beta$ 4GalT7 and the W224H mutant purified from *E. coli*.** Activity of purified wild-type (WT) and mutant enzymes was evaluated in standard conditions using 5 mM acceptor substrate and 0.04 mM UDP-sugar (0.06  $\mu$ Ci of UDP-[U-<sup>14</sup>C]sugar) donor substrate in the presence of 0.2  $\mu$ g of purified protein. The reaction products were separated by thin layer chromatography, visualized by autoradiography (shown in inset), and quantified by scintillation counting. The rate values are the means of three experiments. *A*, activity of wild-type  $\beta$ 4GalT7 and the W224H mutant toward various UDP-sugars and 4-MUX as acceptor substrate; *B*, activity of wild-type  $\beta$ 4GalT7 and the W224H mutant toward various xyloside derivatives using UDP-Gal as donor substrate. Results are the means  $\pm$  S.D. of three assays on two independent experiments. *B*, \*,  $p < 0.05$ , versus activity toward 4-MUX.

key residue for donor but also acceptor substrate binding and that substitution of this residue with His is able to modulate the specificity of  $\beta$ 4GalT7 toward acceptor xyloside molecules.

The mechanism of most inverting glycosyltransferases requires one acidic amino acid that activates the acceptor hydroxyl group by deprotonation (42). In *Drosophila*  $\beta$ 4GalT7, Asp<sup>211</sup> (equivalent to Asp<sup>228</sup> in human  $\beta$ 4GalT7) has been suggested to act as the catalytic base (10). In bovine  $\beta$ 4GalT1, Asp<sup>318</sup> and possibly Glu<sup>317</sup> (equivalent to Asp<sup>228</sup> and Glu<sup>227</sup> in human  $\beta$ 4GalT7) have been proposed as possible candidates to play such a role (40, 41). On the other hand, Glu<sup>317</sup> has also been suggested to interact with the Gal residue of the nucleotide sugar donor substrate (41). To further define the role of these carboxylate residues in  $\beta$ 4GalT7, Glu<sup>227</sup> and Asp<sup>228</sup> were subjected to conservative and nonconservative substitutions, and the resulting recombinant proteins were purified from the *E. coli* expression system, as was the wild-type protein (Fig. 5). Replacement of Glu<sup>227</sup> with Ala yielded an inactive enzyme,

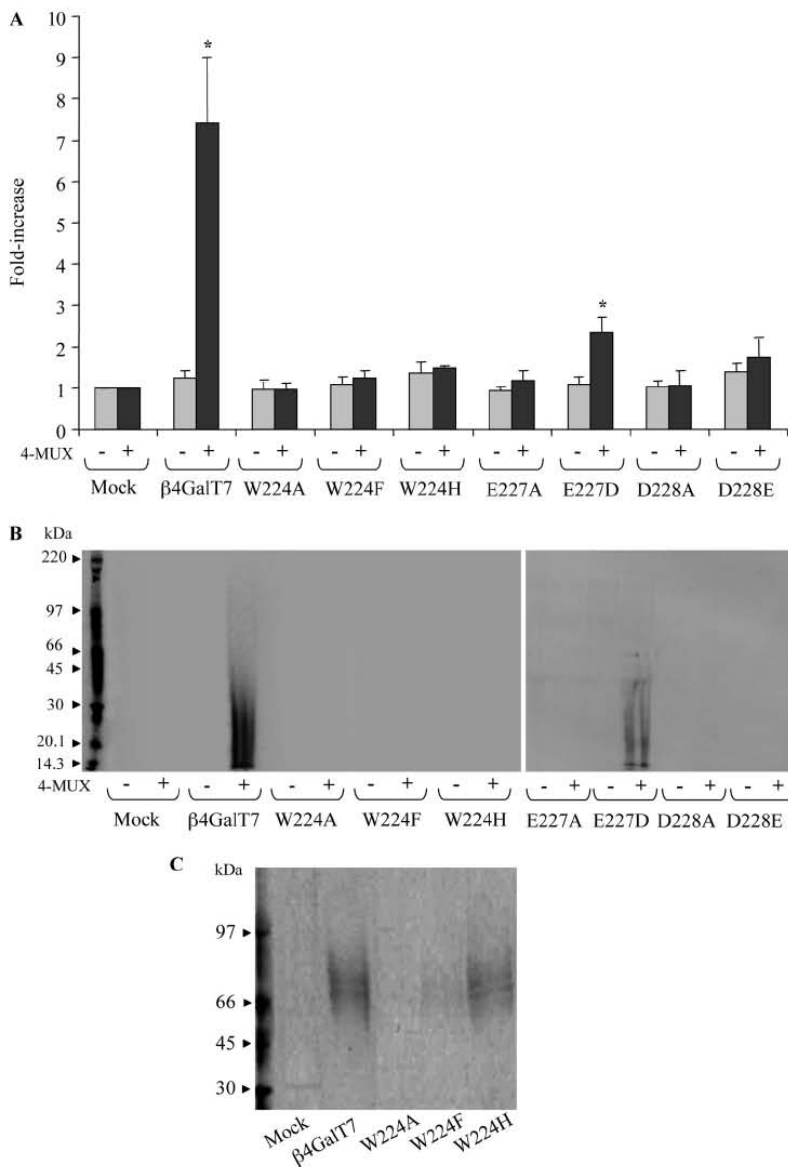
whereas E227D displayed detectable activity. Determination of the kinetic parameters of this mutant revealed about a 3-fold decrease in  $k_{cat}$  value with no modification in affinity, as evidenced by no significant change in  $K_m$  toward donor or acceptor substrates.

We next examined the activity of recombinant Asp<sup>228</sup> mutants expressed and purified from *E. coli*. Replacement of this residue with Ala resulted in a complete loss of activity for the purified enzyme. Although a residual activity could be detected for the D228E mutant (data not shown), the high  $K_m$  value toward 4-MUX did not allow determination of kinetic parameters. This result highlighted the functional importance of Asp<sup>228</sup> and indicates that the conservative mutation to Glu also severely affected enzyme activity. Altogether, this set of data confirms that the presence of these acidic residues is essential at the catalytic center.

**Effect of Trp<sup>224</sup>, Glu<sup>227</sup>, and Asp<sup>228</sup> Mutations on *ex Vivo* Galactosyltransferase Activity of  $\beta$ 4GalT7 toward 4-MUX**—We next designed a series of analyses to determine whether Trp<sup>224</sup>, Glu<sup>227</sup>, and Asp<sup>228</sup> residues were important for functionality *ex vivo*. For this purpose, CHO pgsB-618 cells, defective in galactosyltransferase activity and exhibiting dramatically impaired GAG synthesis (43), were transfected with wild-type and mutant  $\beta$ 4GalT7 cDNAs. In the first instance, we evaluated the capacity of the recombinant wild-type and mutant enzymes to prime GAG chain synthesis on the exogenous  $\beta$ -D-xyloside, 4-MUX, in deficient cells. For this purpose, [<sup>35</sup>S]SO<sub>4</sub> incorporation into GAG chains assembled onto 4-MUX was quantified by liquid scintillation counting, and radiolabeled GAGs were subjected to SDS-PAGE analysis. Results shown in Fig. 7A indicate that expression of wild-type  $\beta$ 4GalT7 in CHO pgsB-618 cells was able to promote GAG chain synthesis from 4-MUX *ex vivo*, as indicated by a 7.5-fold increase in sulfate incorporation compared with untreated or mock-transfected cells. In confirmation of this finding, SDS-PAGE analysis revealed the presence of radiolabeled GAG chains in  $\beta$ 4GalT7-transfected cells upon treatment with 4-MUX, whereas no radioactivity could be detected in mock-transfected cells or when addition of the xyloside was omitted (Fig. 7B).

To test the functional importance of the Trp<sup>224</sup> residue in this *ex vivo* model system, we evaluated the ability of the  $\beta$ 4GalT7 mutated at this position to initiate GAG synthesis from 4-MUX in CHO pgsB-618 cells. As expected, W224A and W224F mutants that were inactive or exhibited low *in vitro* activity failed to prime GAG synthesis in deficient cells (Fig. 7, A and B). Interestingly, sulfate incorporation as well as SDS-PAGE analysis of GAG chains in  $\beta$ 4GalT7-deficient CHO pgsB-618 cells transfected with W224H showed that this mutant did not initiate GAG synthesis from 4-MUX *ex vivo*. Although this mutant exhibited *in vitro* activity, this observation is consistent with the large increase in the affinity of  $\beta$ 4GalT7 toward the acceptor substrate 4-MUX.

We also compared the consequences of mutating the Glu<sup>227</sup> and Asp<sup>228</sup> residues on  $\beta$ 4GalT7 galactosyltransferase toward 4-MUX in the *ex vivo* and *in vitro* model systems. In agreement with *in vitro* studies, the inactive E227A and D228A mutants were not able to prime GAG synthesis in  $\beta$ 4GalT7-deficient cells, confirming that these residues are essential for the func-



**FIGURE 7. Effect of wild-type and mutated  $\beta$ 4GalT7 expression on GAG synthesis initiated from 4-MUX or on decorin core protein in CHO pgsB-618 cells.** For GAG analysis initiated from exogenous xyloside, cells were transiently transfected with wild-type (WT) or mutated  $\beta$ 4GalT7 cDNA or mock-transfected (empty vector) and then incubated in medium containing [ $^{35}$ S]SO $_4$  in the presence (+) or absence (-) of 4-MUX (5  $\mu$ M) for 12 h prior to GAG analysis. **A**, 1 ml of medium was applied to a G-50 column to separate nonincorporated radiolabeled sulfur, and radiolabeled GAGs present in the eluates were quantified by scintillation counting. Results are the mean  $\pm$  S.D. of three separate experiments; \*,  $p < 0.05$ , versus untreated; **B**, radiolabeled GAG chains were resolved by gel electrophoresis and visualized by autoradiography. **C**, CHO pgsB-618 cells stably expressing the recombinant human His-tagged decorin were transiently transfected with wild-type or mutant  $\beta$ 4GalT7 cDNA, or mock-transfected, and incubated in the presence of [ $^{35}$ S]SO $_4$  for 12 h prior to analysis. His-tagged decorin was purified on nickel column, analyzed by SDS-PAGE, and radiolabeled sulfated GAG chain of the recombinant protein was visualized by autoradiography.

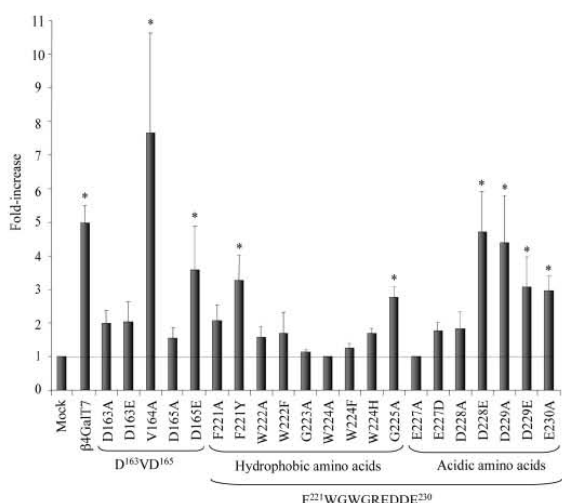
tion of this enzyme (Fig. 7A). On the other hand, results indicated that expression of the mutant in which Glu<sup>227</sup> was converted to Asp produced a 2-fold increase in sulfate incorpo-

ration in the culture medium of transfected cells upon addition of 4-MUX (Fig. 7A). As expected, radiolabeled GAGs could be detected by SDS-PAGE analysis in cells transfected by the E227D mutant, although the rate of sulfate incorporation was lower than that produced upon wild-type enzyme expression. Furthermore, substitution of Asp<sup>228</sup> with Glu, which strongly impaired *in vitro* activity, led to a mutant that was not able to prime GAG synthesis *ex vivo*. Altogether, these *ex vivo* studies are in good agreement with the *in vitro* findings showing the functional importance of these two carboxylate residues for the function of  $\beta$ 4GalT7.

**Effect of Trp<sup>224</sup> Mutations on *ex Vivo* Initiation of GAG Chain Synthesis onto the Core Protein of Decorin**—We compared the capacity of wild-type and Trp<sup>224</sup> mutants to assemble GAG chain synthesis onto the core protein of decorin, taken as a model proteoglycan. For this purpose, a His-tagged form of human decorin was stably expressed in CHO pgsB-618 cells defective in  $\beta$ 4GalT7. This cell line was transiently transfected with wild-type and Trp<sup>224</sup> substitution forms of  $\beta$ 4GalT7, and the recombinant decorin was purified by nickel chelating chromatography from the culture medium following GAG chain labeling by sulfate incorporation. Gel electrophoresis of purified material from cells expressing wild-type  $\beta$ 4GalT7 showed the presence of a GAG-substituted form of decorin migrating as a polydisperse smear centered at about 70 kDa, whereas mock-transfected cells showed no radiolabeling in agreement with the absence of GAG substitution in the  $\beta$ 4GalT7-deficient cell line (Fig. 7C). When the W224A-substituted  $\beta$ 4GalT7 cDNA was transfected into the cell line stably expressing recombinant decorin, no sulfated GAG chains could be detected. [ $^{35}$ S]SO $_4$  labeling

of decorin was also low when cells were transfected with the W224F mutant, whereas expression of W224H led to a profile of decorin GAG sulfation that was comparable with that of

## $\beta$ 4GalT7 Key Functional Amino Acids



**FIGURE 8. Effect of wild-type and mutated  $\beta$ 4GalT7 expression on total proteoglycan synthesis in CHO pgsB-618 cells.** Cells were transiently transfected with wild-type or mutated  $\beta$ 4GalT7 cDNAs or mock-transfected (empty vector) for 48 h and placed in [ $^{35}$ S]SO $_4$  labeling medium 8 h prior to proteoglycan extraction and GAG quantification, as described under "Experimental Procedures." Results are expressed as the ratio of radioactivity associated with GAGs extracted from wild-type or mutated  $\beta$ 4GalT7 transfected cells on radioactivity associated with GAGs extracted from mock-transfected cells. The results are the mean values  $\pm$  S.D. of at least three independent experiments each performed in triplicate. \*,  $p < 0.05$ , versus mock-transfected.

wild-type  $\beta$ 4GalT7 (Fig. 7C). These results corroborated the importance of Trp at position 224 and indicate that replacement of Trp<sup>224</sup> with Phe or His yielded enzymes that were able to sustain GAG chain substitution in  $\beta$ 4GalT7-deficient cells.

**Effect of Wild-type and Mutant  $\beta$ 4GalT7 Expression on GAG Synthesis *ex Vivo***—In a final set of experiments, we established the activity profile of the library of  $\beta$ 4GalT7 mutants in the *ex vivo* model system and compared it with their *in vitro* activity. For this purpose, we determined the capacity of wild-type and mutated  $\beta$ 4GalT7 enzymes to restore GAG synthesis in defective CHO pgsB-618 cells, by radioactive sulfate incorporation. As shown in Fig. 8, transfection of CHO pgsB-618 cells with wild-type  $\beta$ 4GalT7 cDNA was able to restore GAG synthesis, producing up to a 5-fold increase in the rate of [ $^{35}$ S]SO $_4$  incorporation, compared with mock-transfected cells. Examination of the consequences of mutations within the  $^{163}\text{DVD}^{165}$  motif indicated that expression of the mutants in which the Asp residues were converted to Ala did not produce a significant increase in sulfate incorporation compared with mock-transfected cells ( $\leq 2$ -fold). Similarly, the D163E mutant (inactive *in vitro*) was not able to support GAG synthesis when expressed in CHO-deficient cells (Fig. 8). By contrast, the GAG synthesis rate following expression of D165E was in the same range as the wild-type enzyme, indicating that Glu at this position was able to sustain significant activity, both *in vitro* and *ex vivo*. Furthermore, substitution of Val<sup>164</sup> with Ala, which did not affect *in vitro*  $\beta$ 4GalT7 activity, resulted in a mutant enzyme able to increase GAG synthesis in defective cells, to about the same extent as that observed for the wild-type. Altogether, these GAG

synthesis studies are in agreement with our *in vitro* findings showing the functional importance of both Asp residues of the  $^{163}\text{DVD}^{165}$  motif. Furthermore, in agreement with the *in vitro* assays, our results indicate that substitution of Asp with Glu at position 165, but not position 163, yielded an enzyme able to restore GAG synthesis, as the wild-type enzyme in  $\beta$ 4GalT7-deficient cells.

The consequences of mutations within the  $^{221}\text{FWGWG-REDDE}^{230}$  motif on GAG synthesis were next investigated in CHO pgsB-618-defective cells. Expression of the F221A mutant did not induce a significant increase in the rate of [ $^{35}$ S]SO $_4$  incorporation, compared with mock-transfected cells (Fig. 8), indicating that this nonconservative mutation exerts a more severe effect on total GAG synthesis than on *in vitro* galactosyltransferase activity toward the exogenous xyloside 4-MUX. On the other hand, the conservative change of Phe<sup>221</sup> to Tyr resulted in an enzyme that was able to restore significant GAG synthesis in deficient cells, concordant with the *in vitro* properties of this mutant. The nonconservative mutation of the neighboring Trp<sup>222</sup> residue to Ala led to reduction of *in vitro* enzyme activity, as well as GAG synthesis *ex vivo* (Fig. 8). Furthermore, the W222F mutant was not able to restore GAG synthesis, although this conservative mutation at position 222 did not affect *in vitro* galactosyltransferase activity toward 4-MUX.

Investigation of the *in vitro* kinetic properties of several mutants of Trp at position 224 suggested a key role for this residue that is located at a strategic position in the  $\beta$ 4GalT7 active site. In agreement with the lack or very low galactosyltransferase activity of W224A and W224F, these mutants were not able to restore GAG synthesis when transfected in  $\beta$ 4GalT7-deficient cells (Fig. 8). Expression of the W224H mutant did not significantly increase the rate of [ $^{35}$ S]SO $_4$  incorporation, compared with mock-transfected cells. The low activity of this mutant appears to be unable to restore the production of significant amounts of total GAGs, but this activity was sufficient to initiate some glycanation of specific proteoglycans, such as decorin, as indicated above. Altogether, analysis of the consequences of substituting Trp<sup>224</sup> on GAG synthesis in defective cells supports a key functional role of this residue.

The effects of expression of Ala mutants of Gly<sup>223</sup> or Gly<sup>225</sup> that surround Trp<sup>224</sup> to form a short flexible loop were also examined. Substitution of Gly<sup>223</sup> with Ala, which moderately impaired *in vitro* activity toward 4-MUX, produced a more important effect *ex vivo*, because this mutant was unable to promote GAG synthesis upon transfection in deficient cells (Fig. 8). This observation may be explained by the presence of the bulky side chain of Ala compared with Gly that may be tolerated in the presence of a small acceptor exogenous xyloside, but not when the glycopeptide primer of proteoglycans is the substrate. On the other hand, G225A, which was inactive toward 4-MUX *in vitro* and *ex vivo* (data not shown), was able to restore GAG synthesis when expressed in CHO pgsB-618 cells, although less than the wild-type protein. These results indicate that mutation of Gly to Ala at position 225 affects binding of 4-MUX more than that of the core protein of proteoglycans. In agreement, we found that this mutant was able to initiate decorin glycanation *ex vivo* (data not shown). However, more information on the organization of the acceptor binding site

based on the three-dimensional structure of  $\beta$ 4GalT7 in complex with xyloside substrate is required to propose a structure-based explanation of the effects of these mutations.

Examination of the consequences of the mutations of the C-terminal acidic area of the  $^{221}$ FWGWGREDD $^{230}$  motif indicated that mutation of Glu $^{227}$  to Ala abolished not only *in vitro* but also *ex vivo*  $\beta$ 4GalT7 activity (Fig. 8). Furthermore, transfection of the E227D mutant, which exhibited a reduced *in vitro* galactosyltransferase activity compared with wild-type  $\beta$ 4GalT7, was not able to promote GAG synthesis in CHO-defective cells. Mutation of Asp $^{228}$  to Ala strongly impaired *ex vivo* GAG synthetic activity of  $\beta$ 4GalT7, whereas transfection of the D228E mutant restored GAG synthesis in deficient cells to the same rate as the wild-type enzyme. On the other hand, examination of the activity of nonconservative or conservative mutations of Asp $^{229}$  in GAG biosynthesis led to unexpected results. Whereas D229A and D229E mutants were found completely devoid of *in vitro* galactosyltransferase activity, both mutants were able to produce an increase in GAG synthesis up to that of the wild-type enzyme in deficient cells. A possible explanation for this discrepancy would be that mutations at this position affect binding of the aglycone part of the exogenous xyloside 4-MUX but not of the core protein of proteoglycans, which may involve additional interactions with  $\beta$ 4GalT7 enzyme. In agreement with this hypothesis, we found that both mutants were unable to prime GAG synthesis from 4-MUX *ex vivo* but were able to sustain decorin glycanation (data not shown). However, elucidation of the three-dimensional organization of the acceptor substrate-binding site is required to precisely define the role of Asp $^{229}$ . Finally, mutation of Glu $^{230}$ , the last residue of the motif, to Ala did not impair GAG synthetic activity, consistent with the effect of this mutation observed on galactosyltransferase activity *in vitro*.

## DISCUSSION

The role of GAGs in a wide range of biological processes has led to considerable interest in understanding the properties of the glycosyltransferases involved in their biosynthesis.  $\beta$ 4GalT7 is essential for the formation of the tetrasaccharide linkage region that initiates both glucosamino- and galactosaminoglycan synthesis (6); it therefore plays a key role in the early steps of GAG chain formation built on cellular and extracellular proteoglycans.  $\beta$ 4GalT7 has also received much attention because it is able to prime GAG chains from exogenous  $\beta$ -D-xylosides. These compounds have been widely employed for examining the contribution of proteoglycans to biological processes and are currently attracting interest as potential drugs (12, 16, 17).

Here, we have focused on the organization of the functional peptide regions of human  $\beta$ 4GalT7 and the mechanisms of Gal transfer from the UDP-Gal donor substrate onto the Xyl residue of exogenous xylosides and of the glycopeptide primer of GAG chains of proteoglycans. To pinpoint potential domains involved in substrate recognition and/or catalysis, we carried out a phylogenetic analysis of the family combined with molecular modeling of the active site of human  $\beta$ 4GalT7, based on the crystal structure available for *Drosophila*  $\beta$ 4GalT7 (10). The three-dimensional structure of bovine  $\beta$ 4GalT1 complexed

with UDP-Gal was also considered (31). The canonical motifs,  $^{163}$ DVD $^{165}$  and  $^{221}$ FWGWGREDD $^{230}$  were subjected to systematic site-directed mutagenesis, and we assessed the functional consequences of these in two model systems designed to assess *in vitro* kinetic properties and GAG synthesis *ex vivo*. For the latter, decorin was chosen as the model core protein as it is a member of the small leucine-rich proteoglycan family and an important constituent of interstitial extracellular matrices (44). Decorin is covalently linked with one GAG chain of the chondroitin/dermatan sulfate type, and previous reports showed that genetic defects of  $\beta$ 4GalT7 activity affect decorin glycanation with severe clinical consequences linked to the progeroid form of Ehlers-Danlos syndrome (45).

Multiple alignment of 48 sequences of  $\beta$ 4GalT7 plus the 6 human  $\beta$ 4GalT1 to  $\beta$ 4GalT6 highlighted the presence of eight conserved motifs. Among those, motifs 3 and 5 contain two functional domains, *i.e.* the  $^{163}$ DVD $^{165}$  canonical motif of many glycosyltransferases and the  $^{221}$ FWGWGREDD $^{230}$  domain corresponding to the so-called  $\beta$ 4GT motif (46). The latter domain encompasses an Asp/Glu catalytic dicarboxylate found in many glycosyltransferases (47, 48). Our investigations showed the functional importance of both aspartate residues of the  $^{163}$ DVD $^{165}$  domain. In this respect,  $\beta$ 4GalT7 exhibits typical features of glycosyltransferases of the GTA-fold group that require Mn $^{2+}$  as cofactor, in which a three-residue motif Asp-Xaa-Asp participates in metal ion binding (42). Examination of the organization of human  $\beta$ 4GalT7 active site revealed that Asp $^{165}$  (as its counterpart Asp $^{147}$  in *Drosophila* and Asp $^{254}$  in bovine  $\beta$ 4GalT1) is in a position to establish a coordination bond with the divalent metal cation, and we could demonstrate that replacing this residue with Glu maintains some of this function. By contrast, the longer side chain of Glu at position 163, when substituted to Asp, is not in a favorable position to provide interactions with the Gal unit of the donor substrate, as proposed for Asp $^{163}$  in the modeled organization of the human  $\beta$ 4GalT7 donor-binding site. Furthermore, the Val $^{164}$  to Ala substitution did not affect  $\beta$ 4GalT7 activity and led to a mutant enzyme able to restore GAG synthesis in deficient cells to a similar extent as that observed for the wild-type enzyme, indicating that this residue does not play a prevalent role in UDP-Gal binding. This result is in agreement with the proposed location of this Val residue in the structure of the *Drosophila* and the modeled human  $\beta$ 4GalT7, pointing out of the donor-binding site. Altogether, we show that the structure of the  $^{163}$ DVD $^{165}$  motif of  $\beta$ 4GalT7 is similar to that of glycosyltransferases of the GT-A fold group in which the first acidic residue of the sequence interacts directly with the sugar donor and the second acidic residue coordinates the metal divalent cation (21).

The conserved  $^{221}$ FWGWGREDD $^{230}$  peptide region can be divided in two subdomains with an N-terminal part consisting of five hydrophobic residues located in a flexible loop and a C-terminal part containing a stretch of acidic amino acids belonging to an  $\alpha$ -helix. Conservative and nonconservative substitutions at Phe $^{221}$  and Trp $^{222}$  produced only a moderate impairment of enzyme activity and function and did not change the  $K_m$  value toward the donor substrate. Our results suggest that these residues are not directly involved in UDP-Gal bind-

## $\beta$ 4GalT7 Key Functional Amino Acids

ing, in agreement with their position out the donor substrate binding pocket of the modeled human  $\beta$ 4GalT7. Thus, the fact that these residues are highly conserved may indicate a role in maintaining the structural integrity of the  $\beta$ 4GalT7 active site. Examination of the consequences of mutating Trp<sup>224</sup> confirmed that this residue holds key functional importance in  $\beta$ 4GalT7. Substitution with Ala resulted in a completely inactive enzyme. No galactosyltransferase activity could be detected when W224A was expressed in eukaryotic cells or when produced and purified from recombinant *E. coli*. Similarly, expression of this mutant was not able to restore GAG synthesis and decorin glycanation in  $\beta$ 4GalT7-deficient CHO pgsB-618 cells. Replacement of the large Trp side chain by Phe, a smaller aromatic residue, was able to restore *in vitro* activity to a minor extent, and the low level of activity displayed by the W224F mutant did not allow kinetic parameters to be determined. A most striking finding of this study was revealed by analyzing the consequences of mutating Trp<sup>224</sup> to His, both on activity and specificity of  $\beta$ 4GalT7. Indeed, this substitution produced an enzyme exhibiting a 4.4-fold decrease in  $K_m$  value toward UDP-Gal compared with the wild-type enzyme whose affinity toward acceptor 4-MUX was strongly impaired. Characterization of the properties of the purified W224H mutant confirmed a large increase in affinity and showed enhanced selectivity toward UDP-Gal. Furthermore, we showed that the acceptor specificity profile of  $\beta$ 4GalT7 was modified when Trp<sup>224</sup> was replaced with His. Indeed, whereas MN-Xyl was a preferred substrate for the wild-type enzyme, W224H exhibited higher activity toward 4-MUX. Molecular dynamic simulations of the human  $\beta$ 4GalT7 structure, based on the x-ray crystal structure of *Drosophila*  $\beta$ 4GalT7, indicate that the Ne1 atom of the indole side chain of Trp<sup>224</sup> forms a hydrogen bond with the  $\beta$ -phosphate group of UDP-Gal. Furthermore, in the crystal structure of bovine  $\beta$ 4GalT1, it has been proposed that Trp<sup>314</sup> (equivalent to Trp<sup>224</sup> in  $\beta$ 4GalT7) also aids in keeping the acceptor anchored by means of hydrophobic interactions (49). Thus, it can be hypothesized that in the  $\beta$ 4GalT7 W224H mutant, the Ne2 atom of the imidazole side chain of His is able to sustain hydrogen bond formation with the pyrophosphate of UDP-Gal. However, the lack of an aromatic moiety in the imidazole side chain of His (compared with indole side chain of Trp) may impair stacking interactions with the acceptor substrate 4-MUX. In line with our results, a key role for a Trp residue at the catalytic center of several glycosyltransferases has previously been reported. For example, Trp<sup>314</sup> in  $\beta$ 4GalT1 (50), Trp<sup>284</sup> in EXTL2 (51), and Trp<sup>314</sup> in  $\alpha$ 3GT (52) have all been shown to be located at a critical position for interacting with both sugar nucleotide and acceptor substrate. Thus, our study, in common with these other examples, highlights the presence of a Trp residue as a critical feature at the catalytic center of several glycosyltransferases. Supporting this idea, a consensus sequence containing the short amino acid stretch WGGE is present in human GalNAc-T4 transferase (<sup>334</sup>WGGE<sup>337</sup> (53)) and in *Pasteurella multocida* hyaluronan synthase (<sup>366</sup>WGGE<sup>369</sup>, (54)).

Similar to what is observed in the structure of *Drosophila*  $\beta$ 4GalT7 and bovine  $\beta$ 4GalT1, Trp<sup>224</sup> (equivalent to Trp<sup>207</sup> in *Drosophila* and to Trp<sup>314</sup> in  $\beta$ 4GalT1) lies in a flexible loop and

is flanked by Gly residues at positions 223 and 225. Mutation of Gly<sup>223</sup> to Ala had a minor effect on 4-MUX enzyme activity *in vitro* and on decorin GAG synthesis (data not shown), although the G223A substituted enzyme was not able to restore GAG synthesis in deficient cells. On the other hand, replacement of Gly<sup>225</sup> abolished *in vitro* activity toward 4-MUX, whereas G225A sustained decorin glycanation (data not shown) and, to some extent, GAG synthesis in CHO pgsB-618 cells. Based on structural data and molecular modeling analyses, it is likely that mutation of these residues, which are predicted to allow a rotary motion of the Trp during the conformational change induced by UDP-Gal binding, affects positioning of Trp<sup>224</sup> within the catalytic center (10, 31). These mutations may also impact the organization of the acceptor-binding site with different consequences on  $\beta$ 4GalT7 activity when Xyl is linked to a hydrophobic aglycone such as 4-MU or to the core protein of proteoglycans. Elucidation of the organization of the human  $\beta$ 4GalT7 acceptor-binding site by x-ray crystallography is required to assess these assumptions.

Finally, we investigated the functional importance of the C-terminal acidic residues of the <sup>221</sup>FWGWGRE<sup>230</sup> motif, which is highly conserved among all members of the  $\beta$ 4GalT7 family. Whereas a crucial role could be demonstrated for Glu<sup>227</sup>, Asp<sup>228</sup>, and Asp<sup>229</sup>, the glutamate residue at position 230 appears much less important in terms of function. Indeed, analysis of the *in vitro* consequences of Glu<sup>227</sup> mutations showed that Ala substitution completely abrogated enzyme activity, whereas replacement by Asp led to only a partial reduction in  $\beta$ 4GalT7 activity, and the enzyme was able to sustain GAG synthesis activity in CHO pgsB-618 cells. Furthermore, our study indicated that a nonconservative mutation of Asp<sup>228</sup> had a severe impact on the galactosyltransferase activity of  $\beta$ 4GalT7. However, replacement of Asp to Glu at that position produced a mutant with significant residual activity *in vitro* and *ex vivo*. Altogether, our data are concordant with a carboxylate residue acting as catalytic base in  $\beta$ 4GalT7, facilitating the deprotonation of the acceptor substrate and the nucleophilic attack on the C1 of the sugar residue of UDP-Gal. In the three-dimensional structure of bovine  $\beta$ 4GalT1 complexed with UDP-Gal, Glu<sup>317</sup> was proposed as the catalytic base (40). On the other hand, subsequent analysis of the structure of bovine W314A- $\beta$ 4GalT1 complexed with UDP, Mn<sup>2+</sup>, and GlcNAc predicted that Asp<sup>318</sup> (equivalent to Asp<sup>228</sup> in human  $\beta$ 4GalT7) should form a hydrogen bond with the hydroxyl group at C4 of GlcNAc and be involved in deprotonating the hydroxyl group (49). When Glc was docked into the acceptor-binding site of *D. melanogaster*  $\beta$ 4GalT7-UDP structure, it was predicted to adopt a similar position to xylose, and residue Asp<sup>211</sup> was predicted to act as catalytic base (10). Molecular modeling of human  $\beta$ 4GalT7 active site using *Drosophila*  $\beta$ 4GalT7 as template favors the role of Asp<sup>228</sup> as catalytic base, whereas Glu<sup>227</sup> is predicted to interact with O4 of Gal. Interestingly, the presence of a functional Glu/Asp motif has been identified in several glycosyltransferases such as the cyclin glucan synthase (47) and the peptidoglycan glycosyltransferase module of class A penicillin-binding protein. The latter contains a conserved dicarboxylate Glu<sup>233</sup>-Asp<sup>234</sup> (48) and glycan chain elongation catalyzed by this enzyme was proposed to function with the Glu

residue as active site general base to deprotonate the GlcNAc 4-OH group of lipid II acting as an acceptor. Further investigation of the interactions between human  $\beta$ 4GalT7 and xylose is underway to decipher the respective role of each residue of the conserved dicarboxylate in the  $\beta$ 4GalT7 family.

In conclusion, this work has identified two functional regions,  $^{163}\text{DVD}^{165}$  (in motif 3) and  $^{221}\text{FWGWGREDD}^{230}$  (in motif 5), critical for the organization of UDP-Gal-binding site of human  $\beta$ 4GalT7, and it suggests the existence of two  $\beta$ 4GalT7-specific conserved peptide motifs (motifs 6 and 8). We also demonstrated the central role of Trp $^{224}$  in governing interactions with both donor and acceptor substrates by mutation of this amino acid, and we showed that it is thus possible to modulate the activity of  $\beta$ 4GalT7 toward exogenous xylosides and proteoglycan core protein. Finally, our results support the critical role for a dicarboxylate motif in the vicinity of the key Trp $^{224}$  in the glycosyl transfer catalyzed by  $\beta$ 4GalT7. This study brings us closer to an understanding of the enzymology of human  $\beta$ 4GalT7, and it provides useful information toward the design of molecules targeting this enzyme, which have potential therapeutic applications.

*Acknowledgment*—Matthieu Chabel is gratefully acknowledged for expert technical assistance.

#### REFERENCES

- Bishop, J. R., Schuksz, M., and Esko, J. D. (2007) *Nature* **446**, 1030–1037
- Kreuger, J., Spillmann, D., Li, J. P., and Lindahl, U. (2006) *J. Cell Biol.* **174**, 323–327
- Prydz, K., and Dalen, K. T. (2000) *J. Cell Sci.* **113**, 193–205
- Kusche-Gullberg, M., and Kjellén, L. (2003) *Curr. Opin. Struct. Biol.* **13**, 605–611
- Bui, C., Ouzzine, M., Talhaoui, I., Sharp, S., Prydz, K., Coughtrie, M. W., and Fournel-Gigleux, S. (2010) *FASEB J.* **24**, 436–450
- Almeida, R., Levery, S. B., Mandel, U., Kresse, H., Schwientek, T., Bennett, E. P., and Clausen, H. (1999) *J. Biol. Chem.* **274**, 26165–26171
- Gulberti, S., Lattard, V., Fondeur, M., Jacquinet, J. C., Mulliert, G., Netter, P., Magdalou, J., Ouzzine, M., and Fournel-Gigleux, S. (2005) *J. Biol. Chem.* **280**, 1417–1425
- Quentin, E., Gladen, A., Rodén, L., and Kresse, H. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1342–1346
- Nakamura, Y., Haines, N., Chen, J., Okajima, T., Furukawa, K., Urano, T., Stanley, P., Irvine, K. D., and Furukawa, K. (2002) *J. Biol. Chem.* **277**, 46280–46288
- Ramakrishnan, B., and Qasba, P. K. (2010) *J. Biol. Chem.* **285**, 15619–15626
- Okajima, T., Yoshida, K., Kondo, T., and Furukawa, K. (1999) *J. Biol. Chem.* **274**, 22915–22918
- Lugemwa, F. N., Sarkar, A. K., and Esko, J. D. (1996) *J. Biol. Chem.* **271**, 19159–19165
- Martin, N. B., Masson, P., Sepulchre, C., Theveniaux, J., Millet, J., and Bellamy, F. (1996) *Semin. Thromb. Hemost.* **22**, 247–254
- Kisilevsky, R., Szarek, W. A., Ancsin, J., Vohra, R., Li, Z., and Marone, S. (2004) *J. Mol. Neurosci.* **24**, 167–172
- Mani, K., Havsmark, B., Persson, S., Kaneda, Y., Yamamoto, H., Sakurai, K., Ashikari, S., Habuchi, H., Suzuki, S., Kimata, K., Malmström, A., Westergren-Thorsson, G., and Fransson, L. A. (1998) *Cancer Res.* **58**, 1099–1104
- Belting, M., Borsig, L., Fuster, M. M., Brown, J. R., Persson, L., Fransson, L. A., and Esko, J. D. (2002) *Proc. Natl. Acad. Sci. U.S.A.* **99**, 371–376
- Schuksz, M., Fuster, M. M., Brown, J. R., Crawford, B. E., Ditto, D. P., Lawrence, R., Glass, C. A., Wang, L., Tor, Y., and Esko, J. D. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13075–13080
- Amado, M., Almeida, R., Schwientek, T., and Clausen, H. (1999) *Biochim. Biophys. Acta* **1473**, 35–53
- Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2008) *Curr. Drug Targets* **9**, 292–309
- Kitayama, K., Hayashida, Y., Nishida, K., and Akama, T. O. (2007) *J. Biol. Chem.* **282**, 30085–30096
- Breton, C., and Imberty, A. (1999) *Curr. Opin. Struct. Biol.* **9**, 563–571
- Hennet, T. (2002) *Cell. Mol. Life Sci.* **59**, 1081–1095
- Huang, X., and Madan, A. (1999) *Genome Res.* **9**, 868–877
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410
- Harduin-Lepers, A., Mollicone, R., Delannoy, P., and Oriol, R. (2005) *Glycobiology* **15**, 805–817
- Combet, C., Blanchet, C., Geourjon, C., and Deléage, G. (2000) *Trends Biochem. Sci.* **25**, 147–150
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680
- Castresana, J. (2000) *Mol. Biol. Evol.* **17**, 540–552
- Galtier, N., Gouy, M., and Gautier, C. (1996) *Comput. Appl. Biosci.* **12**, 543–548
- Sali, A., and Blundell, T. L. (1993) *J. Mol. Biol.* **234**, 779–815
- Ramakrishnan, B., Balaji, P. V., and Qasba, P. K. (2002) *J. Mol. Biol.* **318**, 491–502
- Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Wang, B., Pearlman, D. A., Crowley, M., Brozell, S., Tsui, V., Gohlke, H., Mongan, J., Hornak, V., Cui, G., Beroza, P., Schafmeister, C., Caldwell, J. W., Ross, W. S., and P.A., K. (2004) *AMBER 8*, University of California, San Francisco
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Andres, J. L., Head-Gordon, M., Replogle, E. S., and Pople, J. A. (1998) *Gaussian 98*, Revision A 11.2, Gaussian, Inc., Pittsburgh
- Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolosváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M. J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2010) *AMBER 11*, University of California, San Francisco
- Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254
- Fournel-Gigleux, S., Jackson, M. R., Wooster, R., and Burchell, B. (1989) *FEBS Lett.* **243**, 119–122
- Fondeur-Gelinotte, M., Lattard, V., Oriol, R., Mollicone, R., Jacquinet, J. C., Mulliert, G., Gulberti, S., Netter, P., Magdalou, J., Ouzzine, M., and Fournel-Gigleux, S. (2006) *Protein Sci.* **15**, 1667–1678
- Bansal, S. K., and Gessner, T. (1980) *Anal. Biochem.* **109**, 321–329
- Li, D., Fournel-Gigleux, S., Barré, L., Mulliert, G., Netter, P., Magdalou, J., and Ouzzine, M. (2007) *J. Biol. Chem.* **282**, 36514–36524
- Gastinel, L. N., Cambillau, C., and Bourne, Y. (1999) *EMBO J.* **18**, 3546–3557
- Ramakrishnan, B., Ramasamy, V., and Qasba, P. K. (2006) *J. Mol. Biol.* **357**, 1619–1633
- Breton, C., Snajdrová, L., Jeanneau, C., Koca, J., and Imberty, A. (2006) *Glycobiology* **16**, 29R–37R
- Esko, J. D., Weinke, J. L., Taylor, W. H., Ekborg, G., Rodén, L., Anantharamaiah, G., and Gawish, A. (1987) *J. Biol. Chem.* **262**, 12189–12195
- Roughley, P. J. (2006) *Eur. Cell Mater.* **12**, 92–101
- Seidler, D. G., Faiyaz-Ul-Haque, M., Hansen, U., Yip, G. W., Zaidi, S. H., Teebi, A. S., Kiesel, L., and Götte, M. (2006) *J. Mol. Med.* **84**, 583–594

## $\beta$ 4GalT7 Key Functional Amino Acids

46. Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2005) *Trends Biochem. Sci.* **30**, 53–62
47. Ciochini, A. E., Roset, M. S., Briones, G., Iñón de Iannino, N., and Ugalde, R. A. (2006) *Glycobiology* **16**, 679–691
48. Terrak, M., Sauvage, E., Derouaux, A., Dehareng, D., Bouhss, A., Breukink, E., Jeanjean, S., and Nguyen-Distèche, M. (2008) *J. Biol. Chem.* **283**, 28464–28470
49. Ramasamy, V., Ramakrishnan, B., Boeggeman, E., and Qasba, P. K. (2003) *J. Mol. Biol.* **331**, 1065–1076
50. Ramakrishnan, B., and Qasba, P. K. (2001) *J. Mol. Biol.* **310**, 205–218
51. Pedersen, L. C., Dong, J., Taniguchi, F., Kitagawa, H., Krahn, J. M., Pedersen, L. G., Sugahara, K., and Negishi, M. (2003) *J. Biol. Chem.* **278**, 14420–14428
52. Zhang, Y., Swaminathan, G. J., Deshpande, A., Boix, E., Natesh, R., Xie, Z., Acharya, K. R., and Brew, K. (2003) *Biochemistry* **42**, 13512–13521
53. Bennett, E. P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M. A., Merx, G., van Kessel, A. G., Eiberg, H., Steffensen, R., and Clausen, H. (1998) *J. Biol. Chem.* **273**, 30472–30481
54. DeAngelis, P. L., Jing, W., Drake, R. R., and Achyuthan, A. M. (1998) *J. Biol. Chem.* **273**, 8454–8458

## Supplemental data Tables



## S-T1

Sequences of sense and antisense primers used for site-directed mutagenesis.

Mutated codons are underlined.

Mutant	Primer	Sequence
Asp163Ala	Sense	5'-GAC TAC ATT GCC ATG CAC <u>GCA</u> GTT GCA CTG CTC CCT CTC-3'
	Antisense	5'-GAG AGG GAG CAG TGC AAC <u>TGC</u> GTG CAT GGC AAT GTA GTC-3'
Asp163Glu	Sense	5'-GAC TAC ATT GCC ATG CAC <u>GAG</u> GTT GCA CTG CTC CCT CTC-3'
	Antisense	5'-GAG AGG GAG CAG TGC AAC <u>CTC</u> GTG CAT GGC AAT GTA GTC-3'
Val164Ala	Sense	5'-TAG ATT GCC ATG CAC GAC <u>GCT</u> GAC CTG CTC CCT CTC AAC-3'
	Antisense	5'-GTT GAG AGG GAG CAG GTC <u>AGC</u> GTC GTG CAT GGC AAT CTA-3'
Asp165Ala	Sense	5'-ATT GCC ATG CAC GAC GTT <u>GCA</u> CTG CTC CCT CTC AAC GAG-3'
	Antisense	5'-CTC GTT GAG AGG GAG CAG <u>TGC</u> AAC GTC GTG CAT GGC AAT-3'
Asp165Glu	Sense	5'-ATT GCC ATG CAC GAC GTT <u>GAG</u> CTG CTC CCT CTC AAC GAG-3'
	Antisense	5'-CTC GTT GAG AGG GAG CAG <u>CTC</u> AAC GTC GTG CAT GGC AAT-3'
Phe221Ala	Sense	5'-ATG TCC AAC CGC <u>GCG</u> TGG GGC TGG GGC-3'
	Antisense	5'-GCC CCA GCC CCA <u>CGC</u> GCG GTT GGA CAT-3'
Phe221Tyr	Sense	5'-ATG TCC AAC CGC <u>TAC</u> TGG GGC TGG GGC-3'
	Antisense	5'-GCC CCA GCC CCA <u>GTA</u> GCG GTT GGA CAT-3'
Trp222Ala	Sense	5'-TCC AAC CGC TTC <u>GCG</u> GGC TGG GGC CGC-3'
	Antisense	5'-GCG GCC CCA GCC <u>CGC</u> GAA GCG GTT GGA-3'
Trp222Phe	Sense	5'-TCC AAC CGC TTC <u>TTT</u> GGC TGG GGC CGC-3'
	Antisense	5'-GCG GCC CCA GCC <u>AAA</u> GAA GCG GTT GGA-3'
Gly223Ala	Sense	5'-AAC CGC TTC TGG <u>GCG</u> TGG GGC CGC GAG-3'
	Antisense	5'-CTC GCG GCC CCA <u>CGC</u> CCA GAA GCG GTT-3'
Trp224Ala	Sense	5'-CGC TTC TGG GGC <u>GCG</u> GGC CGC GAG GAC-3'
	Antisense	5'-GTC CTC GCG GCC <u>CGC</u> GCC CCA GAA GCG-3'
Trp224His	Sense	5'-CGC TTC TGG GGC <u>CAC</u> GGC CGC GAG GAG-3'
	Antisense	5'-CTC CTC GCG GCC <u>GTG</u> GCC CCA GAA GCG-3'
Trp224Phe	Sense	5'-CGC TTC TGG GGC <u>TTC</u> GGC CGC GAG GAG-3'
	Antisense	5'-CTC CTC GCG GCC <u>GAA</u> GCC CCA GAA GCG-3'
Gly225Ala	Sense	5'-TTC TGG GGC TGG <u>GCG</u> CGC GAG GAC GAC-3'
	Antisense	5'-GTC GTC CTC GCG <u>CGC</u> CCA GCC CCA GAA-3'
Glu227Ala	Sense	5'-GGC TGG GGC CGC <u>GCG</u> GAC GAC GAC TTC-3'
	Antisense	5'-GAA CTC GTC GTC <u>CGC</u> GCG GCC CCA GCC-3'
Glu227Asp	Sense	5'-GGC TGG GGC CGC <u>GAC</u> GAC GAC GAC TTC-3'
	Antisense	5'-GAA CTC GTC GTC <u>GTC</u> GCG GCC CCA GCC-3'
Asp228Ala	Sense	5'-TGG GGC TGG GGC CGC GAG <u>GCA</u> GAC GAG TTC TAC CGG CGC-3'
	Antisense	5'-GCG CCG GTA GAA CTC GTC <u>TGC</u> CTC GCG GCC CCA GCC CCA-3'
Asp228Glu	Sense	5'-TGG GGC TGG GGC CGC GAG <u>GAG</u> GAC GAG TTC TAC CGG CGC-3'
	Antisense	5'-GCG CCG GTA GAA CTC GTC <u>CTC</u> CTC GCG GCC CCA GCC CCA-3'
Asp229Ala	Sense	5'-GGC TGG GGC CGC GAG GAC <u>GCA</u> GAG TTC TAC CGG CGC ATT-3'
	Antisense	5'-AAT GCG CCG GTA GAA CTC <u>TGC</u> TGC CTC GCG GCC CCA GCC-3'
Asp229Glu	Sense	5'-GGC TGG GGC CGC GAG GAC <u>GAG</u> GAG TTC TAC CGG CGC ATT-3'
	Antisense	5'-AAT GCG CCG GTA GAA CTC <u>CTC</u> TGC CTC GCG GCC CCA GCC-3'
Glu230Ala	Sense	5'-CGC GAG GAC GAC <u>GCG</u> TTC TAC CGG CGC-3'
	Antisense	5'-GCG CCG GTA GAA <u>CGC</u> GTC GTC CTC GCG-3'
Glu230Asp	Sense	5'-CGC GAG GAC GAC <u>GAC</u> TTC TAC CGG CGC-3'
	Antisense	5'-GCG CCG GTA GAA <u>GTC</u> GTC GTC CTC GCG-3'

**S-T2**

Protein accession numbers, species name and sequence length of the 48 members of the family 7 of  $\beta$ 1,4-galactosyltransferases (B4GT7) and the six related B4GT1\_HUMAN to B4GT6\_HUMAN enzymes used as outgroup. Sequences are in the same order as in Fig. 1 and 2.

Accession number	Species name	Sequence length (amino acid number)
B4GT7_HUMAN	<i>Homo sapiens</i>	327
NP_001035053	<i>Pan troglodytes</i>	327
XP_001093051	<i>Maccaca mulatta</i>	327
Q4FZU7	<i>Rattus norvegicus</i>	327
Q8R087	<i>Mus musculus</i>	327
NP_001162105	<i>Ovis aries</i>	327
XP_584194	<i>Bos taurus</i>	327
NP_001161894	<i>Sus scrofa</i>	327
XP_538566	<i>Canis familiaris</i>	316
XP_001502185	<i>Equus caballus</i>	327
XP_001371123	<i>Monodelphis domestica</i>	319
NP_001035000	<i>Gallus gallus</i>	319
XP_002196772	<i>Taeniopygia guttata</i>	322
NP_001088501	<i>Xenopus laevis</i>	323
BC158315	<i>Silurana tropicalis</i>	322
CBH40225	<i>Rana catesbeiana</i>	319
NP_001072098	<i>Takifugu rubripes</i>	317
CAJ77199	<i>Tetraodon nigroviridis</i>	317
NP_001098251	<i>Oryzias latipes</i>	316
AAH92934	<i>Danio rerio</i>	317
CBH40227	<i>Squalus acanthias</i>	291
XP_002155259	<i>Hydra magnipapillata</i>	348
XP_002610451	<i>Branchiostoma floridae</i>	313
NP_651319	<i>Drosophila melanogaster</i>	322
CAJ77193	<i>Drosophila simulans</i>	322
CAJ77192	<i>Drosophila sechellia</i>	321
XP_001981894	<i>Drosophila erecta</i>	322
XP_002099136	<i>Drosophila yakuba</i>	322
XP_001955270	<i>Drosophila ananassae</i>	318
XP_002073576	<i>Drosophila willistoni</i>	319
XP_001996949	<i>Drosophila grimshawi</i>	310
XP_002054247	<i>Drosophila virilis</i>	308
CAJ77191	<i>Drosophila mojavensis</i>	311
CBH40223	<i>Glossina morsitans</i>	311
XP_001653000	<i>Aedes aegypti</i>	313
XP_001865270	<i>Culex quinquefasciatus</i>	304
XP_001603688	<i>Nasonia vitripennis</i>	314
XP_624051	<i>Apis mellifera</i>	294
XP_002433968	<i>Ixodes scapularis</i>	290
XP_001944650	<i>Acyrtosipon pisum</i>	302
XP_973612	<i>Tribolium castaneum</i>	308
XP_002429945	<i>Pediculus humanus corporis</i>	294
CAJ77188	<i>Ciona savigny</i>	325
XP_002121762	<i>Ciona intestinalis</i>	319
XP_002641643	<i>Caenorhabditis briggsae</i>	320
NP_499164	<i>Caenorhabditis elegans</i>	289
CAJ77198	<i>Schistosoma mansoni</i>	327
AAW24494	<i>Schistosoma japonicum</i>	324
B4GT1_HUMAN	<i>Homo sapiens</i>	398
B4GT2_HUMAN	<i>Homo sapiens</i>	372
B4GT3_HUMAN	<i>Homo sapiens</i>	393
B4GT4_HUMAN	<i>Homo sapiens</i>	344
B4GT5_HUMAN	<i>Homo sapiens</i>	388
B4GT6_HUMAN	<i>Homo sapiens</i>	382

**S-T3**

**Hydrogen bonds between human  $\beta$ 4GalT7 protein modeled from bovine  $\beta$ 4GalT1 structure and the donor substrate UDP-Gal. UDP-Gal atoms are named according to the PDB nomenclature.**

<b>Proton acceptor</b>	<b>Proton donor</b>
UDP-Gal-O2A	Arg <sup>104</sup> -N $\eta$ 1
UDP-Gal-O4'	Tyr <sup>199</sup> -O $\eta$
UDP-Gal-O1A	Arg <sup>104</sup> -N $\eta$ 1
UDP-Gal-O2*	Val <sup>164</sup> -N
UDP-Gal-O1A	Arg <sup>104</sup> -N $\eta$ 2
UDP-Gal-O2	Arg <sup>102</sup> -N
UDP-Gal-O3*	Asp <sup>165</sup> -N
UDP-Gal-O1B	Trp <sup>224</sup> -N $\epsilon$ 1
UDP-Gal-O3A	Trp <sup>263</sup> -N $\epsilon$ 1
UDP-Gal-O3*	Arg <sup>104</sup> -N $\eta$ 2
UDP-Gal-O5*	Arg <sup>104</sup> -N $\eta$ 1
Asp <sup>163</sup> -O $\delta$ 1	UDP-Gal-O3*
Asp <sup>165</sup> -O $\delta$ 2	UDP-Gal-O2'
Pro <sup>100</sup> -O	UDP-Gal-O2*

**S-T4**

Hydrogen bonds between human  $\beta$ 4GalT7 protein modeled from *Drosophila*  $\beta$ 4GalT7 structure and the donor substrate UDP-Gal. UDP-Gal atoms are named according to the PDB nomenclature.

Proton acceptor	Proton donor
UDP-Gal-O1B	Arg <sup>266</sup> -N $\epsilon$
UDP-Gal-O1B	Arg <sup>224</sup> -N $\epsilon$ 1
UDP-Gal-O6'	Gly <sup>225</sup> -N
UDP-Gal-O2B	Arg <sup>266</sup> -N $\eta$ 2
UDP-Gal-O3'	Tyr <sup>199</sup> -O $\eta$
UDP-Gal-O1A	Arg <sup>104</sup> -N $\eta$ 2
UDP-Gal-O2*	Arg <sup>102</sup> -N
UDP-Gal-O3'	Gly <sup>202</sup> -N
UDP-Gal-O1A	Arg <sup>104</sup> -N $\epsilon$
UDP-Gal-O1B	Tyr <sup>194</sup> -O $\eta$
UDP-Gal-O2'	His <sup>191</sup> -N $\epsilon$ 2
UDP-Gal-O2A	Arg <sup>104</sup> -N $\eta$ 2
UDP-Gal-O2'	Tyr <sup>199</sup> -O $\eta$
Asp <sup>227</sup> -O $\epsilon$ 1	UDP-Gal-O4'
Pro <sup>100</sup> -O	UDP-Gal-O3*
Tyr <sup>199</sup> -O $\eta$	UDP-Gal-O2'
Asp <sup>163</sup> -O $\delta$ 2	UDP-Gal-O3'

## LEGENDS OF SUPPLEMENTAL DATA FIGURES

**S-F1.** Protein sequence alignment of bovine  $\beta$ GalT1 (PDB entry code 1O0R) with human  $\beta$ GalT7 used to generate the model of the 3D structure of human  $\beta$ GalT7 (upper panel). Protein sequence alignment of human with *Drosophila*  $\beta$ GalT7 (PDB entry 3LW6) used to generate the model of the 3D structure of human  $\beta$ GalT7 (lower panel). Identical amino acids between sequences are indicated by asterisks.

**S-F2.** Stereo view of the molecular model of the human  $\beta$ GalT7 structure in complex with the donor substrate based on bovine  $\beta$ GalT1. The UDP-Gal binding domain was modeled by energy minimization and molecular dynamics calculations based on the structure of bovine  $\beta$ GalT1 (PDB entry code 1O0R). Representation of the interactions of D<sup>163</sup>VD<sup>165</sup> motif (A), and of F<sup>221</sup>WGWGREDE<sup>230</sup> (B) region with UDP-Gal molecule. The  $\alpha$ -carbon trace of the protein is shown in green. UDP-Gal is presented in colored "sticks", and the Mn<sup>2+</sup> divalent cation is colored in cyan. The residues of the D<sup>163</sup>VD<sup>165</sup> and F<sup>221</sup>WGWGREDE<sup>230</sup> motifs are shown in "ball and sticks". For simplicity, hydrogen atoms and atoms of the protein main chain other than  $\alpha$ -carbons are not shown. Dashed lines indicate the predicted interactions of UDP-Gal-Mn<sup>2+</sup> with the side chain of Asp<sup>163</sup> and Asp<sup>165</sup> residues in panel A, and between the  $\beta$ -phosphate group of UDP-Gal and Trp<sup>224</sup> in panel B. The atom coordinates are those of the average structure obtained during the production phase of molecular dynamics calculation, as indicated in "Experimental Procedures". Figures are drawn with PyMOL molecular visualization system (<http://www.pymol.org/>).

**S-F1**

```

                                **      *
Bovine      GRYTPMDCISPHKVAIIIP 187
Human      EHWEEDASWGPRLAVLVP 101

* * * * *      * *      * * * * *      * * * * *
Bovine      FRNRQEHLYWLYLHPILQRQQLDYGIYVINQAGESMFNRAKLLNVGFK 237
Human      FRERFEELLVFPVPHMRRFLSRKKIRHHIYVLNQVDHFRFNRAALINVGFL 150

*      * * * * *      *      * *      *
Bovine      EALKDYDYNCFVFSVDVLI PMNDHNTYRCFSQ-PRHISVAMDKFGFSLPY 286
Human      ESSNS--TDYIAMHDVDLLPLNEELDYGFPEAGPFHVASP--ELHPLYHY 196

* * *      * * *      * * * * *      * *      *
Bovine      VQYFGGVSALSQQFLSINGFPNNYWGWGGEDDDIYNRLAFRGMSVSRPN 336
Human      KTYVGGILLLSKQHYRLCNGMSNRFWGWGREDEFYRRIKGAGLQLFRPS 246

*      * * *      *      * * *      * *
Bovine      AVIGKTRMIRHSRDKKNEP-NPQRFDRIAHTKETMLSDG-LNSLTYMVLE 384
Human      GITTGYKTFRHLHDPAWRKRDKQRIAAQKQEQFKVDREGGLNTVKYHVAS 296

                                *
Bovine      VQRYPL-----YTKITVDIGTPS 402
Human      RTALSVGGAPCTVLNIMLDCDKTA 320

                                * * * * *
Drosophila      GASVHKMALLVP 82
Human      SWGPRLAVLVP 100

** * * * * *      * *      * * * * *      * * * * *
Drosophila      FRDRFEELLQFVPHMTAFLKRQGVAAHHIFVLNQVDRFRFNRAALINVGFK 132
Human      FRERFEELLVFPVPHMRRFLSRKKIRHHIYVLNQVDHFRFNRAALINVGFL 150

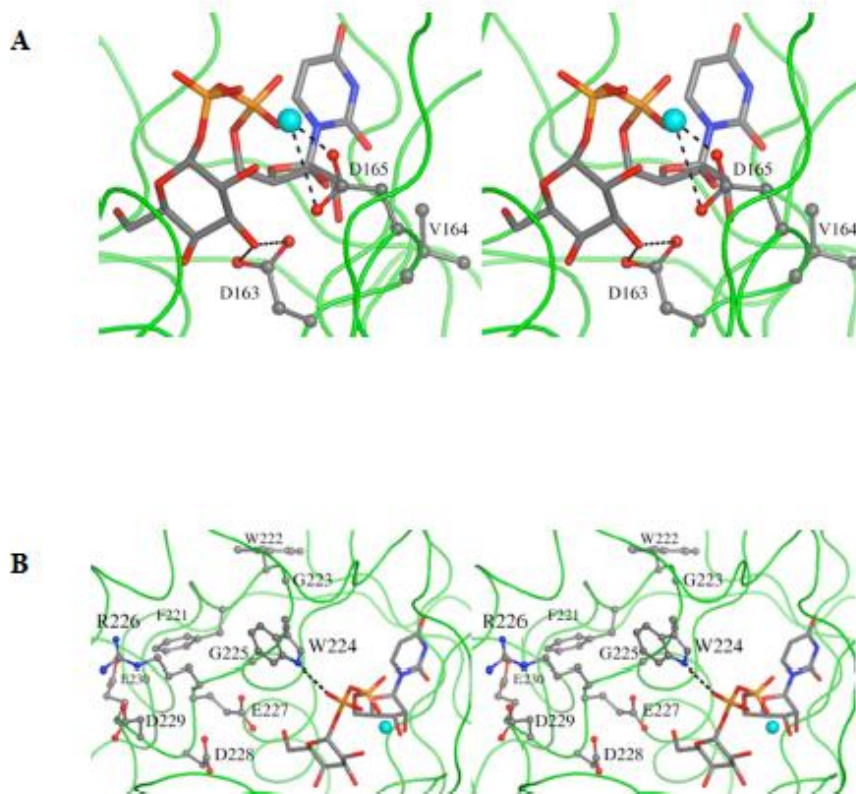
*      * * * * * * * * * *      * * * * *
Drosophila      FASDVYDYIAMHDVDLLPLNDNLLYEYPSLGLPLHIAGPKLHPKYHYDNF 182
Human      ESSNSTDYIAMHDVDLLPLNEELDYGFP-EAGPFHVASPELHPLYHYKTY 199

* * * * *      *      * * * * *      * * * * *      * *
Drosophila      VGGILLVRRREHFQMNGMSNQYWGWGLEDEFFVRI RDAGLQVTRPQNIK 232
Human      VGGILLLSKQHYRLCNGMSNRFWGWGREDEFYRRIKGAGLQLFRPSGIT 249

** * * * * *      * * * * *      * *      * * * * *
Drosophila      TGTNDTFSHIHNRVHRKRDTQKCFNQKEMTRKRDHKTGLDNVVKFKILKVH 282
Human      TGYK-TFRHLHDPAWRKRDKQRIAAQKQEQFKVDREGGLNTVKYHVASRT 298

* * * * *      * * * * *
Drosophila      EMLIDQVPVTILNILLDCDVNKPWCDCS 311
Human      ALSVGGAPCTVLNIMLDCDKTATPWCTFS 327

```



# Caractérisation structurale et fonctionnelle de la $\beta$ 4GalT7 humaine et exploration du site actif de l'enzyme

Mon sujet de thèse a porté sur la caractérisation des enzymes impliquées dans l'initiation de la synthèse et la modification des chaînes de GAGs. Le principal axe de cette étude a consisté à déterminer l'organisation structurale et fonctionnelle de la  $\beta$ 4GalT7 humaine afin d'identifier les résidus d'acides aminés qui jouent un rôle clé dans la reconnaissance des substrats et/ou la catalyse enzymatique. Pour atteindre ces objectifs, nous avons systématiquement réalisé des mutations conservatives et non conservatives des acides aminés situés dans des domaines hautement conservés au sein de la famille des  $\beta$ 4GalTs et qui ont été identifiés comme potentiellement impliqués dans l'organisation du site actif et/ou la catalyse enzymatique. L'effet de chaque mutation sur l'activité GalT et les propriétés cinétiques de l'enzyme a été étudié *in vitro*. Nous avons également analysé la spécificité de la  $\beta$ 4GalT7 sauvage et mutée vis-à-vis de différents substrats donneurs et accepteurs. Afin de compléter cette étude, j'ai analysé, *ex vivo*, les conséquences des mêmes mutations sur l'initiation de la synthèse des GAGs à partir d'un substrat xylosidique exogène et de la protéine "core" de la décorine humaine recombinante, comme substrat accepteur endogène.

Cette étude a fait l'objet d'une publication (**Publication n°1**).

IDENTIFICATION OF KEY FUNCTIONAL RESIDUES IN THE ACTIVE SITE OF HUMAN  $\beta$ 1,4-GALACTOSYLTRANSFERASE 7: A MAJOR ENZYME IN THE GLYCOSAMINOGLYCAN SYNTHESIS PATHWAY

**Ibtissam Talhaoui**, Catherine Bui, Rafael Oriol, Guillermo Mulliert, Sandrine Gulberti, Patrick Netter, Michael W.H. Coughtrie, Mohamed Ouzzine and Sylvie Fournel-Gigleux.

*J. Biol. Chem.* (2010) 285 (sous presse)

## I. ANALYSE PHYLOGENETIQUE DE LA FAMILLE DES $\beta$ 4GalT7s

La  $\beta$ 4GalT7 appartient à la famille des  $\beta$ 4GalTs qui contient six autres membres. L'exploration des bases des données par BLAST a permis d'établir un arbre phylogénétique qui montre la présence d'une relation évolutive entre les membres de la famille des  $\beta$ 4GalTs humaines et entre les  $\beta$ 4GalT7s de 47 autres espèces différentes et la  $\beta$ 4GalT7 humaine (Publication n°1, Figure 1). L'alignement de séquences des différentes protéines  $\beta$ 4GalT7s



identifiées par l'étude phylogénétique, met en évidence la présence de 8 domaines peptidiques fortement conservés entre les espèces et montre plus particulièrement deux motifs, les motifs 3 et 5 qui sont conservés au sein de toutes les  $\beta$ 4GalTs (Publication n°1, Figures 1 et 2).

De façon complémentaire, nous avons développé un modèle moléculaire de la  $\beta$ 4GalT7 basé sur la structure 3D de la  $\beta$ 4GalT1 bovine (code PDB 1O0R) et suite à l'élucidation de la structure 3D de la  $\beta$ 4GalT7 de drosophile (code PDB 3LW6), un 2<sup>ème</sup> modèle moléculaire de l'enzyme humaine basé sur cette dernière a été proposé (Publication n°1, Figure 3). Cette approche a permis la prédiction d'acides aminés potentiellement impliqués dans la reconnaissance des substrats et/ou la catalyse enzymatique. L'ensemble de ces stratégies suggère notamment un rôle potentiel des résidus d'acides aminés des motifs 3 et 5 dans l'organisation du site de fixation de l'UDP-Gal, substrat donneur commun à l'ensemble des membres de cette famille.

## **II. EXPLORATION DU ROLE DES ACIDES AMINES DES MOTIFS <sup>163</sup>DVD<sup>165</sup> ET <sup>221</sup>FWGWGREDD<sup>230</sup> PAR MUTAGENESE DIRIGEE**

### **II.1. Activité galactosyltransférase *in vitro* de la $\beta$ 4GalT7 recombinante humaine et propriétés cinétiques de l'enzyme sauvage et des mutants**

Dans un premier temps, nous avons procédé à la mutagenèse dirigée systématique par PCR des acides aminés des deux motifs sélectionnés (3 et 5), dans l'objectif de déterminer le rôle de chaque résidu. La protéine sauvage et les mutants ont été surexprimés dans un système d'expression hétérologue basé sur la transfection de cellules eucaryotes HeLa en culture. Ainsi, j'ai pu produire des protéines recombinantes humaines à des taux équivalents (Publication n°1, Figure 4A) et comparer l'activité de l'enzyme sauvage et des mutants.

Au sein du domaine 3, nous avons exploré le rôle des résidus aminés du motif <sup>163</sup>DVD<sup>165</sup>. Pour ce faire, j'ai réalisé des substitutions conservatives et non conservatives des Asp163 (D163) et Asp165 (D165) qui ont été remplacés par un résidu Glu (E) ou Ala (A). La mesure de l'activité enzymatique montre que toute mutation de l'acide aminé en position 163 abolit complètement l'activité GalT de la  $\beta$ 4GalT7 recombinante humaine (Publication n°1, Figure 4B). La mutation de l'Asp165 en Ala produit le même effet, tandis que la substitution D165E restaure partiellement l'activité enzymatique par comparaison au mutant D165A, avec une activité qui est de l'ordre de 47% par rapport à la protéine sauvage. Enfin, le remplacement de la Val en position 164 par une Ala n'affecte pas l'activité de la  $\beta$ 4GalT7 humaine (Publication n°1, Figure 4B). A l'issue de ces résultats, nous avons déterminé les constantes cinétiques

apparentes de la protéine sauvage et des mutants actifs, en présence d'UDP-Gal comme substrat donneur et du 4-MOX comme substrat accepteur. Le mutant V163A présente des constantes cinétiques apparentes comparables à celles de l'enzyme sauvage pour les deux substrats, tandis que le mutant D165E montre une forte augmentation des deux  $K_m$ , suggérant une diminution de l'affinité de cette enzyme pour l'UDP-Gal et le 4-MOX, (Publication n°1, Tableau 1). Ce résultat est en accord avec le rôle du motif DVD (région 3) prédit par la structure 3D des  $\beta$ 4GalTs (-T1 bovine et -T7 de drosophile) et par les deux modèles moléculaires de la  $\beta$ 4GalT7 humaine. Ainsi, ces derniers suggèrent que l'Asp163 formerait une liaison hydrogène avec l'oxygène en position 3 du Gal et l'Asp165 formerait une liaison de coordination, par son groupement carboxylique, avec les cations divalents  $Mn^{2+}$ . De plus, il est suggéré qu'à l'inverse des acides aminés Asp163 et Asp165, la Val164 n'entre pas en interaction directe ou indirecte avec le substrat donneur (Publication n°1, Figure 3A).

En second lieu, nous avons exploré par la même approche, le rôle fonctionnel des acides aminés situés dans le domaine 5. J'ai réalisé une série de mutations du motif <sup>221</sup>**FWGWGREDD**<sup>230</sup>. J'ai également initié l'étude du motif <sup>257</sup>**H**XH<sup>259</sup> (région 6), dont l'importance fonctionnelle a été suggérée sur la base de la structure 3D de la  $\beta$ 4GalT7 de drosophile. La substitution du résidu Phe221 (F221) par un résidu Ala n'affecte pas l'activité enzymatique du mutant (Publication n°1, Figure 4C). L'analyse de l'effet de cette mutation sur les paramètres enzymatiques montre une augmentation du  $K_m$  de la protéine vis-à-vis du substrat accepteur 4-MOX, ce qui suggère un rôle potentiel du résidu Phe221 dans la reconnaissance du substrat accepteur et/ou l'organisation du site de fixation de ce substrat. D'un autre côté, la substitution conservative de l'acide aminé Phe221 par un résidu Tyr (Y) ne change ni l'activité de l'enzyme ni les constantes cinétiques vis-à-vis des deux substrats. Ces constatations permettent de conclure qu'une Tyr en position 221 peut remplacer la chaîne latérale de la Phe (Publication n°1, Figure 4C). Le remplacement du résidu Trp222 par une Ala provoque une très forte diminution de l'activité enzymatique, en revanche, sa substitution par une Phe produit un mutant présentant une activité comparable à celle de la  $\beta$ 4GalT7 humaine sauvage. De plus, cette mutation ne modifie pas les constantes cinétiques de manière significative. Ce résultat est en accord avec le modèle moléculaire qui a prédit que le résidu Trp222 est situé loin du centre catalytique de la protéine. La substitution des Gly (G) 223 et 225 par une Ala provoque, respectivement, une diminution et une absence totale de l'activité GalT *in vitro*. Le résidu Trp224 a attiré une attention particulière. En effet, la structure 3D de la  $\beta$ 4GalT1 bovine, de la  $\beta$ 4GalT7 de la drosophile, et le modèle moléculaire de la structure de

la  $\beta$ 4GalT7 humaine indiquent une position clé de cet acide aminé dans le site actif de l'enzyme. Effectivement, il est proposé que le Trp224 interagirait avec le groupement  $\beta$ -phosphate du substrat donneur (Publication n°1, Figure 3B).

Mes résultats montrent que le mutant W224A exprimé dans les cellules HeLa ne présente pas d'activité GalT vis-à-vis de la 4-MOX *in vitro*. La mutation de ce résidu en Phe ou en His permet une restauration modérée de l'activité enzymatique de l'ordre de 9% et 33%, respectivement. L'analyse des constantes cinétiques apparentes du mutant W224H indique des modifications importantes par rapport à la protéine sauvage. En effet, on constate une diminution d'un facteur de 10 du  $K_m$  vis-à-vis de l'UDP-Gal, suggérant une augmentation importante de l'affinité de l'enzyme pour son substrat donneur (UDP-Gal). Cela est accompagné d'une augmentation du  $K_m$  de ce mutant vis-à-vis du substrat accepteur (4-MOX). Ce résultat confirme l'importance fonctionnelle du résidu Trp224 et son implication potentielle à la fois dans l'interaction avec les substrats donneur et accepteur.

J'ai poursuivi cette étude par l'exploration de la deuxième partie du motif **<sup>221</sup>FWGWGREDE<sup>230</sup>**, qui est une région constituée exclusivement d'acides aminés acides. Ce type de résidu aminé (Asp ou Glu) est souvent impliqué dans la catalyse enzymatique des GTs. Nous avons observé que le remplacement du Glu227 et de l'Asp228 de la  $\beta$ 4GalT7 recombinante humaine par une Ala provoque une abolition totale de l'activité enzymatique *in vitro*, tandis que les mutants E227D et D228E présentent une activité détectable, bien qu'inférieure à celle de l'enzyme sauvage. D'un autre côté, la détermination des valeurs de  $K_m$  pour ces mutants montre un changement significatif concernant uniquement le mutant D228E avec une augmentation du  $K_m$  vis-à-vis du substrat accepteur. Cette diminution de l'affinité du mutant D228E envers le 4-MOX pourrait s'expliquer par l'encombrement stérique occasionné par la substitution du résidu Asp228 par un Glu, dont la chaîne latérale est plus longue. Les structures cristallographiques de la  $\beta$ 4GalT1 bovine et de la  $\beta$ 4GalT7 de drosophile proposent un rôle catalytique potentiel des acides aminés équivalents aux résidus 227 et/ou 228. La mutation conservative et non conservative du résidu Asp229 produit des enzymes complètement inactives, démontrant ainsi le rôle crucial de cet acide aminé. En revanche, la mutation du Glu230 en Ala ne semble affecter ni l'activité enzymatique ni les propriétés catalytiques de l'enzyme.

Afin de mieux comprendre le rôle moléculaire des résidus Trp224, Glu227 et Asp228 qui semblent présenter une importance fonctionnelle significative, j'ai développé un système d'expression hétérologue de la  $\beta$ 4GalT7 dans la bactérie *E. coli*. L'enzyme délétée des 60

acides aminés N-terminaux a été exprimée en fusion avec la GST. L'optimisation des conditions de culture des bactéries BL21 et de la lyse cellulaire nous a permis de produire des quantités importantes de la  $\beta$ 4GalT7 sous forme soluble, en fusion avec la GST. J'ai mis au point un protocole de purification basé sur l'utilisation de l'affinité de la protéine de fusion GST pour le glutathion permettant de disposer de quantités d'enzyme compatibles avec les études cinétiques.

Afin de confirmer l'effet, caractérisé *in vitro*, des mutations de la  $\beta$ 4GalT7 sur les constantes cinétiques de l'enzyme, j'ai entrepris des purifications de la protéine sauvage et portant une mutation au niveau du Trp224, du Glu227 et de l'Asp228. Ainsi, j'ai produit et purifié des protéines recombinantes solubles, GST-W224A, GST-W224F, GST-W224H, GST-E227A, GST-E227D, GST-D228A, GST-D228E et la  $\beta$ 4GalT7 sauvage exprimée en fusion avec la GST.

Cette approche nous a permis d'obtenir le matériel protéique indispensable à la réalisation de cette partie du projet dédiée à des études cinétiques approfondies. Les quantités de protéines et les activités enzymatiques évaluées à chaque étape de la procédure de purification des protéines recombinantes mutées ou sauvage sont résumées sous forme de tableau (Tableau 7).

De plus, nous avons pu envisager la réalisation d'études physicochimiques de la  $\beta$ 4GalT7 sauvage et des enzymes portant une mutation au niveau des acides aminés identifiés comme potentiellement impliqués dans l'interaction avec les substrats ou dans la catalyse enzymatique. En effet, nous avons initié un projet de caractérisation structurale et fonctionnelle de la  $\beta$ 4GalT7 plus approfondie, basé principalement sur l'utilisation de la technique de résonance plasmonique de surface sur Biacore. Enfin, nous avons lancé des premières tentatives de cristallisation de la  $\beta$ 4GalT7 humaine en vue de l'élucidation de sa structure 3D en collaboration avec nos collègues grenoblois.

	Type de protéines	Protéines totales (mg)	Activité enzymatique totale (nmole/min)	Activité enzymatique spécifique (nmole/min/mg)	Facteur de purification	Rendement %
<b>GST-β4GalT7</b>	<i>Lysat de bactéries</i>	117.12	18774.0	160.30	1	100
	<i>Protéines solubles</i>	3.71	1454.70	392.40	2.5	7.75
	<i>Protéine purifiée</i>	0.48	822.70	1720.0	10.4	4.40
<b>GST-W224H</b>	<i>Lysat de bactéries</i>	118.45	267.70	2.26	1	100
	<i>Protéines solubles</i>	3.50	64.75	18.50	8.2	24.19
	<i>Protéine purifiée</i>	0.5395	66.70	123.64	54.7	24.90
<b>GST-E227D</b>	<i>Lysat de bactéries</i>	132.21	1072.20	8.11	1	100
	<i>Protéines solubles</i>	4.51	102.74	22.78	2.8	9.58
	<i>Protéine purifiée</i>	0.6642	57.97	87.28	10.8	5.40
<b>GST-D228E</b>	<i>Lysat de bactéries</i>	108.40	581.40	8.50	1	100
	<i>Protéines solubles</i>	3.20	46.51	25.50	3	8
	<i>Protéine purifiée</i>	0.5787	23.26	57.33	6.7	4

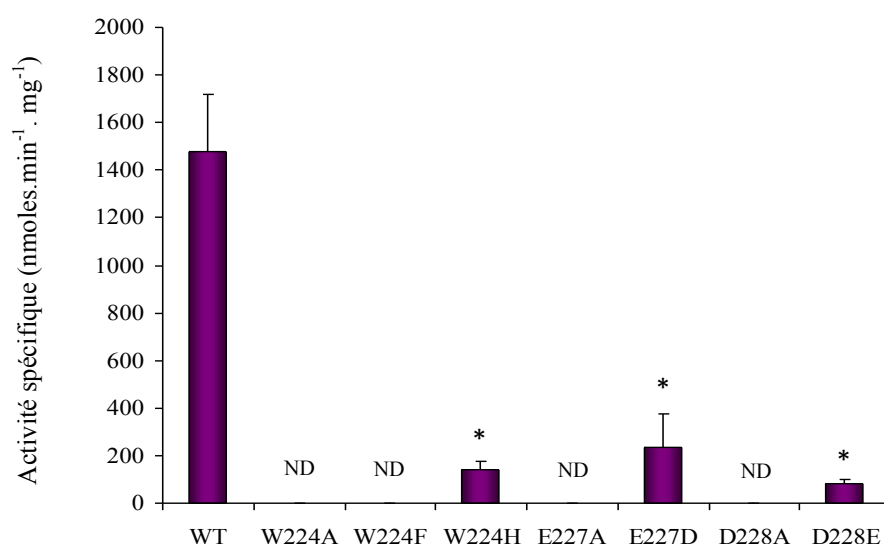
**Tableau 7 :** Tableau de purification de la β4GalT7 recombinante humaine exprimée dans des cellules *E. coli*.

Les protéines sont purifiées à partir d'une culture bactérienne de 200 ml, comme indiqué dans la section "matériels et méthodes". Les mutants GST-W224A, GST-W224F, GST-E227A et GST-D228A étant inactifs, nous n'avons pas pu déterminer le rendement et le facteur de purification.

Les protéines purifiées en fusion avec la GST ainsi obtenues ont été utilisées pour déterminer, *in vitro*, l'activité enzymatique et les constantes cinétiques réelles de la  $\beta$ 4GalT7 sauvage et mutée en position 224, 227 et 228, notamment les valeurs de  $K_m$  et  $k_{cat}$  vis-à-vis des substrats donneur et accepteur (Publication n°1, Tableau 2).

La  $\beta$ 4GalT7 humaine purifiée présente des valeurs de  $K_m$  sensiblement comparables à celles déterminées pour l'enzyme membranaire exprimée dans des cellules HeLa, à la fois pour l'UDP-Gal et le 4-MOX. Ce résultat suggère que la délétion des 60 premiers acides aminés de l'extrémité N-terminale de l'enzyme et la fusion à la GST n'a pas affecté l'organisation du domaine catalytique de la  $\beta$ 4GalT7 humaine, ni ses propriétés cinétiques.

La mesure de l'activité enzymatique des protéines purifiées montrent que les mutants W224H, E227D et D228E sont actifs. En revanche la substitution de chacun des trois acides aminés par une Ala provoque une abolition totale de l'activité GalT de l'enzyme (Figure 35).



**Figure 35. Mesure de l'activité enzymatique de la  $\beta$ 4GalT7 sauvage et mutée en position 224, 227 ou 228, exprimées dans la bactérie et purifiées par chromatographie d'affinité**

Les essais enzymatiques ont été réalisés dans un volume réactionnel de 50  $\mu$ l, en présence de 200 ng de protéines purifiées, 1 mM d'UDP-Gal, (0.05  $\mu$ Ci UDP[U<sup>14</sup>C]Gal), 5 mM de 4-MOX, 100 mM de cacodylate de sodium contenant 10 mM de cations Mn<sup>2+</sup>, comme décrit dans la section "matériels et méthodes". Chaque valeur correspond à la moyenne  $\pm$  SD d'essais en double réalisés sur trois purifications différentes. ND, activité enzymatique non détectée; \*, valeur d'activité significativement différente de l'activité enzymatique mesurée pour l'enzyme sauvage à  $P < 0.05$ .

J'ai poursuivi cette étude par la détermination des constantes cinétiques réelles de l'enzyme purifiée sauvage et des mutants actifs, exprimés dans des bactéries *E. coli*. J'ai démontré que

la substitution du Trp224 par une His augmente l'affinité de l'enzyme envers le substrat donneur UDP-Gal de façon importante, confirmant le résultat obtenu précédemment avec le système d'expression eucaryote. Néanmoins, au cours de la réalisation des cinétiques enzymatiques en vue de la détermination des constantes de l'enzyme W224H vis-à-vis du substrat accepteur 4-MOX, nous avons constaté une augmentation importante de la valeur du  $K_m$  (4-MOX), ce qui n'a pas permis de déterminer la valeur exacte de ce paramètre cinétique de ce mutant, dans les conditions expérimentales utilisées. Ce résultat est en accord avec la diminution de l'affinité du W224H envers le substrat accepteur 4-MOX décrite auparavant. Quant au mutant E227D, il ne montre pas de modifications importantes de l'affinité de l'enzyme vis-à-vis de chacun des deux substrats, tandis que la valeur des  $k_{cat}$  est diminuée de façon significative. Enfin, le remplacement de l'Asp228 par un résidu aminé Glu provoque une baisse importante de l'activité enzymatique. A cela s'ajoute probablement une augmentation de la valeur du  $K_m$  vis-à-vis du substrat accepteur (4-MOX). Par conséquent, nous n'avons pu déterminer les constantes cinétiques de ce mutant dans nos conditions expérimentales.

Cette approche a permis de confirmer les résultats trouvés précédemment, pour les enzymes exprimées sous forme de protéines membranaires dans des cellules HeLa.

## **II.2. Etude de l'effet de la mutation du Trp224 sur la spécificité de substrat de la $\beta$ 4GalT7 humaine**

Les travaux réalisés sur la  $\beta$ 4GalT7 humaine exprimée dans les cellules eucaryotes ou sous forme purifiée suggèrent que le résidu Trp224 est impliqué à la fois dans l'interaction avec le substrat donneur et accepteur. Ce résultat nous a conduit à étudier l'effet de la mutation du Trp en position 224 sur la spécificité de cette enzyme vis-à-vis de substrats de différentes structures.

### **II.2.1. Effet de la mutation W224H sur la spécificité de la $\beta$ 4GalT7 vis-à-vis du substrat donneur**

L'ensemble de l'étude de la spécificité de la  $\beta$ 4GalT7 a été effectué sur les protéines purifiées, exprimées dans la bactérie *E. coli*. Dans une première approche, j'ai tenté d'étudier, par la technique de CCM, la spécificité de la  $\beta$ 4GalT7 recombinante humaine à partir d'homogénats de cellules eucaryotes exprimant l'enzyme sauvage ou mutée (résultats non montrés). J'ai constaté que l'activité enzymatique endogène due à la présence d'autres enzymes dans le lysat cellulaire est supérieure à la quantité du produit de la réaction catalysée par l'enzyme

d'intérêt. C'est-à-dire que pour certains sucres nucléotidiques, l'activité enzymatique évaluée à partir de cellules transfectées avec le plasmide vide, pcDNA, utilisées comme contrôle négatif, était supérieure à celle mesurée à partir de cellules exprimant la  $\beta$ 4GalT7 sauvage ou portant une mutation sur le Trp224. Pour cette raison, la spécificité de la  $\beta$ 4GalT7 a été évaluée sur du matériel purifié.

L'évaluation de l'activité enzymatique de la  $\beta$ 4GalT7 humaine purifiée vis-à-vis de plusieurs substrats donneurs montre que cette enzyme est capable de catalyser le transfert de différents résidus glycosidiques à partir du sucre nucléotidique correspondant. En effet, l'enzyme sauvage est capable de transférer outre le Gal, le Glc et le Xyl à partir des UDP-sucres correspondants, en présence de 4-MOX, bien qu'avec des efficacités beaucoup plus faibles (Publication n°1, Figure 6A). Ceci indique que la  $\beta$ 4GalT7 ne présente pas une spécificité stricte vis-à-vis de l'UDP-Gal. Par contre, nous n'avons pu détecter aucune activité enzymatique avec l'UDP-GalNAc, l'UDP-GlcNAc et l'UDP-GlcA. La substitution du Trp224 par un résidu His augmente la sélectivité de la protéine qui perd toute capacité à utiliser d'autres sucres nucléotidiques que l'UDP-Gal. De plus, de la même façon que l'enzyme sauvage, cette mutation ne permet pas à la  $\beta$ 4GalT7 de prendre en charge les substrats UDP-GalNAc, UDP-GlcNAc et UDP-GlcA.

### **II.2.2. Effet de la mutation W224H sur la spécificité de la $\beta$ 4GalT7 vis-à-vis du substrat accepteur**

Dans la première partie de ce travail, j'ai démontré que la substitution du Trp224 par un résidu His diminue fortement l'affinité de l'enzyme envers le 4-MOX, par rapport à l'enzyme sauvage. Par conséquent, nous avons également cherché à savoir si la mutation de ce résidu aminé modifie la spécificité de la  $\beta$ 4GalT7 humaine vis-à-vis du substrat accepteur. Ainsi, j'ai testé deux groupes de substrats (Figure 36A et 36B) différents de par la nature du résidu glycosidique ou de l'aglycone: 4-MO-Xyl (**4-MOX**), 4-MO-Gal, 4-MO-Glc, 4-MO-GalNAc, 4-MO-GlcNAc et 4-MO-GlcA, et 4-NP-Xyl, MN-Xyl et MN-Xyl-2P.

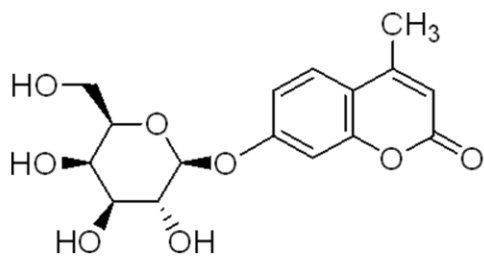
Les résultats obtenus confirment que la  $\beta$ 4GalT7 présente une sélectivité stricte envers le Xyl, aussi bien pour l'enzyme sauvage que pour le mutant W224H. En effet, pour le premier groupe de substrats testés seul le 4-MOX est pris en charge par ces enzymes (Publication n°1, Figure 6B).

Dans un deuxième temps, j'ai mesuré l'activité enzymatique de chacune des deux enzymes en présence de Xyl lié à différents types d'aglycones (Figure 36B). La présence d'un groupement

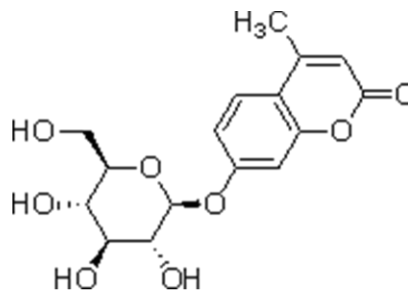


phosphate en position C2 du Xyl provoque une abolition totale de l'activité GalT, ce qui est en accord avec les travaux réalisés précédemment dans notre équipe (Gulberti *et al.*, 2005).

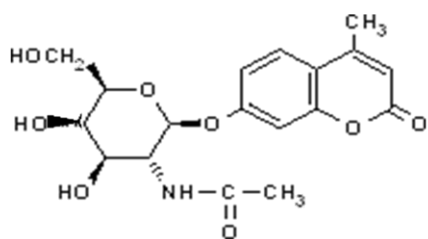
**A**



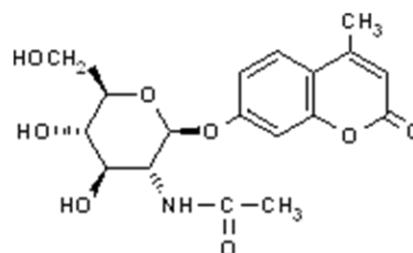
4MO-Gal



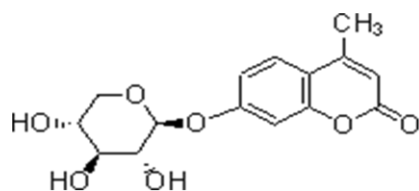
4MO-Glc



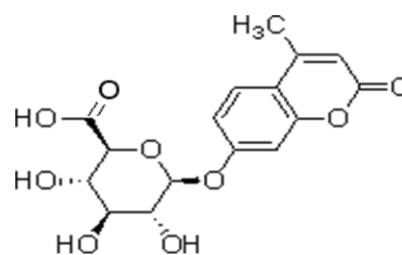
4MO-GalNAc



4MO-GlcNAc

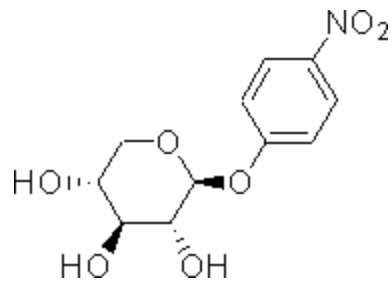


4MO-Xyl

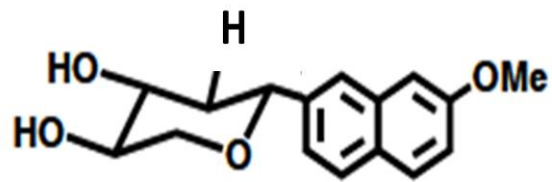


4MO-GlcA

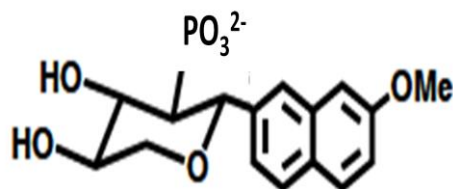
**B**



4-NP-Xyl



MN-Xyl



MN-Xyl-2P

**Figure 36. Structure chimique des substrats accepteurs testés.**

**A.** Différents résidus glycosidiques sont fixés sur la même aglycone (MO). **B.** Différentes aglycones sont fixées sur le Xyl. MN, méthoxynaphtyl.

Ce résultat confirme le rôle clé de la phosphorylation dans la régulation de la synthèse des chaînes de GAGs des PGs. De façon importante, cette étude montre qu'à l'inverse de l'enzyme sauvage, le remplacement du Trp224 par une His produit une protéine qui présente une activité enzymatique plus élevée en présence du 4-MOX qu'en présence du MN-Xyl comme substrat accepteur (Publication n°1, Figure 6B). Ces constatations permettent de suggérer que le résidu 224 de la  $\beta$ 4GalT7 humaine est impliqué dans la reconnaissance du substrat accepteur, et plus particulièrement, avec sa partie aglycone.

### **II.3. Activité galactosyltransférase *ex vivo* de la $\beta$ 4GalT7 humaine sauvage et mutée**

La caractérisation *in vitro* de la  $\beta$ 4GalT7 humaine et d'une série de mutants par la mise en œuvre de deux systèmes d'expression, nous a permis d'identifier des acides aminés qui jouent un rôle crucial dans la reconnaissance des substrats. J'ai démontré par cette approche que l'acide aminé en position 224 est impliqué à la fois dans des interactions avec chacun des deux substrats et que sa substitution par un résidu His pourrait moduler la spécificité de l'enzyme vis-à-vis du substrat accepteur. De plus, le résidu Glu en position 227 et le résidu Asp en position 228 jouent un rôle essentiel dans la  $\beta$ 4GalT7 humaine, c'est pourquoi nous avons cherché à savoir quel est l'impact de la mutation de ces acides aminés sur la capacité de la  $\beta$ 4GalT7 à stimuler la synthèse des chaînes de GAGs *ex vivo*, dans des cellules déficientes CHO-pgsB618.

#### **II.3.1. Substrat accepteur xylosidique exogène: 4-MOX**

Afin d'évaluer l'activité enzymatique de la  $\beta$ 4GalT7 humaine *ex vivo*, j'ai utilisé une méthode consistant à incuber les cellules CHO-pgsB618 recombinantes en culture en présence du substrat xylosidique exogène 4-MOX et de  $\text{Na}_2[^{35}\text{S}]\text{SO}_4$ . Les CHO-pgsB618 sont des cellules qui présentent la particularité d'être déficientes en  $\beta$ 4GalT7 et qui ne synthétisent pas de GAGs endogènes. L'incubation des cellules CHO-pgsB618, préalablement transfectées par l'ADNc codant pour la  $\beta$ 4GalT7 sauvage, en présence de 4-MOX permet d'initier la synthèse de chaînes de GAGs sur le xyloside exogène. Le taux de synthèse a été estimé par la mesure de la quantité de  $[^{35}\text{S}]\text{SO}_4$  incorporée.

Les mutants W224A et W224F ne sont pas capables de restaurer l'activité GalT de la  $\beta$ 4GalT7 après traitement des cellules au 4-MOX, de la même façon que le mutant W224H (Publication n°1, Figure 7A et 7B). Ce dernier résultat est très probablement dû à la diminution de l'affinité de ce mutant envers le substrat 4-MOX par rapport à l'enzyme

sauvage. L'effet de la mutation du résidu Glu227 sur l'initiation de la biosynthèse des GAGs *ex vivo* corrobore les résultats *in vitro*. En effet, le mutant E227D actif *in vitro*, initie la synthèse des GAGs *ex vivo* en présence du substrat xylosidique exogène 4-MOX. En revanche, on constate que le mutant D228E qui présente une activité *in vitro*, est inactif en présence du substrat 4-MOX dans les cellules en culture. (Publication n°1, Figure 7A et 7B).

### **II.3.2. Substrat accepteur endogène: protéine "core" de la décorine**

Comme il a été exposé précédemment, le syndrome progéroïde d'Ehlers-Danlos est une conséquence directe de mutations de la  $\beta$ 4GalT7 chez l'homme. Au niveau moléculaire, cette maladie est caractérisée en particulier par une substitution aberrante des GAGs de la décorine, avec une forte baisse de la synthèse de GAGs sur cette protéine "core". Cette observation nous a conduit à déterminer l'effet de la mutation du Trp224 sur la capacité de l'enzyme à initier la synthèse des GAGs sur la décorine recombinante. Dans cet objectif, j'ai exprimé la décorine humaine recombinante dans les cellules CHO-pgsB618 de façon stable. Ces cellules ont été ensuite transfectées avec la  $\beta$ 4GalT7 humaine sauvage ou mutée en position 224, avant marquage des chaînes de GAGs de la décorine, en présence de soufre radioactif. Enfin, j'ai purifié la décorine recombinante par chromatographie d'affinité basée sur l'interaction entre la séquence poly-His présente sur la protéine "core" de la décorine et les billes de nickel. Les échantillons ainsi obtenus ont été analysés par SDS-PAGE sur gel à gradient, afin de mettre en évidence la présence de chaînes glycosaminoglycaniques radiomarquées (Publication n°1, figure 7C).

Les résultats montrent que l'enzyme sauvage et le mutant W224H sont capables d'initier la synthèse des GAGs sur la décorine. En revanche, aucun signal n'a été détecté dans les cellules exprimant le mutant W224A ou transfectées avec le plasmide vide. En effet, comme attendu, les cellules déficientes en  $\beta$ 4GalT7 ne sont pas capables de former des GAGs sur la décorine. De plus, le mutant W224F présente un signal très modéré, résultat en accord avec la faible activité GalT de ce mutant observée *in vitro*.

### **II.4. Caractérisation *ex vivo* des propriétés fonctionnelles de la $\beta$ 4GalT7 sauvage et des mutants**

Le deuxième axe de cette étude a porté sur la caractérisation de l'effet de l'expression de la  $\beta$ 4GalT7 et des mutants sur la synthèse des GAGs endogènes totaux *ex vivo*. Ainsi, nous avons comparé les deux activités GalT, mesurées *in vitro* et *ex vivo*, de la protéine sauvage et

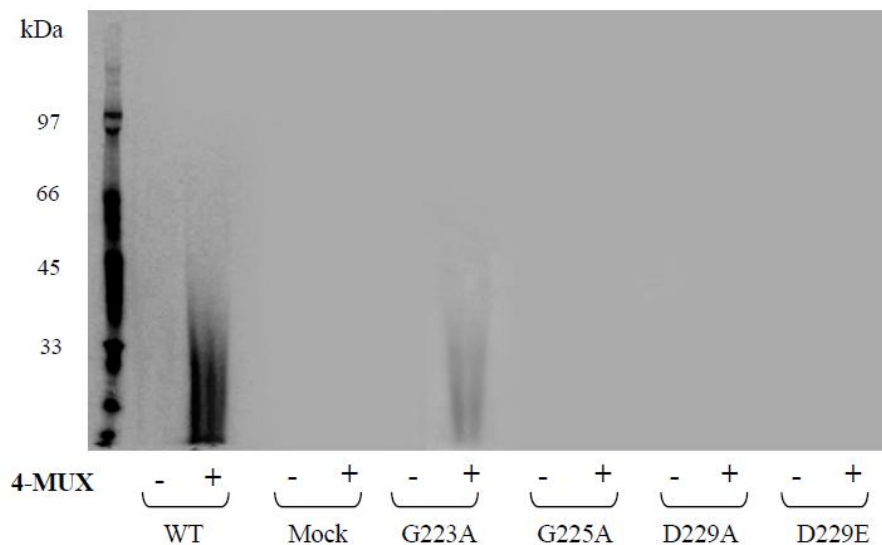
de l'ensemble des mutants générés au cours de cette étude (Publication n°1, Figure 8). L'expression de la  $\beta$ 4GalT7 sauvage dans des cellules CHO-pgsB618 induit une augmentation importante de la synthèse des GAGs, qui se traduit par une augmentation du taux d'incorporation du  $^{35}\text{S}$  d'un facteur de 5, en comparaison avec les cellules déficientes en cette enzyme.

En ce qui concerne le motif  $^{163}\text{DVD}^{165}$ , sa mutation semble avoir des conséquences similaires sur l'activité GalT mesurée *in vitro* et *ex vivo*, confirmant l'importance fonctionnelle des résidus Asp163 et Asp165. La substitution de l'acide aminé Phe221 par une Ala ne permet pas à la  $\beta$ 4GalT7 d'initier la synthèse des GAGs *ex vivo*. Cette mutation a donc des conséquences plus délétères comparées à l'effet constaté *in vitro*. Le mutant F221Y est actif *ex vivo*, ce qui est en accord avec ses propriétés enzymatiques déterminées *in vitro*. La mutation non conservative du Trp222 en Ala provoque une forte baisse de la synthèse des GAGs, comparable à la faible activité mesurée *in vitro*. Cependant, le mutant W222F est incapable de restaurer l'activité GalT *ex vivo*. J'ai également approfondi l'étude du rôle du Trp224 dans l'initiation de la synthèse des GAGs *ex vivo*, et j'ai pu réaffirmer l'importance du rôle fonctionnel de cet acide aminé. Ainsi, de façon concordante aux résultats obtenus *in vitro*, les mutants W224A et W224F sont incapables de synthétiser des chaînes de GAGs. De façon complémentaire, l'évaluation du taux d'incorporation du  $^{35}\text{S}$  *ex vivo* dans des cellules CHO-pgsB618 transfectées avec le mutant W224H montre que cette mutation provoque une diminution drastique de la biosynthèse des GAGs en comparaison avec l'enzyme sauvage et ne permet pas de restaurer, de façon significative, la formation de ces molécules dans des cellules déficientes en  $\beta$ 4GalT7. J'ai également montré que les mutations non conservatives des acides aminés Glu227 et Asp228 ne permettent pas l'initiation de la synthèse des chaînes de GAGs *ex vivo*, ce qui est en accord avec les résultats obtenus *in vitro*. On observe que le mutant E227D, qui présente une activité détectable *in vitro*, ne permet cependant pas de restaurer la synthèse des GAGs totaux de façon significative. Par contre, l'expression du mutant D228E, actif *in vitro*, induit une augmentation importante de la synthèse des GAGs totaux dans les cellules CHO déficientes. Enfin, la substitution du résidu aminé en position 230 de la  $\beta$ 4GalT7 humaine ne semble pas affecter l'activité GalT ni *in vitro* ni *ex vivo*.

De façon inattendue, nous avons constaté que la mutation des Gly en position 223 et 225 ainsi que le résidu Asp229, n'exercent pas le même effet, *in vitro* et *ex vivo*, sur l'activité GalT de la  $\beta$ 4GalT7 recombinante humaine. En effet, la substitution de la Gly223 par une Ala produit une protéine active *in vitro* en présence de 4-MOX comme substrat accepteur. En revanche,

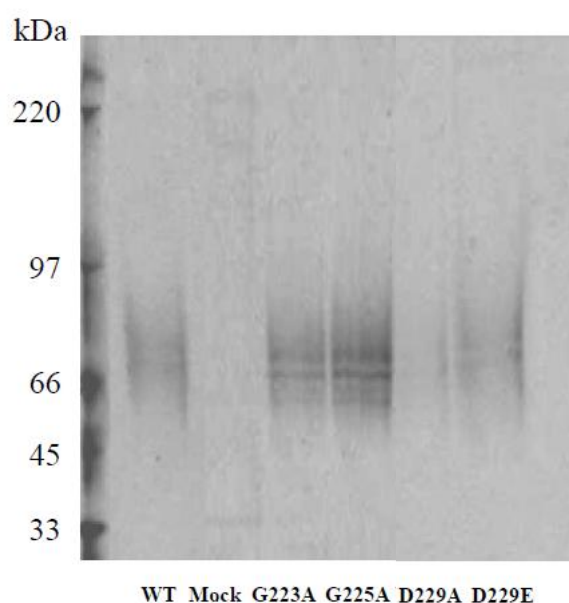
elle est incapable d'initier la synthèse des GAGs *ex vivo*. Ce résultat peut être expliqué par le fait que l'encombrement stérique occasionné par le remplacement de la Gly par l'Ala permet au substrat accepteur exogène de petite taille d'accéder au site actif de l'enzyme, ce qui n'est pas le cas pour le substrat accepteur endogène glycopeptide de taille plus importante. Cela se manifeste donc par l'absence totale de l'activité GalT de la  $\beta$ 4Gal7 humaine *ex vivo*.

Quant aux mutations G225A, D229A et D229E, elles produisent des protéines inactives *in vitro*. De façon intéressante, ces mutants se montrent capables d'initier la synthèse des GAGs *ex vivo*. Afin de pouvoir comprendre ce résultat, nous avons entrepris une approche complémentaire, qui nous a permis de démontrer que ces trois mutants, à l'inverse du G223A, ne peuvent pas synthétiser des GAGs *ex vivo* en présence du 4-MOX comme substrat accepteur, mais restent capable de substituer des chaînes glycosaminoglycaniques sur la décorine et les glycopeptides endogènes (Figure 37 et 38 et Publication n°1, Figure 8).



**Figure 37. Effet de la surexpression de la  $\beta$ 4GalT7 sauvage et mutée sur l'initiation de la biosynthèse des GAGs dans des cellules CHO-pgsB618 traitées au 4-MOX**

Les cellules CHO-pgsB618 ont été transfectées par le plasmide vide ou portant le gène d'intérêt, puis incubées en présence de  $^{35}\text{S}$  pendant 12 h, en absence ou présence de 5  $\mu\text{M}$  de substrat xylosidique 4-MOX (4-MUX). Le  $^{35}\text{S}$  non incorporé a été éliminé sur une colonne de Sephadex G-50, les GAGs radiomarqués présents dans l'éluat ont été analysés par SDS-PAGE, comme décrit dans la section "matériels et méthodes".



**Figure 38. Effet de la surexpression de la  $\beta$ 4GalT7 sauvage et mutée sur la biosynthèse des chaînes de GAGs sur la décorine exprimée dans des cellules CHO-pgsB618**

Les cellules CHO-pgsB618 exprimant la décorine recombinante humaine de façon stable ont été transfectées transitoirement par le plasmide vide ou portant le gène d'intérêt, avant d'être incubées en présence de  $^{35}\text{S}$  pendant 12 h. La décorine recombinante a été purifiée à partir du milieu sur une colonne de nickel, puis analysée par SDS-PAGE comme décrit dans "matériels et méthodes".

L'ensemble de ces résultats nous laisse penser que les mutations de la  $\beta$ 4GalT7 humaine en position 225 et 229, affecte les interactions avec le 4-MOX mais ne semble pas altérer les interactions entre l'enzyme et la protéine "core" des PGs endogènes ou de la décorine.

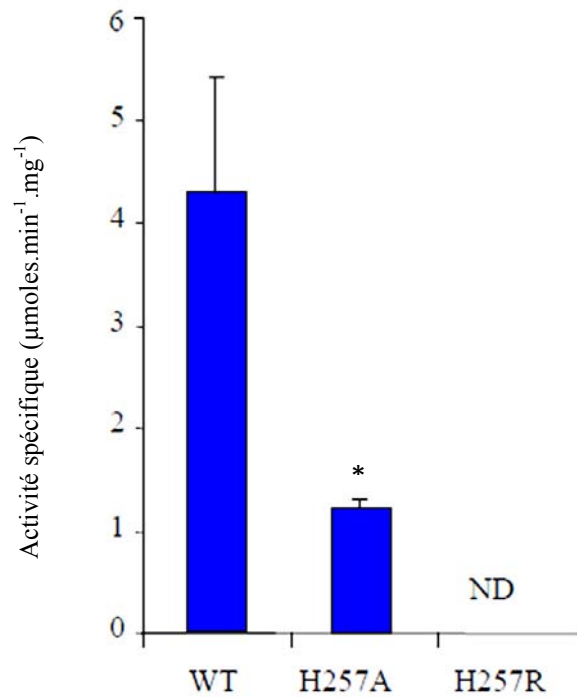
### III. ETUDE DU ROLE DU PREMIER RESIDU HIS (His257) DU MOTIF HXH

Enfin, dans une approche complémentaire, j'ai étudié le rôle du résidu His257 hautement conservé au sein de la famille des  $\beta$ 4GalTs et entre les  $\beta$ 4GalT7s de différentes espèces, situé dans le motif 6 (Publication n°1, Figure 1). Il a été suggéré que cet acide aminé serait impliqué dans l'interaction avec les ions  $\text{Mn}^{2+}$ .

#### III.1. Etudes *in vitro/ex vivo*

La substitution du résidu His257 par une Ala provoque une perte de 80% environ de l'activité enzymatique par rapport à l'enzyme sauvage. De façon intéressante, le remplacement de

l'His257 par une Arg abolit l'activité GalT de façon complète. Ce résultat est probablement dû à l'encombrement stérique occasionné par le groupement guanidine de l'Arg (Figure 39).



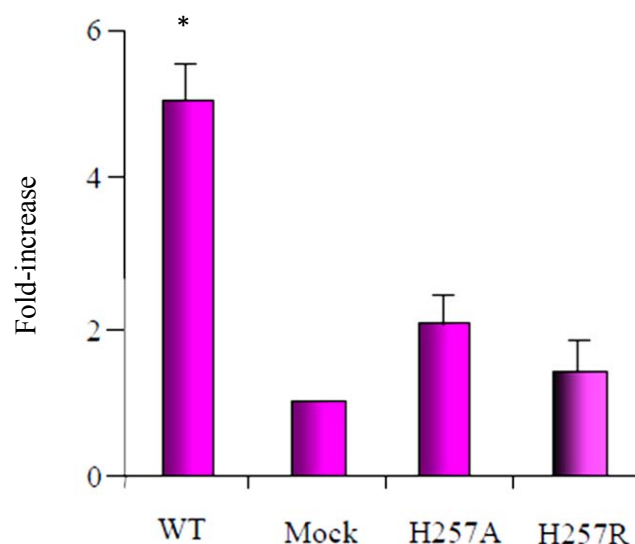
**Figure 39. Détermination de l'effet des mutations conservative et non conservative du résidu His257 sur l'activité GalT *in vitro***

Les essais enzymatiques ont été réalisés avec 40 µg de protéines totales de cellules HeLa transfectées avec le gène codant la β4GalT7 sauvage ou mutée, en présence de 1 mM d'UDP-Gal, (0.05 µCi UDP[U<sup>14</sup>C]Gal), 5 mM de 4-MOX et 100 mM de cacodylate de sodium contenant 10 mM d'ions Mn<sup>2+</sup>. ND, activité enzymatique non détectée. \*, valeur d'activité significativement différente de l'activité enzymatique mesurée pour l'enzyme sauvage à  $P < 0.05$ .

J'ai poursuivi cette étude par l'analyse des conséquences de la mutation conservative et non conservative du résidu His257 sur la capacité de la β4GalT7 à initier la synthèse des chaînes de GAGs *ex vivo*, dans des cellules CHO-pgsB618.

Dans cet objectif, j'ai exprimé la β4GalT7 et les mutants dans des cellules CHO-pgsB618, puis j'ai évalué le taux d'incorporation du <sup>35</sup>S dans les chaînes des GAGs synthétisées. Le résultat obtenu montre que les deux mutants H257A et H257R sont incapables de restaurer la formation des GAGs *ex vivo* de façon significative (Figure 40), confirmant un rôle important de l'acide aminé en position 257 de la β4GalT7 humaine.





**Figure 40. Détermination de l'effet de la mutation de l'His257 sur l'initiation de la synthèse des chaînes de GAGx *ex vivo***

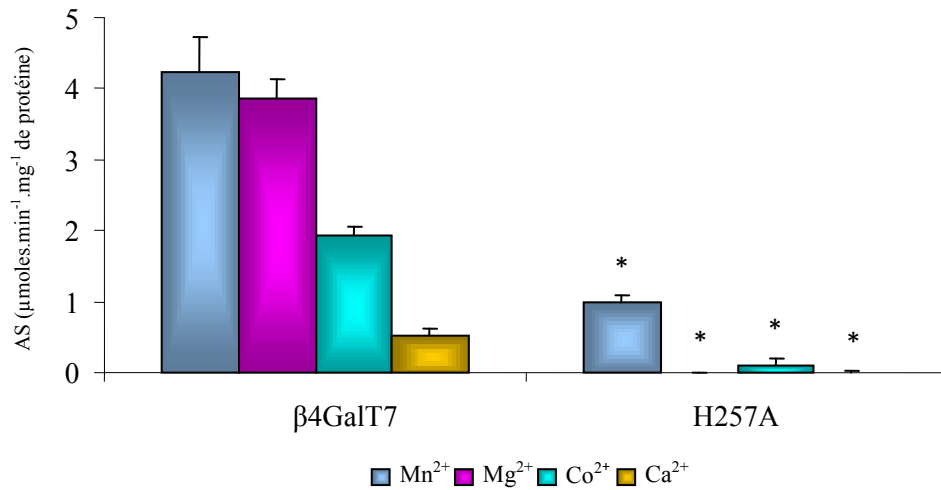
Les cellules CHO-pgsB618 préalablement transfectées par le plasmide vide ou portant le gène d'intérêt, ont été incubées pendant 8 h en présence de 10  $\mu\text{Ci}$  de  $^{35}\text{S}$ . Les GAGs radiomarqués ont été purifiés à partir du milieu comme décrit dans "matériels et méthodes". Les résultats sont présentés sous forme de ratio de la radioactivité associée aux GAGs extraits des cellules transfectées avec le gène codant l'enzyme sauvage ou mutée par rapport à la radioactivité associée aux GAGs purifiés à partir de cellules CHO-pgsB618 transfectées avec le plasmide vide. Chaque valeur correspond à la moyenne de trois expériences indépendantes réalisées en double  $\pm$  SD. \*, valeur significativement différente comparée à la valeur correspondant aux cellules transfectées par le plasmide vide, à  $P < 0.05$ .

### III.2. Activation de la $\beta 4\text{GalT7}$ sauvage et du mutant H257A par les cations divalents

Enfin, j'ai complété cette étude par une analyse de l'activation de la  $\beta 4\text{GalT7}$  par différents cations divalents et de l'effet de la mutation du résidu aminé His257 sur ce mécanisme.

On constate que la  $\beta 4\text{GalT7}$  humaine exprimée dans des cellules eucaryotes peut être activée par différents cations divalents. Ainsi j'ai pu mesurer l'activité GalT de la  $\beta 4\text{GalT7}$  en présence des cations métalliques  $\text{Mn}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Co}^{2+}$  et  $\text{Ca}^{2+}$ . L'activité la plus élevée est obtenue en présence d'ions  $\text{Mn}^{2+}$  suivie des ions  $\text{Mg}^{2+}$ ,  $\text{Co}^{2+}$  et la plus faible activité est mesurée en présence d'ions  $\text{Ca}^{2+}$ . De façon intéressante, le mutant H257A ne semble pas avoir la même préférence envers les cations divalents comparé à l'enzyme sauvage. En effet, cette mutation produit une protéine incapable d'utiliser les cations  $\text{Mg}^{2+}$  et  $\text{Ca}^{2+}$  comme catalyseur de la réaction. Ces résultats confirment que l'acide aminé His257 de la  $\beta 4\text{GalT7}$

humaine est impliqué dans la formation de liaisons de coordination avec les ions  $Mn^{2+}$ , cependant cette étude doit être approfondie pour confirmer ce postulat.



**Figure 41. Activation de la  $\beta 4GalT7$  sauvage humaine et du mutant H257A par les cations divalents**

Les essais enzymatiques ont été réalisés sur des homogénats cellulaires de cellules HeLa, transfectées par le gène codant la  $\beta 4GalT7$  sauvage ou le mutant H257A. Les protéines totales ont été incubées pendant 60 min à 37°C, en présence de 1 mM d'UDP-Gal, (0.05  $\mu Ci$  UDP[ $^{14}C$ ]Gal), 5 mM de 4-MOX, 100 mM de cacodylate de sodium contenant 10 mM de cations divalents ( $Mn^{2+}$ ,  $Mg^{2+}$ ,  $Co^{2+}$  ou  $Ca^{2+}$ ). Les produits de la réaction ont été séparés par CLHP, comme décrit dans la section "matériels et méthodes". Chaque valeur correspond à la moyenne  $\pm$  SD d'essais en double réalisés sur trois transfections différentes. \*, activation de l'enzyme par le cation divalent est significativement différente de l'activation de l'enzyme sauvage à  $P < 0.05$ .

*Publication n° 2*



## Molecular characterization of $\beta$ 1,4-galactosyltransferase 7 genetic mutations linked to the progeroid form of Ehlers–Danlos syndrome (EDS)

Catherine Bui<sup>a,1,2</sup>, Ibtissam Talhaoui<sup>a,1</sup>, Matthieu Chabel<sup>a</sup>, Guillermo Mullier<sup>b</sup>, Michael W.H. Coughtrie<sup>c</sup>, Mohamed Ouzzine<sup>a</sup>, Sylvie Fournel-Gigleux<sup>a,\*</sup>

<sup>a</sup>UMR 7561 CNRS-Université de Nancy 1, Faculté de Médecine, BP 184, 54505 Vandœuvre-lès-Nancy, France

<sup>b</sup>UMR 7036 CNRS-Université de Nancy 1, Faculté de Sciences et Techniques, BP 70239, Vandœuvre-lès-Nancy, France

<sup>c</sup>Division of Medical Sciences, University of Dundee, Ninewells Hospital & Medical School, Dundee DD1 9SY, United Kingdom

### ARTICLE INFO

#### Article history:

Received 3 June 2010

Revised 26 July 2010

Accepted 2 August 2010

Available online 5 August 2010

Edited by Takashi Gajdosh

This work is dedicated to G.J. Dutton who died on the 1st June 2010 and who has been of great inspiration in our study of glucosaminyl- and other glycosyltransferases.

#### Keywords:

Ehlers–Danlos syndrome

Point mutation

Galactosyltransferase defect

Glycosaminoglycan synthesis

Genital glycosyltransferase disorder

### ABSTRACT

$\beta$ 1,4-Galactosyltransferase 7 ( $\beta$ 4GalT7) is a key enzyme initiating glycosaminoglycan (GAG) synthesis. Based on *in vitro* and *ex vivo* kinetics studies and structure-based modelling, we molecularly characterized  $\beta$ 4GalT7 mutants linked to the progeroid form of Ehlers–Danlos syndrome (EDS), a severe connective tissue disorder. Our results revealed that loss of activity upon L206P substitution due to altered protein folding is the primary cause for the GAG synthesis defect in patients carrying the compound A186D and L206P mutations. We showed that R270C substitution strongly reduced  $\beta$ 4GalT7 affinity towards xyloside acceptor, thus affecting GAG chains formation. This study establishes the molecular basis for  $\beta$ 4GalT7 defects associated with altered GAG synthesis in EDS.

© 2010 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

### 1. Introduction

Amino acid exchanges in  $\beta$ 1,4-galactosyltransferase 7 ( $\beta$ 4GalT7, E.C.2.4.1.133), which is involved in the synthesis of the glycosaminoglycan (GAG) linkage region of proteoglycans, are associated with the pathology of progeroid type Ehlers–Danlos syndrome (EDS) [1]. EDS forms a group of inherited connective tissue disorders characterized by defects in various extracellular matrix proteins including collagens and small leucine-rich proteoglycans,

such as decorin. Patients harbouring mutations in the  $\beta$ 4GalT7 gene exhibit aged appearance, developmental delay, dwarfism, craniofacial disproportion and general osteopenia [1,2]. In addition, hypermobile joints, defects in wound healing and loose skin are observed. Aberrant GAG substitution of decorin was initially described in a progeroid EDS patient who carried compound heterozygous amino acid exchanges (A186D/L206P) in  $\beta$ 4GalT7 enzyme [3,4]. More recently, two patients exhibiting typical progeroid EDS features including craniofacial appearance, skeletal abnormalities, and aged appearance were described [2]. Sequence analysis of the  $\beta$ 4GalT7 gene from these patients revealed a missense mutation causing the substitution R270C. This defect was found to be associated with reduced galactosyltransferase activity and aberrant GAG substitution of decorin and biglycan in cultured skin fibroblasts [5]. Interestingly, a recent report indicated that in addition to alterations of the dermatan-sulfate chain of decorin, changes in the structure of heparan-sulfate proteoglycans may also contribute to the phenotypic features observed in this  $\beta$ 4GalT7-deficient form of EDS, such as altered cell migration and delayed wound repair [6].

**Abbreviations:**  $\beta$ 4GalT7,  $\beta$ 1,4-galactosyltransferase 7; EDS, Ehlers–Danlos syndrome; GAG, glycosaminoglycan; Gal, galactose; GlcA, glucuronic acid; 4-MUX, 4-methylumbelliferone- $\beta$ -D-xylopyranoside; PBS, phosphate buffered saline; Xyl, xylose.

\* Corresponding author. Address: Sylvie Fournel-Gigleux, UMR 7561 CNRS-Université de Nancy 1, Faculté de Médecine, BP 184, 54505 Vandœuvre-lès-Nancy, France.

E-mail address: [sf@medecine.univ-nancy.fr](mailto:sf@medecine.univ-nancy.fr) (S. Fournel-Gigleux).

<sup>1</sup> These authors equally contributed to this work.

<sup>2</sup> Present address: Institute of Cellular Medicine, Newcastle University, Medical School, Newcastle upon Tyne NE2 4BQ, United Kingdom.

GAGs are extensively modified linear polysaccharide chains normally attached to core proteins to form proteoglycans. These complex molecules located on the cell surface and in extracellular matrices of virtually every tissue mediate highly diverse key cell events, ranging from mechanical support inside and outside cells to intricate effects on a wide variety of cellular and biochemical processes, such as blood clotting, cell adhesion, differentiation, proliferation and motility [7]. These functions depend on interactions of the GAG chains with a variety of molecules including growth factors, cytokines and their receptors, enzymes such as matrix proteases and coagulation factors, as well as extracellular matrix proteins [8]. The synthesis of GAG chains that govern these functions is initiated by the formation of a tetrasaccharide linkage region composed of glucuronic acid-galactose-galactose-xylose (GlcA $\beta$ 1,3Gal $\beta$ 1,3Gal $\beta$ 1,4Xyl1-) followed by the elongation of the two main types of GAGs (i.e. heparin, heparan-sulfate and chondroitin-/dermatin-sulfate) [9]. The  $\beta$ 4GalT7 enzyme catalyses the transfer of a galactose (Gal) residue provided by UDP- $\alpha$ -D-Gal (UDP-Gal) onto Xyl, a key step in the synthesis of the linkage region of GAG chains [10]. Since the formation of this linker tetrasaccharide is required prior to polymerization of both heparan-sulfate and dermatan-/chondroitin-sulfate chains, genetic defects of  $\beta$ 4GalT7 strongly affect the biological functions mediated by GAGs, including tissue development and differentiation, leading to severe clinical features that are characteristic of the progeria-type variant forms of EDS [1,2].

Although the alterations in GAG synthesis have been well defined from studies performed in fibroblasts of patients affected by progeroid form of EDS, the consequences of these gene mutations on  $\beta$ 4GalT7 properties have not yet been investigated at the molecular level. By site-directed mutagenesis and in vitro and ex vivo kinetic studies, combined to computer-aided modelling of the human protein structure, we deciphered the mechanisms by which the A186D, L206P and K270C mutations affect  $\beta$ 4GalT7 function and disrupt GAG synthesis pathways in the EDS congenital glycosylation disorder.

## 2. Materials and methods

### 2.1. Materials

4-Methylumbelliferyl- $\beta$ -D-xylopyranoside (4-MUX), UDP- $\alpha$ -D-galactose (UDP-Gal) were provided by Sigma-Aldrich, and UDP[ $^{14}$ C]Gal and Na $_2$ [ $^{35}$ S]SO $_4$  was from Perkin-Elmer. The eukaryotic expression vector pcDNA3.1(+) and competent One Shot<sup>TM</sup> Top10 *Escherichia coli* cells were from Invitrogen-Fisher Scientific, and pET-41a(+) vector and *E. coli* BL21(DE3) cells were from Novagen-EMD4Biosciences.

### 2.2. Expression vector construction

The human  $\beta$ 4GalT7 sequence was cloned as previously described [11]. For expression of  $\beta$ 4GalT7 in mammalian cells, the full-length cDNA was modified by PCR to include a KpnI site and a Kozak consensus sequence at the 5' end, and a sequence encoding a myc tag (EQKJEEEDL) and a XbaI site at the 3' end, prior subcloning into the pcDNA3.1(+) to produce pcDNA- $\beta$ 4GalT7. For bacterial expression, a truncated form of  $\beta$ 4GalT7 was expressed as a fusion protein with glutathione-S-transferase (GST). The sequence lacking codons for the 60 N-terminal amino acids was amplified from full-length cDNA using the corresponding primers including NcoI and NotI sites and subcloned into the pET-41a(+) vector (Novagen, EMD4Biosciences) to produce plasmid pET- $\beta$ 4GalT7. Mutations were generated using the QuickChange II XL site-directed mutagenesis kit (Stratagene), employing pcDNA- $\beta$ 4GalT7 or pET-

**Table 1**

Kinetic parameters of wild-type  $\beta$ 4GalT7 and R270C EDS mutant for donor UDP-Gal and the acceptor 4-MUX.

Enzyme	$K_m$ (mM) UDP-Gal	$K_m$ (mM) 4-MUX	$K_{12}$ (mM)	$k_{cat}$ (s $^{-1}$ )
$\beta$ 4GalT7	0.45 $\pm$ 0.09	0.61 $\pm$ 0.08	0.09 $\pm$ 0.03	1.14
R270C	0.46 $\pm$ 0.16	6.12 $\pm$ 0.96	0.29 $\pm$ 0.09	ND

Kinetic parameters towards donor and acceptor substrates were evaluated by double substrate kinetics using 0.2  $\mu$ g purified GST- $\beta$ 4GalT7 protein at concentrations of UDP-Gal and of 4-MUX varying from 0 to 5 mM. Data were fitted to Eq. (1) described in Section 2 by non-linear least square regression analysis using the curve-fitter program of SigmaPlot 9.0. ND, not determined, i.e. no  $k_{cat}$  could not be accurately evaluated for R270C mutant since 4-MUX is poorly soluble at concentration higher than 10 mM required to achieve acceptor substrate saturation.

$\beta$ 4GalT7 plasmids as template, and sense and antisense primers listed in Table 1 (Supplementary data).

### 2.3. Heterologous expression of wild-type and mutated $\beta$ 4GalT7

CHO pgsB-618 cells defective in  $\beta$ 4GalT7 activity established by Esko et al. [12] were purchased from the American Type Culture Collection (ATCC). Cultured cells were individually transfected at 80% confluency with wild-type and mutated pcDNA- $\beta$ 4GalT7 plasmids using ExGen 500 reagent (Euromedex, Souffelweyersheim, France) according to the manufacturer's recommendations. Cells were harvested in phosphate buffered saline (PBS) 48 h after transfection, pelleted by centrifugation at 5000 $\times$ g, resuspended in 0.25 M sucrose, 5 mM HEPES buffer (pH 7.4) and sonicated three-fold 5 s. Protein concentration was determined by the method of Bradford [13] prior to SDS-PAGE or activity analyses. Alternatively, transfected cells were transferred to labelling medium in the presence of 4-MUX prior to GAG isolation, as described below.

To express wild-type and mutated GST- $\beta$ 4GalT7, *E. coli* BL21(DE3) cells harbouring pET- $\beta$ 4GalT7 expression vector were grown in Luria-Bertani (LB) medium and gene expression was induced with 1 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG, Sigma-Aldrich) added to bacteria cultured at 20  $^{\circ}$ C for 16h. GST- $\beta$ 4GalT7 proteins were purified by affinity chromatography using Glutathione Sepharose<sup>TM</sup> 4B columns (GE Healthcare Life Sciences).

### 2.4. Galactosyltransferase activity

The in vitro assay of  $\beta$ 4GalT7 activity was performed as previously described [11]. Reactions were performed in 100 mM sodium cacodylate buffer (pH 7.4) containing 10 mM MnCl $_2$ , 1 mM UDP-Gal, 0.05  $\mu$ Ci UDP[ $^{14}$ C]Gal, 5 mM 4-MUX and 40  $\mu$ g of total cell protein or 0.2  $\mu$ g purified protein. Incubations were carried out at 37  $^{\circ}$ C for 30–60 min in a total volume of 50  $\mu$ l, and reaction products were separated by thin layer chromatography [14,15]. For kinetic studies, reaction products were analyzed by high performance liquid chromatography (HPLC) using a reverse phase C $_{18}$  column attached to a Waters e2695 instrument equipped with a Berthold FlowStar radioactivity monitor.

### 2.5. Kinetic analyses

The kinetic parameters for UDP-Gal and 4-MUX were determined using double substrate kinetics with concentrations of UDP-Gal and 4-MUX varying between 0 and 5 mM by non-linear regression using Sigma Plot<sup>TM</sup>. Data were analyzed for a two-substrate system by fitting to an equation for sequential symmetrical initial velocity pattern, equation below (ordered or random equilibrium mechanism) [16,17]. In this equation  $V_{max}$  is the maximum velocity, [A] is the concentration of UDP-Gal, and [B] the concentration of 4-MUX.  $K_A$  and  $K_B$  are the true  $K_m$  for donor and acceptor substrates and  $K_{12}$  is the dissociation constant for UDP-Gal.

$$v = \frac{V_{max}[A][B]}{K_1K_2 + K_2[A] + K_1[B] + [A][B]} \quad (1)$$

## 2.6. Analysis of GAG synthesis by $\text{Na}_2^{35}\text{S}$ incorporation

To analyze GAG chain synthesis initiated from exogenously added xyloside, CHO pgsB-618 cells transfected with wild-type, mutant pcDNA- $\beta$ 4GalT7, or empty vector (mock) were incubated for 12 h in Fischer's medium containing  $10 \mu\text{Ci/ml}$   $\text{Na}_2^{35}\text{S}$  in the presence or absence of 4-MUX ( $0$ – $10 \mu\text{M}$ ). Medium ( $1 \text{ ml}$ ) was applied to a C-50 column to separate non-incorporated radio-labelled sulfate, and GAGs present in the eluates were quantified by scintillation counting or resolved by SDS-PAGE using Criterion Precast gels (4–15% Bis-Tris, Bio-Rad) and visualized by autoradiography.

## 2.7. Indirect immunofluorescence staining of recombinant CHO pgsB-618 cells

Indirect immunofluorescence staining of cell surface proteoglycan GAG chains in CHO pgsB-618 transfected cells was performed as previously described [18], using mouse anti-heparan-sulfate primary antibodies (10E4 [19], Seikagaku Corporation) and AlexaFluor 488-conjugated goat anti-mouse secondary antibodies (Molecular Probes/Invitrogen). Wild-type and mutated  $\beta$ 4GalT7 proteins expressed in CHO pgsB-618 transfected cells were simultaneously immunostained following Triton-X100-permeabilization using anti-myc primary antibodies and AlexaFluor 594-conjugated goat anti-mouse secondary antibodies, following the same procedure.

## 2.8. Molecular modelling of human $\beta$ 4GalT7 structure

The molecular model of the human  $\beta$ 4GalT7 was built using Modeller version 8v2 [20], based on the structure of monomer A of bovine  $\beta$ 4GalT1 complexed with UDP-Gal (PDB entry 100R) [21]. The energy of the resulting model was minimized using the Amber simulation program version 8 [22]. Charges of the atoms of UDP-Gal were calculated using the GAUSSIAN94 package [23] and the HF/6-31 G\* basis set. Atom-centered charges were fitted with Antechamber module of Amber 8 [22].  $\text{Mn}^{2+}$  parameters available from Amber site (<http://ambermd.org>). Docking of UDP-Gal- $\text{Mn}^{2+}$  complex in the human  $\beta$ 4GalT7 structure was performed manually based on the position of this ligand in bovine  $\beta$ 4GalT1. The whole system was solvated, the energy was minimized and submitted to molecular dynamics at 300K. A model for a O- $\beta$ -xylosylpeptide corresponding to the 15 N-terminal amino acids of human decorin (EDEASGIGPEVPDDR) containing Xyl attached to Ser 0y (N-decorin-Xyl) was generated using the AMBER/GLYCAM Input Configurator server (<http://glycam.crc.uga.edu/AMBER>). Docking N-decorin-Xyl into  $\beta$ 4GalT7 acceptor binding site was performed manually based on the position of the first N-acetyl-o-glucosamine in the structure of bovine  $\beta$ 4GalT1 (PDB entry 1TW5), and this model was submitted to energy minimization and molecular dynamics as described above. Finally, the model of human  $\beta$ 4GalT7 was compared to the crystal structure of Drosophila  $\beta$ 4GalT7 (PDB entry 3LW6) using the Dali program for protein structure comparison available at <http://ekhidna.biocenter.helsinki.fi/dalserver/> [24].

## 2.9. Statistical analyses

All data are presented as mean  $\pm$  S.D. of three determinations on at least two sets of experiments. For comparison among multiple groups, one-way analysis of variance (ANOVA) followed by Bonfer-

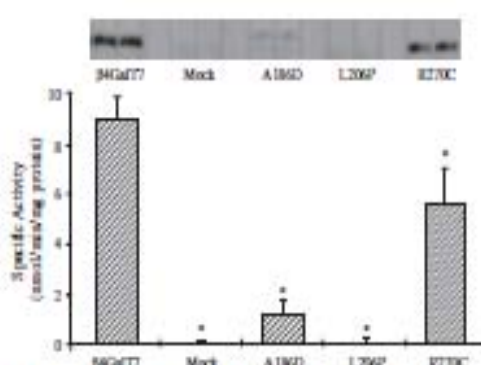


Fig. 1. Galactosyltransferase activity of EDS- $\beta$ 4GalT7 mutants expressed in CHO pgsB-618 towards 4-MUX. Activity of wild-type and mutant enzymes was evaluated using  $5 \text{ mM}$  4-MUX and  $1 \text{ mM}$  UDP-Gal ( $0.05 \mu\text{Ci}$  UDP- $^{14}\text{C}$ -Gal) in the presence of  $40 \mu\text{g}$  cell homogenate. The reaction products were separated by thin layer chromatography, visualized by autoradiography (shown in inset) and quantified by scintillation counting. Data are the mean  $\pm$  S.D.  $n = 3$ , \* $P < 0.05$ .

roni post-hoc correction test was used. Statistical analyses were performed using GraphPad Prism software (version 5.0).

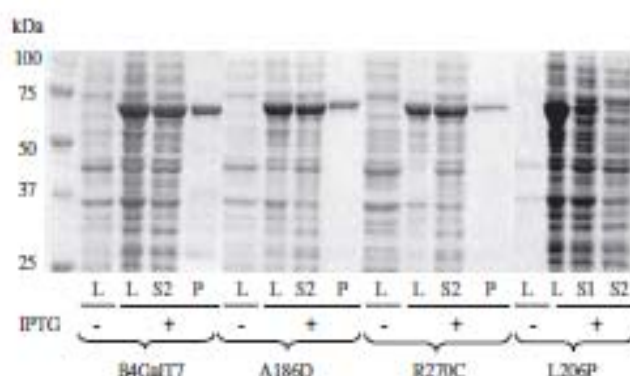
## 3. Results

### 3.1. EDS mutants exhibit large differences in $\beta$ 4GalT7 activity and kinetic properties

Wild-type  $\beta$ 4GalT7, and A186D, L206P and R270C mutants were expressed as full-length proteins in CHO pgsB-618 cells and tested for in vitro galactosyltransferase activity using 4-MUX as acceptor substrate. Fig. 1 revealed large differences in galactosyltransferase activity between wild-type and mutant enzymes, although the recombinant proteins were expressed at similar level, as indicated by immunoblot analysis using anti-myc antibodies (not shown). The  $\beta$ 4GalT7-A186D mutant exhibited nine-times lower activity than wild-type, whereas substitution of Leu206 with Pro led to an inactive enzyme. The  $\beta$ 4GalT7-R270C mutant catalyzed the transfer of Gal onto 4-MUX with a specific activity reaching about 60% that of the wild-type enzyme.

To further analyze the kinetic properties of EDS- $\beta$ 4GalT7 mutants, the catalytic domain of wild-type and mutant enzymes was produced as a soluble protein fused to GST in *E. coli*. The recombinant wild-type enzyme, and  $\beta$ 4GalT7-A186D and R270C mutants were purified to apparent homogeneity by affinity chromatography (Fig. 2). Determination of the kinetic parameters of the purified enzymes showed that the wild-type GST- $\beta$ 4GalT7 fusion protein was highly active with a  $k_{cat}$  up to  $1.1 \text{ s}^{-1}$  (Table 1). Furthermore, the  $K_m$  value of this enzyme towards UDP-Gal and 4-MUX was about  $0.45 \text{ mM}$  and  $0.60 \text{ mM}$ , respectively. These values are in the same range as apparent  $K_m$  determined for the membrane-bound enzyme expressed in HeLa cells (data not shown), indicating that expression of the catalytic domain, i.e. lacking the transmembrane domain and stem region did not substantially affect substrate binding and catalytic properties of the enzyme. The  $K_m$  value for GST- $\beta$ 4GalT7 towards UDP-Gal was also in the same range as reported by Pasek et al. [25], for the human  $\beta$ 4GalT7 expressed in fusion with galectin (i.e.  $0.51 \text{ mM}$ ).

Consistent with the low level of activity detected for the  $\beta$ 4GalT7-A186D mutant expressed in deficient CHO cells, this variant showed little activity when produced in bacteria, preventing further kinetic analyses. The  $\beta$ 4GalT7-L206P mutant, which lacked activity in eukaryotic cells, was mainly produced as insoluble inactive protein when expressed in *E. coli* (Fig. 2). These results indicate that exchange of leucine at position 206 to a proline



**Fig. 2.** Expression and purification of SDS- $\beta$ 4GalT7 mutants in *E. coli*. Wild-type GalT7 and mutant proteins were expressed in *E. coli*, and A186D, and R270C mutants were purified from supernatant of 12000 $\times$ g centrifugation by affinity chromatography. Proteins were analyzed by SDS-PAGE and stained with Coomassie Brilliant Blue. Shown from left to right, cell lysate from uninduced and IPTG-induced bacteria (L, 25  $\mu$ g), 12000 $\times$ g supernatant (S2, 1.2  $\mu$ g) and purified fusion proteins (P, 1–3  $\mu$ g). S1 refers to 5000 $\times$ g supernatant from cell lysates expressing  $\beta$ 4GalT7-L206P mutant.

residue probably affected protein folding and dramatically impaired its function. In contrast, substitution of Arg270 to cysteine resulted in a  $\beta$ 4GalT7 protein which retained activity. Interestingly, whereas the  $K_m$  value of the  $\beta$ 4GalT7-R270C mutant towards UDP-Gal was similar to that of the wild-type enzyme, its affinity for the xyloside acceptor substrate 4-MUX was dramatically reduced (10-times higher  $K_m$  value compared to wild-type).

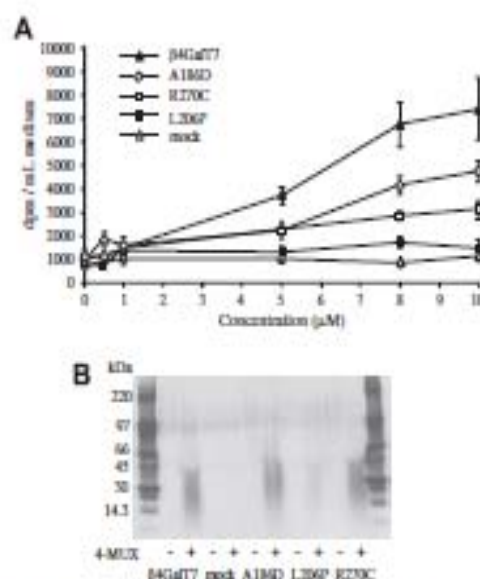
### 3.2. Molecular modelling of human $\beta$ 4GalT7

A molecular model of human  $\beta$ 4GalT7 was built based on the structure of bovine  $\beta$ 4GalT1 in complex with UDP-Gal molecule (PDB 100R). The position of A186, L206 and R270 are highlighted in the structure is shown in Fig. 5, panel A. We also used the crystal structure of the bovine  $\beta$ 4GalT1 in complex with an oligosaccharide acceptor substrate, as template (PDB 1TW5), to generate a model of the human  $\beta$ 4GalT7 acceptor binding site containing *N*-decofin-Xyl. The results shown Fig. 5B suggest a potential interaction between Arg270 and the serine residue of decorin peptide. Furthermore, we carried out a superposition of the modelled human  $\beta$ 4GalT7 and of the crystal structure of *Drosophila*  $\beta$ 4GalT7 (PDB 3LW6 [26]), by structure alignment using Dali program [24] (Supplementary data, Fig. 1). The results show a Z-score of 22.8 and a RMSD (Root mean standard deviation) of 2.6 Å, indicating a good fit between the structures of modelled human and *Drosophila*  $\beta$ 4GalT7. The two structures exhibit a very similar organization of the core protein with Ala186 and Leu206 (numbered in human  $\beta$ 4GalT7) located in a secondary structure environment similar to that of Ala169 and Val189 (numbered in fly  $\beta$ 4GalT7) (Supplementary data, Fig. 1). On the other hand, Arg270 is located at the surface of the protein, in a less conserved structural domain between human and fly  $\beta$ 4GalT7 proteins. In the case of the protein from *Drosophila*, this domain contains a disulfide bond between Cys255 and Cys310 [26], which is not present in human, contributing to a different secondary structure of the C-terminal part of the  $\beta$ 4GalT7 protein compared to its human counterpart domain containing Arg270.

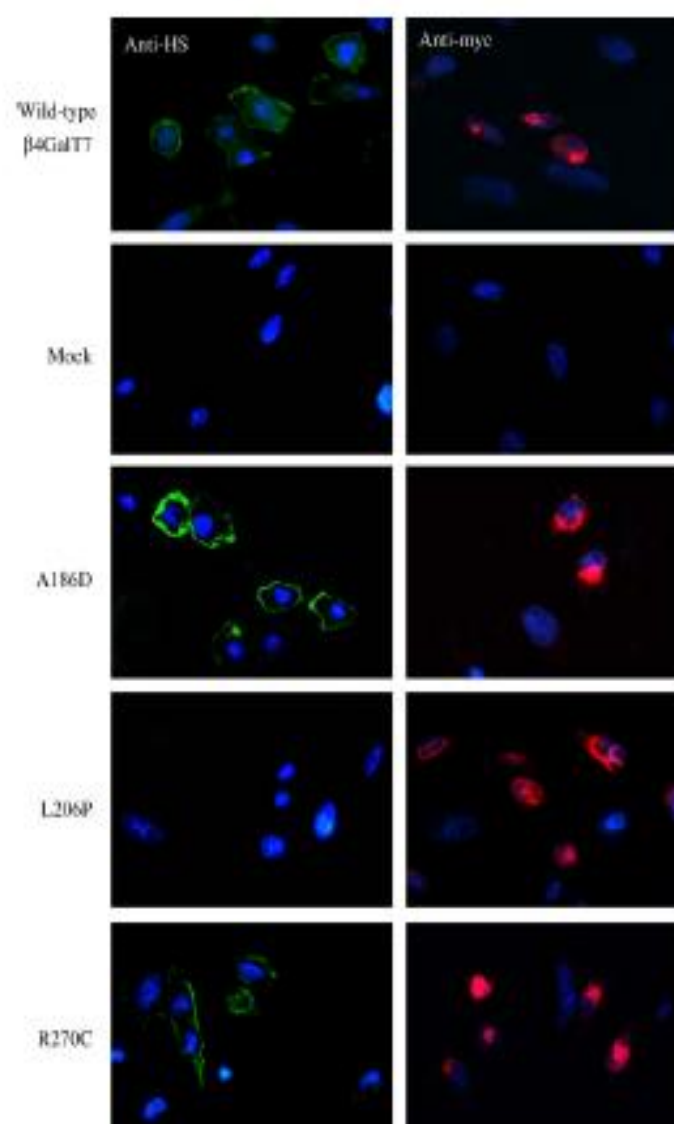
### 3.3. Effect of EDS mutations on GAG synthesis in CHO pgsB-618 cells

We next designed a series of experiments to determine the functional consequences of EDS mutations on  $\beta$ 4GalT7 activity *ex vivo*. For this purpose, we analyzed the capacity of wild-type and mutant enzymes to prime GAG synthesis on the exogenous  $\beta$ -D-xyloside 4-MUX, in  $\beta$ 4GalT7-deficient CHO pgsB-618 cells. The rate of GAG synthesis was evaluated in recombinant cells as

a function of xyloside concentration. Results shown in Fig. 3A indicate that expression of wild-type  $\beta$ 4GalT7 in these cells induced a concentration-dependent increase in GAG synthesis. No detectable [ $^{35}$ S] $_4$ -incorporation was detected in mock-transfected cells, up to 10  $\mu$ M 4-MUX concentration (Fig. 3A). In confirmation of these observations, SDS-PAGE analysis revealed the presence of radiolabelled GAG chains in  $\beta$ 4GalT7-transfected CHO pgsB-618 cells upon addition of 4-MUX, whereas no radioactivity could be detected in mock-transfected cells, or when addition of xyloside was omitted (Fig. 3B). Interestingly, although the  $\beta$ 4GalT7-A186D mutant exhibited low *in vitro* enzyme activity, this variant was able to prime GAG synthesis in CHO deficient cells up to a synthesis rate reaching two-third that of the wild-type enzyme at the highest concentration of 4-MUX (Fig. 3A). The presence of radiolabelled



**Fig. 3.** Kinetics of GAG chain synthesis from 4-MUX in CHO pgsB-618 cells expressing wild-type or EDS- $\beta$ 4GalT7 mutants. Cells seeded at  $6 \times 10^5$  cells in six-well plates were individually transfected with pcDNA- $\beta$ 4GalT7 or A186D, L206P, and R270C mutants (A) Kinetics of GAG synthesis as a function of 4-MUX concentration. 24h following transfection, cells were incubated in the presence of  $\text{Na}_2^{35}\text{SO}_4$  and increasing concentrations of 4-MUX for 12h before GAG isolation and quantification by scintillation counting. (B) Gel electrophoresis of radiolabelled GAG chains isolated from conditioned medium of recombinant CHO pgsB-618 cells incubated in the presence of  $\text{Na}_2^{35}\text{SO}_4$  and 4-MUX (5  $\mu$ M).



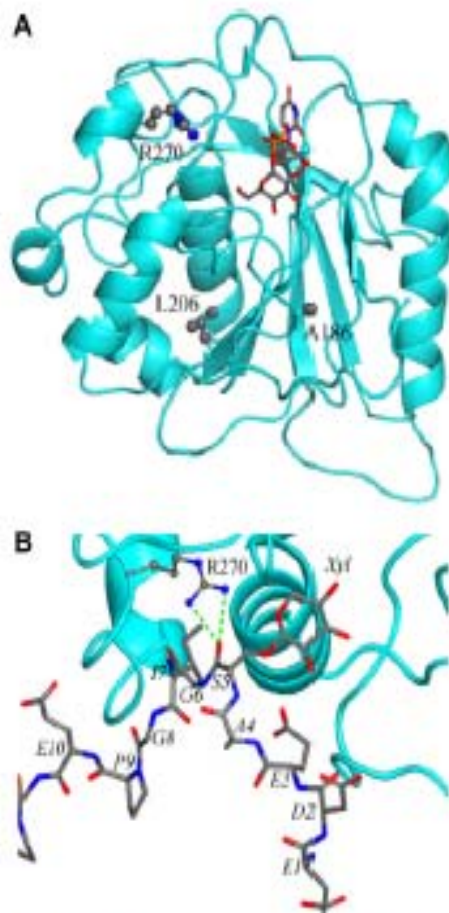
**Fig. 4.** Heparan-sulfate expression in CHO pgsB-618 cells transfected with EDS- $\beta$ 4GalT7 mutants. Cells were transfected with pcDNA- $\beta$ 4GalT7, A186D, L206P, or R270C or mock-transfected and indirect immunofluorescence staining was performed using 10E4 anti-heparan-sulfate (anti-HS, left panels), or following permeabilisation using anti-myc antibodies (right panels). Nuclei were stained by Hoechst reagent prior to mounting slides.

GAG chains was confirmed by SDS-PAGE analysis in the culture medium of cells transfected by this mutant in the presence of 4-MUX (Fig. 3B). Furthermore, the  $\beta$ 4GalT7-L206P mutant exhibited no GAG priming activity towards 4-MUX in transfected cells, consistent with the lack of galactosyltransferase activity observed for this enzyme *in vitro*. In agreement, no radiolabelled GAG chains were observed upon gel electrophoresis analysis in those cells, in the presence of 4-MUX (Fig. 3).  $\beta$ 4GalT7-R270C mutant was able to produce a significant increase in GAG synthesis rate as a function of 4-MUX concentration, although at a lower level compared to the wild-type enzyme. This result is in agreement with the *in vitro* kinetic properties of this mutant, which were characterized by a strong decrease in affinity towards xyloside acceptor substrate and reduced efficiency compared to the wild-type  $\beta$ 4GalT7.

To further determine the consequences of EDS mutations on GAG synthesis, we carried out immunofluorescence analyses of heparan-sulfate expression in  $\beta$ 4GalT7-defective CHO pgsB-618 cells transfected with the corresponding cDNAs, since it was previ-

ously reported that wild-type CHO-K1 cells predominantly produce heparan-sulfate GAG chains [12]. A prominent staining of cell membrane heparan-sulfates by 10E4 antibodies (anti-heparan-sulfate), was observed in CHO pgsB-618 cells transfected with wild-type  $\beta$ 4GalT7, and this staining pattern was associated with typical Golgi pattern of the recombinant protein revealed with anti-myc antibodies (Fig. 4). As expected, no heparan-sulfate expression could be detected in mock-transfected CHO pgsB-618 cells by indirect immunofluorescence analysis. The intracellular location of  $\beta$ 4GalT7-A186D, and of  $\beta$ 4GalT7-R270C mutant was similar to that of the wild-type exhibiting typical Golgi staining (Fig. 4). In addition, these mutants were able to restore heparan-sulfate cell surface expression revealed by 10E4 antibodies, confirming *ex vivo* galactosyltransferase analyses indicating that they were able to sustain, at least partially, GAG chain priming in deficient CHO cells. By contrast, although the  $\beta$ 4GalT7-L206P mutant was expressed, as shown by intracellular staining using anti-myc antibodies, no membrane staining of cellular heparan-sulfate could





**Fig. 5.** Molecular modelling of human  $\beta 4\text{GalT7}$  structure. (A) The model of  $\beta 4\text{GalT7}$  bound on bovine  $\beta 4\text{GalT7}$  as complex with UDP-Gal (PDB entry 1O0R) is shown as "ribbons" and of UDP-Gal as "sticks". Positions of Ala186, Leu206 and Arg270 are highlighted in "balls-and-sticks". (B) The model of the acceptor binding site of human  $\beta 4\text{GalT7}$ . Protein N-decetin-Xyl is "sticks". Hydrogen bonds between Arg270 ( $\beta 4\text{GalT7}$ ) and the main chain of Ser5 attached to Xyl (N-decetin-Xyl) are shown in green dotted line.

be seen, confirming that this mutant exerted the most severe effect on GAG synthesis. Similar to previous report [4], expression of this mutant exhibited a diffuse pattern (Fig. 4), suggesting altogether that L206P substitution results in a conformational change of the  $\beta 4\text{GalT7}$  protein, loss of activity and altered intracellular trafficking.

#### 4. Discussion

Individuals affected by the progeroid-type EDS suffer from severe clinical manifestations including craniofacial and skeletal abnormalities, developmental retardation and aged appearance [1]. Although detailed studies on the mechanisms of GAG alterations contributing to the complex pathology of this genetic disorder have been carried out in skin fibroblasts from affected patients [5,6], the molecular basis underlying  $\beta 4\text{GalT7}$  deficiency is not yet fully understood. Here, we investigated the molecular consequences of progeroid-type EDS mutations A186D, L206P, and R270C on (i)  $\beta 4\text{GalT7}$  enzyme activity and kinetic properties *in vitro*, and (ii) *ex vivo* GAG chain synthesis in defective CHO pgsB-618 cells. These approaches, combined with computer-aided modelling of the structure of human  $\beta 4\text{GalT7}$  and its acceptor binding site, allowed us to elucidate the mechanisms underlying

$\beta 4\text{GalT7}$  functional defects in EDS congenital glycosylation disorder. The  $\beta 4\text{GalT7}$ -deficient CHO pgsB-618 line was transfected with wild-type, A186D, L206P or R270C cDNA with the resulting pattern of activities, (i)  $\beta 4\text{GalT7}$ -A186D mutant exhibited about 10% activity of wild-type; (ii) mutation of leucine at position 206 to proline abolished activity; (iii) galactosyltransferase activity of the  $\beta 4\text{GalT7}$ -R270C mutant was reduced two-fold compared to wild-type. To assess the functional consequences of the mutations in CHO pgsB-618 transfected cells, we developed two approaches. Firstly, we studied the effect of the supplementation of 4-MUX on GAG synthesis. Secondly, we analyzed the restoration of heparan-sulfate expression by indirect immunofluorescence analysis. *Ex vivo* kinetic studies of GAG synthesis revealed that expression of the  $\beta 4\text{GalT7}$ -A186D mutant partially restored GAG priming compared to the wild-type enzyme. Immunostaining of heparan-sulfate chains in  $\beta 4\text{GalT7}$ -A186D transfected cells also revealed a pattern similar to that observed upon expression of the wild-type enzyme. These observations corroborate previous report indicating that this mutant restored GAG synthesis based on flow cytometry analysis [4]. It should be noted that the reduction in  $\beta 4\text{GalT7}$  galactosyltransferase activity *in vitro* due to the A186D substitution was greater than the decrease in GAG priming evaluated in the CHO pgsB-618. Interestingly, it was reported that this mutation caused a larger reduction in apparent  $K_m$  of  $\beta 4\text{GalT7}$  towards UDP-Gal than acceptor substrate [10]. Since UDP-Gal serves as a sugar donor for several galactosyltransferases involved in glycoprotein and GAG synthesis, the concentration of this molecule might not be rate-limiting, thus possibly explaining the moderate effect of the  $\beta 4\text{GalT7}$ -A186D mutation *ex vivo*.

Comparison of the consequences of the three genetic mutations of  $\beta 4\text{GalT7}$  causing the progeroid form of EDS revealed that only substitution of leucine at position 206 with proline, led to a complete loss of enzyme activity. Our results showed that this mutant was unable to prime GAG synthesis from exogenous xyloside 4-MUX, nor restore heparan-sulfate expression in CHO pgsB-618 defective cells. Molecular modelling of the human  $\beta 4\text{GalT7}$  protein, comparison of this model to the crystal structure of the *Drosophila*  $\beta 4\text{GalT7}$  [26], and multiple sequence alignment studies provide clues to explain this effect. Indeed, leu206 is strongly conserved in  $\beta 4\text{GalT7}$ -related sequences among various animal species (Talhaoui, I., Oriol, R. et al. unpublished). In our model of the human  $\beta 4\text{GalT7}$  structure (see Fig. 5A), this residue is positioned in a small conserved  $\beta$ -sheet, similarly to its counterpart Val189 in *Drosophila*  $\beta 4\text{GalT7}$  3D structure (see Fig. 1 Supplementary data). Introduction of a proline residue at this position has been predicted to disrupt the secondary or even the tertiary structure of the protein [26]. In agreement with this assumption, we found that in contrast to other EDS mutants, the catalytic domain of  $\beta 4\text{GalT7}$ -L206P was expressed as insoluble inactive material in *E. coli*.

It was previously shown by Seidler et al. [5] that  $\beta 4\text{GalT7}$  activity towards *para*-nitrophenyl-Xyl in skin fibroblasts of patient harbouring the R270D mutation was reduced compared to fibroblasts from healthy patients. In agreement with this finding, comparison of *in vitro* and *ex vivo* activity of  $\beta 4\text{GalT7}$ -R270C mutant expressed in CHO pgsB-618-defective cells showed a decrease of both *in vitro* galactosyltransferase activity and GAG priming. Investigation of the kinetic properties of the purified mutant enzyme following expression in *E. coli* shed light on the biochemical basis of  $\beta 4\text{GalT7}$ -altered function upon mutation. A large decrease in affinity towards 4-MUX acceptor substrate was noticed for this mutant, suggesting that arginine to cysteine replacement at position 270 may alter the organization of the acceptor binding site. An increase in  $K_m$  value upon R270C mutation was also observed when using a synthetic glycine-serine-Xyl derivative as acceptor substrate (data not shown). Interestingly, molecular modelling of  $\beta 4\text{GalT7}$  bound to Xyl attached to the serine of a decorin peptide predicted that

Arg270 is located in the vicinity of the core protein of this proteoglycan (Fig. 5B), which glycanation is affected in EDS [5]. In the *Drosophila* structure, it has been found that Lys254 (corresponding to Arg270 in primary sequence alignment) is located at the C-terminal end of the long flexible loop (residues 242–255), and the authors predicted that a cysteine substitution may affect the conformational flexibility of this loop [26]. Noteworthy, protein comparison of the C-terminal domain containing Arg270 indicates that human and *Drosophila*  $\beta$ 4GalT7 structures present a different secondary structure organization. Therefore, crystal structure resolution of the human  $\beta$ 4GalT7 in complex with acceptor substrate is required to definitively establish the consequences of R270C-EDS mutation.

In conclusion, based on current results, we propose the following molecular basis underlying pangeroid type EDS. (1) Although the Ala186 to aspartate replacement strongly reduced in vitro  $\beta$ 4GalT7 enzyme activity when expressed in deficient-CHO pgsB-618 cells and in *E. coli*, the consequences of this mutation on GAG synthesis are much less pronounced. This finding suggests that this mutation is not the primary cause for the alteration of GAG synthesis observed in the EDS patient carrying the compound A186D and I206P mutations. (2) In contrast, the dramatic effect of I206P both on in vitro and ex vivo  $\beta$ 4GalT7 activity can be attributed to a major conformational change induced by the leucine to proline substitution at this position. This is likely to account for the alteration of GAG synthesis in the pangeroid EDS patient carrying the compound mutations. (3) The  $\beta$ 4GalT7-R270C mutant exhibited a strong decrease in affinity together with a decreased catalytic efficiency that translated into a reduced ability to prime GAG synthesis upon expression in  $\beta$ 4GalT7-deficient cells. These molecular alterations are likely to cause the altered GAG structure and connective tissue disorder observed in the patient carrying a homozygous mutation of  $\beta$ 4GalT7 at this position.

#### Acknowledgments

Dr. J.-C. Jacquinet (UMR CNRS 6005, University of Orléans, France) is gratefully acknowledged for providing glycine-serine-Xyl. This work was supported by The Agence Nationale de la Recherche (GAGNetwork ANR-08-PCV1-0023-01), C2 INSERM-University of Dundee, Royal Society International Joint Program, Région Lorraine, Communauté Urbaine du Grand Nancy, CNRS and University Henri Poincaré Nancy 1. Part of this work was performed under the auspices of an EAL (European Associated Laboratory) between UMR CNRS 7561-UHP Nancy 1 and the University of Dundee. Ibtissam Talhaoui is recipient of a MRST PhD fellowship.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2010.08.001.

#### References

[1] Kresse, H., Nothig, S., Quentin, E., Hoffmann, J., Giesel, J., Okada, S. and Timpl, T. (1987) Glycosaminoglycan-free small proteoglycan core protein is secreted by fibroblasts from a patient with a syndrome resembling pangeroid. *Am. J. Hum. Genet.* **41**, 436–453.

[2] Fayaz-Ul-Haque, M., Zaidi, S.H., N-Ali, M., Al-Mutairi, M.S., Kennedy, S., N-Thani, G., Tsai, L.C. and Teitel, A.S. (2004) A novel missense mutation in the galactosyltransferase-1 ( $\beta$ 4GalT7) gene in a family exhibiting facioscapular humeral and Ehlers-Danlos syndrome resembling the pangeroid type. *Am. J. Med. Genet. A* **128A**, 39–45.

[3] Quentin, E., Glaser, A., Baden, L. and Kresse, H. (1990) A genetic defect in the biosynthesis of dermatan sulfate proteoglycan: galactosyltransferase I deficiency in fibroblasts from a patient with a pangeroid syndrome. *Proc. Natl. Acad. Sci. USA* **87**, 1342–1346.

[4] Okajima, T., Nakamoto, S., Furukawa, K. and Uraha, T. (1993) Molecular basis for the pangeroid variant of Ehlers-Danlos syndrome. Identification and characterization of two mutations in galactosyltransferase I gene. *J. Biol. Chem.* **274**, 28941–28944.

[5] Seidler, D.G., Fayaz-Ul-Haque, M., Hamam, U., Yip, G.W., Zaidi, S.H., Teitel, A.S., Kiesel, L. and Götte, M. (2006) Defective glycosylation of dermatan and biglycan, altered collagen structure, and abnormal phenotype of the skin fibroblasts of an Ehlers-Danlos syndrome patient carrying the novel Arg270C substitution in galactosyltransferase I ( $\beta$ 4GalT-7). *J. Mol. Med.* **84**, 583–594.

[6] Götte, M., Spillmann, D., Yip, G.W., Venting, E., Schürmeyer, F.G., van Kuppevelt, T.H. and Niesel, L. (2008) Changes in heparan sulfate are associated with delayed wound repair, altered cell migration, adhesion and contractility in the galactosyltransferase I ( $\beta$ 4GalT-7) deficient form of Ehlers-Danlos syndrome. *Hum. Mol. Genet.* **17**, 999–1009.

[7] Eiko, J.D., Kimata, K. and Lindahl, U. (2009) Proteoglycans and sulfated glycosaminoglycans. In: *Encyclopedia of Glycobiology* (Varki, A., Cummings, R.D., Eiko, J.D., Hildebrandt, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W. and Hildebrandt, M.E., Eds.), 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY) (Chapter 16).

[8] Knäuper, J., Spillmann, D., Li, J.P. and Lindahl, U. (2000) Interactions between heparan sulfate and proteins: the concept of specificity. *J. Cell Biol.* **174**, 323–327.

[9] Prydz, K. and Dalen, K.T. (2000) Synthesis and sorting of proteoglycans. *J. Cell Sci.* **113** (Pt 2), 193–205.

[10] Almeida, R., Levery, S.B., Mandel, U., Ennos, H., Schwientek, T., Bennett, E.P. and Claessen, H. (1998) Cloning and expression of a proteoglycan UDP-galactose 3-xylose ( $\beta$ 1,4-galactosyltransferase I A seventh member of the human  $\beta$ 1,4-galactosyltransferase gene family. *J. Biol. Chem.* **273**, 26165–26171.

[11] Gilbert, S., Lattard, V., Fondeur, M., Jacquinet, J.C., Mollereau, G., Netter, P., Magliolas, J., Ouzrine, M. and Fourel-Gigoux, S. (2005) Phosphorylation and mutation of oligosaccharide substrates critically influence the activity of human  $\beta$ 1,4-galactosyltransferase 7 (GalT-7) and  $\beta$ 1,3-glucosyltransferase 1 (GalT-1) involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J. Biol. Chem.* **280**, 14173–14225.

[12] Eiko, J.D., Wronski, J.L., Taylor, W.H., Eiberg, G., Baden, L., Anantharaman, G. and Gewirtz, A. (1987) Inhibition of chondroitin and heparan sulfate biosynthesis in Chinese hamster ovary cell mutants defective in galactosyltransferase I. *J. Biol. Chem.* **262**, 12189–12195.

[13] Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.

[14] Bessal, S.K. and Gerner, T. (1980) A unified method for the assay of uridine diphosphoglucosyltransferase activities toward various aglycones using uridine diphospho-( $^{14}$ C)glucosamine acid. *Anal. Biochem.* **108**, 321–328.

[15] Li, D., Fourel-Gigoux, S., Bami, L., Mollereau, G., Netter, P., Magliolas, J. and Ouzrine, M. (2007) Identification of aspartic acid and histidine residues mediating the reaction mechanism and the substrate specificity of the human UDP-glucosyltransferase 7A. *J. Biol. Chem.* **282**, 36514–36524.

[16] Segel, I.H. (1975) *Rapid equilibrium bioassay and transient-state enzyme kinetics*, pp. 273–345, John Wiley & Sons, New York.

[17] Boegeman, E. and Quidry, P.A. (2002) Studies on the metal binding sites in the catalytic domain of  $\beta$ 1,4-galactosyltransferase. *Glycobiology* **12**, 395–407.

[18] Bai, C., Ouzrine, M., Talhaoui, I., Sharp, S., Prydz, K., Coughtrie, M.W. and Fourel-Gigoux, S. (2010) Epigenetic methylation-associated repression of heparan sulfate 3-O-sulfotransferase gene expression contributes to the invasive phenotype of H-DIC-05 chondrosarcoma cells. *FASEB J.* **24**, 436–450.

[19] David, G., Bai, X.M., Van der Schueren, B., Goutman, J.J. and Van den Berghe, H. (1992) Developmental changes in heparan sulfate expression: in situ detection with mAbs. *J. Cell Biol.* **119**, 961–975.

[20] Sali, A. and Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815.

[21] Ramakrishnan, B., Balaji, P.V. and Quidry, P.A. (2002) Crystal structure of  $\beta$ 1,4-galactosyltransferase complex with UDP-Gal reveals an oligosaccharide acceptor binding site. *J. Mol. Biol.* **318**, 491–502.

[22] Cox, D.A. et al. (2004) *AMBER 8*, University of California, San Francisco.

[23] Frack, M.J. et al. (1998) *Cambridge 98 (Revision A.11.2)*, Cambridge, Inc., Pittsburgh, PA.

[24] Holm, L. and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38** (Suppl.), W545–W548.

[25] Paek, M., Boegeman, E., Ramakrishnan, B. and Quidry, P.A. (2010) Galactin-1 as a fusion partner for the production of soluble and folded human  $\beta$ 1,4-galactosyltransferase-7 in *E. coli*. *Biochem. Biophys. Res. Commun.* **394**, 679–684.

[26] Ramakrishnan, B. and Quidry, P.A. (2010) Crystal structure of the catalytic domain of *Drosophila*  $\beta$ 1,4-Galactosyltransferase-7. *J. Biol. Chem.* **285**, 15619–15626.

Supplementary Data  
Figure 1

A

```

LLL LLEEEEEEEE LLLHHHHHHH HHHHHHHHHH LLEEEEEEEE EELLLLLLLH
HHHHHHHHHH
Drosophila -----GAS VHKMALLVPF RDRFEELLQF VPHMTAFLKR QGVAHHIFVL NQVDRFRFNR ASLINVGFQF
Human   EHWEEDASWG PHRLAVLVPF RERFEELLVF VPHMRRFLSR KKIRHHIYVL NQVDHFRFNR AALINVGFLF
      LLLLLLLLLL LLEEEEEEEE LLLLLHHHLH HHHHHHHHHH LLEEEEEEEE EELLLLLLLL HHHHHHHHHH

      169      189
      |      |
      LLLLLLEEEE ELLLEEEELL LLLLLLLLLL LLEELLIL LLL LLL LLLLLLEEEE HHHHHHLLL
Drosophila ASDVYDYIAM HDVDLLPLND NLLYEyPSSL GPLHIAgpkL HPK---YHYD NFVGGILLVR REHFQKMNGM
Human   SSNSTDYIAM HDVDLLPLNE ELDYG-FPEA GPFHVA---S PELHPLYHYK TYVGGILLLS KQHRYLCNGM
      LLLLLLEEEE LLLLEEEELL LLLL-LLLL LLEELL L LLLLLLLLLL LLLLLLEEEE HHHHHHLLL

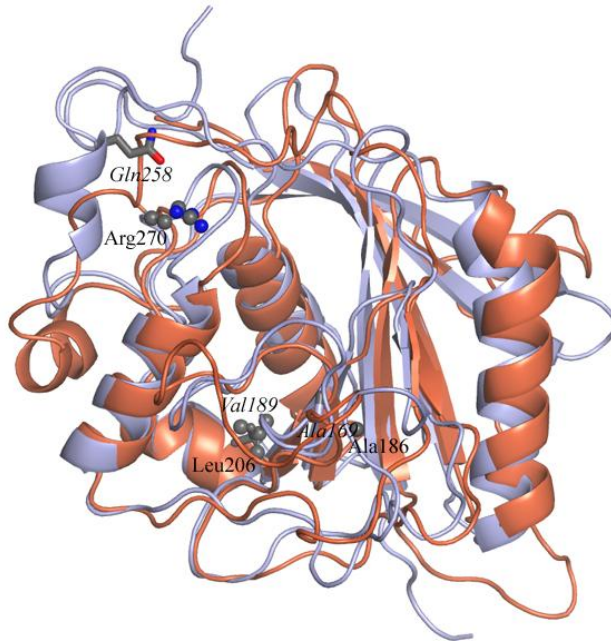
      186      206
      |      |
      258
      |
      LLLLLLLLLL HHHHHHHHHH LLLLLLLLLL LLLLILLEE ELLLIILL LLLLLLH HHHhLLL
Drosophila SNQYWGWGLE DDEFFVRIKD AGLQVTRPQn IKTGtdTFS HIHNryHRKR --DTQKcfnQ KEMtRKRD--
Human   SNRFWGWGRE DDEFYRRIKG AGLQLFRPS- GITT---GYK TFRH--LHDP AWRKRDQK-R IAA-QKQEQQ
      LLLLLLLLLL HHHHHHHHHH LLLLLLLLLL LLLL LEE ELLL LLLL LLLLLLLL L LEE-ELLHHH

      270

      LLLLH-H LLEEEEEEEE EEEL lEE EEEEEELL Lllhhlll
Drosophila --HKTGLD-N VKYKILKVHE MLID--qvPV TILNILLDCD Vnktpwcdcs
Human   KVDREGGLNT VKYHVASRTA LSVGAPCTV LNIMLDCDKT A-----
      HHHHLLLLLL LLEEEELL LLLLLLLLLL LEEEEELL L

```

B



## **Caractérisation enzymatique des mutations de la $\beta$ 4GalT7 humaine associées au syndrome progéroïde d'Ehlers-Danlos**

Bien que les mutations de la  $\beta$ 4GalT7 humaine à l'origine de la forme progéroïde du syndrome d'Ehlers-Danlos aient été identifiées depuis plusieurs années, on ne dispose que de données limitées sur les caractéristiques enzymatiques des mutants associés à cette maladie génétique. La deuxième partie de ma thèse a été consacrée à la caractérisation moléculaire des mutants A186D, L206P et R270C, afin de comprendre le mécanisme par lequel les mutations génétiques de la  $\beta$ 4GalT7 humaine altèrent le processus de synthèse des chaînes de GAGs des PGs.

Cette étude a fait l'objet de la **Publication n° 2**

MOLECULAR CHARACTERIZATION OF  $\beta$ 1,4-GALACTOSYLTRANSFERASE 7 GENETIC MUTATIONS LINKED TO THE PROGEROID FORM OF EHLERS-DANLOS SYNDROME (EDS)

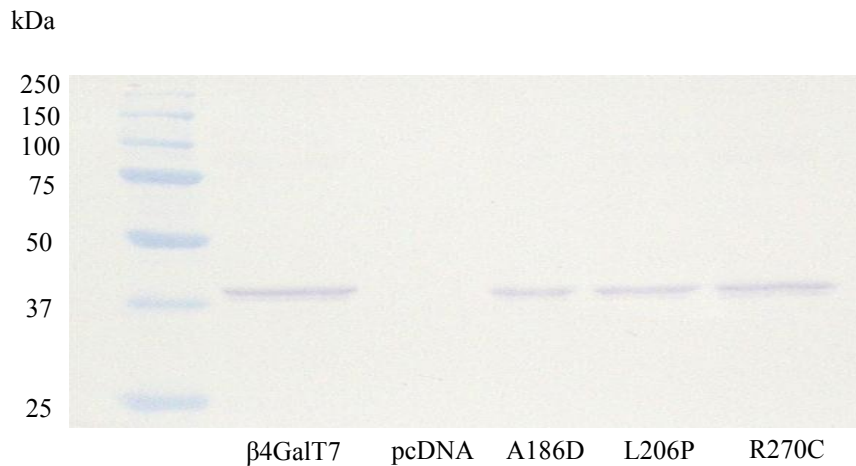
**Ibtissam Talhaoui\***, Catherine Bui\*, Matthieu Chabel, Guillermo Mulliert, Michael W.H. Coughtrie, Mohamed Ouzzine, Sylvie Fournel-Gigleux.

\* These authors equally contributed to this work

*FEBS Lett.* (2010) 584; 3962-8

### **1. Analyse de l'effet des mutations associées au syndrome progéroïde d'Ehlers-Danlos sur l'activité enzymatique de la $\beta$ 4GalT7 humaine et ses propriétés cinétiques**

Dans un premier temps, nous avons exprimé la  $\beta$ 4GalT7 humaine sauvage et les mutants A186D, L206P et R270C dans les cellules CHO-pgsB618 déficientes, et nous avons évalué l'activité enzymatique GalT, *in vitro*, en présence de 4-MOX comme substrat accepteur. Simultanément, nous avons analysé le taux d'expression de la  $\beta$ 4GalT7 et des trois mutants par immunoblot.



**Figure 42 . Analyse du taux d'expression de la  $\beta$ 4GalT7 sauvage et des mutants Ehlers-Danlos par immunoblot**

20  $\mu$ g de protéines totales ont été déposés par puit. Les protéines recombinantes exprimées dans les cellules CHO-pgsB618 ont été révélées au moyen d'anticorps anti-Myc, comme décrit dans "matériels et méthodes".

La Figure 42 montre que le niveau d'expression de ces protéines recombinantes humaines produites dans les cellules CHO-pgsB618 est comparable.

La mutation de l'Ala186 en Asp réduit fortement l'activité de la  $\beta$ 4GalT7 mesurée *in vitro*. La substitution de la Leu206 par une Pro produit une protéine complètement inactive. Enfin, le remplacement de l'acide aminé en position 270 de la  $\beta$ 4GalT7 humaine par une Cys induit une diminution modérée de l'activité de cette enzyme *in vitro*, d'environ 30% comparée à l'enzyme sauvage (Publication n°2, Figure 1).

Dans un second temps, nous avons construit un vecteur plasmidique permettant l'expression hétérologue de la protéine dans la bactérie *E. coli*. La  $\beta$ 4GalT7 recombinante humaine sauvage, ainsi que les mutants A186D, L206P et R270C, ont été exprimés en fusion avec la GST, puis purifiés par chromatographie d'affinité comme décrit dans la partie "matériels et méthodes". Les différentes fractions protéiques collectées ont été analysées par SDS-PAGE et colorées au bleu de Coomassie (Publication n°2, Figure 2).

Les résultats montrent que la GST-β4GalT7, la GST-A186D et la GST-R270C sont exprimées sous forme soluble et ont pu être purifiées avec un degré de pureté et une qualité satisfaisants (Tableau 8 et Publication n°2, Figure 2). Il est important de noter que la substitution de la Leu206 par une proline conduit à la production d'une protéine GST-L206P insoluble qui forme des corps d'inclusions. Cette observation suggère que cette mutation affecte les propriétés physicochimiques de la protéine, et vraisemblablement altère son repliement et son organisation 3D.

	Type de protéines	Protéines totales (mg)	Activité enzymatique totale (nmoles/min)	Activité enzymatique spécifique (nmoles/min/mg)	Facteur de purification	Rendement %
<b>GST-R270C</b>	<i>Lysat de bactéries</i>	135	1902.9	56.1	1	100
	<i>Protéines solubles</i>	31	1328.3	170.3	3.0	69.8
	<i>Protéine purifiée</i>	0.515	75	580	10.3	4

**Tableau 8: Tableau de purification de la protéine GST-R270C recombinante humaine exprimée dans la bactérie *E.coli*.**

La protéine a été purifiée par chromatographie d'affinité à partir d'une culture bactérienne de 200 ml. L'activité enzymatique a été déterminée par CLHP, comme décrit dans "matériels et méthodes".

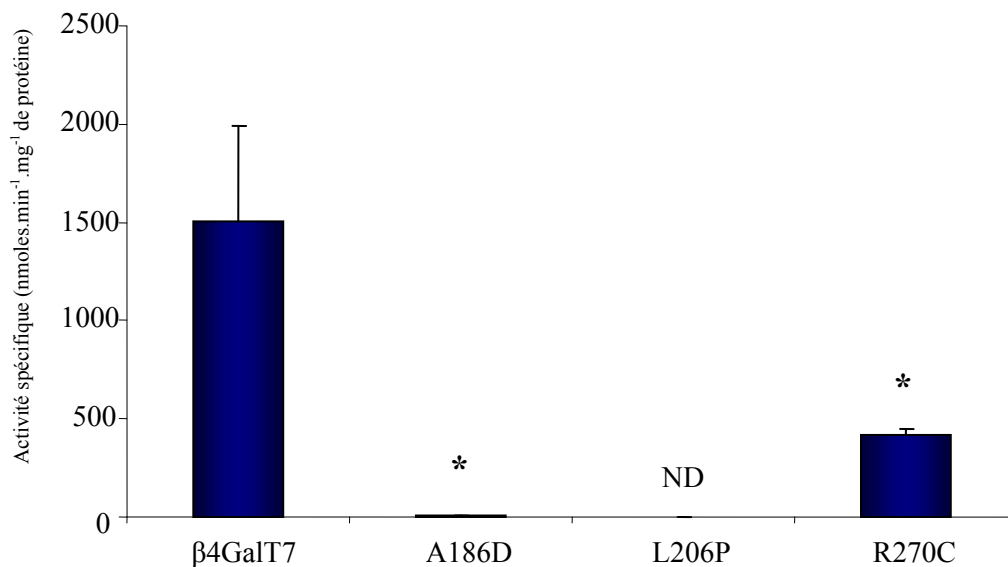
Au cours de cette étude, j'ai produit et purifié des quantités relativement importantes de protéines β4GalT7 sauvage et mutées (Tableau 8). Cela nous a permis de comparer les constantes cinétiques réelles de ces enzymes en faisant varier simultanément la concentration en substrat donneur (UDP-Gal) et accepteur (4-MOX).

Les protéines purifiées ont été utilisées pour déterminer l'effet des différentes mutations associées au syndrome progéroïde d'Ehlers-Danlos sur les propriétés cinétiques de la β4GalT7 humaine. Les données expérimentales ont été ajustées à une équation caractéristique d'un mécanisme séquentiel.

$$v = \frac{V_{\max}[A][B]}{K_{ia}K_B + K_B[A] + K_A[B] + [A][B]}$$

Ainsi, nous avons pu déterminer la constante de Michaelis-Menten de chacun des substrats, et les constantes d'inhibition  $K_i$ .

Les conséquences des mutations sur l'activité GalT des protéines exprimées dans les cellules CHO-pgsB618 et de celles exprimées dans la bactérie *E. coli* sont similaires. Ainsi la mutation de l'Ala186 en Asp provoque une diminution majeure de l'activité comparée à l'enzyme sauvage, tandis que la substitution de l'Arg270 par un résidu Cys permet de préserver une partie importante de l'activité GalT de l'enzyme. En outre, comme il est indiqué précédemment, la mutation L206P conduit à la production d'une protéine insoluble en fusion avec la GST, raison pour laquelle nous n'avons pas pu déterminer l'activité enzymatique de ce mutant (Figure 43).



**Figure 43. Mesure de l'activité enzymatique, *in vitro*, de la  $\beta$ 4GalT7 sauvage et des mutants Ehlers-Danlos exprimés dans la bactérie *E. coli***

Les essais enzymatiques ont été réalisés dans un volume réactionnel de 50  $\mu$ l, en présence de 200 ng de protéines purifiées, 1 mM d'UDP-Gal, (0.05  $\mu$ Ci UDP[U<sup>14</sup>C]Gal), 5 mM de 4-MOX, 100 mM de cacodylate de sodium contenant 10 mM de cations Mn<sup>2+</sup>, comme décrit dans la section "matériels et méthodes". Chaque valeur correspond à la moyenne  $\pm$  SD d'essais en double réalisés sur trois purifications différentes. ND, non déterminé; \* valeur d'activité significativement différente de l'activité enzymatique mesurée pour l'enzyme sauvage à  $P < 0.05$ .

Le tableau 1 (Publication n°2) montre que la protéine sauvage purifiée est très active avec un  $k_{cat}$  de l'ordre de 1s<sup>-1</sup>. Les valeurs de  $K_m$  vis-à-vis de l'UDP-Gal et du 4-MOX sont

comparables à celles obtenues pour la protéine exprimée dans des cellules HeLa et CHO-pgsB618. Le très faible niveau d'activité du mutant A186D n'a pas permis la détermination des paramètres cinétiques de cette protéine. Quant au mutant R270C, l'affinité de cette enzyme pour le substrat donneur est inchangée, toutefois, la valeur de son  $K_m$  vis-à-vis du substrat accepteur 4-MOX est fortement augmentée, indiquant une réduction d'affinité de ce mutant envers le xyloside.

## **2. Modélisation moléculaire de la $\beta$ 4GalT7 humaine**

Afin de comprendre les effets moléculaires des mutations associées au syndrome progéroïde d'Ehlers-Danlos sur l'activité GalT, nous avons développé en collaboration avec le Dr. G. Mulliert, un modèle moléculaire de la  $\beta$ 4GalT7 humaine basé sur la structure 3D de la  $\beta$ 4GalT1 bovine cristallographiée en présence d'UDP-Gal (Publication n°2, Figure 5A). Parallèlement, la  $\beta$ 4GalT1 bovine cristallisée avec le substrat accepteur a été utilisée pour développer un modèle moléculaire de la  $\beta$ 4GalT7 humaine du site actif de l'enzyme, contenant un peptide de décorine sur lequel est attaché le Xyl (Publication n°2, Figure 5B). D'après ce modèle moléculaire, le résidu Arg270 serait localisé à proximité de la Ser de la décorine et pourrait potentiellement entrer en interaction avec cet acide aminé de la protéine "core". De façon complémentaire, nous avons réalisé une superposition du modèle moléculaire de la  $\beta$ 4GalT7 humaine et de la structure cristallographique de la  $\beta$ 4GalT7 de drosophile récemment cristallisée, par le serveur "Dali Program" (Publication n°2, Supplementary data).

Le résultat obtenu montre une bonne superposition entre le modèle moléculaire du domaine N-terminal de la protéine humaine et la structure 3D de la  $\beta$ 4GalT7 de drosophile. L'Ala186 et la Leu206 de la  $\beta$ 4GalT7 humaine correspondent respectivement à l'Ala169 et la Val189 de la  $\beta$ 4GalT7 de drosophile, et sont situées dans les domaines peptidiques bien superposés. Quant à l'Arg270 de la protéine humaine, la Figure 1B (Supplementary data) montre que ce résidu aminé est situé dans une région qui présenterait une organisation structurale différente de celle de la  $\beta$ 4GalT7 de drosophile.

## **3. Analyse des conséquences fonctionnelles des mutations associées à la forme progéroïde du syndrome d'Ehlers-Danlos sur la synthèse des GAGs *ex vivo* dans les cellules CHO-pgsB618**

La dernière partie de cette étude a été consacrée à la détermination de l'effet des mutations Ehlers-Danlos sur la capacité de la  $\beta$ 4GalT7 humaine à initier la synthèse des GAGs *ex vivo*, à



partir du substrat accepteur 4-MOX. Dans cet objectif, nous avons étudié la cinétique de synthèse des chaînes de GAGs par les cellules CHO-pgsB618 transfectées avec le plasmide vide ou portant le gène codant pour la protéine sauvage ou mutée, en fonction de la concentration du substrat xylosidique. Le traitement des cellules incubées ou non en présence de 4-MOX a été couplé à un radiomarquage au  $^{35}\text{S}$ . En accord avec les résultats obtenus *in vitro* (Publication n°1, Figure 7A), l'expression de la  $\beta 4\text{GalT7}$  humaine restaure la synthèse des chaînes glycosaminoglycaniques sur le substrat xylosidique, dans les cellules déficientes en cette enzyme. De façon inattendue, le mutant A186D qui présente une très faible activité GalT *in vitro* est capable d'initier la synthèse des GAGs *ex vivo* à des taux significatifs. La substitution de l'Arg270 par une Cys induit une diminution significative de l'activité GalT mesurée *ex vivo* qui se traduit par une baisse de la quantité du  $^{35}\text{S}$  incorporée dans les chaînes de GAGs. Ce résultat est en accord avec la diminution de l'affinité de l'enzyme vis-à-vis du substrat accepteur mise en évidence par les études cinétiques. Enfin, le mutant L206P est inactif *in vitro* et est incapable de restaurer la synthèse des GAGs sur le substrat xylosidique *ex vivo*. L'ensemble de ces résultats a été également vérifié par SDS-PAGE. Ainsi, nous avons confirmé, par radiomarquage, la présence de chaînes de GAGs dans les cellules CHO-pgsB618 déficientes incubées en présence de substrat 4-MOX après transfection avec le plasmide portant le gène de la  $\beta 4\text{GalT7}$ , le mutant A186D et R270C (Publication n°2, Figure 3B).

De façon complémentaire, j'ai réalisé un marquage par immunofluorescence indirecte des cellules CHO-pgsB618, transfectées avec le plasmide vide ou portant le gène codant pour la  $\beta 4\text{GalT7}$  humaine sauvage ou pour les mutations Ehlers-Danlos, afin de déterminer l'effet de ces mutations sur la synthèse des chaînes de GAGs. Esko *et al.* (1987) ont démontré que les cellules CHO non déficientes en  $\beta 4\text{GalT7}$  produisent majoritairement des GAGs membranaires de type HS. En accord avec ces observations, l'analyse par immunofluorescence montre la présence de chaînes d'HS sur la membrane des cellules CHO-pgsB618 exprimant la  $\beta 4\text{GalT7}$  humaine sauvage et les mutants A186D et R270C révélées par l'anticorps monoclonal 10E4. Nous avons simultanément étudié la localisation subcellulaire de l'enzyme sauvage et des mutants. On constate que la  $\beta 4\text{GalT7}$  sauvage et les mutants présentent une localisation caractéristique des GTs au niveau de l'appareil de Golgi. Les cellules transfectées avec le plasmide vide n'ont montré aucune fluorescence et ne synthétisent donc pas de chaînes d'HS. Cette approche expérimentale corrobore les résultats *in vitro* et *ex vivo* obtenus précédemment. En effet, aucune présence de HS n'a été détectée

pour le mutant L206P (Publication n°2, Figure 4). Enfin, l'analyse de l'expression de ce mutant montre une localisation diffuse et différente par rapport à celle observée dans le cas de l'enzyme sauvage et des deux autres mutants. Cette observation pourrait être liée à la perturbation du processus d'adressage cellulaire de cette protéine mutée en accord avec l'altération de sa conformation.

*Publication n° 3*

## Epigenetics: methylation-associated repression of heparan sulfate 3-O-sulfotransferase gene expression contributes to the invasive phenotype of H-EMC-SS chondrosarcoma cells

Catherine Bui,\* Mohamed Ouzzine,\* Ibtissam Talhaoui,\* Sheila Sharp,<sup>†</sup> Kristian Prydz,<sup>‡</sup> Michael W. H. Coughtrie,<sup>†</sup> and Sylvie Fournel-Gigleux<sup>\*•1</sup>

\*UMR 7561 CNRS-University of Nancy I, Faculty of Medicine, Vandoeuvre-lès-Nancy, France;

<sup>†</sup>Division of Medical Sciences, University of Dundee, Ninewells Hospital and Medical School,

Dundee, UK; and <sup>‡</sup>Department of Molecular Biosciences, University of Oslo, Oslo, Norway

**ABSTRACT** Heparan sulfate proteoglycans (HSPGs), strategically located at the cell-tissue-organ interface, regulate major biological processes, including cell proliferation, migration, and adhesion. These vital functions are compromised in tumors, due, in part, to alterations in heparan sulfate (HS) expression and structure. How these modifications occur is largely unknown. Here, we investigated whether epigenetic abnormalities involving aberrant DNA methylation affect HS biosynthetic enzymes in cancer cells. Analysis of the methylation status of glycosyltransferase and sulfotransferase genes in H-EMC-SS chondrosarcoma cells showed a typical hypermethylation profile of 3-OST sulfotransferase genes. Exposure of chondrosarcoma cells to 5-aza-2'-deoxycytidine (5-Aza-dc), a DNA-methyltransferase inhibitor, up-regulated expression of 3-OST1, 3-OST2, and 3-OST3A mRNAs, indicating that aberrant methylation affects transcription of these genes. Furthermore, HS expression was restored on 5-Aza-dc treatment or reintroduction of 3-OST expression, as shown by indirect immunofluorescence microscopy and/or analysis of HS chains by anion-exchange and gel-filtration chromatography. Notably, 5-Aza-dc treatment of EMC cells or expression of 3-OST3A cDNA reduced their proliferative and invading properties and augmented adhesion of chondrosarcoma cells. These results provide the first evidence for specific epigenetic regulation of 3-OST genes resulting in altered HSPG sulfation and point to a defect of HS-3-O-sulfation as a factor in cancer progression.—Bui, C., Ouzzine, M., Talhaoui, I., Sharp, S., Prydz, K., Coughtrie, M. W. H., Fournel-Gigleux, S. 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.* 24, 000–000 (2010). www.fasebj.org

**Key Words:** tumor pathogenesis • proteoglycan synthesis • epigenetic regulation • glycosyltransferase

ACCUMULATING EVIDENCE INDICATES that cell function and phenotype are highly influenced by heparan sul-

fate proteoglycans (HSPGs) present in the extracellular matrix and on the surface of virtually every animal cell type. Located at the cell-tissue-organ interface, HSPGs interact with numerous molecules, including growth factors, cytokines, extracellular matrix components, and coagulation factors, thereby regulating a large number of biological processes (1). These interactions are important not only in physiological processes, such as organogenesis, angiogenesis, blood coagulation, growth factor signaling, and lipid metabolism, but are also implicated in pathological circumstances such as viral/bacterial adherence and cancer (2). There is now compelling evidence that HSPGs regulate several aspects of cancer biology, including tumorigenesis, tumor invasion, and metastasis (3). Tumor HSPGs differ in composition from those in corresponding normal tissue, which might affect the efficiency of growth factor stimulation of cancer cells (4). Similarly, the tumor microenvironment might affect the supporting vasculature, which differs in structure and integrity compared with vessels in normal tissue (5). Thus, it appears that the cellular roles of HSPGs are often misregulated in cancer, as a result of alterations in HSPG structure and expression level.

The synthesis of the glycosaminoglycan (GAG) chains that govern these functions is initiated by the formation of a tetrasaccharide linkage region composed of glucuronic acid-galactose-galactose-xylose (GlcA $\beta$ 1,3Gal $\beta$ 1,3Gal $\beta$ 1,4Xyl $\beta$ 1-), where xylose is attached to specific serine residues of the core protein of PGs. This linkage tetrasaccharide is built up by the stepwise addition of each sugar residue derived from the corresponding UDP-sugar and catalyzed by O-xylosyltransferase I (XT-I) (6),  $\beta$ 1,4-galactosyltransferase 7 ( $\beta$ 4GalT7) (7),  $\beta$ 1,3-galactosyltransferase 6 ( $\beta$ 3GalT6) (8), and  $\beta$ 1,3-glucuronosyltransferase I (GlcAT-I) (9). Following the transfer of  $\alpha$ -N-acetylglucosamine (GlcNAc), polymer-

<sup>1</sup> Correspondence: UMR CNRS 7561-University Henri Poincaré Nancy I, Faculty of Medicine, BP 184, 54505 Vandoeuvre-lès-Nancy, France. E-mail: sfg@medecine.uhp-nancy.fr  
doi: 10.1096/fj.09-136291

ization of HS chains proceeds *via* the addition of  $\beta$ 1,4-GlcA and  $\alpha$ 1,4-GlcNAc units in alternating sequence to the nonreducing end of the growing polymer. The elongation reaction is carried out by HS-copolymerases encoded by the *EXT1/EXT2* tumor suppressor genes (10). Concomitant with chain elongation, modification of the chain occurs, initiated by *N*-deacetylation/*N*-sulfation of GlcNAc units. Further modification reactions include C5 epimerization of some GlcA into L-iduronic acid (IdoUA), 2-*O*-sulfation of IdoUA, and 6-*O*-sulfation of GlcNAc residues. *O*-sulfation also occurs at position 3 of *N*-sulfated and *N*-unsubstituted GlcNAc units, the latter modifications being the hallmarks for the binding domains of antithrombin and glycoprotein gD of herpes simplex virus type 1 (HSV-1). These modifications proceed *via* the action of HS-biosynthetic enzymes, including *N*-deacetylase/*N*-sulfotransferases (NDSTs), 2-*O*-sulfotransferases (HS2STs), 3-*O*-sulfotransferases (3-OSTs), and 6-*O*-sulfotransferases (HS6STs) (11).

A number of studies highlight the potential of carbohydrate-based therapeutics in several diseases (12). However, reaching this step requires advanced understanding of the mechanisms involved in the synthesis of these macromolecules. Although the basic biosynthetic scheme for HS has been outlined (13), only a limited number of studies have addressed the mechanisms underlying the assembly process of GAG chains (14–16). We recently showed that modulation of XT-I expression regulates GAG-chain synthesis during cartilage destruction and repair in rheumatoid conditions (17). It is well established that genetic defects in HS biosynthetic enzymes have severe biological consequences. In humans, mutations in the *EXT1* and *EXT2* genes coding for HS-polymerases led to a bone disorder known as hereditary multiple exostoses (HME), characterized by the formation of benign cartilage-capped tumors. This condition is associated with a significantly higher risk of developing malignancies such as chondrosarcomas and osteosarcomas (18). Interestingly, mice deficient in *Ext2* lack HS and also develop exostoses (19). Mutations of the  $\beta$ 4*Galt7* gene coding for a glycosyltransferase involved in the initial step of GAG synthesis are associated with the progeroid form of Ehlers-Danlos syndrome, characterized by aged appearance, developmental delay, dwarfism, and connective tissue abnormalities (20).

Epigenetic phenomena are increasingly implicated in the pathogenesis of various disorders (21), and the role of epigenetic alterations in carcinogenesis is clearly established (22). One main epigenetic change associated with tumorigenesis is aberrant DNA methylation. Malignant cells show major disruptions in their DNA methylation profiles, which manifest as hypermethylation of gene promoters, global hypomethylation, and increased rates of mutation at methylated CpG dinucleotides. Hypermethylation of CpG-enriched regions, also called “CpG islands,” generally located in the promoter and downstream proximal sequence region, has been the most extensively studied area of research of DNA methylation. In cancer cells, hypermethylation of these regions results in the loss of gene expression. Some of these susceptible genes include tumor suppres-

or genes, cell cycle regulator genes, DNA repair genes, and genes associated with apoptosis. In addition, a correlation between aberrant methylation of promoter regions and transcriptional activity of various genes coding for enzymes involved in matrix remodeling, such as matrix metalloproteinases and heparanase, has also been described (23). Interestingly, inspection of the 5′ region of the genes coding for the glycosyltransferases and most sulfotransferases involved in HS synthesis and modification reveals the presence of typical CpG islands, raising the possibility that expression of these enzymes might be subject to epigenetic silencing *via* DNA methylation in tumor cells. Supporting this assumption, recent findings highlight the importance of *EXT1* epigenetic inactivation leading to abrogation of HS synthesis in the process of tumor onset and progression (24). Furthermore, aberrant methylation of the *3-OST2* gene has been described in several types of tumors (25).

The critical role of HSPGs in human cancer emphasizes the importance of understanding the regulation of these macromolecules at the molecular genetic level. Here, we have investigated the role of epigenetics in the regulation of HS-biosynthetic enzymes in a human chondrosarcoma cell line, since gene disruption in the HS pathway is associated with bone-cartilage tumors. Specifically, we aimed to determine whether the glycosyltransferase and sulfotransferase genes responsible for HS chain synthesis and modification are affected by an epigenetic mechanism involving DNA methylation. Our results provide strong evidence for aberrant methylation and down-regulation of the *3-OST* gene family in the H-EMC-SS (HEMC) cell line, which is associated with altered HS glycosaminoglycan chains, as demonstrated both at biochemical and cellular levels. Furthermore, we show that treatment of the cells with the DNA-methyltransferase inhibitor 5-aza-2′-deoxycytidine (5-Aza-dc) or reintroduction of 3-OSTs into the chondrosarcoma cell line by cDNA expression modified cell behavior in terms of proliferation, adhesion, and migration. Notably, 5-Aza-dc exposure or 3-OST3A cDNA expression resulted in a decrease in the proliferative and migration properties of chondrosarcoma cells, and augmented their adhesion capacity. These findings underline the significance of HS alterations in the phenotype of invasive HEMC chondrosarcoma cells and identify 3-*O*-sulfation as a contributing factor to this process. To our knowledge, this is the first report showing down-regulation of a specific set of HS-*O*-sulfotransferase enzymes in chondrosarcoma cells involving an epigenetic mechanism. Our observation may be of particular relevance in understanding the process of tumor progression and metastasis and in supporting the development of novel therapies targeting DNA methylation.

## MATERIALS AND METHODS

### Cell lines, culture conditions, and 5-Aza-dc treatment

The human chondrosarcoma cell line HEMC was obtained from the European Collection of Animal Cell Cultures

(ECACC; Salisbury, UK) and was cultured in F12/Dulbecco's modified Eagle's medium (F12/DMEM) (1:1) supplemented with 10% (v/v) fetal bovine serum (FBS; Gibco-Invitrogen, Cergy-Pontoise, France), penicillin (100 U/ml)/streptomycin (100 µg/ml), and 1 mM glutamine. The HL-60 acute promyelocytic leukemia-derived cell line was obtained from ECACC and cultured in RPMI-1640 supplemented with 10% (v/v) FBS, penicillin (100 U/ml)/streptomycin (100 µg/ml), and 1 mM glutamine. HeLa cells were cultured in DMEM supplemented with 10% (v/v) FBS, penicillin (100 U/ml)/streptomycin (100 µg/ml), and 1 mM glutamine. Cells were maintained in a 37°C incubator in a 5% CO<sub>2</sub> humidified atmosphere.

Chondrosarcoma cells were seeded in 25-cm<sup>2</sup> flasks at 5 × 10<sup>5</sup> cells, allowed to attach overnight, and treated with 5-Aza-dc (Sigma-Aldrich, Gillingham, UK) dissolved in phosphate-buffered saline (PBS) at 2 and 4 µM final concentration for 7 d. Control cultures contained the vehicle only. During the course of treatment, medium was changed every day; cells were passaged twice and placed in fresh medium containing the same concentration of 5-Aza-dc or vehicle. At the end of the treatment period, cells were washed once with PBS and harvested for genomic DNA or mRNA analysis. Alternatively, cells were subjected to the appropriate medium for metabolic radiolabeling prior to GAG analysis or for other cellular assays (see below).

#### CpG island analysis, genomic DNA preparation, bisulfite modification, and methylation pattern analysis

The DNA sequences around the transcription start site (−1000 bp) and the first exon of each glycosyltransferase and sulfotransferase gene were obtained by a GenBank database search (<http://genome-www.stanford.edu/genecards>), and the presence of CpG islands was predicted using the MethPrimer online search tool (<http://www.urogene.org/methprimer/>) (26). The criteria used for the definition of CpG islands included regions of DNA >500 bp with G+C content ≥55%. The following genes contain a typical CpG island in the proximal promoter/first exon sequence: *XT-1*, *β4GalT7*, *β3GalT6*, *GlcAT-1*, *EXT1*, *NDST2*, *NDST3*, *HS2ST1*, *3-OST1*, *3-OST2*, *3-OST3A*, *3-OST3B*, *3-OST6*, *HS6ST1*, *HS6ST2*, and *HS6ST3*. The *NDST1* gene contains a cluster of CpG nucleotides encompassing 366 bp (+302 to +668) located in the 5' untranslated region of the first exon, and the *NDST4* sequence contains no region meeting the defined criteria of a CpG island.

The methylation status of the glycosyltransferase and sulfotransferase genes was determined by the genomic DNA bisulfite-conversion method, which induces chemical modification of cytosine residues to uracil in unmethylated DNA, whereas those that are methylated are resistant to the treatment and remain as cytosines (27). Genomic DNA was extracted from untreated or 5-Aza-dc-treated HEMC cells using GenomicPrep DNA isolation kit (GE Healthcare, Amersham, UK). The amount and purity of DNA were determined using a Nanodrop spectrophotometer (Fisher Scientific, Loughborough, UK), and 0.2 µg DNA was used for bisulfite conversion using the EpiTect Bisulfite kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. Genomic DNA from human whole blood (Clontech, Mountain View, CA, USA) was used as a source of control DNA. Following bisulfite conversion, two sequential PCR reactions were performed to amplify the modified fragments of interest. The first round of amplification was carried out with 0.1 µg of DNA template using primer sets ("Out sens" and "Out rev"). Nested primers ("In sens" and "In rev") were used in the subsequent PCR reaction. Primer sets were designed, including the use of Primer3 Input 0.4.0 software

([primer3-web/htdocs/input-040.htm](http://primer3-web/htdocs/input-040.htm)) for the prediction of primer hairpins, dimers, and false primers and are listed in Supplemental Table 1. PCR conditions for the first PCR reaction were 94°C for 8 min followed by 20 cycles at 94°C for 30 s, 52°C for 30 s, 72°C for 30 s, ending with a final extension of 10 min at 72°C. PCR conditions for the second step were 94°C for 8 min followed by 35 cycles at 94°C for 30 s, 55°C for 30 s, 72°C for 30 s, and finally 10 min at 72°C. The resultant PCR products were purified with the QIAquick PCR Purification kit (Qiagen). Initial screening experiments were performed by direct sequencing PCR products from bisulfite converted DNA samples. In a second step, the PCR amplicons were subcloned into pGEM-T-easy TA (Promega, Madison, WI, USA), and at least two recombinant clones were sequenced for each gene of interest. Nonmethylated cytosines were successfully converted by the bisulfite reaction in the PCR amplicons or recombinant clones.

#### Reverse transcription-PCR analysis of mRNA expression level

Total RNA was isolated from untreated and 5-Aza-dc-treated HEMC cells using the SV RNA Isolation System (Promega). RNA (0.5 µg) was reversed-transcribed into cDNA using oligodT primer and PrimeScript reverse transcriptase (Clontech). The resultant cDNA was used for PCR amplification with Bio-X-ACT DNA polymerase (Bioline, London, UK) following optimization of MgCl<sub>2</sub> concentrations and annealing temperatures. To determine the linear amplification segment for each gene of interest, different amplification steps were initially carried out between 20 and 40 cycles, and the number of cycles for quantification was established to remain in the exponential phase of amplification. The primers and specific PCR conditions used are listed in Supplemental Table 2. The PCR products were electrophoresed on 2.0% (w/v) agarose gels and stained with ethidium bromide, and the expression levels were evaluated following image capture using GeneFlash (Syngene, Cambridge UK) and Scion Image 1.63 (Scion Corp., Frederick, MD, USA) software. The mRNA expression levels of *EXT1*, *NDST1*, *NDST2*, *HS2ST1*, *HS6ST1*, *HS6ST2*, *3-OST1*, *3-OST2*, *3-OST3A*, *3-OST3B*, and *3-OST6* were quantified relative to glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) as an internal control. For *NDST3*, *NDST4*, and *HS6ST3*, no transcript could be detected in control or 5-Aza-dc-treated HEMC cells when amplification was performed up to 40 cycles. However, these transcripts were successfully amplified from brain or placenta total mRNA (Clontech) used as positive control (not shown).

#### 3-OST1 and 3-OST3A cDNA expression in HEMC cells

3-OST1 and 3-OST3A cDNA were amplified by PCR from IMAGE clones (Geneservice, Cambridge, UK) as templates, using a 5' primer containing a *Hind*III site and a Kozak sequence, and a 3' primer containing the sequence encoding YPYDVPDYA-hemagglutinin (HA) peptide tag and an *Xho*I site. The amplified fragments were inserted into *Hind*III-*Xho*I sites of pCDNA3.1 expression vector (Invitrogen) to generate pCDNA-3OST1 and pCDNA-3OST3A plasmids. Each vector was transiently transfected into HEMC cells, as described previously (28). Briefly, HEMC cells plated into 6-well plates, grown at 80% confluency, were transfected with 2.5 µg of plasmid using ExGen 500 reagent (Euromedex, Souffelweyersheim, France). The empty vector was used for mock-transfection, as a control. Forty-eight hours after transfection, cells were trypsinized and plated for cellular assays, as described below. Simultaneously, expression of recombinant proteins 3-OST1 and 3-OST3A was analyzed by Western

blotting using anti-HA antibodies (Abcam, Cambridge, UK), as described previously (29).

#### Indirect immunofluorescence staining of chondrosarcoma and HeLa cells

For indirect immunofluorescence staining of cell/matrix HS chains, untreated, 5-Aza-dc-treated cells, or cells transfected with pCDNA-3OST1 and pCDNA-3OST3A were washed with PBS and fixed with 3% paraformaldehyde/PBS for 20 min at room temperature. For HS staining, coverslips were rinsed 3 times in PBS and blocked in 10% (v/v) goat serum/PBS. Cells were then incubated with mouse anti-human HS antibody (10E4 epitope) at 1:50 dilution (Seikagaku Corp., Tokyo, Japan) for 1 h, washed 3 times with PBS, incubated with AlexaFluor 594-conjugated goat anti-mouse antibodies (Molecular Probes/Invitrogen) for 30 min and finally washed with PBS. For 3-OST1 and 3-OST3A-tagged protein staining, cells were permeabilized with a PBS solution containing 0.1% (w/v) Triton-X100 for 4 min and incubated in 10% (v/v) goat serum/PBS. Recombinant cells were incubated with FITC-conjugated mouse anti-HA antibody (Sigma-Aldrich) for 30 min, and washed with PBS. Coverslips were treated with a 5  $\mu$ g/ml Hoechst/PBS solution for nucleus visualization and were finally mounted with Hydromount or Moviol (National Diagnostics, Hull, UK) containing 2.5% 1,4-diazabicyclo[2.2.2]octane (DABCO; Sigma-Aldrich).

#### Metabolic labeling and ion-exchange and gel-filtration chromatography analysis of cell HSPGs

Subconfluent untreated and 5-Aza-dc-treated HEMC cells were transferred to appropriate labeling media containing 2% FBS with [<sup>35</sup>S]SO<sub>4</sub> (PerkinElmer Life Sciences, Courtabœuf, France) in Fisher's medium without sulfate (GIBCO-Invitrogen), or [<sup>3</sup>H]glucosamine (PerkinElmer, Life Sciences, Courtabœuf, France) in RPMI 1640 (Gibco-Invitrogen) without glutamine, or [<sup>35</sup>S] Easytag Express Protein Labeling Mix (73% L-[<sup>35</sup>S]methionine and 22% L-[<sup>35</sup>S]cysteine) in DMEM without L-methionine and L-cysteine for 22 h. Cells were washed twice with PBS, harvested by scraping, and pelleted by centrifugation. Cell pellets were resuspended in 50 mM Tris-HCl buffer (pH 7.5) and 1% (w/v) Triton X-100, and total DNA amounts were quantified using the Quant-iT dsDNA HS Assay Kit (Molecular Probes/Invitrogen). Nonincorporated radioactive label was removed from cell extracts by Sephadex G-50 fine chromatography, as described previously (30), and radiolabeled macromolecules were analyzed by anion-exchange chromatography or gel filtration based on previously described methods (31).

For ion-exchange chromatography, [<sup>35</sup>S]SO<sub>4</sub> or L-[<sup>35</sup>S]methionine/L-[<sup>35</sup>S]cysteine radiolabeled samples corresponding to the same amount of treated or untreated cells (based on DNA quantification) were incubated with chondroitinase ABC from *Proteus vulgaris* (Sigma-Aldrich) (32) and then subjected to DEAE-anion-exchange chromatography (Bio-LogicLP system; Bio-Rad, Hercules, CA, USA), as described previously (33). Briefly, samples were diluted in 1 ml of buffer A (8 M urea, 0.02 M Bis-tris, and 0.05 M NaCl, pH 6.0) and applied onto a 2.5-ml Macro-Prep DEAE support (Bio-Rad). After 20 ml of buffer A, a linear gradient mixing buffer A with buffer B (8 M urea, 0.02 M Bis-tris, and 0.8 M NaCl, pH 6.0) was applied at a rate of 0.5 ml/min. Fractions of 1 ml were collected and analyzed for radioactivity by scintillation counting.

For gel-filtration chromatography, [<sup>35</sup>S]SO<sub>4</sub> or [<sup>3</sup>H]glucosamine-radiolabeled macromolecules were treated with proteinase K, as described previously (34). Following heat

inactivation, samples were submitted to chondroitinase ABC digestion, as described above. Samples corresponding to 5-Aza-dc-treated or untreated cells were applied to a column (1×40 cm) of Sepharose CL-6B (Sigma-Aldrich) along with blue dextran and K<sub>2</sub>CrO<sub>4</sub> as internal standards, and elution was performed with 0.15 M NaCl in 0.05 M Tris, HCl buffer (pH 8.0), and 0.1% Triton X-100 at a rate of 6 ml/h, as described previously (35). Fractions of 1 ml were collected and analyzed for radioactivity by scintillation counting.

#### Cell proliferation, adhesion, and migration assays

For proliferation assays, 5-Aza-dc-treated or transfected cells were seeded onto 24-well plates (Costar; Dutscher, Brumath, France) at 25 × 10<sup>3</sup> cells/well. Cells were harvested each day during a 5-d period by scraping in lysis buffer [PBS containing 1% (w/v) sodium dodecyl sulfate]. Cell lysates were sonicated in a total volume of 1 ml PBS buffer, and cell proliferation was evaluated by measuring the DNA content of each sample by spectrofluorimetry on a fluorimeter (Qubit; Molecular Probes-Invitrogen, Cergy Pontoise, France) using the Quant-iT dsDNA HS Assay kit.

For adhesion assays, 96-well plates were coated by adsorption with 100 or 200  $\mu$ g fibronectin (from human foreskin fibroblasts, Sigma-Aldrich) per well for 1 h at 37°C, or left untreated for control experiments. After washing with medium containing 0.1% BSA, wells were saturated by incubation with F12/DMEM containing 0.5% BSA for 1 h at 37°C. Plates were then washed once with medium containing 0.1% BSA and placed on ice before cells were harvested by trypsinization, counted and resuspended in complete F12/DMEM at a density of 5 × 10<sup>5</sup> cells/ml. Each well was inoculated with 25 × 10<sup>3</sup> cells and incubated for 30 min at 37°C, and nonadherent cells were removed by washing. The extent of adhesion was determined after fixing cells with 4% paraformaldehyde for 15 min, staining adherent cells with 1% crystal violet for 10 min. Plates were rinsed 3 times in water, and 10% (v/v) acetic acid solution was added to each well to allow staining quantification by absorbance measurement at 550 nm on a spectrophotometer (SAFAS Ix-Xenius, Monaco, France).

HEMC cell migration was analyzed using a Boyden chamber chemotaxis assay (36). Following 5-Aza-dc treatment or 3-OST cDNA expression, cells were incubated with 1  $\mu$ g/ml mitomycin C (Sigma-Aldrich) in complete F12/DMEM for 2 h and seeded at 5 × 10<sup>4</sup> cells/well on top of 24-well plate inserts (Corning Costar; Sigma-Aldrich) containing a polycarbonate filter (6.5-mm diameter, 8- $\mu$ m pores). Cells were incubated for 8, 16, or 24 h for 5-Aza-dc treated cells, and for 18 h in the case of transfected cells at 37°C in the upper chamber in F12/DMEM containing 0.1% BSA, whereas the lower chamber contained F12/DMEM supplemented with 5% FBS. At the end of the incubation period, nonmotile cells were removed from the upper side of the filters with cotton-tipped swabs. Migration of cells was quantified by staining with methyl blue, and measuring the optical density at 663 nm of cells that had migrated to the underside of the filters.

#### Statistical analysis

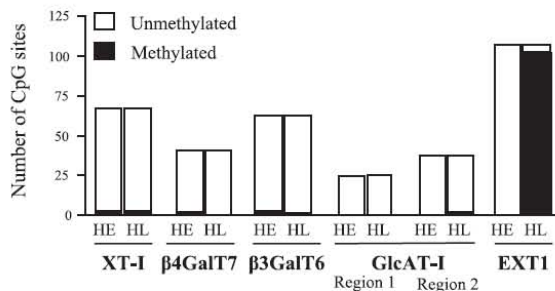
All data are presented as means  $\pm$  SE of 3 to 6 observations. For comparison among multiple groups, either 1-way or 2-way ANOVA followed by Bonferroni *post hoc* correction test was used. Statistical analyses were performed using GraphPad Prism 4.0 software (GraphPad, San Diego, CA, USA).

## RESULTS

### DNA methylation profiling of *N*- and *O*-sulfotransferase genes in HEMC chondrosarcoma cells

To test whether the HS synthesis pathway is regulated *via* epigenetic mechanisms involving DNA methylation in chondrosarcoma cells, we systematically searched the genomic sequences encoding the biosynthetic enzymes for the presence of clustered CpG dinucleotides, representing putative hypermethylation loci.

A region encompassing 1000 bp upstream of the transcription start site and the first exon of *XT-I*,  $\beta 4GalT7$ ,  $\beta 3GalT6$ , *GlcAT-I*, and *EXT1* genes coding for the glycosyltransferases involved in HS chains synthesis was analyzed. We found that each of these genes contains a typical 5' CpG island, indicating that they have the potential for hypermethylation-associated silencing. We determined the methylation status of the CpG islands associated with the *XT-I*,  $\beta 4GalT7$ ,  $\beta 3GalT6$ , and *GlcAT-I* genes by the bisulfite-sequencing method, using primers presented in Supplemental Table 1. These genes encode the glycosyltransferases responsible for the formation of the tetrasaccharide linkage region, which is common to the biosynthetic pathways of both HS and chondroitin sulfate (CS) chains, and is a prerequisite step for GAG-chain elongation. We also examined the sequence of the promoter CpG island of *EXT1*, which encodes a glycosyltransferase specific for HS-chain polymerization, which has been reported to be hypermethylated in leukemia cell lines (24). Our results show that all glycosyltransferase DNA sequences analyzed within the proximal promoter region and first exon were unmethylated or exhibited a low frequency of methylated cytosine residues within their CpG islands in HEMC cells (Fig. 1). The methylation state of these genes in chondrosarcoma cells was similar to that of normal whole blood



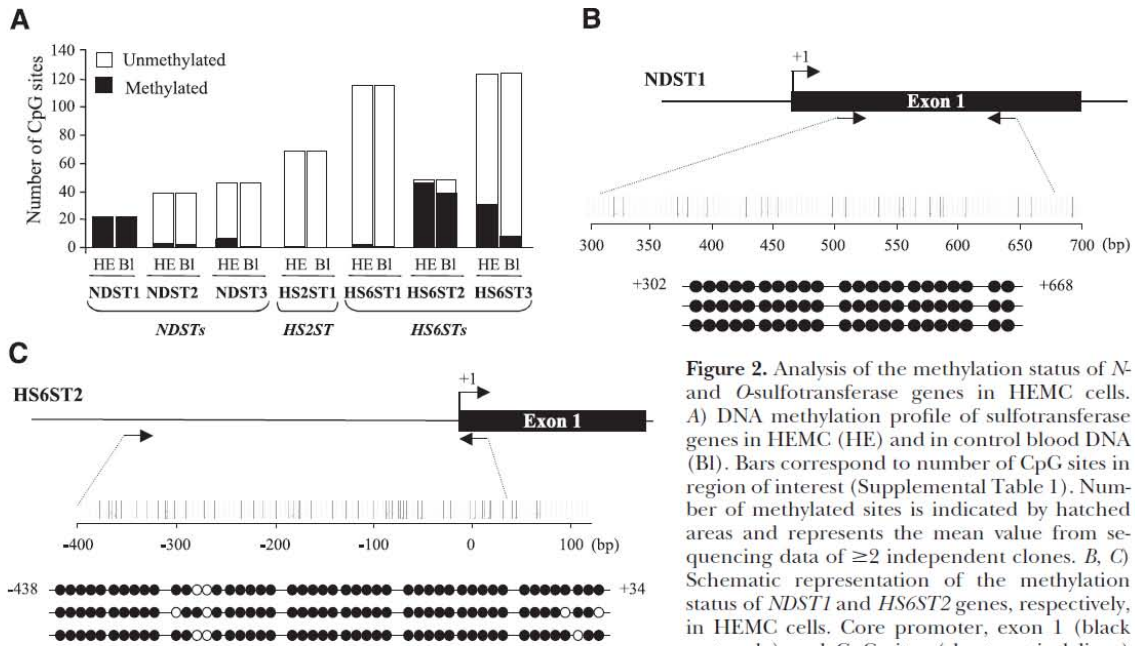
**Figure 1.** Analysis of the methylation status of glycosyltransferase genes in HEMC and HL-60 cells. Genomic DNA was extracted from chondrosarcoma HEMC (HE) or HL-60 (HL) cells, and the methylation profile of the CpG island of the glycosyltransferase genes was analyzed by bisulfite sequencing. Number of methylated CpG sites was determined by counting unconverted cytosine residues. Bars correspond to number of CpG sites in region of interest (Supplemental Table 1). Number of methylated sites is indicated by hatched areas and represents the mean value from sequencing data of 2 independent clones.

DNA (data not shown). For comparison purposes, we also performed bisulfite genomic sequencing of the CpG island regions of the *XT-I*,  $\beta 4GalT7$ ,  $\beta 3GalT6$ , *GlcAT-I*, and *EXT1* genes in HL-60 leukemia cells. The results showed a lack or low number of methylated loci in *XT-I*,  $\beta 4GalT7$ ,  $\beta 3GalT6$ , and *GlcAT-I* glycosyltransferase genes, similar to our observations in HEMC cells. By contrast, in the HL-60 leukemia cell line, the *EXT1*-promoter region analyzed (corresponding to positions  $-536$  to  $+203$  relative to the transcription start site) contained 98% methylated cytosines among 106 CpG sites (Fig. 1). This is indicative of the typical hypermethylation profile of the *EXT1*-promoter region in certain types of leukemia cell lines, in agreement with a previous report (24). Together, these data provide no evidence for methylation differences between HEMC and normal cells in the glycosyltransferase genes involved in the initiation and polymerization of HS chains, whereas the *EXT1* CpG island exhibits a hypermethylation profile in HL-60 cells.

Concomitant with the synthesis of HS chains, substitution of the saccharide residues with sulfate groups occurs *via* the action of several families of *N*- and *O*-sulfotransferases governing key biological functions (11). The first modification enzymes, NDSTs, remove *N*-acetyl groups from selected GlcNAc residues and replace them with sulfate groups. Sulfation of IdoUA and GlcA residues is carried out by 2-*O*-sulfotransferases, followed by 6-*O*-sulfation and 3-*O*-sulfation of glucosamine units. To determine whether sulfation of HS chains might be epigenetically regulated through DNA methylation, we examined the promoter region/first exon of NDST and HS-*O*-sulfotransferase genes in HEMC cells. Although the promoter of the *NDST1* gene contained no canonical CpG island, a cluster of CpG dinucleotides (22 sites) located in the 5' untranslated region of the first exon of this gene ( $+302$  to  $+668$ ) was identified. Bisulfite conversion analysis showed that this region was fully methylated both in HEMC cells and in blood DNA (Fig. 2A, B). Analysis of the methylation pattern of the CpG island present in the *NDST2* and *NDST3* genes indicated a low methylation frequency in the case of *NDST2* (0–4 of 39 sites), as well as in the *NDST3* sequence (6–7 of 46 sites) (Fig. 2A). Analysis of the *NDST4* gene revealed the presence of a small number of CpG sites (5 in a 453-bp region located in the proximal promoter sequence; Supplemental Table 1).

Interestingly, analysis of HS-*O*-sulfotransferase genes, including *HS2ST*, *HS6ST*, and *3-OST* families, revealed that a majority of these genes contain a canonical CpG island. In HEMC cells, as in control genomic DNA, the predicted CpG island present in the *HS2ST1* and *HS6ST1* genes exhibited a very low methylation frequency (1–1 of 68 CpG sites and 0–2 of 115 methylated CpG sites, respectively) (Fig. 2A). On the other hand, the *HS6ST2* gene showed a typical hypermethylation profile in HEMC cells (44–45–46 methylated cytosines of 48 CpG sites,  $45 \pm 1$ ), which was also observed in blood samples (Fig. 2A, C). Analysis of the *HS6ST3* gene showed a highly dense CpG dinucleotide region ( $-721$  to  $+8$ ) comprising 123 CpG sites. About 30 and





**Figure 2.** Analysis of the methylation status of *N*- and *O*-sulfotransferase genes in HEMC cells. *A*) DNA methylation profile of sulfotransferase genes in HEMC (HE) and in control blood DNA (BI). Bars correspond to number of CpG sites in region of interest (Supplemental Table 1). Number of methylated sites is indicated by hatched areas and represents the mean value from sequencing data of  $\geq 2$  independent clones. *B*, *C*) Schematic representation of the methylation status of *NDST1* and *HS6ST2* genes, respectively, in HEMC cells. Core promoter, exon 1 (black rectangle) and CpG sites (short vertical lines) are shown in the top part of each panel. Predicted transcription start site is marked by a curved arrow; primers are indicated by arrows. Methylated cytosines (solid circles) and unmethylated cytosines (open circles) in the region of interest are represented in the bottom part of each panel. Results are from bisulfite sequencing of 3 different clones.

5.7% of these sites were found to be methylated in HEMC cells and blood DNA, respectively (Fig. 2A).

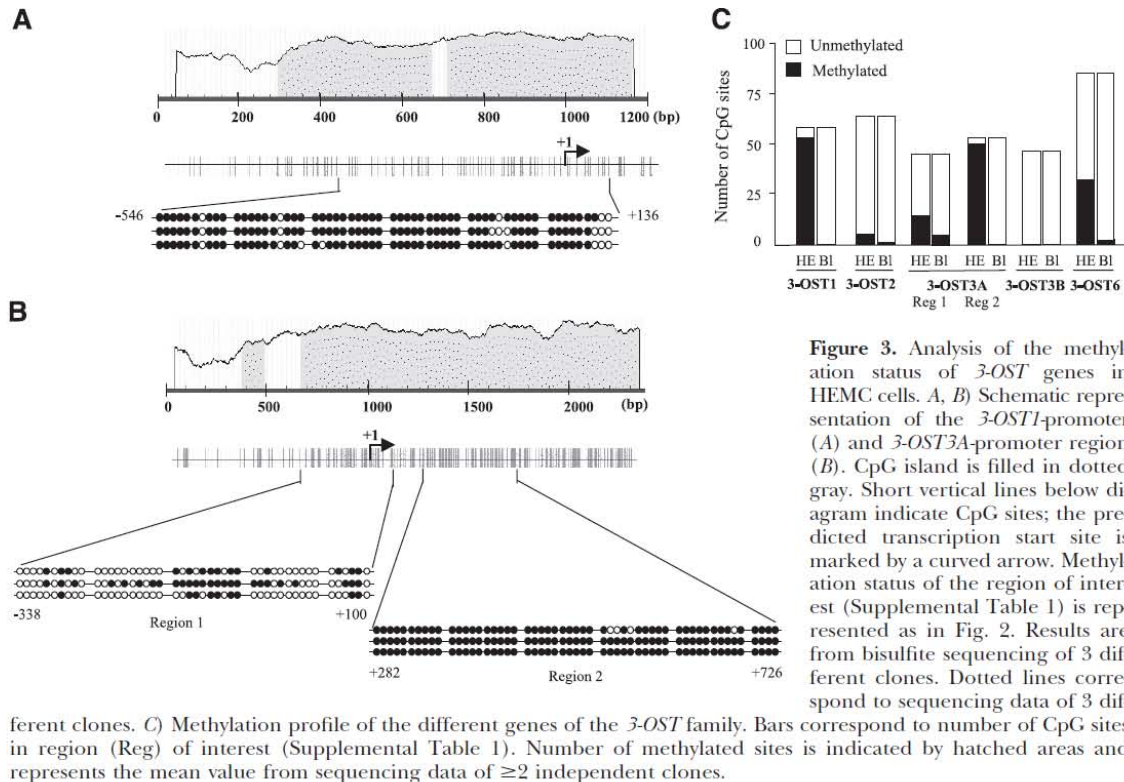
We finally investigated the methylation state of the largest HS-*O*-sulfotransferase gene family, *3-OST*, which comprises 7 members known to biosynthesize key structures involved in several important physiological processes and possibly in the development and/or progression of tumors (25). The CpG island encompassing 682 bp located in the proximal promoter and 5' region of the first exon (-546 to +136) of the *3-OST1* gene is illustrated in Fig. 3A. Analysis of the methylation status of this fragment showed ~88% methylated sites (51–51–54 methylated cytosines among 59 CpG dinucleotides,  $52.0 \pm 1.7$ ), revealing a typical hypermethylation profile in HEMC cells (Fig. 3A, C). No methylated site was found when the same region was studied in blood DNA (Fig. 3C). Surprisingly, bisulfite sequencing of the *3-OST2* gene, previously reported to be hypermethylated in a range of human cancers (25), showed only a moderate methylation frequency in HEMC cells, with ~10% methylated sites in a 612 bp-fragment (-520 to +92, 5–8 methylated cytosines among 64 CpG sites) (Fig. 3C). The 5' domain of the *3-OST3A* gene contains a high density of CpG dinucleotides typical of a CpG island, as illustrated in Fig. 3B. Examination of the methylation state of two regions corresponding to positions -338 to +100 and +282 to +726 relative to the transcription start site revealed 35.5% (9–14–26 among 46 sites) and 97% (50–54–54 among 54 CpG sites,  $52.7 \pm 2.3$ ) methylated cytosines, respectively (Fig. 3B, C). By contrast, bisulfite sequencing of the *3-OST3B* gene showed a lack of methylated loci from position -354 to +218, despite the presence of a typical CpG

cluster (47 CpG sites) (Fig. 3C). Finally, analysis of the methylation profile of the *3-OST6* CpG island revealed that 38% of CpG sites in a 739 bp-fragment (-647 to +92) were methylated (28–37 methylated cytosines among 86 CpG sites, Fig. 3C). Analysis of the methylation profile of the *3-OST* genes performed on blood DNA sample showed <2% methylated sites (Fig. 3C). Altogether, the *3-OST* gene family revealed a typical aberrant methylation pattern specific to HEMC cells, with the most striking hypermethylation profile observed in the case of *3-OST1* and *3-OST3A*.

#### Hypermethylated sulfotransferase gene expression is elevated by 5-Aza-dc treatment in HEMC cells

To study the potential role of CpG methylation on the transcriptional activity of glycosyltransferase and sulfotransferase genes, we analyzed the effect of an inhibitor of DNA-methyltransferase, 5-Aza-dc (37), on mRNA expression. HEMC cells were treated with two concentrations of 5-Aza-dc and mRNA expression levels were compared in treated and untreated cells by semiquantitative RT-PCR. Expression analysis of *EXT1* in the chondrosarcoma cell line showed that HEMC cells expressed a high level of *EXT1* transcript and that mRNA expression was unchanged on 5-Aza-dc treatment (Fig. 4A, C). This result is consistent with the unmethylated status of the CpG island of the *EXT1* proximal promoter/first exon in HEMC cells.

Because aberrant methylation was essentially observed in the case of sulfotransferase genes, further investigation of the consequences of 5-Aza-dc treatment

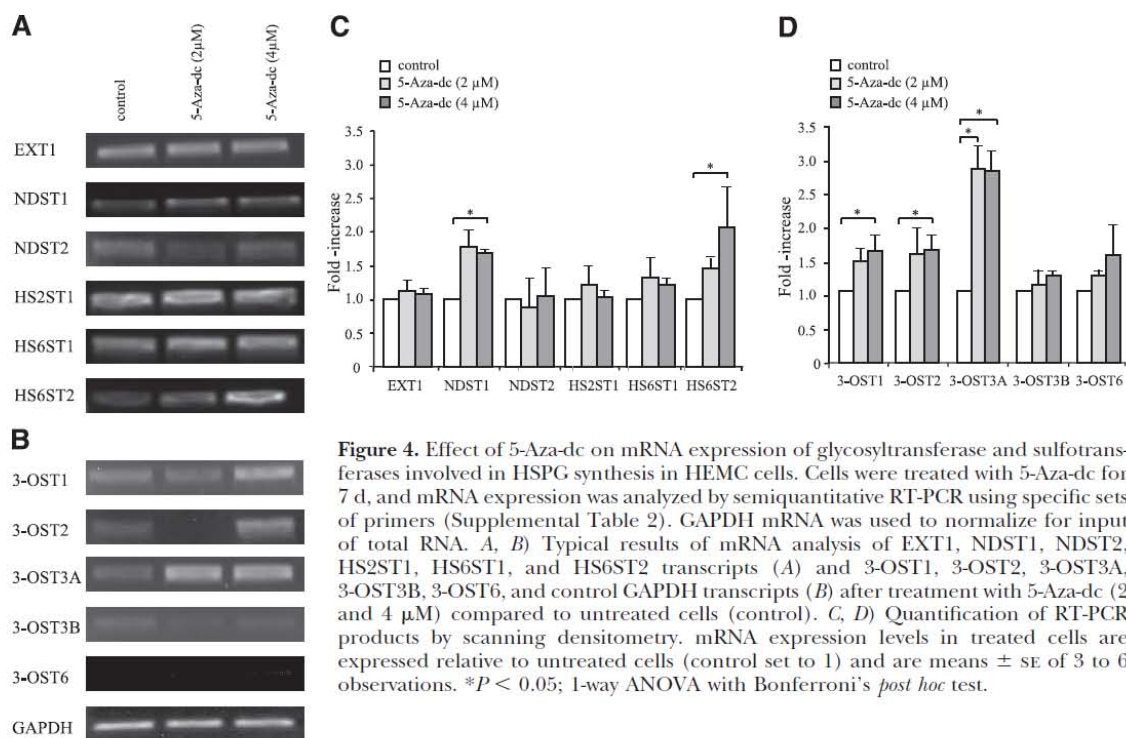


**Figure 3.** Analysis of the methylation status of 3-OST genes in HEMC cells. *A, B*) Schematic representation of the 3-OST1-promoter (*A*) and 3-OST3A-promoter region (*B*). CpG island is filled in dotted gray. Short vertical lines below diagram indicate CpG sites; the predicted transcription start site is marked by a curved arrow. Methylation status of the region of interest (Supplemental Table 1) is represented as in Fig. 2. Results are from bisulfite sequencing of 3 different clones. Dotted lines correspond to sequencing data of 3 different clones. *C*) Methylation profile of the different genes of the 3-OST family. Bars correspond to number of CpG sites in region (Reg) of interest (Supplemental Table 1). Number of methylated sites is indicated by hatched areas and represents the mean value from sequencing data of  $\geq 2$  independent clones.

on sulfotransferase mRNA expression was carried out. RT-PCR analysis showed that NDST1 mRNA level was significantly increased (up to 1.7-fold) on exposure of HEMC cells to the highest dose of 5-Aza-dc (Fig. 4A, C), whereas NDST2 mRNA expression was unchanged, in agreement with the unmethylated state of the promoter CpG island of this gene in chondrosarcoma cells. By contrast, no NDST3 or NDST4 transcripts were detected in control or 5-Aza-dc-treated cells, although these transcripts could be amplified from total brain or placenta mRNA, respectively (data not shown). Interestingly, the NDST3 CpG island was unmethylated in HEMC cells, and the NDST4 gene exhibited no typical CpG island. Furthermore, the results show that 5-Aza-dc treatment did not modify HS2ST1 and HS6ST1 mRNA expression levels in HEMC cells (Fig. 4A, C). This could be expected, since the CpG island of these genes exhibited no apparent hypermethylation pattern in chondrosarcoma cells. By contrast, a significant increase (up to 2-fold) in HS6ST2 expression level was observed in HEMC cells when treated with the highest dose of demethylating agent. Despite the presence of methylated sites in the HS6ST3 gene, the corresponding transcript could not be detected in treated or untreated HEMC cells (data not shown). In summary, our results show that NDST1 and HS6ST2 mRNA expression only was upregulated in HEMC cells on 5-Aza-dc treatment, suggesting that methylation events are involved in the regulation of these genes. Bisulfite sequencing data indicate that 5-Aza-dc treatment of HEMC cells reduced the number of methylated sites by

30% in NDST1 gene and by 35% in HS6ST2 gene, which may contribute, at least in part, to transcription activation (Supplemental Fig. 1). However, since 5-Aza-dc would certainly hypomethylate many other genes coding for transcription factors involved in the regulation of NDST1 and HS6ST2, it is possible that the increase in gene expression may also be due to other mechanisms of promoter activation. Further work is required to establish that point.

Because several members of the 3-OST gene family exhibited aberrant DNA methylation patterns, it was interesting to analyze the effect of the demethylating agent on expression of these genes. RT-PCR analysis of 3-OST mRNA transcripts revealed different effects of the demethylating agent among the 3-OST members in HEMC cells. 5-Aza-dc treatment failed to significantly increase the expression of 3-OST3B mRNA transcript, concordant with the low methylation state of the CpG island of this gene. By contrast, 3-OST1 and 3-OST2 mRNA expression levels were significantly increased (1.5-fold) at the higher dose of the demethylating agent compared to untreated cells (Fig. 4B, D). The transcriptional activity of the 3-OST3A gene was also stimulated by treatment with both doses of 5-Aza-dc, which produced  $\sim 2.7$ -fold increase in mRNA expression levels (Fig. 4B, D). Analysis of the methylation status of this gene following 5-Aza-dc treatment of HEMC cells showed  $\sim 54\%$  promoter demethylation, suggesting that this effect could account, at least in part, for the increase in 3-OST3A mRNA expression. Altogether, the results indicate that treatment with



**Figure 4.** Effect of 5-Aza-dc on mRNA expression of glycosyltransferase and sulfotransferases involved in HSPG synthesis in HEMC cells. Cells were treated with 5-Aza-dc for 7 d, and mRNA expression was analyzed by semiquantitative RT-PCR using specific sets of primers (Supplemental Table 2). GAPDH mRNA was used to normalize for input of total RNA. *A, B*) Typical results of mRNA analysis of EXT1, NDST1, NDST2, HS2ST1, HS6ST1, and HS6ST2 transcripts (*A*) and 3-OST1, 3-OST2, 3-OST3A, 3-OST3B, 3-OST6, and control GAPDH transcripts (*B*) after treatment with 5-Aza-dc (2 and 4 μM) compared to untreated cells (control). *C, D*) Quantification of RT-PCR products by scanning densitometry. mRNA expression levels in treated cells are expressed relative to untreated cells (control set to 1) and are means ± SE of 3 to 6 observations. \**P* < 0.05; 1-way ANOVA with Bonferroni's *post hoc* test.

5-Aza-dc restores the expression of 3-OST3A and, to a lesser extent, of 3-OST1 and 3-OST2, supporting the view that DNA methylation is a mechanism affecting the transcription of these genes in HEMC cells.

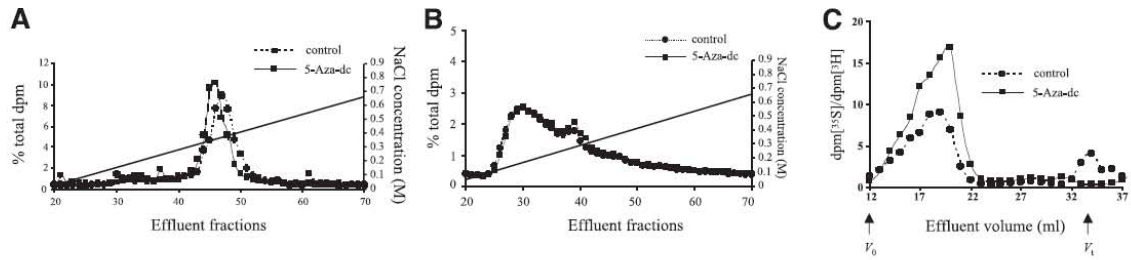
#### The HS sulfation pattern is modified by up-regulation of sulfotransferase expression

Unlike mutational inactivation, epigenetic alterations can be reversed by pharmacological means, such as the use of DNA demethylating agents, for example 5-Aza-dc. The next series of experiments was designed to determine whether this compound was able to rescue gene functionality. To provide evidence for possible changes in HS expression on 5-Aza-dc treatment in HEMC cells, we examined the effect of the demethylating agent on GAG-chain structure using biochemical approaches. To assess possible changes in the overall charge density of GAG chains on 5-Aza-dc treatment, [<sup>35</sup>S]sulfate-labeled and [<sup>35</sup>S]methionine/cysteine-labeled macromolecules isolated from treated and untreated HEMC cells were subjected to anion-exchange chromatography. **Figure 5A** shows the elution pattern of [<sup>35</sup>S]sulfate-radiolabeled macromolecules following treatment with chondroitinase ABC. This endoglycosidase degrades galactosaminoglycan chains, leaving intact HSPGs bound to the column. As shown in Fig. 5A, HSPGs from untreated cells emerged as a single peak when the salt gradient was applied (at ~0.45 M NaCl). Comparison of the gel chromatography profiles of the chondroitinase digests of [<sup>35</sup>S]sulfate-labeled macromolecules extracted from untreated and 5-Aza-dc-

treated HEMC cells showed a shift of the resistant peak corresponding to HSPGs (fractions 45–55), suggesting a charge increase of HSPGs on 5-Aza-dc treatment. By contrast, [<sup>35</sup>S]methionine/cysteine-labeled macromolecules from untreated and treated cells exhibited a similar profile on anion-exchange chromatography (Fig. 5B), indicating no major modification in protein synthesis on 5-Aza-dc exposure.

To further investigate changes in GAG synthesis induced by 5-Aza-dc, treated and untreated cells were metabolically labeled with [<sup>35</sup>S]sulfate or [<sup>3</sup>H]glucosamine, and PGs isolated from cell/matrix extracts were treated with proteinase K to release GAG chains from their core proteins. GAG chains were then subjected to digestion with chondroitinase ABC and analyzed by gel exclusion chromatography. Analysis of the length of HS chains released by proteinase K treatment did not show a marked difference, as indicated by comparison of the chromatography profiles for HS chains from treated and untreated cells (Fig. 5C). However, these HS chains contained more [<sup>35</sup>S]sulfate relative to incorporated [<sup>3</sup>H]glucosamine, corroborating the idea that the sulfate density of HS chains is increased on 5-Aza-dc treatment, whereas the chain length is not affected (Fig. 5C). Together, these results indicate that treatment with the demethylating agent 5-Aza-dc affects the extent of HS sulfation, rather than HS-chain length in HEMC cells.

To further examine the effect of 5-Aza-dc on HS-GAG chains, we carried out indirect immunofluorescence analysis of the chondrosarcoma cells, using 10E4 anti-HS monoclonal antibody, which is commonly used



**Figure 5.** Effect of 5-Aza-dc treatment on cellular HSPG negative charge density and chain length in HEMC cells. *A, B*) Anion-exchange chromatography analysis of HSPGs. Untreated and 5-Aza-dc-treated HEMC cells (4  $\mu$ M) were radiolabeled with [ $^{35}$ S]SO $_4$  or [ $^{35}$ S]methionine/cysteine for 22 h. After removal of free [ $^{35}$ S] by Sephadex G-50 gel filtration, cellular samples were subjected to specific digestion by chondroitinase ABC and resolved by anion-exchange chromatography. Elution was carried out with a NaCl gradient (0.05–0.8 M). Fractions were collected, and radioactivity was evaluated by scintillation counting. *A*) Elution profile of [ $^{35}$ S]SO $_4$ -radiolabeled macromolecules. *B*) Elution profile of [ $^{35}$ S]methionine/cysteine-radiolabeled macromolecules. *C*) Gel-filtration analysis of HS chains. Untreated and 5-Aza-dc-treated HEMC cells (4  $\mu$ M) were radiolabeled with [ $^{35}$ S]SO $_4$  or [ $^3$ H]glucosamine for 22 h. After removal of free radioactive label by Sephadex G-50, samples were treated with proteinase K for 2 h prior to digestion by chondroitinase ABC. Samples were subjected to gel-filtration chromatography on Sepharose Cl-6B, and radioactivity associated with the collected fractions was evaluated by scintillation counting. Results are expressed as [ $^{35}$ S]SO $_4$ /[ $^3$ H]glucosamine ratio. Moving mobile phase column volume ( $V_0$ ) and total liquid phase column volume ( $V_t$ ), corresponding to the volume of elution of blue dextran and K $_2$ CrO $_4$ , respectively, are indicated by arrows. Each chromatogram shows a profile that is a representative of 3 separate experiments.

to detect HS chains of PGs (38–40). Prominent staining of the cell membrane was observed in HEMC cells treated with the demethylating agent, whereas only a very low signal could be observed in untreated cells (Fig. 6A, B). Since 10E4 antibodies were initially described to recognize *N*-sulfated glucosamine residues normally present on HS chains (38), the absence of staining in HEMC cells was unexpected (Fig. 6A).

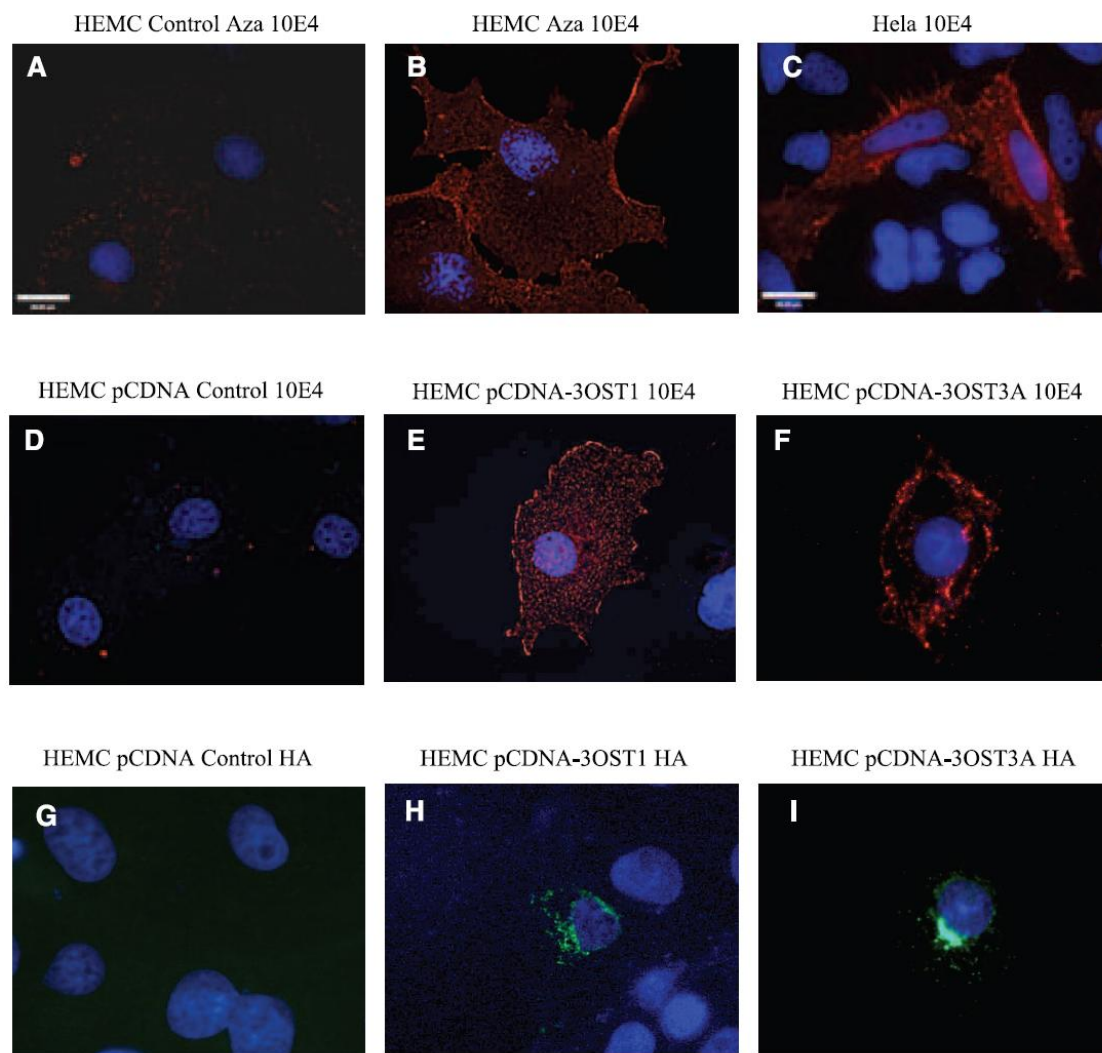
Indeed, HeLa cells used as positive controls exhibited a strong plasma membrane signal (Fig. 6C). However, several studies indicate that the 10E4 epitope associated with HSPGs may be inaccessible in some cell lines (39) and highlight that further characterization of the epitope is needed (38). Recent evidence indicates that removal of *O*-sulfate groups from HS chains by expression of recombinant sulfatase (Sulf) isoforms in hepatocarcinoma (SNU 449) (41) or pancreatic (Panc-1) cell lines (42), abolished 10E4 staining, suggesting that *O*-sulfation might contribute to the antigenicity of HS chains toward 10E4 antibodies. In agreement, we show that transfection of HEMC cells with 3-OST1 cDNA, followed by immunostaining using the 10E4 antibody, gave a strong punctate pattern (Fig. 6E), similar to that observed on 5-Aza-dc treatment, whereas mock transfected cells exhibited low or no staining (Fig. 6D). When recombinant cells were probed with anti-HA antibodies, efficient expression of 3-OST1 was revealed (Fig. 6H), whereas mock-transfected cells showed no staining (Fig. 6G). Similarly, 3-OST3A cDNA expression in HEMC cells was shown using anti-HA antibodies (Fig. 6I) and was associated with a pronounced staining of cell surface HS probed with anti-10E4 antibodies (Fig. 6F). Altogether, these observations indicate that 3-OST1 or 3-OST3A expression produced HS epitopes containing 3-*O*-sulfated glucosamine residues that are recognized by the 10E4 antibody. Furthermore, our data suggest that 5-Aza-dc treatment induces modifications of HS chains that are

similar or identical to those resulting from induction of 3-OST expression.

#### Effect of restoration of 3-OST expression on proliferation, adhesion, and migration in HEMC cells

The next series of experiments were designed to assess the functional consequences of exposing chondrosarcoma cells to the demethylating agent 5-Aza-dc and of expression of 3-OST isoforms by cDNA transfection on proliferation, adhesion, and migration properties. As shown in Fig. 7A, 5-Aza-dc produced a significant inhibitory effect on HEMC cell proliferation, as rapidly as 2 d after drug exposure, that reached 43% of inhibition after 5 d. No cytotoxicity of 5-Aza-dc at the concentration used (4  $\mu$ M) was observed as verified using trypan blue (not shown). Since the antiproliferative effect of the demethylating agent may result from up-regulation of several genes, it was important to determine the specific effect of 3-*O*-sulfation on the proliferation rate of cells transfected with 3-OST cDNAs. Interestingly, expression of the 3-OST3A isoform significantly reduced HEMC cell proliferation up to 20%, whereas 3-OST1 did not exert an antiproliferative effect compared to mock-transfected cells (Fig. 7B). Protein expression of each isoform was verified by immunoblot analysis using anti-HA antibodies (not shown).

Next, the effect of 5-Aza-dc on HEMC cell adhesion onto fibronectin-coated plates or plates coated with BSA only was assayed. No significant difference between 5-Aza-dc-treated and control cells was observed when cells were seeded on plates coated with BSA only or low-dose fibronectin. However, a moderate but significant increase in adhesion was seen when 5-Aza-dc-treated cells were seeded on plates coated with 200  $\mu$ g of fibronectin (Fig. 7C). Interestingly, this effect was enhanced when HEMC cells were transfected with the



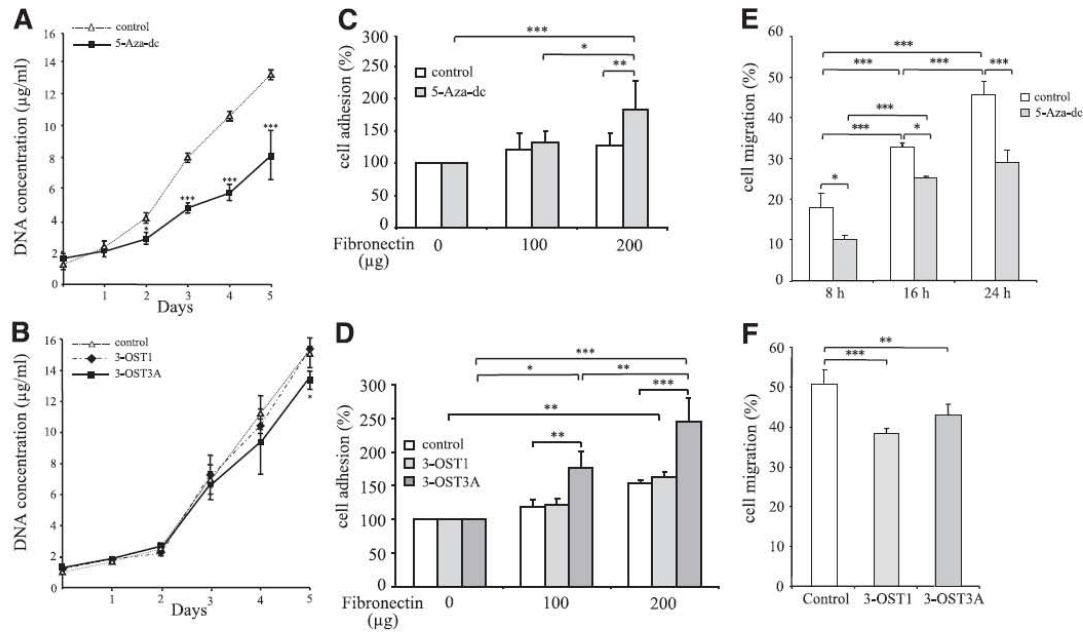
**Figure 6.** 5-Aza-dc treatment or expression of recombinant human 3-OST1 or 3-OST3A increases cell surface/matrix HSPG expression in HEMC cells. Cells were treated with 5-Aza-dc (4  $\mu$ M) for 7 d or transfected with 3-OST1 or 3-OST3A, and indirect immunofluorescence staining was performed as described in Materials and Methods with 10E4 anti-HSPGs (A–F) or following permeabilization with anti-HA antibodies (G–I). A) Untreated HEMC cells. B) 5-Aza-dc-treated HEMC cells. C) HeLa cells. D, G) Mock-transfected HEMC cells. E, H) pCDNA-3OST1 HEMC transfected cells. F, I) pCDNA-3OST3A-transfected HEMC cells.

3-OST3A expression vector. Indeed, 3-OST3A-expressing cells showed a dose-dependent increase of adhesion when seeded onto fibronectin-coated plates, reaching 65% over control mock-transfected cells (Fig. 7D). However, no significant effect was observed when cells were transfected with the 3-OST1 expression vector.

Finally, we analyzed the effect of 5-Aza-dc on cell motility using an *in vitro* migration assay. Untreated HEMC cells were found to be highly motile, reaching a 45–50% cell migration rate after 24 h (Fig. 7E). However, cells treated with 5-Aza-dc exhibited much less migration potency through the polycarbonate membrane using 5% serum as a chemoattractant over a period of 8, 16, or 24 h (Fig. 7E). The consequences of 3-OST1 and 3-OST3A cDNA expression in HEMC cells

on the migration rate were next examined. The expression of 3-OST1 and 3-OST3A reduced the migration rate by 22 and 15%, respectively, compared to mock-transfected cells (Fig. 7F). These results indicate that overexpression of either 3-OST1 or 3-OST3A cDNA resulted in a less motile phenotype with a larger effect in the case of 3-OST1.

Altogether, our data showed that 5-Aza-dc inhibits the proliferative rate and migratory ability of HEMC cells and produces a moderate increase of cell adhesion on fibronectin-coated plates. 3-OST3A cDNA expression acted in a comparable way to that of the demethylating agent, whereas 3-OST1 only produced an antimigratory effect. These results indicate that reversing the epigenetic loss of 3-OST expression or restoring expres-



**Figure 7.** Effect of 3-OST expression on HEMC cell proliferation, adhesion, and migration. *A*) Untreated and 4  $\mu\text{M}$  5-Aza-dc-treated HEMC cells were seeded onto 24-well plates, and proliferation was evaluated by measurement of DNA content after 1 to 5 d culture time. *B*) HEMC cells were transiently transfected with 3-OST1 or 3-OST3A cDNA for 48 h and analyzed for proliferation as described for *A*. Results are expressed as DNA concentration and are means  $\pm$  SE of 3 to 6 observations.  $*P < 0.05$ ,  $***P < 0.001$  vs. untreated or mock-transfected cells; 2-way ANOVA with Bonferroni's multiple comparison *post hoc* test. *C*) Untreated and 4  $\mu\text{M}$  5-Aza-dc-treated HEMC cells were plated onto wells coated with BSA only or with 100 or 200  $\mu\text{g}$  of fibronectin for 30 min at 37°C. After washing, adherent cells were quantified after fixation by crystal violet staining and absorbance measurement at a wavelength of 550 nm. Attachment of cells to plates coated with BSA only was assigned a value of 100%. *D*) HEMC cells were transiently transfected with 3-OST1 or 3-OST3A cDNA for 48 h and assayed for adhesion as described for *C*. Attachment of cells to plates coated with BSA only was assigned a value of 100%. Results are expressed as means  $\pm$  SE of 3 to 6 observations.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$  vs. corresponding control; 1-way ANOVA with Bonferroni's multiple comparison *post hoc* test. *E*) Untreated and 4  $\mu\text{M}$  5-Aza-dc-treated HEMC cells were incubated with mitomycin C to block cell proliferation. Cells were plated on Transwell inserts in serum-free medium containing 0.1% BSA and left to migrate through the polycarbonate membrane to the lower chamber, containing medium with 5% FBS. After 8, 16, and 24 h, untreated and 5-Aza-dc-treated cells were washed, fixed, and stained with methyl blue. Quantification was performed by optical density reading at a wavelength of 663 nm. *F*) HEMC cells were transiently transfected with 3-OST1 or 3-OST3A cDNA for 48 h and analyzed for migration potency after 18 h as described for *E*. Results are expressed as percentage of migration and are means  $\pm$  SE of 3 to 6 observations.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$  vs. corresponding control; 1-way ANOVA with Bonferroni's multiple comparison *post hoc* test.

sion by transient 3-OST3A expression is accompanied by a reduction of the proliferation and migration capacity of HEMC cells, together with an increased adhesion ability that could counteract the invasiveness of chondrosarcoma cells.

## DISCUSSION

A growing body of evidence supports an important role for GAGs at key pathophysiological stages of tumorigenesis and progression. HS-GAGs regulate tumor growth, tumor-cell adhesion, invasion, hematogenous metastasis, and angiogenesis (3). These effects are thought to be mediated by multiple interactions with growth factors, cytokines, chemokines, extracellular matrix proteins, and coagulation factors (1). In some cases, these interactions are highly specific in terms of the degree and position of sulfation required to achieve

high-affinity binding. The classic example is antithrombin III, in which a specific pentasaccharide, including a 3-*O*-sulfated glucosamine residue, is essential for high-affinity interactions (43, 44). Specific patterns of sulfation are required for binding members of the fibroblast growth factor (FGF) family and signal transduction mediated by these growth factors (45). *In vivo* studies in transgenic mice clearly indicate that changes in the fine structure of cell surface HS have profound effects on tumor cell growth and metastasis (46). However, although it is well documented that the profile of cellular HS-GAGs is critical for tumor behavior, our understanding of the regulation of the biosynthetic machinery governing the structure of HS chains in cancer cells remains limited. In this study, we show for the first time that a specific expression pattern of HS-3-*O*-sulfotransferases is driven by an epigenetic mechanism involving DNA methylation in a chondrosarcoma cell line, selected as a model of invasive tumors involving connec-

tive tissue. We demonstrate that CpG islands within the proximal region of the *3-OST* genes were hypermethylated, an event associated with down-regulation of the expression of several members of the *3-OST* family in HEMC cells. Notably, exposure of chondrosarcoma cells to a demethylating agent or overexpression of *3-OST3A* cDNA led to the reversion of certain characters associated with invasiveness, namely cell proliferation, attachment to extracellular matrix proteins, and cell migration. Overexpression of *3-OST1* produced a moderate but significant antimigratory effect on HEMC cells. To our knowledge, this study provides the first evidence for the effect of epigenetic regulation of HS-*O*-sulfotransferase gene expression on cancer cell phenotype.

These findings are the result of an extensive DNA methylation analysis of the genes coding for the biosynthetic enzymes involved in HS-chain formation and modification in HEMC cells, where we first observed that the CpG island of the glycosyltransferase genes involved in HS synthesis exhibit a low number of methylated sites in the chondrosarcoma cells. The biosynthetic pathway of HS chains is initiated by the sequential addition of 4 monosaccharides, forming a linker sequence that serves as a primer for further polymerization of the polysaccharide chains *via* the successive and coordinate action of an array of glycosyltransferase enzymes. The formation of the tetrasaccharide linker region catalyzed by XT-I,  $\beta$ 4GalT7,  $\beta$ 3GalT6, and GlcAT-I is common to HS- and CS-chain synthesis pathways. The observation that this process is unaffected by inhibition of CpG methylation is not unexpected, since the initiation steps are absolutely required prior to the elongation of both types of GAG chains and thus are predicted to be tightly regulated (47). Indeed, a defect in initiation of GAG synthesis would compromise not only HS but also CS expression and would be expected to be deleterious to cell and tissue functions.

Our results also provide strong evidence that DNA methylation-associated gene silencing alters neither the initiation nor the polymerization of HS chains in HEMC cells. This is based on the following: the *EXT1* gene promoter displayed no hypermethylation profile, no change in *EXT1* mRNA transcript level was observed on 5-Aza-dc treatment, and analysis of cellular HSPGs showed no modification of HS-chain length on cell treatment with a DNA-methyltransferase inhibitor. These data, based on genetic and biochemical analysis, indicate that HS-chain formation catalyzed by *EXT1*, the glycosyltransferase playing a primordial role in HS chain polymerization, is not regulated by an epigenetic mechanism involving DNA methylation in chondrosarcoma cells. The lack of hypermethylation of the *EXT1* gene in HEMC cells is consistent with other studies indicating that neither chondrosarcomas nor osteosarcomas exhibited aberrant methylation at the *EXT1* locus (24, 48). Interestingly, Ropero *et al.* (24) provided evidence for aberrant methylation of the *EXT1* promoter in acute types of leukemia, suggesting that the defect in HS chains resulting from epigenetic silencing of an HS biosynthetic glycosyltransferase is a critical step in the development of this group of malignancies.

However, in chondrosarcoma, our results and other studies indicate that although genetic alteration of the *EXT1* locus leads to cartilaginous and bone malignancies in HME, silencing of the *EXT* gene is not a characteristic feature of these tumor cells. Altogether, our results show that in HEMC cells, regulation of the glycosyltransferases that are responsible for the initiation and polymerization of the saccharide backbone of HS chains does not involve an epigenetic mechanism driven by DNA methylation.

A strong body of evidence has demonstrated that HS sequence patterns generated by sulfotransferases have profound consequences on the cellular functions of proteoglycans (11). This led us to examine the potential effect of epigenetic regulation on properties and functions of these enzymes, which modify the polysaccharide GAG chains, in a model tumor cell line. Interestingly, examination of the methylation status of the *NDST* and *OST* gene families revealed aberrant methylation status of a specific set of genes in chondrosarcoma cells. Indeed, establishment of the methylation pattern of the genes that contain CpG islands showed extensive DNA methylation of the *3-OST1*, *3-OST3A*, and *3-OST6* gene family members in HEMC cells, and a lower number of methylated sites in the proximal promoter region of the *3-OST2* gene. By contrast, low-level or no promoter methylation was observed in the *NDST2*, *NDST3*, *HS2ST1*, and *HS6ST1* genes. The role of DNA methylation as a regulatory mechanism influencing HSPG 3-*O*-sulfation was confirmed by the observation that treatment with 5-Aza-dc, widely used as a DNA-methyltransferase inhibitor, enhanced *3-OST1*, *3-OST2*, and to the largest extent, *3-OST3A* mRNA expression level in HEMC cells. Further analysis of the methylation profile of the *3-OST3A* CpG island following treatment of HEMC cells with 5-Aza-dc indicated that the drug significantly reduced the number of methylated sites, suggesting that the increase in *3-OST3A* gene expression could be attributed, at least in part, to 5-Aza-dc-induced demethylation. However, it is also possible that demethylation of transcription factor genes involved in *3-OST3A* regulation on 5-Aza-dc treatment could contribute to sulfotransferase promoter activation.

The lack of effect of 5-Aza-dc on *3-OST3B*, *HS2ST1*, and *HS6ST1* transcript levels is correlated with the low number or absence of methylated sites at the CpG island of these genes. On the other hand, *NDST1* and *HS6ST2* were found to be hypermethylated in chondrosarcoma cells, and exposure to 5-Aza-dc reduced their methylation content and induced their mRNA levels, indicating that expression of these genes might be regulated at the epigenetic level. However, since *NDST1* and *HS6ST2* exhibited a similar hypermethylation profile in blood cells, no exclusive link between the methylation status of these genes and the pathogenesis of chondrosarcoma could be established at that stage.

Since a methylation event was found to be involved in the regulation of several members of the *3-OST* family in HEMC cells, a key issue was next to determine whether up-regulation of *3-OST* expression had an effect on HSPG structure and function. Following [ $^{35}$ S]sulfate metabolic labeling, analysis of cellular

HSPGs by anion-exchange chromatography provided evidence for an increase in charge on 5-Aza-dc treatment. This effect was likely due to an increase in HS-chain sulfation, based on the following: the degree of HS-GAG chain sulfation was increased, whereas the chain length was unchanged, as shown by gel-filtration chromatography; the synthesis and charge of proteins were not affected by 5-Aza-dc treatment; and a strong increase in HEMC cell membrane/pericellular HS expression on 5-Aza-dc treatment, as well as on 3-OST cDNA transfection, was evidenced by indirect immunofluorescence microscopy. Altogether, our results suggest that reversing the hypermethylation state of HS-biosynthetic genes by 5-Aza-dc treatment in HEMC cells mainly affects HS-chain sulfation, an effect that can be attributed, at least in part, to enhanced 3-OST expression.

Epigenetic changes play predominant roles in the promotion or suppression of tumor growth, invasion, and metastasis. It was thus essential to evaluate the potential biological significance of aberrant methylation of HS-biosynthetic enzymes in the chondrosarcoma cells. To this aim, we determined the effects of 5-Aza-dc treatment, on one hand, and those of 3-OST1 and 3-OST3A cDNA expression, on the other hand, on several parameters linked to the invasive character of HEMC cells. Our results show that not only 5-Aza-dc treatment but also expression of 3-OST3A cDNA was able to produce a significant decrease in chondrosarcoma cell proliferation. In addition, 3-OST3A overexpression promoted cell adhesion to a fibronectin substrate, an important finding since attachment of tumor cells to the extracellular matrix is required for cell motility and invasion. Concordantly, it was recently shown that the fine structure of HS is critical in binding the HepII domain of fibronectin for optimal cell attachment (49). Our results suggest that 3-O-sulfation catalyzed by 3-OST3A may also be a determinant for fibronectin binding, although further studies are required to confirm this assumption. Finally, an *in vitro* migration assay clearly showed that either treatment of HEMC cells with the demethylating agent or expression of 3-OST1 and 3-OST3A cDNA resulted in a less motile cell phenotype. Although the mechanisms underlying these observations warrant further investigation, our findings suggest that dysregulation of 3-OST expression, with the consequent effects on HS-3-O-sulfation, contributes to several characteristics associated with tumor behavior, at least in HEMC cells. Should this be found to be a more widespread effect, it would have significant implications for our understanding of the molecular basis of tumor cell invasiveness and metastasis. The regulation of 3-OST expression in different types of tumor cells is currently under investigation. The key structure for the 3-O-sulfated functional domains of heparin and HSPGs are produced *via* the action of 7 distinct enzymes belonging to the largest HS-O-sulfotransferase family. 3-OST1 transfers sulfate groups to the 3-OH position of the N-sulfated glucosamine residue that is linked to a GlcA residue at the nonreducing end, to provide a binding site for anti-thrombin (43). HS modified by 3-OST2, OST3A, OST3B, OST4, and OST6 binds to HSV-1 gD and assists

viral entry (50, 51). Nonetheless, the biological significance of 3-O-sulfate other than regulation of blood coagulation and HSV-1 entry remains unknown at the present time, although several recent reports support additional functions, such as in glomerular filtration or FGF-7 [keratinocyte growth factor (KGF)] binding (52). Surprisingly, *Hs3st1*<sup>-/-</sup> knockout mice did not exhibit a procoagulant phenotype, but showed postnatal lethality and intrauterine growth retardation (53). Reduction of the expression of *Drosophila melanogaster* 3OSTB, a homologue of mammalian 3-OST enzymes, by transgenic RNAi resulted in a phenotype of Notch signaling deficiency (54). Interestingly, Mochizuki *et al.* (55) recently demonstrated that several 3-OST enzymes produce HS-tetrasulfated disaccharides present in most tissues, suggesting a broader, although unknown, physiological role. Recently, Vanpouille *et al.* (56) demonstrated that 3-O-sulfated glucosamine residues were responsible for HS interactions with the inflammatory mediator cyclophilin B. An intriguing possibility being raised is that the promigratory activity of this inflammatory mediator might be dependent on interactions with specific binding HS domains present at the cell surface of responsive T lymphocytes. Our investigations of the functional properties triggered by 3-OST expression support the possibility of a novel role for HS-3-O-sulfation as an antimigratory/antiinvasiveness modulator, at least in chondrosarcoma cells. In this regard, the importance of 3-O-sulfate groups in the specificity of interactions between heparin/HS and FGF-7 and FGF-10 is of special interest (52), taking into account the effect of these growth factors in the regulation of the proliferation of several cell types (57) and cell-stroma interactions (58). Our study reinforces the concept that regulation of sulfotransferases, which in turn determines specific carbohydrate structural modifications of HS chains, has a vital influence in the modulation of glycan function in cancer (59).

## CONCLUSIONS

We have demonstrated that aberrant methylation of a specific set of sulfotransferase genes responsible for the modification of HS chains alters GAG-chain sulfation and contributes to alteration of the cell phenotype and invasive properties of HEMC chondrosarcoma cells, an important aspect of cancer growth and metastasis control. Epigenetic abnormalities, such as aberrant methylation of CpG islands, have a major effect on carcinogenesis, and methylation of many genes is found to be associated with poor response to chemotherapy and with poor prognosis (60). However, epigenetic alterations are potentially reversible. The use of DNA demethylating agents has turned out to be effective for hematological malignancies and is being used to treat solid tumors (60). Chondrosarcomas remain challenging malignancies both in terms of therapeutic options and prognosis, since they are resistant to conventional chemotherapy and radiation therapy (61). Our results uncover a new mechanism that might play a role in chondrosarcoma pathogenesis



and lead to the identification of novel therapeutic approaches. **[E]**

This work was supported by an Institut National de la Santé et de la Recherche Médicale (INSERM)-University of Dundee International Collaboration Contract (C2I), the Agence Nationale de la Recherche (GAG Network ANR-08-PCVI-0023-01), the Ligue Régionale contre le Cancer, a Royal Society International Joint Grant (to M.W.H.C., and S.F.-G.), and the Région Lorraine. C.B. was the recipient of a Stevenson Exchange Scholarship. We are grateful to Theodora Tucker (University of Dundee, Dundee, UK) for her expertise and assistance with the indirect immunofluorescence microscopy studies, to Dr. Céline Pourreyaon (University of Dundee, Dundee, UK) for her helpful advice regarding the cell migration assays, and Sandrine Gulberti for critical reading of the manuscript. Dr. Narayanan Venkatesan (McGill University, Montréal, Québec) is gratefully acknowledged for his help in performing statistical analysis and for critical reading of the manuscript.

## REFERENCES

- Bishop, J. R., Schuksz, M., and Esko, J. D. (2007) Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature* **446**, 1030–1037
- Bernfield, M., Götte, M., Park, P. W., Reizes, O., Fitzgerald, M. L., Lincecum, J., and Zako, M. (1999) Functions of cell surface heparan sulfate proteoglycans. *Annu. Rev. Biochem.* **68**, 729–777
- Sasisekharan, R., Shriver, Z., Venkataraman, G., and Narayanasami, U. (2002) Roles of heparan-sulphate glycosaminoglycans in cancer. *Nat. Rev. Cancer* **2**, 521–528
- Su, G., Blaine, S. A., Qiao, D., and Friedl, A. (2007) Shedding of syndecan-1 by stromal fibroblasts stimulates human breast cancer cell proliferation via FGF2 activation. *J. Biol. Chem.* **282**, 14906–14915
- Kadenhe-Chiweshe, A., Papa, J., McCrudden, K. W., Frischer, J., Bae, J. O., Huang, J., Fisher, J., Lefkowitz, J. H., Feirt, N., Rudge, J., Holash, J., Yancopoulos, G. D., Kandel, J. J., and Yamashiro, D. J. (2008) Sustained VEGF blockade results in microenvironmental sequestration of VEGF by tumors and persistent VEGF receptor-2 activation. *Mol. Cancer Res.* **6**, 1–9
- Götting, C., Kuhn, J., Zahn, R., Brinkmann, T., and Kleesiek, K. (2000) Molecular cloning and expression of human UDP-d-Xylose:proteoglycan core protein  $\beta$ -D-xylosyltransferase and its first isoform XT-II. *J. Mol. Biol.* **304**, 517–528
- Almeida, R., Levery, S. B., Mandel, U., Kresse, H., Schwientek, T., Bennett, E. P., and Clausen, H. (1999) Cloning and expression of a proteoglycan UDP-galactose:beta-xylose  $\beta$ 1,4-galactosyltransferase I. A seventh member of the human beta4-galactosyltransferase gene family. *J. Biol. Chem.* **274**, 26165–26171
- Bai, X., Zhou, D., Brown, J. R., Crawford, B. E., Hennes, T., and Esko, J. D. (2001) Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth member of the beta 1,3-galactosyltransferase family (beta 3GalT6). *J. Biol. Chem.* **276**, 48189–48195
- Kitagawa, H., Tone, Y., Tamura, J., Neumann, K. W., Ogawa, T., Oka, O., Kawasaki, T., and Sugahara, K. (1998) Molecular cloning of glucuronosyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J. Biol. Chem.* **273**, 6615–6618
- McCormick, C., Leduc, Y., Martindale, D., Mattison, K., Esford, L. E., Dyer, A. P., and Tufaro, F. (1998) The putative tumour suppressor EXT1 alters the expression of cell surface heparan sulfate. *Nat. Genet.* **19**, 158–161
- Kusche-Gullberg, M., and Kjellen, L. (2003) Sulfotransferases in glycosaminoglycan biosynthesis. *Curr. Opin. Struct. Biol.* **13**, 605–611
- Urbaniak, M. D., Tabudravu, J. N., Msaki, A., Matera, K. M., Brenk, R., Jaspars, M., and Ferguson, M. A. (2006) Identification of novel inhibitors of UDP-Glc 4'-epimerase, a validated drug target for african sleeping sickness. *Bioorg. Med. Chem. Lett.* **16**, 5744–5747
- Prydz, K., and Dalen, K. T. (2000) Synthesis and sorting of proteoglycans. *J. Cell Sci.* **113**, 193–205
- Barré, L., Venkatesan, N., Magdalou, J., Netter, P., Fournel-Gigleux, S., and Ouzzine, M. (2006) Evidence of calcium-dependent pathway in the regulation of human beta1,3- glucuronosyltransferase-1 (GlcAT-I) gene expression: a key enzyme in proteoglycan synthesis. *FASEB J.* **20**, 1692–1694
- Gulberti, S., Lattard, V., Fondeur, M., Jacquinet, J. C., Mulliert, G., Netter, P., Magdalou, J., Ouzzine, M., and Fournel-Gigleux, S. (2005) Phosphorylation and sulfation of oligosaccharide substrates critically influence the activity of human  $\beta$ 1,4-galactosyltransferase 7 (GalT-I) and  $\beta$ 1,3-glucuronosyltransferase I (GlcAT-I) involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J. Biol. Chem.* **280**, 1417–1425
- Busse, M., Feta, A., Presto, J., Wilen, M., Gronning, M., Kjellen, L., and Kusche-Gullberg, M. (2007) Contribution of EXT1, EXT2, and EXT3 to heparan sulfate chain elongation. *J. Biol. Chem.* **282**, 32802–32810
- Venkatesan, N., Barré, L., Magdalou, J., Mainard, D., Netter, P., Fournel-Gigleux, S., and Ouzzine, M. (2009) Modulation of xylosyltransferase I expression provides a mechanism regulating glycosaminoglycan chain synthesis during cartilage destruction and repair. *FASEB J.* **23**, 813–822
- McCormick, C., Duncan, G., and Tufaro, F. (1999) New perspectives on the molecular basis of hereditary bone tumours. *Mol. Med. Today* **5**, 481–486
- Stückens, D., Zak, B. M., Rougier, N., Esko, J. D., and Werb, Z. (2005) Mice deficient in Ext2 lack heparan sulfate and develop exostoses. *Development* **132**, 5055–5068
- Quentin, E., Gladen, A., Roden, L., and Kresse, H. (1990) 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. U. S. A.* **87**, 1342–1346
- Weissmann, G. (2008) Fashions in science: from philosophers' cAMP to epigenetics. *FASEB J.* **22**, 4033–4037
- Vucic, E. A., Brown, C. J., and Lam, W. L. (2008) Epigenetics of cancer progression. *Pharmacogenomics* **9**, 215–234
- Sanderson, R. D., Yang, Y., Kelly, T., MacLeod, V., Dai, Y., and Theus, A. (2005) Enzymatic remodeling of heparan sulfate proteoglycans within the tumor microenvironment: growth regulation and the prospect of new cancer therapies. *J. Cell. Biochem.* **96**, 897–905
- Ropero, S., Setien, F., Espada, J., Fraga, M. F., Herranz, M., Asp, J., Benassi, M. S., Franchi, A., Patino, A., Ward, L. S., Bovee, J., Cigudosa, J. C., Wim, W., and Esteller, M. (2004) Epigenetic loss of the familial tumor-suppressor gene *exostosin-1* (EXT1) disrupts heparan sulfate synthesis in cancer cells. *Hum. Mol. Genet.* **13**, 2753–2765
- Miyamoto, K., Asada, K., Fukutomi, T., Okochi, E., Yagi, Y., Hasegawa, T., Asahara, T., Sugimura, T., and Ushijima, T. (2003) Methylation-associated silencing of heparin sulfate D-glucosaminyl 3-O-sulfotransferase-2 (3-OST-2) in human breast, colon, lung and pancreatic cancers. *Oncogene* **22**, 274–280
- Li, L. C., and Dahiya, R. (2002) MethPrimer: designing primers for methylation PCRs. *Bioinformatics* **18**, 1427–1431
- Clark, S. J., Statham, A., Stürzaker, C., Molloy, P. L., and Frommer, M. (2006) DNA methylation: bisulphite modification and analysis. *Nat. Protoc.* **1**, 2353–2364
- Ouzzine, M., Magdalou, J., Burchell, B., and Fournel-Gigleux, S. (1999) Expression of functionally active human hepatic UDP-glucuronosyltransferase (UGT1A6) lacking the N-terminal signal sequence in the endoplasmic reticulum. *FEBS Lett.* **454**, 187–191
- Gulberti, S., Fournel-Gigleux, S., Mulliert, G., Aubry, A., Netter, P., Magdalou, J., and Ouzzine, M. (2003) The functional glycosyltransferase signature sequence of the human beta 1,3-glucuronosyltransferase is a XDD motif. *J. Biol. Chem.* **278**, 32219–32226
- Svennevig, K., Kolset, S. O., and Bangstad, H. J. (2006) Increased syndecan-1 in serum is related to early nephropathy in type 1 diabetes mellitus patients. *Diabetologia* **49**, 2214–2216
- Yanagishita, M., Midura, R. J., and Hascall, V. C. (1987) Proteoglycans: isolation and purification from tissue cultures. *Methods Enzymol.* **138**, 279–289

32. Dick, G., Grondahl, F., and Prydz, K. (2008) Overexpression of the 3'-phosphoadenosine 5'-phosphosulfate (PAPS) transporter 1 increases sulfation of chondroitin sulfate in the apical pathway of MDCK II cells. *Glycobiology* 18, 53–65
33. Tveit, H., Dick, G., Skibeli, V., and Prydz, K. (2005) A proteoglycan undergoes different modifications en route to the apical and basolateral surfaces of Madin-Darby canine kidney cells. *J. Biol. Chem.* 280, 29596–29603
34. Venkatesan, N., Barre, L., Benani, A., Netter, P., Magdalou, J., Fournel-Gigleux, S., and Ouzzine, M. (2004) Stimulation of proteoglycan synthesis by glucuronosyltransferase-I gene delivery: A strategy to promote cartilage repair. *Proc. Natl. Acad. Sci. U. S. A.* 101, 18087–18092
35. Vuong, T. T., Prydz, K., and Tveit, H. (2006) Differences in the apical and basolateral pathways for glycosaminoglycan biosynthesis in Madin-Darby canine kidney cells. *Glycobiology* 16, 326–332
36. Valster, A., Tran, N. L., Nakada, M., Berens, M. E., Chan, A. Y., and Symons, M. (2005) Cell migration and invasion assays. *Methods* 37, 208–215
37. Jones, P. A. (1985) Altering gene expression with 5-azacytidine. *Cell* 40, 485–486
38. David, G., Bai, X. M., Van der Schueren, B., Cassiman, J. J., and Van den Berghe, H. (1992) Developmental changes in heparan sulfate expression: in situ detection with mAbs. *J. Cell Biol.* 119, 961–975
39. Mani, K., Cheng, F., Sandgren, S., Van Den Born, J., Havsmark, B., Ding, K., and Fransson, L. A. (2004) The heparan sulfate-specific epitope 10E4 is NO-sensitive and partly inaccessible in glypican-1. *Glycobiology* 14, 599–607
40. Leteux, C., Chai, W., Nagai, K., Herbert, C. G., Lawson, A. M., and Feizi, T. (2001) 10E4 antigen of scrapie lesions contains an unusual nonsulfated heparan motif. *J. Biol. Chem.* 276, 12539–12545
41. Lai, J. P., Chien, J. R., Moser, D. R., Staub, J. K., Aderca, I., Montoya, D. P., Matthews, T. A., Nagorney, D. M., Cunningham, J. M., Smith, D. I., Greene, E. L., Shridhar, V., and Roberts, L. R. (2004) hSulf1 sulfatase promotes apoptosis of hepatocellular cancer cells by decreasing heparin-binding growth factor signaling. *Gastroenterology* 126, 231–248
42. Li, J., Kleeff, J., Abiatari, I., Kayed, H., Giese, N. A., Felix, K., Giese, T., Buchler, M. W., and Friess, H. (2005) Enhanced levels of Hsulf-1 interfere with heparin-binding growth factor signaling in pancreatic cancer. *Mol. Cancer* 4, 14
43. Shworak, N. W., Liu, J., Fritze, L. M., Schwartz, J. J., Zhang, L., Logeart, D., and Rosenberg, R. D. (1997) Molecular cloning and expression of mouse and human cDNAs encoding heparan sulfate D-glucosaminyl 3-O-sulfotransferase. *J. Biol. Chem.* 272, 28008–28019
44. Petitou, M., Casu, B., and Lindahl, U. (2003) 1976–1983, a critical period in the history of heparin: the discovery of the antithrombin binding site. *Biochimie (Paris)* 85, 83–89
45. Allen, B. L., and Rapraeger, A. C. (2003) Spatial and temporal expression of heparan sulfate in mouse development regulates FGF and FGF receptor assembly. *J. Cell Biol.* 163, 637–648
46. Yang, Y., Macleod, V., Miao, H. Q., Theus, A., Zhan, F., Shaughnessy, J. D., Jr., Sawyer, J., Li, J. P., Zcharia, E., Vlodaysky, I., and Sanderson, R. D. (2007) Heparanase enhances syndecan-1 shedding: a novel mechanism for stimulation of tumor growth and metastasis. *J. Biol. Chem.* 282, 13326–13333
47. Sugahara, K., and Kitagawa, H. (2000) Recent advances in the study of the biosynthesis and functions of sulfated glycosaminoglycans. *Curr. Opin. Struct. Biol.* 10, 518–527
48. Tsuchiya, T., Osanai, T., Ogose, A., Tamura, G., Chano, T., Kaneko, Y., Ishikawa, A., Orui, H., Wada, T., Ikeda, T., Namba, M., Takigawa, M., Kawashima, H., Hotta, T., Tsuchiya, A., Ogino, T., and Motoyama, T. (2005) Methylation status of EXT1 and EXT2 promoters and two mutations of EXT2 in chondrosarcoma. *Cancer Genet. Cytogenet.* 158, 148–155
49. Mahalingam, Y., Gallagher, J. T., Couchman, J. R. (2007) Cellular adhesion responses to the heparin-binding (HepII) domain of fibronectin require heparan sulfate with specific properties. *J. Biol. Chem.* 282, 3221–3230
50. Shukla, D., Liu, J., Blaiklock, P., Shworak, N. W., Bai, X., Esko, J. D., Cohen, G. H., Eisenberg, R. J., Rosenberg, R. D., and Spear, P. G. (1999) A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell* 99, 13–22
51. Xu, D., Tiwari, V., Xia, G., Clement, C., Shukla, D., and Liu, J. (2005) Characterization of heparan sulphate 3-O-sulphotransferase isoform 6 and its role in assisting the entry of herpes simplex virus type 1. *Biochem. J.* 385, 451–459
52. Ye, S., Luo, Y., Lu, W., Jones, R. B., Linhardt, R. J., Capila, I., Toida, T., Kan, M., Pelletier, H., and McKeehan, W. L. (2001) Structural basis for interaction of FGF-1, FGF-2, and FGF-7 with different heparan sulfate motifs. *Biochemistry* 40, 14429–14439
53. HajMohammadi, S., Enjoji, K., Princivalle, M., Christi, P., Lech, M., Beeler, D., Rayburn, H., Schwartz, J. J., Barzegar, S., de Agostini, A. I., Post, M. J., Rosenberg, R. D., and Shworak, N. W. (2003) Normal levels of anticoagulant heparan sulfate are not essential for normal hemostasis. *J. Clin. Invest.* 111, 989–999
54. Kamimura, K., Rhodes, J. M., Ueda, R., McNeely, M., Shukla, D., Kimata, K., Spear, P. G., Shworak, N. W., and Nakato, H. (2004) Regulation of Notch signaling by *Drosophila* heparan sulfate 3-O-sulfotransferase. *J. Cell Biol.* 166, 1069–1079
55. Mochizuki, H., Yoshida, K., Shibata, Y., and Kimata, K. (2008) Tetrasulfated disaccharide unit in heparan sulfate: Enzymatic formation and tissue distribution. *J. Biol. Chem.* 283, 31237–31245
56. Vanpouille, C., Deligny, A., Delehedde, M., Denys, A., Melchior, A., Lienard, X., Lyon, M., Mazurier, J., Fernig, D. G., and Allain, F. (2007) The heparin/heparan sulfate sequence that interacts with cyclophilin B contains a 3-O-sulfated N-unsubstituted glucosamine residue. *J. Biol. Chem.* 282, 24416–24429
57. Uzan, B., Figeac, F., Portha, B., Movassat, J. (2009) Mechanisms of KGF-mediated signaling in pancreatic duct cell proliferation and differentiation. *PLoS One* 4, e4734
58. Kaminski, A., Hahne, J. C., Haddouti, M., Florin, A., Wellmann, A., Wernert, N. (2006) Tumour-stroma interactions between metastatic prostate cancer cells and fibroblasts. *Int. J. Mol. Med.* 18, 941–950
59. Chandrasekaran, E. V., Xue, J., Neelamegham, S., Matta, K. L. (2006) The pattern of glycosyl- and sulfotransferase activities in cancer cell lines: a predictor of individual cancer-associated distinct carbohydrate structures for the structural identification of signature glycans. *Carbohydr. Res.* 341, 983–994
60. Gronback, K., Hother, C., and Jones, P. A. (2007) Epigenetic changes in cancer. *Apmis.* 115, 1039–1059
61. Chow, W. A. (2007) Update on chondrosarcomas. *Curr. Opin. Oncol.* 19, 371–376

Received for publication May 18, 2009.  
Accepted for publication September 17, 2009.

## Supplemental data

**Table 1.** Sequence of primers used for amplification of bisulfite converted genomic DNA sequences by nested-PCR. First round of PCR was performed with primers “Out sens” and “Out rev” and second round of PCR was performed with primers “In sens” and “In rev”. Primers sets 1 and 2 for *GlicAT-I* and *3-OST3A* were used to amplified the corresponding regions of the same gene. TSS, transcription start site.

Gene	Primer name and sequence		Location relative to TSS	Size	CG sites	
<i>XT-1</i>	Out sens	5'-GGATGAAGGGATTAGATTGTG-3'	In sens	5'-TTTAGTGGGATAGGGGTG-3'	+21 to 358 bp	65
	Out rev	5'-CTTACAAAATCCCAATCTAATAAAC-3'	In rev	5'-CTCCTCTCAAACACTACAAC-3'	+379	
<i><math>\beta</math>4GalT7</i>	Out sens	5'-TTTAGTTTAGTGTATTGTTGGAGG-3'	In sens	5'-GTTAGGAAGGAGTCGTTT-3'	-174 to 247 bp	40
	Out rev	5'-AAACCACACGCCCTCACAAAAC-3'	In rev	5'-ACCTACCGTCTCCCAAAA-3'	+73	
<i><math>\beta</math>3GalT6</i>	Out sens	5'-GTGGGAAATAGTTTAGGTAGAG-3'	In sens	5'-GTAGGTTTGGTTAGTAGGAGATT-3'	-585 to 417 bp	62
	Out rev	5'-CCC GCCCAACAACCTTCATAAC-3'	In rev	5'-CACCTTATCATTAATCAACC-3'	-168	
<i>GlicAT-I</i>	Out sens 1	5'-GTATTATAGTAGGATTTTAGGAG-3'	In sens1	5'-GATATTTAGAGGATTAATGTTGGG-3'	-458 to 328 bp	24
	Out rev 1	5'-TTCTTCAACTTCAACTTCATAACC-3'	In rev 1	5'-CCTACCAACTCCCAACTAAAATAA-3'	-130	
<i>EXT1</i>	Out sens 2	5'-GGATGGATGAATAAGGTGAAGATT-3'	In sens 2	5'-GTGATGAGGGGCGGTTT-3'	-323 to 348 bp	37
	Out rev 2	5'-CTATCAACAACAACCTTAATAACAAC-3'	In rev 2	5'-TTCTTCAACTCAACTCATAAAC-3'	+25	
<i>NDST-1</i>	Out sens	5'-GTATGTAGAGGAGGAGTAAG-3'	In sens	5'-GTAGGTAGTTAGGAGTTAG-3'	-536 to 739 bp	106
	Out rev	5'-TTACCTCTCGAATTCCTCTC-3'	In rev	5'-CAACTTCAACTCATCCACCAC-3'	+203	
<i>NDST-2</i>	Out sens	5'-GTTTGTATGGTGATATAAAGGG-3'	In sens	5'-TTTGTGAATTTGTTGGTTAGTGG-3'	+302 to 366 bp	22
	Out rev	5'-CACAATACTTATCCACAACCTC-3'	In rev	5'-CCACCTCTAACCCAAATTAC-3'	+668	
<i>NDST-3</i>	Out sens	5'-AGTGTGTTGGAGTTGGTTG-3'	In sens	5'-GAGTGGATGAGTTAGGTTTG-3'	-281 to 245 bp	39
	Out rev	5'-TACCCCAACTTCCCTACTTT-3'	In rev	5'-CTACCCACAACCTAACCCAAC-3'	-36	
<i>NDST-4</i>	Out sens	5'-GGTTTGGGTAGTTAGATTTTITAG-3'	In sens	5'-GTTTTAGGGTTGTGGTTTTTITAG-3'	-787 to -326	46
	Out rev	5'-AACAACAACCTTATAACATCTCTC-3'	In rev	5'-TATTTACAACCCACAATCACTC-3'	326	
<i>HS6ST-1</i>	Out sens	5'-GTTTAGGGAATTAGTATTGG-3'	In sens	5'-GAATGAGGAGAAGTATTTGAAAAG-3'	-776 to 453 bp	5
	Out rev	5'-AATTCCTTCAACTAATCATCCTC-3'	In rev	5'-CAGTTTCAACTATAACACTAATCC-3'	-323	
<i>HS6ST-2</i>	Out sens	5'-GAAAGTGTAGTTAGTGGTTAG-3'	In sens	5'-GTTTGTAGTTAAGGTTAAGG-3'	-671 to 747 bp	115
	Out rev	5'-CCAAAACCTAATCTTAAACC-3'	In rev	5'-CACCAACACGAACCTTACTAAC-3'	+76	
<i>HS6ST-3</i>	Out sens	5'-GAGAAAATCGGATAGAGTG-3'	In sens	5'-TAA AGA GGA TAG GGG TAG AAG-3'	-438 to 472 bp	48
	Out rev	5'-CCTAACAAAACCTCATCTC-3'	In rev	5'-CATCACCAAAACCTAACCAAC-3'	+34	
<i>HS2ST-1</i>	Out sens	5'-GTATTTAGGGTAGGTTTTTITG-3'	In sens	5'-GGGTTTTTGTATATCGGTTGG-3'	-721 to 729 bp	123
	Out rev	5'-CAACAACAACCTTATAAACCTTTC-3'	In rev	5'-ATCCATAATCCCACTCAAAAC-3'	+8	
<i>HS2ST-2</i>	Out sens	5'-GATATTCGTTAAGTAGTTAAGAAAG-3'	In sens	5'-GATTGAGTTTGGGTAGGAC-3'	-207 to 591 bp	68
	Out rev	5'-GATTTAGAAAATTGGAGGAGTTT-3'	In rev	5'-CACCAACAACACTACAACCTTAAAC-3'	+384	

<i>3-OST1</i>	Out sens	5'-AGG AAA AGG AGT GTT TAA TGG-3'	In sens	5'-TTGGAGAGGATGGTATTTTTTIG-3'	-546 to	682 bp	59
	Out rev	5'-CCTAAATAAACTCCTAAACTAACC-3'	In rev	5'-CATAAACAACACTACCGAACAC-3'	+136		
<i>3-OST2</i>	Out sens	5'-TTGGGGTAGTAGAGTTTTGTAA-3'	In sens	5'-ATGGTGGGTGTAGAGATTTT-3'	-520 to	612 bp	64
	Out rev	5'-ACTACTTCATTACGCTCCAAC-3'	In rev	5'-CAAAAATCCCAAAAACCCCTCAA-3'	+92		
<i>3-OST3A</i>	Out sens 1	5'-TAGTGTTAAGAGGGGAGTAG-3'	In sens 1	5'-ATTAAAGTTAGTGGCGTTTGGG-3'	-338 to	438 bp	46
	Out rev 1	5'-CTTTCTCTAATCCCCTTACC-3'	In rev 1	5'-CTACTCTTAAACCCACCTAT-3'	+100		
	Out sens 2	5'-ATAGGTGGGTTTAAAGAGTAG-3'	In sens 2	5'-GGTAAGGGGATTAGAGAAAAG-3'	+282 to	444 bp	54
	Out rev 2	5'-CCGACCATACTAAACCTAAAC-3'	In rev 2	5'-CCGACAAAATACCAAAACATCC-3'	+726		
<i>3-OST3B</i>	Out sens	5'-GAAAATTTGGATGGGGAITG-3'	In sens	5'-GGGAGCGTATTTAGTAGG-3'	-354 to	572 bp	47
	Out rev	5'-AACTCAACACATAACGTAATCC-3'	In rev	5'-CTACTACTTCCCCAAAATAAC-3'	+218		
<i>3-OST6</i>	Out sens	5'-GTTAGGTTGTGATTCGATGTG-3'	In sens	5'-ATTGTTAGTTGGGAGTGTAGG-3'	-647 to	739 bp	86
	Out rev	5'-CCAAAAACAACCCGAATACTCC-3'	In rev	5'-ACCAAAAACCCACGAAACATC-3'	+92		

**Table 2.** Sequence, size of primers, and conditions used for RT-PCR analysis. NA, not applicable.

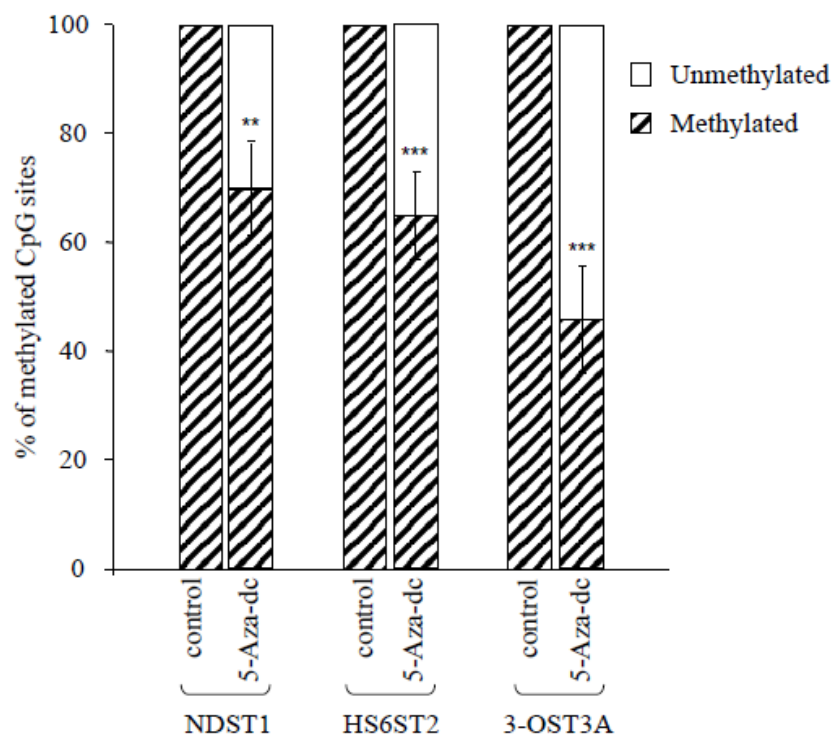
Gene	Primer name	Primer sequence	Product size	Pre-PCR	Denature Step	Annealing	Extension	Cycle	Final incubation
<i>GAPDH</i>	sens	5'-ACCACAGTCCATGCCATCAC-3'	452 bp	94°C, 3 min	94°C, 45 s	60°C, 45 s	72°C, 45 s	24	72°C, 5 min
	reverse	5'-TCCACCACCCCTGTTGCTGT-3'							
<i>NDST1</i>	sens	5'-ACCACAGCCAGACTGAACGTTTGTG-3'	443 bp	94°C, 3 min	94°C, 45 s	58°C, 45 s	72°C, 45 s	32	72°C, 5 min
	reverse	5'-ACAGGGAAATGCCAGTCTGTCTCC-3'							
<i>NDST2</i>	sens	5'-TTCCAGTCTGGCTGATGCCAGC-3'	426 pb	94°C, 3 min	94°C, 45 s	58°C, 45 s	72°C, 45 s	32	72°C, 5 min
	reverse	5'-GTAGATCCCACCTGCCTGGCTAAAGG-3'							
<i>NDST3</i>	sens	5'-CAGGAAAATGAACTCTCTGAGACGG-3'	631 pb	94°C, 3 min	94°C, 45 s	60°C, 45 s	72°C, 45 s	NA	72°C, 5 min
	reverse	5'-AGATGGAAGGAGAAAGGTTTCTGG-3'							
<i>NDST4</i>	sens	5'-CCACCTCTCCACAAACGAGT-3'	330 pb	94°C, 3 min	94°C, 45 s	60°C, 45 s	72°C, 45 s	NA	72°C, 5 min
	reverse	5'-ATCCAGTCTTGGGGTCTC-3'							
<i>HS2ST1</i>	sens	5'-ATGGGGCTCCTCAGGATATG-3'	297 pb	94°C, 3 min	94°C, 45 s	58°C, 45 s	72°C, 45 s	32	72°C, 5 min
	reverse	5'-CTTTGCACACAGGTCATAGGC-3'							
<i>HS6ST1</i>	sens	5'-AAGAAGTGCACCTGCTACCC-3'	358 pb	94°C, 3 min	94°C, 45 s	58°C, 45 s	72°C, 45 s	32	72°C, 5 min
	reverse	5'-GGCAGTCCATGAATCCCTG-3'							
<i>HS6ST2</i>	sens	5'-TCTTTCGGCTCACTGAGTTTC-3'	302 pb	94°C, 3 min	94°C, 45 s	60°C, 45 s	72°C, 45 s	34	72°C, 5 min
	reverse	5'-TGAAAATGGGTCTGAAGGAGC-3'							
<i>HS6ST3</i>	sens	5'-GCACCTGGTGAAGAATCC-3'	339 bp	94°C, 3 min	94°C, 45 s	63°C, 45 s	72°C, 45 s	NA	72°C, 5 min
	reverse	5'-GGGCTTCTCCATCACACATATG-3'							
<i>EXT1</i>	sens	5'-TTGGGTCTTCAGATTCCTG-3'	329 bp	94°C, 3 min	94°C, 45 s	54°C, 45 s	72°C, 45 s	28	72°C, 5 min
	reverse	5'-TCCTCCAGGATGTTTGTTC-3'							
<i>3-OST1</i>	sens	5'-AGCTTCTCGGAAAGCGG-3'	484 bp	94°C, 3 min	94°C, 45 s	58°C, 45 s	72°C, 45 s	34	72°C, 5 min
	reverse	5'-GCCTTGATCCACATTGAGC-3'							
<i>3-OST2</i>	sens	5'-GAATAAGCCACGAAAGAAA-3'	272 bp	94°C, 3 min	94°C, 45 s	58°C, 45 s	72°C, 45 s	35	72°C, 5 min
	reverse	5'-TATGGCAGAAAGGAGAAATGA-3'							
<i>3-OST3A</i>	sens	5'-GAGCGGCCCTGGAAAGAAAGT-3'	247 bp	94°C, 3 min	94°C, 45 s	63°C, 45 s	72°C, 45 s	35	72°C, 5 min
	reverse	5'-AGCTCGGTTCGAAAGATGGG-3'							
<i>3-OST3B</i>	sens	5'-TCATCATCGGCTGAAGAAG-3'	299 bp	94°C, 3 min	94°C, 45 s	63°C, 45 s	72°C, 45 s	35	72°C, 5 min
	reverse	5'-CTGCGTGTAGTCCGAGATG-3'							
<i>3-OST6</i>	sens	5'-CGAAGCTGATCGTGGTGTG-3'	334 bp	94°C, 3 min	94°C, 45 s	54°C, 45 s	72°C, 45 s	32	72°C, 5 min
	reverse	5'-CCCTTGGTGGCTTGAAGTA AAG-3'							

### Supplemental data

**Figure 1.** Analysis of the effect of 5-Aza-dc treatment on the methylation status of sulfotransferase *NDST1*, *HS6ST2* and *3OST3A* (region 2) genes in HEMC cells. Genomic DNA was extracted from untreated and 5-Aza-dc (4 $\mu$ M) treated HEMC cells as described in “Materials and Methods”, and analyzed by bisulfite conversion. Hatched bars bar correspond to the percentage of methylated cytosines in CpG sites of the gene of interest in untreated or treated cells. Data represents the mean value  $\pm$  SEM from sequencing data of three independent clones. One-way analysis of variance followed by Bonferroni’s multiple comparison *post-hoc* correction test was used for analyzing statistical significance (\*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$  versus untreated or mock-transfected).

Supplemental data

Figure 1



## **Etude des mécanismes de régulation des voies de biosynthèse des GAGs**

Un aspect complémentaire de ma thèse a été consacré à la compréhension des mécanismes de régulation de la synthèse des PGs. Dans ce cadre, j'ai participé, en collaboration avec Catherine Bui, à une étude originale qui a consisté à déterminer le rôle des modifications épigénétiques dans la régulation de l'expression des gènes codant pour les enzymes impliquées dans l'initiation, la polymérisation et la modification des chaînes d'HS.

Il est établi que la régulation épigénétique par hyperméthylation de l'ADN constitue un événement majeur du processus de cancérogenèse. De plus, des études récentes ont montré que la méthylation de l'ADN et d'autres modifications épigénétiques pourraient également être impliquées dans certaines arthropathies telles que l'arthrose et les arthrites rhumatoïdes (Takami *et al.*, 2006; Nile *et al.*, 2008). Dans ce contexte, nous nous sommes intéressés à l'étude d'un modèle cellulaire de chondrosarcome humain H-EMC-SS (HEMC), qui constitue une lignée à caractère invasif des tissus conjonctifs.

Cette étude a donné lieu à la **Publication n°3**

EPIGENETICS: METHYLATED-ASSOCIATED REPRESSION OF HEPARAN 3-O-SULFOTRANSFERASE GENE EXPRESSION CONTRIBUTES TO THE INVASIVE PHENOTYPE OF H-EMC-SS CHONDROSARCOMA CELLS

Catherine Bui, Mohamed Ouzzine, **Ibtissam Talhaoui**, Sheila Sharp, Kristian Prydz, Michael W.H. Coughtrie and Sylvie Fournel-Gigleux.

*FASEB J.* (2010) 24; 436-450

### **1. Analyse du profil de méthylation des gènes codant pour les GTs et les STs impliquées dans la synthèse et la modification des HS dans les cellules HEMC**

Dans un premier temps, nous avons analysé la séquence des régions promotrices des GTs intervenant dans la biosynthèse des HS, notamment, les gènes codants pour la XT-I, la  $\beta$ 4GalT7, la  $\beta$ 3GalT6, la GlcAT-I et l'EXT1, ainsi que les gènes codant pour les enzymes impliquées dans la modification des chaînes d'HS par N-sulfatation (NDSTs) et O-sulfatation



(HS2ST, HS6STs et 3-OSTs). Cette approche a permis d'identifier la présence de régions riches en nucléotides cytidine (C) et guanosine (G), appelées îlots CpG, dans la plupart des gènes étudiés. Ces îlots CpG sont localisés au niveau de la région promotrice proximale et du premier exon des gènes cibles. Ils constituent des sites potentiels de méthylation de l'ADN et par conséquent, un mécanisme de régulation de l'expression de ces gènes dans les cellules d'origine cancéreuse HEMC a été envisagé. Nous avons étudié le profil de méthylation de chacun des gènes mentionnés précédemment par la méthode de conversion au bisulfite. C'est une technique qui permet d'analyser la méthylation de l'ADN par conversion spécifique des nucléotides C non méthylés en uracile (U), puis amplification de la région d'intérêt par PCR introduisant des résidus thymidine (T), suivi du séquençage du produit de PCR qui donne une carte de méthylation.

L'analyse du profil de méthylation pour chacun des quatre gènes codant pour les GTs impliquées dans la synthèse de l'amorce tétrasaccharidique montre qu'ils sont très peu ou pas méthylés. Par conséquent, cette régulation épigénétique ne serait pas impliquée dans le contrôle de l'initiation de la synthèse des chaînes de GAGs dans le modèle étudié de chondrosarcome humain (Publication n°3, Figure 1). De façon intéressante, nous avons démontré que les gènes codants pour la NDST1, l'HS6ST2, les 3-OST1, 3-OST3A et 3-OST6 sont hyperméthylés dans un ordre décroissant de 100% à 38% pour la 3-OST6, dans les cellules HEMC (Publication n°3, Figure 2A et 3C).

## **2. Analyse des conséquences de la déméthylation des gènes d'intérêt par évaluation de leur taux d'expression dans des cellules de chondrosarcome humain après traitement par la 5-Aza-dc**

Dans un deuxième temps, nous avons étudié l'effet de la déméthylation des gènes d'intérêt sur leur niveau d'expression, par traitement des cellules HEMC à la 5-Aza-2'-déoxycytidine (5-Aza-dc). Pour ce faire, nous avons donc traité les cellules en culture à la 5-Aza-dc, qui est un agent inhibiteur de la méthylation de l'ADN *via* la répression irréversible des ADN-méthyltransférases (DNMTs). Après extraction et purification des ARN totaux, le taux d'expression des gènes identifiés comme hyperméthylés dans les cellules HEMC a été évalué par RT-PCR semi-quantitative (Publication n°3, Figure 4).

La comparaison du niveau d'expression de l'ARNm du gène EXT1 dans les cellules traitées et non traitées à la 5-Aza-dc ne montre aucune différence. Cette constatation est en accord avec le résultat précédent (Publication n°3, Figure 1), qui a démontré que ce gène n'est pas

hyperméthylé et par conséquent, ne serait pas soumis à un processus de régulation épigénétique par méthylation de l'ADN (Publication n°3, Figure 4A et 4C). En revanche, le traitement par la 5-Aza-dc semble augmenter la transcription des gènes codant pour la NDST1, HS6ST2, 3-OST1, 3-OST2 et 3-OST3A de façon significative (Publication n°3, Figure 4). Ce résultat est également en accord avec la présence d'une hyperméthylation importante au niveau de ces gènes et corrobore l'implication d'une régulation épigénétique *via* un mécanisme de méthylation de l'ADN, de la sulfatation des chaînes d'HS, dans les cellules de chondrosarcome humain.

### **3. Analyse de l'effet du traitement des cellules HEMC par la 5-Aza-dc sur la sulfatation des chaînes de GAGs des PGs cellulaires**

Nous avons poursuivi cette étude par la détermination des conséquences de la méthylation de l'ADN sur la sulfatation des GAGs de ces cellules de chondrosarcome humain. Pour atteindre cet objectif, nous avons développé des approches biochimiques qui permettent d'analyser la taille et la charge ionique de ces molécules par chromatographie d'exclusion et échangeuse d'anions, respectivement.

Les cellules HEMC traitées ou non par la 5-Aza-dc ont été incubées en présence de sulfate radiomarqué [<sup>35</sup>S] ou [<sup>35</sup>S]Cys-Met. Les lysats cellulaires ont été ensuite soumis à une digestion à la chondroïtinase ABC afin de permettre une analyse spécifique des HSPGs produits par les cellules de chondrosarcome humain. En effet, il a été démontré que ces cellules produisent majoritairement des HSPGs cellulaires (résultats non publiés).

Les modifications du degré de sulfatation des chaînes de GAGs s'accompagnent d'une modification de leur charge et par conséquent, l'analyse par chromatographie échangeuse d'anions montrera des changements du temps de rétention de ces molécules.

La comparaison du profil d'élution de l'extrait des cellules traitées à la 5-Aza-dc à celui des cellules non traitées, après radiomarquage au [<sup>35</sup>S], montre une nette différence. En effet, le traitement à la 5-Aza-dc provoque un décalage du pic d'élution vers la droite traduisant une augmentation de la charge des HSPGs produits dans les cellules HEMC (Publication n°3, Figure 5A). Cet effet peut résulter de différents facteurs, soit de l'augmentation du nombre de groupements sulfates, soit de l'augmentation de la quantité des HSPGs produits et/ou encore de l'augmentation de la longueur des chaînes d'HS. Pour vérifier chacune de ces hypothèses, nous avons réalisé différents radiomarquages et analyses chromatographiques.

L'analyse par chromatographie de gel-filtration, après radiomarquage à la [<sup>35</sup>S]Cys-Met, montre que les cellules traitées et non traitées à la 5-Aza-dc présentent le même profil avec des pics superposés (Publication n°3, Figure 5B). Ce résultat permet d'exclure l'hypothèse selon laquelle l'augmentation de la charge des HSPGs est due à l'augmentation de la quantité des PGs produit dans les cellules traitées par la 5-Aza-dc. Ensuite, nous avons procédé à la réalisation d'une deuxième série de radiomarquage en présence de [<sup>35</sup>S]SO<sub>4</sub> et [<sup>3</sup>H]GlcN afin de suivre, respectivement, le degré de sulfatation et la formation des chaînes polysaccharidiques de GAGs. Les PGs isolés à partir de cellules HEMC traitées et non traitées et préalablement radiomarquées ont été soumis à une digestion à la protéinase K, afin de libérer les chaînes glycosaminoglycaniques de leur protéine "core". Les GAGs ainsi obtenus ont été digérés à la chondroïtinase ABC, avant d'être analysés par chromatographie d'exclusion. Le chromatogramme correspondant au rapport du taux d'incorporation [<sup>35</sup>S]/[<sup>3</sup>H] permet d'estimer la longueur des chaînes d'HS. Ainsi, le traitement des cellules à la 5-Aza-dc ne semble pas modifier le profil du chromatogramme obtenu, mais produit une augmentation de la quantité de [<sup>35</sup>S] incorporé par rapport à la quantité de [<sup>3</sup>H] dans les HSPGs des cellules HEMC (Publication n°3, Figure 5C). Cette expérience montre une augmentation du degré de sulfatation des chaînes d'HS et non pas de leur longueur par le traitement déméthylant.

Pour résumer, nous avons démontré par cette approche que la déméthylation de l'ADN, induite par le traitement à la 5-Aza-dc, provoque une augmentation du degré de sulfatation des chaînes HS dans les cellules du chondrosarcome humain. Par conséquent, un mécanisme de méthylation aberrante de l'ADN est potentiellement impliqué dans le contrôle de la sulfatation des HSPGs dans les cellules HEMC.

#### **4. Comparaison entre l'effet de la surexpression des 3-OSTs et le traitement par la 5-Aza-dc sur l'expression des HSPGs dans les cellules HEMC**

De façon complémentaire, nous avons étudié l'expression des HSPGs dans les cellules de chondrosarcome humain, traitées par la 5-Aza-dc ou transfectées avec le plasmide vide ou portant le gène codant pour la 3-OST1 ou 3-OST3A (pcDNA-3-OST1-HA et pcDNA-3-OST3A-HA), par une technique d'immunomarquage par fluorescence indirecte au moyen d'un anticorps anti-HSPG (épitope 10E4).

Les cellules HEMC non traitées ainsi que celles transfectées avec le plasmide vide montrent la présence d'un signal très faible des chaînes d'HS (Publication n°3, Figure 6A et 6D). En

revanche, le traitement de ces cellules avec la 5-Aza-dc ou leur transfection avec la 3-OST1 et 3-OST3A augmente de façon importante l'expression des chaînes d'HS (Publication n°3, Figure 6B, 6E et 6F). Simultanément, l'expression des STs recombinantes dans les cellules transfectées a été suivie par un immunomarquage au moyen d'un anticorps anti-HA (Publication n°3, Figure 6H et 6I).

L'ensemble de ces résultats confirme la présence d'un processus de régulation épigénétique, *via* la méthylation de l'ADN, impliqué dans la régulation de l'expression des gènes codant pour la 3-OST1 et la 3-OST3A et qui par conséquent, influence le taux de sulfatation des HSPGs dans les cellules de chondrosarcome humain.

## **5. Effet des modifications de la sulfatation des HSPGs sur le comportement cellulaire des HEMC**

La dernière partie de ce travail a été consacrée à la caractérisation des conséquences de modifications de la sulfatation des HSPGs, sur les propriétés phénotypiques des cellules HEMC, potentiellement impliquées dans le processus cancéreux tels que la prolifération, l'adhésion, et la migration.

Le traitement par la 5-Aza-dc diminue, comme attendu, la prolifération des cellules HEMC de façon importante (45%), tandis que la surexpression de la 3-OST3A a un effet inhibiteur plus modéré (20%) sur l'inhibition de la prolifération cellulaire. Quant à la surexpression de la 3-OST1, elle ne semble pas affecter ce paramètre cellulaire dans les cellules HEMC (Publication n°3, Figure 7A et 7B). Ces résultats suggèrent que la prolifération des cellules de chondrosarcome humain semble être régulée par la méthylation du gène codant pour la 3-OST3A.

Nous avons ensuite étudié l'effet du traitement par la 5-Aza-dc ou la transfection de la 3-OST1 et 3-OST3A dans les cellules de chondrosarcome humain, sur l'adhésion de ces cellules en utilisant la fibronectine comme ligand. Les cellules traitées par la 5-Aza-dc présentent une augmentation significative de leur capacité d'adhésion en présence de 200 µg de fibronectine (Publication n°3, Figure 7C). Après transfection des cellules HEMC avec le plasmide codant pour la 3-OST3A, leur capacité d'adhésion semble être augmentée en fonction de la dose de la fibronectine utilisée. Enfin, l'expression de la 3-OST1 n'affecte pas le taux d'adhésion des cellules HEMC (Publication n°3, Figure 7D).

De façon complémentaire, le dernier objectif de cette étude a été de déterminer l'implication de la méthylation de l'ADN ou de la surexpression des gènes 3-OST1 et 3-OST3A dans le processus de la migration cellulaire des cellules de chondrosarcome humaine et dans le caractère invasif de ce type de cancer. Pour atteindre cet objectif, nous avons utilisé un système appelé chambre de Boyden qui permet d'évaluer le potentiel migratoire des cellules. Dans un premier temps, nous avons pu démontrer que la déméthylation de l'ADN par le traitement des cellules HEMC par la 5-Aza-dc diminue de façon significative le taux de la migration cellulaire (Publication n°3, Figure 7E). Dans un deuxième temps, nous avons prouvé que la surexpression de la 3-OST1 et 3-OST3A provoque une baisse significative du pouvoir migratoire des cellules du chondrosarcome humain, HEMC (Publication n°3, Figure 7F).

L'ensemble de ces résultats suggère fortement l'implication de l'altération de la sulfatation des HSPGs dans le phénotype invasif des cellules de chondrosarcome humain.

# *Discussion*

## Discussion

Les PGs sont des macromolécules multifonctionnelles complexes, dont les chaînes de GAGs sont impliquées dans de nombreux processus biologiques, physiologiques et pathologiques. Du fait de l'altération de la biosynthèse et/ou de la modification des chaînes glycosaminoglycaniques observées au cours de multiples affections, l'étude de ces molécules constitue un enjeu médical majeur. Notamment, dans le cas du processus cancéreux, la dérégulation de l'expression des GAGs, et plus particulièrement des HS, est associée à la progression tumorale et métastatique. Ainsi, la diminution de la sulfatation des HSPGs modifie les propriétés d'adhésion cellulaire des cellules tumorales et pourrait participer à l'acquisition d'un phénotype invasif (Robinson *et al.*, 1984). Le suivi de l'altération de la sulfatation des chaînes d'HS aurait une valeur pronostique dans le cas de tumeurs malignes telles que le mélanome humain (Sasisekharan *et al.*, 2002). A la lumière de ces travaux, il est devenu évident que la caractérisation des enzymes impliquées dans la machinerie de biosynthèse, notamment les GTs et les STs, ouvrirait des perspectives vers le développement de stratégies innovantes dans le traitement des pathologies associées à la perturbation de la synthèse des GAGs.

La  $\beta$ 4GalT7 est une enzyme qui intervient à un stade précoce de l'initiation des voies de biosynthèse des GAGs. Sa présence est essentielle à la formation des chaînes de CS/DS et d'Hp/HS (Almeida *et al.*, 1999). Chez l'homme, il a été démontré que trois mutations de la  $\beta$ 4GalT7 sont à l'origine d'une maladie héréditaire connue sous le nom de syndrome progéroïde d'Ehlers-Danlos. Par ailleurs, cette enzyme possède la capacité d'initier la synthèse des chaînes de GAGs à partir de substrats xylosidiques exogènes, à la fois *in vitro* et *in vivo*. Ainsi, la  $\beta$ 4GalT7 constitue une cible pharmacologique potentielle pour le développement de molécules thérapeutiques destinées à remédier au déficit de synthèse des chaînes de GAGs.

Dans ce contexte, mon travail de thèse a été organisé selon deux axes : 1- Exploration des mécanismes d'initiation de la synthèse des GAGs; 2- Etude de la régulation de la biosynthèse et des modifications des chaînes d'HS.

Nous nous sommes en premier lieu attachés à la caractérisation structurale et fonctionnelle de la  $\beta$ 4GalT7 recombinante humaine. De façon complémentaire, nous avons cherché à déterminer les mécanismes moléculaires par lesquels les mutations génétiques de la  $\beta$ 4GalT7 contribuent au syndrome progéroïde d'Ehlers-Danlos.

Dans le cadre de ces études, nous avons établi un arbre phylogénétique qui a permis d'identifier 48 membres de la famille des  $\beta$ 4GalT7s au sein de différentes espèces et 6  $\beta$ 4GalTs humaines (1 à 6) dont les gènes présentent une relation évolutive. L'alignement des séquences peptidiques de ces 54 enzymes montre la présence de 8 motifs fortement conservés, notamment les motifs 3 et 5 qui contiennent respectivement les séquences <sup>163</sup>DVD<sup>165</sup> et <sup>221</sup>FWGWGREDDE<sup>230</sup> et le motif 6 contenant la séquence <sup>257</sup>HLH<sup>259</sup>, qui ont fait l'objet de mon étude. Cette approche a été complétée par la réalisation d'un modèle moléculaire de la  $\beta$ 4GalT7 humaine basé sur la structure cristallographique de la  $\beta$ 4GalT7 de drosophile en présence d'UDP (Ramakrishnan and Qasba, 2010) et l'organisation 3D de la  $\beta$ 4GalT1 bovine cristallisée en complexe avec l'UDP-Gal (Ramakrishnan *et al.*, 2002).

A la lumière des informations fournies par ces approches, nous avons procédé par une étude de mutagenèse dirigée, à l'analyse des conséquences fonctionnelles de substitutions conservatives et non conservatives des acides aminés situés dans les séquences <sup>163</sup>**DVD**<sup>165</sup> <sup>221</sup>**FWGWGREDDE**<sup>230</sup> et <sup>257</sup>**HLH**<sup>259</sup> de la  $\beta$ 4GalT7 humaine. Ainsi, j'ai caractérisé, *in vitro* et *ex vivo*, les conséquences de ces mutations sur la reconnaissance des substrats et/ou la catalyse enzymatique, ou encore sur l'organisation structurale de la  $\beta$ 4GalT7 humaine recombinante. La spécificité de l'enzyme sauvage et mutée vis-à-vis des substrats donneur et accepteur a également été analysée. Enfin, nous avons développé une approche complémentaire afin d'étudier la capacité de l'enzyme sauvage et mutée à initier: 1- La synthèse des GAGs endogènes dans des cellules déficientes en  $\beta$ 4GalT7; 2- La synthèse des chaînes de GAGs à partir d'un substrat xylosidique exogène (4-MOX); 3- La synthèse des chaînes de GAGs sur la décorine choisie comme modèle de protéine "core". La décorine est un membre de la famille des SLRPs. Le choix de ce PG est fondé sur le fait que l'altération de la synthèse des chaînes de GAGs de la décorine est caractéristique du syndrome progéroïde d'Ehlers-Danlos (Seidler *et al.*, 2006).

Les approches *in vitro* et *ex vivo* ont confirmé l'importance fonctionnelle des deux résidus Asp du motif <sup>163</sup>DVD<sup>165</sup>. La mutation de l'Asp163 ou de l'Asp165 en Ala conduit à une perte complète de l'activité enzymatique de la  $\beta$ 4GalT7, évaluée *in vitro*. De plus, la substitution conservative de l'Asp163 par un acide aminé Glu ne restaure pas l'activité GalT de l'enzyme. Le modèle moléculaire de l'organisation 3D de la  $\beta$ 4GalT7 humaine suggère une interaction entre ce résidu et l'UDP-Gal. Ainsi, on peut émettre l'hypothèse selon laquelle le remplacement de l'Asp163 par un résidu Glu, dont la chaîne latérale est plus longue, produit un encombrement stérique empêchant l'interaction entre l'enzyme et le substrat donneur.



D'un autre côté, le remplacement de l'Asp165 par un résidu Glu permet une restauration partielle de l'activité GalT *in vitro* et *ex vivo* comparée à celle de l'enzyme sauvage. Ainsi, un résidu Glu en position 165 serait capable de préserver la liaison de coordination avec le cation  $Mn^{2+}$ . Enfin, la mutation de la Val164 en Ala n'a pas de répercussions importantes sur les constantes cinétiques de l'enzyme vis-à-vis de l'UDP-Gal, dont la valeur de  $K_m$  reste inchangée par rapport à la protéine sauvage. Ce résultat suggère que la Val163 n'interviendrait pas dans la reconnaissance ou la fixation du substrat donneur. A l'appui de ce résultat, le modèle moléculaire de la  $\beta$ 4GalT7 suggère que la chaîne latérale de cet acide aminé est située hors du site de fixation de l'UDP-Gal.

L'analyse structurale du domaine catalytique de différentes GTs a révélé que ces enzymes sont des protéines globulaires qui présentent majoritairement un repliement de type GT-A ou GT-B (Bourne *et al.*, 2001; Qasba *et al.*, 2005). Chez les mammifères, la  $\beta$ 4GalT7 ainsi que les GalTs et les GlcNAc-Ts (N-acétylglucosaminyltransférases) sont des enzymes caractérisées par un repliement de type GT-A, qui possèdent généralement un motif DXD impliqué dans la liaison des cations divalents (Breton *et al.*, 1999). La structure cristallographique de la  $\beta$ 4GalT1 bovine établie par Ramakrishnan *et al.* (2002) indique que l'Asp254 (équivalent de l'Asp165 de la  $\beta$ 4GalT7 humaine) forme une liaison de coordination avec le cofacteur métallique  $Mn^{2+}$ , de manière identique à l'Asp147 de la  $\beta$ 4GalT7 de drosophile (Ramakrishnan and Qasba, 2010). Ainsi, la  $\beta$ 4GalT7 humaine présenterait un repliement GT-A de type classique, avec un motif DXD dont le premier résidu Asp interagirait avec l'UDP-Gal, tandis que le deuxième Asp formerait une liaison de coordination avec le cation  $Mn^{2+}$ .

La deuxième partie de cette étude a été consacrée à l'exploration du motif  $^{221}FWGWGREDD^{230}$  de la  $\beta$ 4GalT7 humaine. Cette séquence est fortement conservée au sein des  $\beta$ 4GalTs humaines et des  $\beta$ 4GalT7s de différentes espèces. En effet, ce motif peptidique appartient au domaine catalytique de l'enzyme et contient des acides aminés qui jouent un rôle fonctionnel clé. Dans la  $\beta$ 4GalT1 bovine, ils seraient directement impliqués dans la fixation du substrat donneur et/ou accepteur (Gastinel *et al.*, 1999; Ramakrishnan *et al.*, 2002; Ramakrishnan and Qasba, 2010). Cette séquence peptidique comprend également des résidus carboxylates qui joueraient vraisemblablement un rôle catalytique (Ramakrishnan and Qasba 2003). Mes résultats indiquent que la mutation de la Phe221 et du Trp222 n'a pas de conséquences majeures sur l'activité de l'enzyme ni sur l'affinité envers le substrat donneur, comme l'indiquent les valeurs de  $K_m$  inchangées. Ces résultats suggèrent que ces

acides aminés ne seraient pas directement impliqués dans la reconnaissance de l'UDP-Gal. Le modèle moléculaire de l'organisation 3D de la  $\beta$ 4GalT7 humaine basé sur la structure cristallographique de la  $\beta$ 4GalT1 bovine ou de la  $\beta$ 4GalT7 de drosophile, suggère que la Phe221 et le Trp222 seraient localisés à l'extérieur du site de fixation de l'UDP-Gal. Néanmoins, la conservation de ces acides aminés au sein des  $\beta$ 4GalTs nous laisse supposer qu'ils seraient importants pour l'intégrité structurale de l'enzyme.

Les données bibliographiques, l'étude phylogénétique et l'alignement de séquences des  $\beta$ 4GalTs ainsi que notre modèle moléculaire de la  $\beta$ 4GalT7 humaine, suggèrent que le Trp224 serait un acide aminé dont la localisation est centrale au niveau du site actif de l'enzyme. J'ai donc analysé l'effet des mutations de ce résidu sur l'activité de la  $\beta$ 4GalT7 humaine *in vitro* et *ex vivo*. La substitution du Trp224 par une Ala produit une protéine complètement inactive, aussi bien pour l'enzyme exprimée dans les cellules eucaryotes que la protéine purifiée exprimée dans la bactérie *E. coli*. De plus, le mutant W224A n'est pas capable de restaurer la synthèse des GAGs endogènes dans les cellules CHO-pgsB618 déficientes en  $\beta$ 4GalT7. De façon similaire, cette mutation ne permet pas à l'enzyme d'initier la formation des chaînes de GAGs sur la protéine "core" de la décorine, ni sur le substrat accepteur exogène 4-MOX. Le remplacement du Trp224 par une Phe, acide aminé aromatique de plus petite taille, provoque une diminution importante de l'activité enzymatique *in vitro* et cette protéine mutée ne peut pas initier la biosynthèse des GAGs *ex vivo* dans des cellules CHO-pgsB618 de façon significative. En outre, la faible activité du mutant W224F n'a pas permis de déterminer ses paramètres cinétiques.

De façon intéressante, en générant le mutant W224H, nous avons produit une enzyme active qui présente des différences importantes, à la fois en termes d'affinité et de spécificité comparée à l'enzyme sauvage. Notamment, le remplacement du Trp224 par une His augmente fortement l'affinité de la  $\beta$ 4GalT7 envers le substrat donneur (valeur de  $K_m$  augmentée de 10 fois) et accroît sa sélectivité vis-à-vis du sucre nucléotidique UDP-Gal. En revanche, l'affinité de l'enzyme envers le substrat 4-MOX diminue d'une façon importante. En outre, l'enzyme sauvage et le mutant W224H ne présentent pas la même capacité à prendre en charge les différents substrats accepteurs testés. En effet, la  $\beta$ 4GalT7 sauvage présente une plus forte activité vis-à-vis du substrat MN-Xyl, tandis que le mutant W224H utilise préférentiellement le 4-MOX. L'étude *ex vivo* a montré que l'expression de ce mutant ne permet pas de stimuler la biosynthèse des GAGs totaux de façon significative dans les cellules déficientes CHO-pgsB618 et qu'il est également incapable d'initier la synthèse des

chaînes de GAGs en présence du substrat xylosidique 4-MOX. L'ensemble de ces résultats souligne l'importance fonctionnelle du Trp224 de la  $\beta$ 4GalT7 humaine et montre une implication de cet acide aminé dans les interactions avec les substrats donneur et accepteur.

La structure 3D de la  $\beta$ 4GalT7 de drosophile et le modèle moléculaire de la  $\beta$ 4GalT7 humaine développé au cours de ce travail, indiquent que la molécule d'azote N $\epsilon$ 1 du Trp224 formerait une liaison avec le  $\beta$ -phosphate du pyrophosphate de l'UDP-Gal. De la même façon, dans la  $\beta$ 4GalT1 bovine cristallisée avec l'UDP-Gal et le cofacteur métallique Mn<sup>2+</sup>, le Trp314 (équivalent du Trp224 de la  $\beta$ 4GalT7 humaine) est situé dans le site de fixation du substrat donneur, lié à l'UDP-Gal *via* des interactions avec le pyrophosphate (Ramasamy *et al.*, 2003). Mes résultats expérimentaux permettent de suggérer que, dans le cas du mutant W224H, l'atome N $\epsilon$ 2 de l'imidazole du résidu His pourrait remplacer le noyau indole du Trp224 pour former une liaison hydrogène avec le groupement pyrophosphate de l'UDP-Gal. Cependant, cette mutation ne permet pas d'assurer l'interaction de l'enzyme avec le 4-MOX de façon similaire à l'enzyme sauvage. Ainsi, cela se traduit par une forte diminution de l'affinité de l'enzyme mutée envers ce substrat accepteur. En ce sens, il a été démontré qu'un résidu aminé Trp de différentes GTs, notamment le Trp314 de la  $\beta$ 4GalT1 (Ramakrishnan and Qasba, 2001), le Trp284 de l'EXTL2 (Pedersen *et al.*, 2003) et le Trp314 de l' $\alpha$ 3GT (Zhang *et al.*, 2003), serait localisé dans le site catalytique de l'enzyme de façon à interagir à la fois avec le substrat donneur et le substrat accepteur.

Le Trp314 de la  $\beta$ 4GalT1 bovine et le Trp207 de la  $\beta$ 4GalT7 de drosophile sont entourés de résidus Gly qui permettraient un changement de conformation du site actif de l'enzyme lors de la fixation du substrat donneur (Ramakrishnan *et al.*, 2002; Ramakrishnan and Qasba, 2010). La  $\beta$ 4GalT7 humaine présente une organisation similaire à ces deux enzymes, ainsi le Trp224 est entouré des Gly223 et Gly225. Nous avons donc analysé les conséquences de la mutation de chacun de ces deux résidus sur l'activité enzymatique de la protéine, *in vitro* et *ex vivo*. De façon intéressante, il s'est avéré que la mutation de ces deux acides aminés produit des conséquences différentes sur l'activité GalT évaluée *in vitro* et sur la capacité de l'enzyme à stimuler la biosynthèse des chaînes de GAGs dans les cellules CHO-pgsB618 en culture. Ainsi, le remplacement de la Gly223 par un résidu Ala induit une baisse modérée de l'activité GalT, *in vitro* et *ex vivo*, en présence du 4-MOX ou de la protéine "core" de la décorine comme substrat accepteur. Cependant, l'expression de ce mutant dans les cellules CHO-pgsB618 déficientes en  $\beta$ 4GalT7 ne permet pas de restaurer la formation des GAGs totaux.

De plus, nous avons montré que la mutation de la Gly225 de la  $\beta$ 4GalT7 en Ala, conduit à une perte totale de l'activité de cette enzyme, *in vitro* et *ex vivo*, en présence de 4-MOX comme substrat accepteur. Cependant, ce mutant reste capable de stimuler la formation de chaînes de GAGs sur la protéine "core" de la décorine dans les cellules CHO-pgsB618.

L'analyse de la structure 3D des  $\beta$ 4GalT1s bovine et humaine et de la  $\beta$ 4GalT7 de drosophile nous a aidé à apporter des hypothèses permettant d'expliquer ces résultats. La structure cristallographique de la  $\beta$ 4GalT1 bovine, de la  $\beta$ 4GalT1 humaine et de la  $\beta$ 4GalT7 de drosophile, montre respectivement que chacun des peptides  $^{313}\text{GWG}^{315}$  (Ramakrishnan and Qasba, 2001),  $^{309}\text{GWG}^{311}$  (Ramasamy *et al.*, 2005) et  $^{206}\text{GWG}^{208}$  (Ramakrishnan and Qasba, 2010) forme une courte boucle qui joue un rôle clé dans l'organisation du site actif des  $\beta$ 4GalTs. Ainsi, les Gly permettraient au résidu Trp central d'effectuer un mouvement rotationnel, provoquant un changement de la conformation de l'enzyme pour passer d'une conformation de type I (ouverte) à une conformation de type II (fermée) et inversement. Cette modification est indispensable au déroulement de la réaction catalysée par ces enzymes. On peut émettre l'hypothèse selon laquelle les Gly223 et Gly225 de la  $\beta$ 4GalT7 humaine joueraient un rôle similaire et que la mutation de ces acides aminés affecterait le positionnement du Trp224 dans la poche catalytique de l'enzyme, en diminuant l'accessibilité de ce résidu au substrat accepteur. D'un autre côté, mes résultats montrent que la substitution de chacun des résidus Gly par une Ala ne produit pas les mêmes effets *in vitro* et *ex vivo*, en fonction de la structure de l'accepteur. Ainsi le mutant G225A ne peut catalyser la réaction GalT et son expression ne permet pas de stimuler la formation de chaînes de GAGs qu'en présence d'une protéine "core" (décorine ou protéines "core" endogènes exprimées par les cellules CHO-pgsB618). Cela est probablement dû au fait que la présence d'une Ala en position 225 de la  $\beta$ 4GalT7 humaine empêche l'interaction entre le site actif de l'enzyme et la partie aglycone du substrat accepteur (4-MO). Enfin, la substitution de la Gly223 par une Ala ne permettrait pas à la partie peptidique du substrat accepteur d'accéder au site actif de l'enzyme, tandis que l'aglycone est tolérée, ce qui explique l'activité *in vitro* et *ex vivo* mesurée en présence du substrat 4-MOX. Cependant, le mutant G223A est capable d'initier la biosynthèse de chaînes de GAGs sur la protéine "core" de la décorine mais pas sur les autres protéines "core" endogènes exprimées dans les cellules CHO-pgsB618, ce qui se manifeste par l'incapacité du mutant G223A de la  $\beta$ 4GalT7 humaine à stimuler la biosynthèse des GAGs totaux des cellules CHO-pgsB618. L'élucidation de l'organisation 3D de la  $\beta$ 4GalT7 humaine en complexe avec le substrat accepteur restera l'approche la plus pertinente pour

expliquer précisément l'effet différentiel des mutations des résidus Gly223 et Gly225 sur l'activité GalT en fonction de ces substrats.

De façon similaire, nous avons étudié le rôle des acides aminés à caractère acide, situés à l'extrémité C-terminale du motif <sup>221</sup>FWGWGREEDDE<sup>230</sup>. Les résultats obtenus ont montré l'importance fonctionnelle majeure des résidus Glu227, Asp228 et Asp229, alors que le rôle du résidu Glu230 apparaît moins essentiel. En effet, l'analyse de l'effet, *in vitro* et *ex vivo*, de la substitution du résidu Glu227 par une Ala montre que cette mutation produit une perte totale de l'activité enzymatique. En revanche, la mutation E227D permet de conserver une partie de l'activité de la  $\beta$ 4GalT7 humaine *in vitro*, et l'expression de ce mutant stimule, de façon modérée, la biosynthèse des GAGs dans les cellules déficientes CHO-pgsB618 et sur le substrat xylosidique exogène. D'un autre côté, le remplacement du résidu Asp228 de la  $\beta$ 4GalT7 par une Ala produit une abolition complète de son activité GalT. En outre, le mutant D228E est actif *in vitro* et est capable de restaurer, *ex vivo*, la formation de GAGs dans les cellules CHO-pgsB618 à un niveau comparable à l'enzyme sauvage, mais ne permet pas d'initier cette biosynthèse à partir de substrat xylosidique 4-MOX. Ce résultat est dû à la diminution de l'affinité de ce mutant envers ce substrat accepteur, que montre la détermination des constantes cinétiques de l'enzyme.

L'analyse des propriétés cinétiques des mutations conservative et non conservative des résidus Glu227 et Asp228 sont en accord avec une fonction catalytique, mais mes résultats ne permettent pas de distinguer complètement lequel des deux acides aminés (Glu227 ou Asp228) joue le rôle de base catalytique dans la  $\beta$ 4GalT7 humaine. L'analyse de la première structure de la  $\beta$ 4GalT1 bovine cristallographiée en complexe avec l'UDP-Gal a proposé que le résidu Glu317 (équivalent du Glu227 dans la  $\beta$ 4GalT7 humaine) jouerait le rôle de base catalytique (Gastinel *et al.*, 1999). Par ailleurs, la cristallisation de l'enzyme bovine W314A- $\beta$ 4GalT1 avec le cofacteur métallique, l'UDP et le GlcNAc, indique que le résidu Asp318 (équivalent de l'Asp228 dans l'enzyme humaine) formerait une liaison hydrogène avec le groupement hydroxyle en position C4 du substrat accepteur et qu'il serait donc impliqué dans la déprotonation de ce groupement. Par conséquent, Ramasamy *et al.* (2003) suggèrent que l'Asp318 jouerait un rôle de base catalytique dans la  $\beta$ 4GalT1 bovine. De plus, le "docking" du Glc dans le site de fixation du substrat accepteur de la  $\beta$ 4GalT7 de drosophile a permis d'aboutir à une conclusion similaire suggérant que l'Asp211 de l'enzyme de drosophile serait la base catalytique. L'organisation 3D du modèle moléculaire de l'enzyme humaine, basé sur la structure cristallographique du site catalytique de la  $\beta$ 4GalT7 de drosophile, propose

l'Asp228 comme base catalytique, tandis que le Glu227 interagirait avec l'oxygène en position 4 du Gal. Des investigations supplémentaires sont indispensables afin de confirmer cette proposition. En ce sens, des études physicochimiques par résonance plasmonique de surface, ainsi que des études structurales par cristallographie et diffraction des rayons X de la  $\beta$ 4GalT7 humaine ont été initiées avec nos collaborateurs grenoblois.

Enfin, les mutations conservative et non conservative de l'Asp229 provoquent une perte d'activité enzymatique de la  $\beta$ 4GalT7 *in vitro*. De façon intéressante, ces mutations ne semblent pas affecter la capacité de l'enzyme à initier la synthèse des chaînes de GAGs sur la protéine "core". En outre, les mutants D229A et D229E ne sont pas capables de corriger la déficience de synthèse des chaînes glycosaminoglycaniques sur le substrat accepteur exogène 4-MOX dans les cellules CHO-pgsB618. Il est suggéré que la substitution de l'acide aminé Asp229 de la  $\beta$ 4GalT7 humaine pourrait affecter l'interaction de l'enzyme avec la partie aglycone du 4-MOX et pas la protéine "core" du PG. Néanmoins, des études plus approfondies sont indispensables à la compréhension du rôle de cet acide aminé dans le déroulement de la réaction catalysée par la  $\beta$ 4GalT7 humaine.

Enfin, j'ai initié l'étude du rôle du motif 6 identifié au cours de ce travail. La séquence HXH est fortement conservée entre les  $\beta$ 4GalT7s de différentes espèces. Néanmoins, seul le premier résidu His de cette séquence peptidique est également présent au sein des  $\beta$ 4GalTs humaines. Récemment, l'élucidation de la structure 3D de la  $\beta$ 4GalT7 de drosophile (Ramakrishnan and Qasba, 2010) a mis en évidence un nouveau motif HXH, dont chacun des deux résidus His formerait une liaison de coordination avec le cofacteur métallique  $Mn^{2+}$ . Nous avons donc étudié, par une approche *in vitro/ex vivo*, le rôle du premier résidu His du motif <sup>257</sup>**H**HLH<sup>259</sup> de la  $\beta$ 4GalT7 humaine. La substitution de l'His257 par une Arg produit une enzyme complètement inactive et l'expression du mutant H257R ne permet pas la restauration de la synthèse des GAGs dans les cellules déficientes CHO-pgsB618. La mutation du résidu His257 en Ala a des conséquences plus modérées sur l'activité GalT et maintient la capacité de la  $\beta$ 4GalT7 humaine à stimuler la formation des chaînes de GAGs des PGs endogènes. Ces résultats s'expliquent vraisemblablement par la taille de la chaîne latérale du résidu Arg en position 257 à la place de l'His qui produirait un encombrement stérique ne permettant pas la liaison aux cations  $Mn^{2+}$ . En revanche la mutation de ce résidu His par une Ala est tolérée par l'enzyme. De façon intéressante, cette mutation augmente la sélectivité de l'enzyme envers le type du cation métallique par comparaison avec l'enzyme sauvage. Afin d'expliquer ce résultat, nous projetons de déterminer l'affinité de l'enzyme sauvage ou portant une mutation

sur l'acide aminé His257 vis-à-vis des différents cations divalents, ou encore de déterminer la distance entre le cation métallique et ce résidu His dans la  $\beta$ 4GalT7 humaine cristallisée.

Dans la  $\beta$ 4GalT1 bovine, le résidu His347 (équivalent de l'His257 de la  $\beta$ 4GalT7 humaine) est localisé au niveau d'une région charnière située dans une grande boucle flexible (345-353). L'élucidation de la structure 3D de cette enzyme a montré que l'acide aminé His257 participe à la formation de liaisons de coordination avec le cofacteur métallique  $Mn^{2+}$  (Ramasamy *et al.*, 2005; Ramakrishnan *et al.*, 2006). Plus précisément, il a été proposé que le résidu His347 forme une liaison de coordination avec le cation métallique  $Mn^{2+}$  en remplaçant une molécule d'eau (Ramakrishnan *et al.*, 2006). La comparaison de la structure cristallographique de la  $\beta$ 4GalT7 de drosophile en présence d'UDP et du cation  $Mn^{2+}$  et celle de la  $\beta$ 4GalT1 bovine en présence de l'UDP-Gal et de ce même cofacteur montre une organisation similaire, à l'exception de la localisation du résidu aminé His241 de drosophile (équivalent de l'His 257 de la  $\beta$ 4GalT7 humaine) et de la Met344 de la  $\beta$ 4GalT1 bovine. De façon importante, la superposition de l'enzyme mutée M344H- $\beta$ 4GalT1 bovine et de la  $\beta$ 4GalT7 de drosophile montre une organisation 3D comparable. Cette approche a montré que la Met334 ( $\beta$ 4GalT1 bovine) est superposée avec l'His241 ( $\beta$ 4GalT7 de drosophile) et que l'His347 (bovine) est superposée avec l'His243 (drosophile). Ainsi, il a été suggéré que les résidus His257 et His259 interviendraient, de façon similaire, dans la formation des liaisons de coordinations avec les cations  $Mn^{2+}$ . Cette étude constitue une approche préliminaire de la caractérisation structurale et fonctionnelle du motif  $^{257}HXH^{259}$  spécifique des  $\beta$ 4GalT7s. Un travail complémentaire est à réaliser afin de déterminer le rôle de l'His259 de la  $\beta$ 4GalT7 humaine.

Le deuxième axe de la caractérisation de la  $\beta$ 4GalT7 réalisé au cours de cette thèse a porté sur l'étude des mutations génétiques de cette enzyme, responsables du syndrome progéroïde d'Ehlers-Danlos. Cette maladie est caractérisée par un tableau clinique particulièrement sévère associé aux anomalies de glycosylation des PGs. Parmi celles-ci, les altérations touchant les chaînes de GAGs de la décorine ont été les mieux caractérisées, notamment dans les fibroblastes de peau des patients atteints. Dans ce contexte, nous avons cherché à déterminer les conséquences moléculaires des mutations A186D, L206P et R270C, sur l'activité enzymatique de la  $\beta$ 4GalT7 *in vitro* et *ex vivo*, dans des cellules déficientes CHO-pgsB618. Les répercussions de ces mutations sur les paramètres cinétiques de l'enzyme ont été également analysées avec les protéines purifiées exprimées dans la bactérie *E. coli*. Le développement d'un modèle moléculaire de la  $\beta$ 4GalT7 humaine basé sur la structure

cristallographique de la  $\beta$ 4GalT1 bovine, nous a permis d'interpréter les résultats obtenus et de cerner l'importance fonctionnelle de chacun des trois acides aminés porteur de mutations (A186, L206 et R270).

La mutation A186D provoque une diminution importante de l'activité GalT de l'enzyme recombinante, tandis que la substitution de la Leu206 par une Pro l'abolit totalement. De façon intéressante, la perte d'activité de la  $\beta$ 4GalT7 occasionnée par la mutation de l'Arg270 en Cys est beaucoup plus modérée. Ainsi, la protéine mutée présente une activité d'environ 50% de celle de l'enzyme sauvage. Parallèlement, l'analyse de la capacité de la  $\beta$ 4GalT7 et des mutants Ehlers-Danlos à initier la synthèse des chaînes de GAGs sur le substrat accepteur 4-MOX démontre que malgré sa faible activité *in vitro*, le mutant A186D est capable de restaurer la synthèse des GAGs *ex vivo* dans des cellules déficientes CHO-pgsB618 à un niveau comparable à celui de la protéine sauvage. Ces constatations ont été confirmées par l'analyse des chaînes d'HS par technique d'immunofluorescence et sont en accord avec les travaux réalisés précédemment par Okajima *et al.* (1999b). Quant à l'absence de synthèse des chaînes de GAGs dans des cellules CHO-pgsB618 transfectées avec le gène codant pour la  $\beta$ 4GalT7 humaine portant la mutation L206P, elle a été également confirmée par l'analyse de l'expression des chaînes d'HS en immunofluorescence indirecte, qui montre une absence totale d'expression de ces GAGs dans les cellules exprimant ce mutant. La modélisation moléculaire de la  $\beta$ 4GalT7 humaine ainsi que l'analyse de la structure 3D de la  $\beta$ 4GalT7 de drosophile suggèrent que la Leu206 serait localisée dans un feuillet  $\beta$  fortement conservé dans les  $\beta$ 4GalT7s de différentes espèces. Ainsi, cet acide aminé jouerait un rôle important dans le maintien de la structure secondaire de la protéine, ce qui expliquerait l'altération du repliement de l'enzyme lors de la substitution de la Leu206 par un résidu Pro. Enfin, la mutation R270C provoque une forte diminution de l'affinité de l'enzyme envers le substrat accepteur 4-MOX, qui se traduit par une augmentation importante de la valeur du  $K_m$ . La modélisation moléculaire du site de fixation du substrat accepteur de la  $\beta$ 4GalT7 humaine, contenant le Xyl lié à la sérine d'un peptide de décorine, suggère que l'Arg270 serait située à proximité de la protéine "core" du PG. De façon complémentaire, l'analyse de l'organisation 3D du site de fixation du substrat accepteur dans la  $\beta$ 4GalT7 de drosophile indique que la Lys254 (équivalent de l'Arg270 de l'enzyme humaine) serait située du côté C-terminal d'une longue boucle flexible. Ainsi, il a été suggéré que le remplacement de cet acide aminé par une Cys pourrait affecter la flexibilité de cette boucle et par conséquent, altérer l'organisation du site de fixation du substrat accepteur. Néanmoins, la comparaison des structures secondaires



des  $\beta$ 4GalT7s humaine et de drosophile montre que chacune de ces deux enzymes présente des différences importantes au niveau de la région C-terminale contenant l'Arg270. Ces constatations renforcent l'importance de l'élucidation de la structure 3D de l'enzyme  $\beta$ 4GalT7 humaine en complexe avec le substrat accepteur.

J'ai participé au cours de la dernière partie de ce travail de thèse à l'étude de la régulation des voies de biosynthèse des chaînes de GAGs. Plus précisément, nous avons voulu déterminer si un mécanisme épigénétique, par méthylation d'ADN, intervient dans la régulation de la synthèse des GAGs et/ou dans leur modification par sulfatation. Simultanément, nous avons cherché à savoir dans quelle mesure un tel mécanisme pourrait concourir à l'altération du profil de sulfatation des HS et du phénotype chondrocytaire dans le chondrosarcome humain, en utilisant les cellules HEMC comme modèle de tumeur maligne du cartilage.

Dans une première approche, nous avons mis en évidence la présence de régions riches en CpG au niveau de la séquence promotrice des gènes codant les GTs impliquées dans la formation de l'amorce tétrasaccharidique commune aux CS/DS et Hp/HS. L'établissement du profil de méthylation a montré que ces gènes sont peu méthylés, indiquant que l'expression des quatre GTs de l'amorce tétrasaccharidique ne semble pas soumise à un mécanisme de régulation épigénétique impliquant la méthylation de l'ADN dans les cellules de chondrosarcome humain HEMC. L'analyse du profil de méthylation du gène EXT1 écarte l'hypothèse selon laquelle la polymérisation des chaînes d'HS serait contrôlée par ce processus de régulation épigénétique. Par conséquent, dans les cellules de chondrosarcome humain, la méthylation de l'ADN ne serait impliquée ni dans la régulation de la formation de l'amorce tétrasaccharidique ni dans la polymérisation des chaînes d'HS. Ces résultats sont concordants avec la littérature (Hameetman *et al.*, 2007). En effet, d'autres groupes montrent qu'aucun phénomène de méthylation aberrante du gène EXT1 ne semble être associé aux chondrosarcomes ni aux ostéosarcomes (Ropero *et al.*, 2004; Tsuchiya *et al.*, 2005).

Par ailleurs, de nombreuses études ont démontré que des modifications de la sulfatation des chaînes d'HS sont impliquées dans l'altération des fonctions et du comportement cellulaire caractéristiques des cellules cancéreuses. Pour cette raison, nous avons voulu déterminer si les gènes impliqués dans la sulfatation des chaînes HS sont contrôlés par un mécanisme de régulation épigénétique impliquant une méthylation de l'ADN, dans les cellules HEMC. De façon importante, l'analyse de la carte de méthylation de la famille des NDSTs et des OSTs a montré que les gènes codant pour la 3-OST1, 3-OST3A et 3-OST6 sont hyperméthylés, ainsi

que les gènes codant pour la NDST1, l'HS6ST2 et l'HS6ST3. En revanche, la 3-OST2 présente une méthylation plus modérée au niveau de la région proximale du promoteur. L'utilisation de la 5-Aza-dc, agent inhibiteur des DNMTs, a permis de confirmer que la 3-O-sulfatation des HSPGs dans les cellules HEMC est régulée par un mécanisme épigénétique, *via* la méthylation de l'ADN. Plus précisément, le traitement de ces cellules à la 5-Aza-dc, induit une forte augmentation du taux d'expression de l'ARNm de la 3-OST3A et une augmentation plus modérée du niveau des ARNm de la 3-OST1 et la 3-OST2. Parallèlement, l'analyse de la carte de méthylation de la 3-OST3A dans des cellules HEMC après traitement à la 5-Aza-dc a montré une diminution significative du nombre de sites CpG méthylées, suggérant que l'augmentation de la transcription du gène codant la 3-OST3A est due, au moins en partie, à la déméthylation de ce gène. D'une façon complémentaire, le traitement des cellules HEMC à la 5-Aza-dc ne semble pas affecter la transcription des gènes codant pour la 3-OST3B, HS2ST1 et HS6ST1, en accord avec le faible taux de méthylation des îlots CpG de ces gènes. Quant aux gènes codant pour la NDST1 et HS6ST2, ils sont sensibles à l'effet de l'inhibiteur de la méthylation de l'ADN. En effet, le traitement des cellules HEMC à la 5-Aza-dc augmente le taux de l'ARNm de façon significative, suggérant que ces gènes sont soumis à un processus de régulation épigénétique. Néanmoins, une forte méthylation de la NDST1 et HS6ST2 a été également observée dans des cellules sanguines, utilisées comme contrôle dans cette étude. Par conséquent, l'hyperméthylation des gènes codant pour ces STs ne semble pas spécifique du chondrosarcome humain.

A la lumière de ces résultats, nous avons analysé les conséquences fonctionnelles de la régulation de la 3-O-sulfatation (plus particulièrement les 3-OST1 et -3A) sur les HSPGs, par un mécanisme épigénétique, dans le chondrosarcome humain. Dans cet objectif, nous avons procédé dans un premier temps par traitement des cellules HEMC avec la 5-Aza-dc, suivi d'un radiomarquage  $\text{Na}_2[^{35}\text{S}]\text{SO}_4$ ,  $[^3\text{H}]\text{GlcN}$  et  $[^{35}\text{S}]\text{Cys-Met}$ , combiné à des analyses biochimiques des PGs cellulaires par différentes techniques chromatographiques. Cette approche nous a permis de démontrer que la déméthylation du gène codant pour la 3-OST3A induit une augmentation de l'expression de ce gène, qui à son tour provoque une augmentation du taux de sulfatation des chaînes d'HS. Nous avons démontré également que la longueur des chaînes HS reste inchangée. De façon complémentaire, les analyses d'immunomarquage en fluorescence indirecte montrent une surexpression des HSPGs cellulaires dans les cellules HEMC traitées avec la 5-Aza-dc et dans les cellules exprimant la 3-OST1 ou la 3-OST3A recombinantes.

Enfin, la dernière partie de ce projet a été consacrée à la détermination des conséquences de l'augmentation de la 3-O-sulfatation des HSPGs cellulaires, sur le comportement des cellules HEMC, notamment en termes de prolifération, d'adhésion et de migration cellulaires. En effet, il est établi que les régulations épigénétiques jouent un rôle clé dans les processus cellulaires liés à la cancérogenèse. Nous avons donc déterminé l'impact du traitement des cellules HEMC à la 5-Aza-dc sur la prolifération cellulaire et le résultat obtenu a été comparé au taux de prolifération des cellules HEMC surexprimant la 3-OST1 ou la 3-OST3A. Nous avons procédé de même pour déterminer l'effet du traitement avec l'inhibiteur de méthylation de l'ADN, ainsi que l'effet de la surexpression de ces deux 3-OSTs sur l'adhésion et la migration cellulaires. De façon intéressante, nous avons montré par cette approche que les cellules traitées avec la 5-Aza-dc ou transfectées avec le vecteur pcDNA-3OST3A présentent une diminution significative du taux de la prolifération et une augmentation importante de la capacité de ces cellules à adhérer à un support traité avec la fibronectine. Ce résultat suggère un rôle important de cette modification des chaînes de GAGs, qui pourrait être associée à l'altération des propriétés cellulaires constatée dans certains types de cancers (Sanderson, 2001). Enfin, les tests de migration cellulaire, *in vitro*, ont montré que le traitement des cellules HEMC à la 5-Aza-dc ainsi que la surexpression de la 3-OST1 et 3-OST3A, inhibent la migration de ces cellules d'une manière significative.

L'ensemble de ces résultats suggère que, dans le modèle étudié de chondrosarcome humain, le profil de sulfatation des PGs, et plus particulièrement celui des HSPGs cellulaires, constituent un élément clé dans le contrôle des propriétés cellulaires notamment la prolifération, l'adhésion et la migration cellulaires. De façon importante, la 3-OST1 et la 3-OST3A, qui sont deux enzymes impliquées dans la modification des chaînes de GAGs par 3-O-sulfatation, semblent être soumises à un processus de régulation épigénétique dans ces cellules HEMC.

*Conclusions et  
perspectives*

## Conclusions et perspectives

Les GAGs sont des macromolécules complexes, impliquées dans de nombreux processus biologiques. Leur biosynthèse nécessite l'intervention de plusieurs familles d'enzymes qui assurent la formation de l'amorce tétrasaccharidique commune aux CS/DS et Hp/HS, l'élongation et la modification des chaînes de GAGs. Du fait de leur rôle clé dans de multiples processus physiologiques et pathologiques, l'intérêt suscité par ces molécules ne cesse de croître. En effet, l'altération de la biosynthèse des chaînes de GAGs constitue un événement majeur dans le développement de certaines situations pathologiques, notamment au cours des arthropathies. De plus, la perturbation de la biosynthèse de ces molécules a été constatée aussi bien dans des pathologies sporadiques que dans des maladies génétiques. Le syndrome progéroïde d'Ehlers-Danlos est caractérisé par une diminution de la formation des chaînes glycosaminoglycaniques. Celle-ci a été étudiée en particulier sur la protéine "core" des SLRPs, et plus particulièrement la décorine, au niveau des tissus conjonctifs. A l'heure actuelle, il est clair que la compréhension des mécanismes de biosynthèse et de modification des GAGs est primordiale. En effet, malgré l'identification des enzymes impliquées dans cette machinerie cellulaire, nos connaissances sur leur organisation structurale et fonctionnelle et leur régulation demeurent limitées.

Dans ce contexte, la première partie de ma thèse a été consacrée à la caractérisation structurale et fonctionnelle de la  $\beta$ 4GalT7 humaine. En effet, cette enzyme intervient à un stade précoce de l'initiation de la formation de l'amorce tétrasaccharidique et sa présence est indispensable à la fois pour la synthèse des chaînes de CS/DS et d'Hp/HS. De plus, il est établi que la forme progéroïde du syndrome d'Ehlers-Danlos est une conséquence directe de mutations du gène codant pour la  $\beta$ 4GalT7 chez l'homme. Pour atteindre les objectifs fixés, nous avons réalisé une série de mutations d'acides aminés potentiellement impliqués dans la reconnaissance des substrats et/ou la catalyse enzymatique. L'identification de ces résidus est basée sur l'établissement d'un arbre phylogénétique et la réalisation d'un alignement de séquences des  $\beta$ 4GalTs. Le développement d'un modèle moléculaire de la  $\beta$ 4GalT7 humaine, basé sur la structure cristallographique de la  $\beta$ 4GalT1 bovine, puis de la  $\beta$ 4GalT7 de drosophile, nous a aidé à mieux comprendre le rôle de chaque acide aminé étudié dans le cadre de ce projet.

La caractérisation *in vitro* des conséquences des mutations sur l'activité enzymatique et sur les constantes cinétiques nous a permis d'identifier des acides aminés qui jouent un rôle clé

dans le déroulement de la réaction enzymatique catalysée par la  $\beta$ 4GalT7 humaine. Nous avons démontré l'importance fonctionnelle des deux résidus Asp du motif <sup>163</sup>DVD<sup>165</sup>. En effet, mes résultats suggèrent qu'ils seraient impliqués, respectivement, dans la fixation de l'UDP-Gal et la formation de liaisons de coordination avec le cofacteur Mn<sup>2+</sup>. L'une des principales découvertes de cette étude est l'identification du Trp224 impliqué à la fois dans des interactions avec le substrat donneur et accepteur. De façon intéressante, nous avons pu moduler l'affinité, la spécificité de la  $\beta$ 4GalT7 et sa capacité à initier la synthèse des GAGs *ex vivo* en générant un mutant de cette enzyme par substitution du Trp224 en His.

D'après la littérature et la structure 3D de la  $\beta$ 4GalT7 de drosophile, l'Asp228 constituerait vraisemblablement la base catalytique de l'isoforme humaine. Néanmoins, les approches expérimentales mises en œuvre n'ont pas complètement permis d'éliminer un rôle possible du Glu227. Par conséquent, des investigations supplémentaires, notamment la cristallisation de cette protéine en présence du substrat accepteur, demeurent nécessaires pour la détermination du rôle de chacun de ces deux acides aminés. Ces études sont en cours.

Parallèlement aux études *in vitro*, nous avons développé une approche pluridisciplinaire qui nous a permis de déterminer les conséquences fonctionnelles de la mutation de la  $\beta$ 4GalT7 sur la capacité de l'enzyme à restaurer dans des cellules déficientes CHO-pgsB618, la synthèse des GAGs endogènes totaux ou à partir de la protéine "core" de la décorine, ou encore à partir d'un substrat xylosidique exogène, le 4-MOX. Les résultats obtenus ont corroboré, en général, les approches *in vitro*, soulignant en particulier le rôle crucial du Trp224 de la  $\beta$ 4GalT7 humaine.

Dans la deuxième partie de l'étude structurale et fonctionnelle de cette enzyme, nous nous sommes intéressés à la caractérisation biochimique et moléculaire des mutations de la  $\beta$ 4GalT7, associées au syndrome progéroïde d'Ehlers-Danlos chez l'homme. j'ai, en premier lieu, analysé les conséquences de chacune des mutations (A186D, L206P et R270C) sur l'activité de l'enzyme et sur ses constantes cinétiques. Cette étude a été associée à une caractérisation *ex vivo* de la restauration de la synthèse des chaînes de GAGs dans des cellules déficientes CHO-pgsB618.

Cette approche nous a permis de montrer que la mutation A186D de la  $\beta$ 4GalT7 ne serait pas la principale cause de l'altération de la biosynthèse des chaînes glycosaminoglycaniques, constatée chez les patients atteints du syndrome progéroïde d'Ehlers-Danlos qui sont porteurs de la double mutation A186D/L206P. En revanche, la mutation L206P a des effets très délétères sur l'activité enzymatique évaluée *in vitro* et *ex vivo* et cela est vraisemblablement

dû à une altération majeure de la conformation de la protéine. Enfin, nous avons également démontré que la mutation R270C diminue de façon importante l'affinité de l'enzyme envers le substrat accepteur 4-MOX.

L'ensemble des résultats obtenus au cours de cette étude constitue une avancée importante dans la compréhension des mécanismes de fonctionnement de la  $\beta$ 4GalT7 humaine, notamment, l'identification d'acides aminés impliqués dans l'organisation du site actif de l'enzyme et dans la reconnaissance du substrat donneur et/ou accepteur et de la base catalytique.

Outre son intervention à un stade précoce de l'initiation de la synthèse des GAGs, la  $\beta$ 4GalT7 est une enzyme qui a la particularité de pouvoir utiliser des substrats synthétiques exogènes, pour initier la biosynthèse des GAGs. Ainsi, l'élucidation de son organisation structurale et fonctionnelle devrait permettre de développer des molécules glycomimétiques, qui seront destinées au traitement des pathologies associées à la perturbation de la synthèse des chaînes de GAGs.

La  $\beta$ 4GalT7 est une cible pharmacologique potentielle et l'élucidation de sa structure 3D devrait constituer un grand pas vers le développement de thérapies originales, basées sur le transfert de gène et sur l'addition de molécules glycomimétiques. C'est pourquoi, ayant développé un système performant d'expression dans la bactérie, nous avons lancé les premières tentatives de cristallisation de la  $\beta$ 4GalT7 humaine. De plus, cette approche est complétée par une caractérisation physico-chimique des interactions enzyme-substrat, basée sur l'utilisation de la résonance plasmonique de surface, sur Biacore. Ce travail sera poursuivi par des collaborations avec mon équipe.

Dans une partie complémentaire de mon travail de thèse, j'ai contribué à une étude originale qui a consisté à caractériser les mécanismes de régulation de l'expression des enzymes impliquées dans la biosynthèse des GAGs. Nous nous sommes intéressés, plus particulièrement, à l'étude d'un modèle cellulaire de chondrosarcome humain. La cartographie de méthylation des régions promotrices des gènes codant les enzymes intervenant dans la synthèse et la modification, par N- et O-sulfatation des chaînes de HS, nous a permis de montrer pour la première fois une hyperméthylation des îlots CpG situés dans la région promotrice proximale spécifique de la famille des 3-OSTs dans les cellules HEMC. Cette méthylation serait, très vraisemblablement, la cause de la répression de ces gènes dans la lignée HEMC du chondrosarcome humain, qui se manifeste par un faible taux d'expression des ARNm dans ces cellules.

L'extinction de l'expression des gènes hyperméthylés serait liée au caractère invasif de ces cellules cancéreuses. En effet, le traitement de ces cellules à la 5-Aza-dc a permis de corriger le comportement cellulaire des cellules HEMC. De façon intéressante, l'analyse de l'effet de la surexpression de la 3-OST1 et la 3-OST3A nous a permis de montrer que l'hyperméthylation de ces gènes contribue au phénotype invasif de ce chondrosarcome. Notre étude suggère que la diminution de la sulfatation, en position 3, des HSPGs cellulaires est liée à l'altération des propriétés cellulaires des HEMC et au développement du caractère invasif. Notre investigation confirme également que les gènes responsables de la 3-O-sulfatation seraient soumis à un mécanisme de régulation épigénétique, *via* la méthylation de l'ADN.

Ce processus de contrôle de l'expression des gènes par la méthylation de l'ADN est largement identifié dans différents types de cancers. De façon intéressante, ce mécanisme de régulation de l'expression est associé au développement d'une résistance des cellules cancéreuses à la chimiothérapie, ainsi qu'à un pronostic défavorable. Cependant, l'altération de la méthylation de l'ADN est un phénomène potentiellement réversible. Ainsi, il a été possible au cours des dernières années de développer des approches thérapeutiques dans le traitement des cancers, par l'utilisation de molécules médicamenteuses capables d'induire la déméthylation de ces gènes. De nombreuses études sont actuellement au stade préclinique ou clinique du développement du médicament. Des agents inhibiteurs des DNMTs ont donné des résultats satisfaisants dans le traitement d'hémopathies, notamment dans le cas de la leucémie myéloïde ainsi que dans d'autres types de cancers. Le chondrosarcome reste un cancer particulièrement difficile à diagnostiquer et à traiter, notre étude constitue une avancée importante dans la compréhension des mécanismes de régulation par méthylation de l'ADN, de l'expression des gènes responsables de la sulfatation des chaînes d'HS cellulaires et ouvre des perspectives dans la prise en charge des tumeurs malignes du cartilage.



*Références  
bibliographiques*

## Références bibliographiques

- Abe, N., Muragaki, Y., Yoshioka, H., Inoue, H. and Ninomiya, Y.** (1993). Identification of a novel collagen chain represented by extensive interruptions in the triple-helical region. *Biochem Biophys Res Commun.* **196**, 576-582.
- Aikawa, J., Grobe, K., Tsujimoto, M. and Esko, J. D.** (2001). Multiple isozymes of heparan sulfate/heparin GlcNAc N-deacetylase/GlcN N-sulfotransferase. Structure and activity of the fourth member, NDST4. *J Biol Chem.* **276**, 5876-5882.
- Almeida, R., Amado, M., Davis, L., Levery, S. B., Holmes, E.H., Merks, G., Van Kessel, A.G., Rygard, E., Hassan, H., Bennett, E. and Claussen, H.** (1997). A family of human beta-4-galactosyltransferase. Cloning and expression of two novel UDP-galactose: beta-N-acetylglucosamine beta1,4-galactosyltransferase, beta4Gal-T2 and beta4Gal-T3. *J Biom Chem.* **272**, 31979-31991.
- Almeida, R., Levery, S. B., Mandel, U., Kresse, H., Schwientek, T., Bennett, E.P. and Clausen, H.** (1999). Cloning and expression of a proteoglycan UDP-galactose: beta-xylose beta1,4-galactosyltransferase I. A seventh member of the human beta4-galactosyltransferase gene family. *J Biol Chem.* **274**, 26165-26171.
- Altinbas, M., Coskun, H.S., Er, O., Ozkan, M., Eser, B., Unal, A., Cetin, M. and Soyuer, S.** (2004). A randomized clinical trial of combination chemotherapy with and without low-molecularweight heparin in small cell lung cancer. *J Thromb Haemost.* **2**, 1266-71.
- Amado, M., Almeida, R., Schwientek, T. and Clausen, H.** (1999). Identification and characterization of large galactosyltransferase gene families: galactosyltransferases for all functions. *Biochim Biophys Acta.* **1473**, 35-53.
- Antonsson, P., Heinegard, D., and Oldberg, A.** (1991). Posttranslational modifications of fibromodulin. *J Biol Chem.* **266**, 16859-16861.
- Aviezer, D., Safran, M. and Yayon, A.** (1999). Heparin differentially regulates the interaction of fibroblast growth factor-4 with FGF receptors 1 and 2. *Biochem Biophys. Res Commun.* **263**, 621- 626.
- Awai, K., Maréchal, E., Block, M.A., Brun, D., Masuda, T., Shimada, H., Takamiya, K., Ohta, H. and Joyard, J.** (2001). Two types of MGDG synthase genes, found widely in both 16:3 and 18:3 plants, differentially mediate galactolipid syntheses in photosynthetic and nonphotosynthetic tissues in *Arabidopsis thaliana*. *Proc Natl Acad Sci.* **98**, 10960-5.
- Baffi, M.O., Moran, M.A., and Serra, R.** (2006). Tgfbr2 regulates the maintenance of boundaries in the axial skeleton. *Dev Biol.* **296**, 363-374.
- Baker, J.R., Cifonelli, J.A. and Roden, L.** (1975). The linkage of corneal keratan sulfate to protein. *Connect Tissue Res.* **3**, 149-156.

- Bandtlow, C.E. and Zimmermann, D.R.** (2000). Proteoglycans in the Developing Brain: New Conceptual Insights for Old Proteins. *Physiol Rev.* **80**, 1267-90.
- Bansal, S.K. and Gessner, T.** (1980). A unified method for the assay of uridine diphosphoglucuronyl-transferase activities toward various aglycones using uridine diphospho[U-<sup>14</sup>C] glucuronic acid. *Anal Biochem.* **109**, 321-329.
- Baumann, U., Bauer, M., Letoffe, S., Delepelaire, P., and Wandersman, C.** (1995). Crystal structure of a complex between *Serratia marcescens* metallo-protease and an inhibitor from *Erwinia chrysanthemi*. *J Mol Biol.* **248**, 653-661.
- Bayer-Garner, I.B., Sanderson, R.D., Dhodapkar, M.V., Owens, R.B. and Wilson, C.S.** (2001). Syndecan-1 (CD138) immunoreactivity in bone marrow biopsies of multiple myeloma: shed syndecan-1 accumulates in fibrotic regions. *Mod Pathol.* **14**, 1052-1058.
- Béchar, D.T., Gentina, Delehedde, M., Scherpereel, A., Lyon, M., Aumercier, M., Vazeux, R., Richet, C., Degand, P., Brigitte Jude, B., Janin, A., David G. Fernig, Tonnel, A.B., and Lassalle. P.** (2001). Endocan is a novel chondroitin sulfate/dermatan sulfate proteoglycan that promotes hepatocyte growth factor/scatter factor mitogenic activity. *J Biol Chem.* **276**, 48341-483419.
- Bech-Hansen, N.T., Naylor, M.J., Maybaum, T.A., Sparkes, R.L., Koop, B., Birch, D.G., Bergen, A.A., Prinsen, C.F., Polomeno, R.C., Gal, A., Drack, A.V., Musarella, M.A., Jacobson, S.G., Young, R.S., and Weleber, R.G.** (2000). Mutations in NYX, encoding the leucine-rich proteoglycan nyctalopin, cause X-linked complete congenital stationary night blindness. *Nat Genet.* **26**, 319-323.
- Belting, M., Borsig, L., Fuster, M.M., Brown, J.R., Persson, L., Fransson, L.A., and Esko, J.D.** (2002). *Proc Natl Acad Sci USA.* **99**, 371-376.
- Bennett, K.L., Modrell, B., Greenfield, B., Bartolazzi, A., Stamenkovic, I., Peach, R., Jackson, D.G., Spring, F., Aruffo, A.** (1995). Regulation of CD44 binding to hyaluronan by glycosylation of variably spliced exons. *J Cell Biol.* **131**, 1623-33.
- Bernfield, M., Götte, M., Park, P.W., Reizes, O., Fitzgerald, M.L., Linceum, J., Zako, M.** (1999). Functions of cell surface heparan sulfate proteoglycans. *Annu Rev Biochem.* **68**, 729-777.
- Bertozi, C.R. and Kiessling, L.L.** (2001). Chemical glycobiology. *Science.* **291**, 2357-2364.
- Biglari, A., Bataille, D., Naumann, U., Weller, M., Zirger, J., Castro, M.G. and Lowenstein, P. R.** (2004). Effects of ectopic decorin in modulating intracranial glioma progression in vivo, in a rat syngeneic model. *Cancer Gene Ther.* **11**, 721-732.
- Binari, R.C., Staveley, B.E., Johnson, W.A., Godvarti, R., Sasisekharan, R., and Manoukian, A.S.** (1997). Genetic evidence that heparin-like glycosaminoglycans are involved in wingless signaling. *Development.* **124**, 2623-2632.

**Blochberger, T.C., Vergnes, J.P., Hempel, J. and Hassell, J.R.** (1992). cDNA to chick lumican (corneal keratan sulfate proteoglycan) reveals homology to the small interstitial proteoglycan gene family and expression in muscle and intestine. *J Biol Chem.* **267**, 347-352.

**Boeggman, E. and Qasba, P.K.** (2002). Structure on the metal binding sites in the catalytic domain of  $\beta$ 1,4-galactosyltransferase. *Glycobiology.* **12**, 395-407.

**Borczuk, A.C., Shah, L., Pearson, G.D., Walter, K.L., Wang, L., Austin, J.H., Friedman, R.A. and Powell, C.A.** (2004). Molecular signatures in biopsy specimens of lung cancer. *Am J Respir Crit Care Med.* **170**, 167-174.

**Bourne, Y. and Henrissat, B.** (2001). Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr Opin Struct Biol.* **11**, 593-600.

**Bradford, M.M.** (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* **72**, 248-254.

**Breton, C., Bettler, E., Joziase, D.H., Geremia, R.A. and Imberty, A.** (1998). Sequence-function relationships of prokaryotic and eukaryotic galactosyltransferases. *J Biochem.* **123**, 1000-9.

**Breton, C., Heissigerová, H., Jeanneau, C., Moravcová, J. and Imberty, A.** (2002). Comparative aspects of glycosyltransferases. *Biochem Soc Symp.* **69**, 23-32.

**Breton, C. and Imberty, A.** (1999). Structure-function studies of glycosyltransferases. *Curr Opin Struct Biol.* **9**, 563-571.

**Breton, C., Snajdrová, L., Jeanneau, C., Koca, J. and Imberty, A.** (2006). Structures and mechanisms of glycosyltransferases. *Glycobiology.* **16**, 29-37.

**Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreulx, D., Cambillau, C. and Bourne, Y.** (1999). Crystal structure of the bifunctional N-acetylglucosamine 1-phosphate uridylyltransferase from *Escherichia coli*: a paradigm for the related pyrophosphorylase superfamily. *EMBO J.* **18**, 4096-4107.

**Campbell, J.A., Davies, G.J., Bulone, V. and Henrissat, B.** (1997). A classification of nucleotide-diphosphosugar glycosyltransferases based on amino acid sequence similarities. *Biochem J.* **326**, 929-939.

**Campbell, J.A., Davies, G.J., Bulone, V.V. and Henrissat, B.** (1998). A classification of nucleotidediphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* **329**, 719.

**Campbell, R.E., Mosimann, S.C., Tanner, M.E. and Strynadka, N.C.** (2000). The structure of UDP-N-acetylglucosamine 2-epimerase reveals homology to phosphoglycosyl transferases. *Biochemistry.* **39**, 14993-5001.

- Cantor, J.O., Shteyngart, B., Cerreta, J.M., Liu, M., Armand, G. and Turino, G.M.** (2000). The effect of hyaluronan on elastic fiber injury in vitro and elastase-induced airspace enlargement *in vivo*. *Proc Soc Exp Biol Med.* **225**, 65-71.
- Carey, D.J., Evans, D.M., Stahl, R.C., Asundi, V.K., Conner, K.J., Carbes, P. and Cizmeci-Smith, G.** (1992). Molecular cloning and characterization of *N*-syndecan, a novel transmembrane heparan sulfate proteoglycan. *J Cell Biol.* **117**, 191-192.
- Carey, D.J.** (1997). Syndecans: multifunctional cell-surface coreceptors. *Biochem J.* **327**, 1-16.
- Cassaro, C.M. and Dietrich, C.P.** (1977). Distribution of sulfated mucopolysaccharides in invertebrates. *J Biol Chem.* **252**, 2254-2261.
- Chappell, T., Nasser Hajibagheri, M., Ayscough, K., Pierce, M., and Warren, G.** (1994). Localization of an  $\alpha$  1,2 galactosyltransferase activity to the Golgi apparatus of *Schizosaccharomyces pombe*. *Mol Biol Cell.* **5**, 519-528.
- Christopher, D.O. and Deepak, S.** (2008). The Importance of Heparan Sulfate in Herpesvirus Infection. *Virol Sin.* **23**, 383-393.
- Clark, E.A., Golub, T.R., Lander, E.S., and Hynes, R.O.** (2000). Genomic analysis of metastasis reveals an essential role for RhoC. *Nature.* **406**, 532-535.
- Clausen, H. and Bennett, A.** (1996). A family of UDP-GalNAc: polypeptide N-acetyl galactosaminyl-transferases control the initiation of mucin-type O-linked glycosylation. *Glycobiology.* **6**, 635-646.
- Cobbold S.** (1987). Human leukocyte differentiation antigens: monoclonal antibody computer databases as a tool for the future. *Mol Cell Probes.* **1**, 61-72.
- Couchman, J.R. and Woods A.** (1993). Structure and biology of pericellular proteoglycans. *In Cell Surface and Extracellular Glycoconjugates: Structure and Function.* D. D. Roberts and R. P. Mecham, editors. Academic Press, San Diego. 33-82.
- Couchman, J.R. and Woods, A.** (1996). Syndecans, signaling, and cell adhesion. *J Cell Physiol.* **61**, 578-584.
- Couchman, J.R., Kapoor, M. Sthanam, and Wu, R.R.** (1996). Perlecan and BM-CSPG (Bamacan) are two basement membrane chondroitin/ dermatan sulfate proteoglycans in the Engelbreth-Holm-Swarm tumor matrix. *J Biol Chem.* **271**, 9595-9602.
- Couchman, J.R.** (2003). Syndecans: proteoglycan regulators of cell surface microdomains? *Nat Rev Mol Cell Biol.* **4**, 926-937.
- Coutinho, P.M., Deleury, E., Davies, G.J. and Henrissat, B.** (2003). An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol.* **328**, 307-17.

- Dai, Y., Yang, Y., MacLeod, V., Yue, X., Rapraeger, A.C., Shriver, Z., Venkataraman, G., Sasisekharan, R. and Sanderson, R.D.** (2005). HSulf-1 and HSulf-2 are potent inhibitors of myeloma tumor growth in vivo. *J Biol Chem.* **280**, 40066-40073.
- Danielson, K.G., Baribault, H., Holmes, D.F., Graham, H., Kadler, K.E., and Iozzo, R.V.** (1997). Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. *J Cell Biol.* **136**, 729-743.
- David, G., van der Schueren, B., Marynen, P., Cassiman, J.J. and van den Berghe, H.** (1992). Molecular cloning of amphiglycan, a novel integral membrane heparan sulfate proteoglycan expressed by epithelial and fibroblastic cells. *J Cell Biol.* **118**, 961-969.
- David, G.** (1993). Integral membrane heparan sulfate proteoglycans. *FASEB J.* **7**, 1023-1030.
- De Medeiros Matsushita, M., da Silva, L.F., dos Santos, M.A., Fernezlian, S., Schrupf, J.A., Roughley, P., Hiemstra, P.S., Saldiva, P.H., Mauad, T. and Dolhnikoff, M.** (2005). Airway proteoglycans are differentially altered in fatal asthma. *J Pathol.* **207**, 102-110.
- Depontieu, F., Grigoriu, B.D., Scherpereel, A., Adam, E., Delehedde, M., Gosset, P., and Lassalle, P.** (2008). Loss of Endocan tumorigenic properties after alternative splicing of exon 2. *BMC Cancer.* **8**, 1471-2407.
- Doerge, K.J., Coulter, S.N., Meek, L.M., Maslen, K. and Wood, J.G.** (1997). A humanspecific polymorphism in the coding region of the aggrecan gene. Variable number of tandem repeats produce a range of core protein sizes in the general population. *J Biol Chem.* **271**, 13974-13979.
- Donahue, J.E., Berzin, T.M., Rafii, M. S., Glass, D.J., Yancopoulos, G.D., Fallon, J.R., and Stopa, E.G.** (1999). Agrin in Alzheimer's disease: Altered solubility and abnormal distribution within microvasculature and brain parenchyma. *Proc Natl Acad Sci.* **96**, 6468-6472.
- Duenas Gonzales, A., Kaya, M., Shi, W., Song, H., Testa, J.R., Penn, L.Z., and Filmus, J.** (1998) OCI-5/GPC3, a glypican encoded by a gene that is mutated in the Simpson-Golabi-Behmel overgrowth syndrome, induces apoptosis in a cell line-specific manner. *J Cell Biol.* **141**, 1407-1414.
- Edelgard, Q., Achim, G., Lennart, R. and Hans, K.** (1990). 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.* **87**, 1342-1346.
- Edwards, M.E., Dickson, C.A., Chengappa, S., Sidebottom, C., Gidley, M.J. and Reid, J.S.G.** (1999). Molecular characterisation of a membrane-bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis. *Plant J.* **19**, 691-697.

- Ergun, S., Kilic, N., Wurmbach, J.H., Ebrahimnejad, A., Fernando, M., Sevinc, S., Kilic, E., Chalajour, F., Fiedler, W., Lauke, H. Lamszus, K., Hammerer, P., Weil, J., Herbst, H. and Folkman, J.** (2001). Endostatin inhibits angiogenesis by stabilization of newly formed endothelial tubes. *Angiogenesis* **4**, 193-206.
- Erickson, A.C. and Couchman, J.R.** (2000). Still more complexity in mammalian basement membranes. *J Histochem Cytochem.* **48**, 1291-1306.
- Esko, J.D., Stewart, T.E. and Taylor, W.H.** (1985). Animal cell mutants defective in glycosaminoglycan biosynthesis. *Proc Nat Acad Sci USA.* **82**, 3197-3201.
- Esko, J.D., Weinke, J.L., Taylor, W.H., Ekborg, G., Roden, L., Anantharamaiah, G. and Gawish, A.** (1987). Inhibition of chondroitin and heparan sulfate biosynthesis in Chinese hamster ovary cell mutants defective in galactosyltransferase I. *J Biol Chem.* **262**, 12189-12195.
- Esko, J.D. and Selleck, S.B.** (2002). Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu Rev Biochem.* **71**, 435-471.
- Ezura, Y., Chakravarti, S., Oldberg, A., Chervoneva, I. and Birk, D.E.** (2000). Differential expression of lumican and fibromodulin regulate collagen fibrillogenesis in developing mouse tendons. *J Cell Biol.* **151**,779-787.
- Faiyaz-Ul-Haque, M., Zaidi, S.H., Al Ali, M., Al Mureikhi, M.S., Kennedy, S., Al Thani, G., Tsui, L.C. and Teebi, A.** (2004). A novel missense mutation in the galactosyltransferase-I (B4GALT7) gene in a family exhibiting facioskeletal anomalies and Ehlers-Danlos syndrome resembling the progeroid type. *Am J Med Genet.* **128A**, 39-45.
- Ferro, D.R., Provasoli, A., Ragazzi, M., Casu, B., Torri, G., Bossennec, V., Perly, B., Sinay, P., Petitou, M., and Choay, J.** (1990). Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences *Carbohydr Res.* **195**, 157-167.
- Filla, M.S., Dam, P. and Rapraeger, A. C.** (1998). The cell surface proteoglycan syndecan-1 mediates fibroblast growth factor-2 binding and activity. *J Cell Physiol.* **174**, 310-321.
- Filmus, J.** (2000). Glypicans in growth control and control. *Glycobiology.***11**, 19-23.
- Filmus, J. and Selleck, S.B.** (2001). Glypicans: Proteoglycans with a surprise. *J Clin Invest.* **108**, 497 -501.
- Fine, J.D. and J.R. Couchman.** (1989). Chondroitin 6-sulfate proteoglycan but not heparan sulfate proteoglycan is abnormally expressed in skin basement membrane from patients with dominant and recessive dystrophic epidermolysis bullosa. *J Invest Dermatol.* **92**,611-616.
- Fisher, L.W., Termine, J.D. and Young, M.F.** (1989) Deduced protein sequence of bone small proteglycan I (biglycan) shows homolog with proteoglycan II (decorin) and several nonconnective tissue proteins in variety of spices. *J Biol Chem.* **264**, 4571-4576.

- Font, B., Eichenberger, D., Goldschmidt, D., Boutillon, M.M., and Hulmes, D.I.** (1998). Structural requirements for fibromodulin binding to collagen and the control of type I collagen fibrillogenesis—critical roles for disulphide bonding and the C-terminal region. *Eur J Biochem.* **254**, 580-587.
- Fransson, L.Å., Silverberg, I. and Carlstedt, I.** (1985). Structure of the heparan sulfate-protein linkage region. Demonstration of the sequence galactosyl-galactosyl-xylose-2-phosphate. *J Biol Chem.* **260**, 14722- 14726.
- Fraser, J.R., Laurent T.C. and Laurent, U.B.G.** (1997). Hyaluronan: its nature, distribution, functions and turnover. *J Intern Med.* **242**, 27-33.
- Friedrich, M.V., Gohring, W., Morgelin, M., Brancaccio, A., David, G. and Timpl, R.** (1999). Structural basis of glycosaminoglycan modification and of heterotypic interactions of perlecan domain V. *J Mol Biol.* **294**, 259-270.
- Funderburgh, J.L.,** (2000). Keratan sulfate: structure, biosynthesis, and function. *Glycobiology.* **10**, 951-8
- Furukawa, K. and Sato, T.** (1999).  $\beta$ -1,4-Galactosylation of N-glycans is a complex process. *Biochim Biophys Acta.* **1473**, 54-66.
- Gastinel, L.N., Cambillau, C. and Burne, Y.** (1999). Crystal structure of the bovine  $\beta$ 4Galactosyltransferase catalytic domain and its complex with uridine diphosphogalactose. *EMBO Journal.* **18**, 3546-3557.
- Gerdin, B. and Hallgren, R.** (1997). Dynamic role of hyaluronan (HYA) in connective tissue activation and inflammation. *J Intern Med.* **242**, 49-55.
- Glinskii, A., Hoffman, R., Gerald, W., Glinsky, G. and Kimmel, S.** (2003). GeneChip analysis reveals multiple clinically relevant genetic signatures of metastatic prostate cancer. Washington. *Volume 44.* Edited by: Research AAC. American Association for Cancer Research; Abstract Number: LB-106.
- Gonzalez, A.D., Kaya, M., Shi, W., Song, H., Testa, J. R., Penn, L.Z., Filmus, J.** (1998). OCI-5/GPC3, a glypican encoded by a gene that is mutated in the Simpson-Golabi-Behmel overgrowth syndrome, induces apoptosis in a cell line-specific manner. *J Cell Biol.* **141**, 1407-1414.
- Götte, M., Spillmann, D., Yip, G.W., Versteeg, E., Echtermeyer, F.G., van Kuppevelt, T.H. and Kiesel, L.** (2008). Changes in heparan sulfate are associated with delayed wound repair, altered cell migration, adhesion and contractility in the galactosyltransferase I (beta4GalT-7) deficient form of Ehlers-Danlos syndrome. *Hum Mol Genet.* **17**, 996-1009.
- Gould, S.E., Upholt, W.B. and Kosher, R.A.** (1992). Syndecan 3: A member of the syndecan family of membrane-intercalated proteoglycans that is expressed in high amounts at the onset of chicken limb cartilage differentiation. *Proc Nat Acad Sci. USA.* **89**, 3271-3275.



- Green, W.F., Konnaris, K. and Woolcock, A.J.** (1993). Effect of salbutamol, fenoterol, and sodium cromoglycate on the release of heparin from sensitized human lung fragments challenged with *Dermatophagoides pteronyssinus* allergen. *Am J Respir Cell Mol Biol.* **8**, 518-521.
- Grigoriu, B.D., Depontieu, F., Scherpereel, A., Gourcerol, D., Devos, P., Ouatas, T., Lafitte, J.J., Copin, M.C., Tonnel, A.B., Lassalle, P.** (2006). Endocan expression and relationship with survival in human nonsmall cell lung cancer. *Clin Cancer Res.* **12**, 4575-4582.
- Grobe, K., Ledin, J., Ringvall, M., Holmborn, K., Forsberg, E., Esko, J.D. and Kjellen, L.** (2002). Heparan sulfate and development: differential roles of the N-acetylglucosamine N-deacetylase/N-sulfotransferase isozymes. *Biochim Biophys Acta.* **1573**, 209-215.
- Groffen, A.J.A., Buskens, C.A.F, Van Kuppevelt, T.H., Veerkamp, J.H., Monnens, L.A.H. and van den Heuvel, L.P.W.J.** (1998). Primary structure and high expression of human agrin in basement membranes of adult lung and kidney. *Eur J Biochem.* **254**, 123-128.
- Grover, A., Edwards, S.A., Bourdon, M. and Adamson, E.D.** (1987). Proteoglycan-19, laminin and collagen type IV production is correlated with the levels of mRNA in F9 cell aggregates differentiating in the presence or absence of cyclic AMP. *Diff.* **36**, 138-144.
- Gulberti, S., Lattard, V., Fondeur, M., Jacquinet, J.C., Mulliert, G., Netter, P., Magdalou, J., Ouzzine, M. and Fournel-Gigleux, S.** (2005). Phosphorylation and sulfation of oligosaccharide substrates critically influence the activity of human beta1,4-galactosyltransferase 7 (GalT-I) and beta1,3-glucuronosyltransferase I (GlcAT-I) involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J Biol Chem.* **280**, 1417-25.
- Ha, Y.W., Jeon, B.T., Moon, S.H., Toyoda, H., Toida, T., Linhardt, R.J. and Kim, Y.S.** (2005). Characterization of heparan sulfate from the unossified antler of *Cervus elaphus*. *Carbohydr Res.* **340**, 411-416.
- Halfter, W., Dong, S., Schurer, B. and Cole, G.J.** (1998). Collagen XVIII is a basement membrane heparan sulfate proteoglycan. *J Biol Chem.* **273**, 25404-25412.
- Hameetman, L., David, G., Yavas, A., White, S.J., Taminiau, A.H., Cleton-Jansen, A.M., Hogendoorn, P.C. and Bovee, J.V.** (2007). Decreased EXT expression and intracellular accumulation of heparan sulphate proteoglycan in osteochondromas and peripheral chondrosarcomas. *J Pathol.* **211**, 399-409.
- Handel, T.M., Johnson, Z., Crown, S.E., Lau, E.K. and Proudfoot, A.E.** (2005). Regulation of protein function by glycosaminoglycans - as exemplified by chemokines. *Annu Rev Biochem.* **74**, 385-410.

**Hassell, J.R., Robey, P.G., Barrack, H.J., Wilczek, J., Rennard, S.I. and Martin, G.R.** (1980). Isolation of heparan sulfate-containing proteoglycan from basement membrane. *Proc Natl Acad Sci USA*. **77**, 4494-4498.

**Hassell, J.R., Yamada, Y. and Arikawa-Hirasawa, E.** (2003). Role of perlecan in skeletal development and diseases. *Glycoconj J*. **19**, 263-267.

**Hatanaka, Y., Ishiguro, M., Hashimoto, M., Gastinel, L.N. and Nakagomi, K.** (2001). A model of photoprobe docking with beta1, 4-galactosyltransferase identifies a possible carboxylate involved in glycosylation steps. *Bioorg Med Chem Lett*. **11**, 411-413.

**Heathfield, T.F., Onnerfjord, P., Dahlberg, L. and Heinegard, D.** (2004). Cleavage of Fibromodulin in Cartilage Explants Involves Removal of the N-terminal Tyrosine Sulfate-rich Region by Proteolysis at a Site That Is Sensitive to Matrix Metalloproteinase-13. *J Biol Chem*. **279**, 6286-6295.

**Heegaard, A.M., Robey, P.G., Vogel, W., Just, W., Windom, R.L., Schøller, J., Fisher, L.W. and Young M.F.** (1997). Functional characterization of the human biglycan 5'-flanking DNA and binding of the transcription factor c-Krox. *J Bone Miner Res*. **12**, 2050-2060.

**Heinegård, D. and Saxne, T.** (2010). The role of the cartilage matrix in osteoarthritis. *Nat Rev Rheumatol*. doi:10.1038/nrrheum.2010.198.

**Hemp, N., How, T., Cooper, S.J., Green, T.R., Dong, M., Copland, J.A., Wood, C.G. and Blobe, G.C.** (2008). Expression of the type III TGF-beta receptor is negatively regulated by TGF-beta. *Carcinogenesis* **29**, 905-12.

**Hennet, T.** (2002). The galactosyltransferase family. *Cell Mol Life Sci*. **59**, 1081-1095.

**Henrissat, B.** (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*. **280**, 309-316.

**Henrissat, B. and Bairoch, A.** (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*. **293**, 781-788.

**Henrissat, B. and Bairoch, A.** (1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochem J*. **316**, 695-696.

**Henrissat, B. and Davies, G. J.** (1997). Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol*. **7**, 637-644.

**Ho, H-CK., McGrath, K.E., Brodbeck, K.C., Palis, J. and Schick, B.P.** (2001). Serglycin proteoglycan synthesis in the murine uterine decidua and early embryo. *Biol Reprod*. **64**, 1667-1676.

- Hopf, M., Gohring, W., Mann, K. and Timpl, R.** (2001). Mapping of binding sites for nidogens, fibulin-2, fibronectin and heparin to different IG modules of perlecan. *J Mol Biol.* **311**, 529-541.
- Igura, M., Maita, N., Kamishikiryo, J., Yamada, M., Obita, T., Maenaka, K. and Kohda, D.** (2008). Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J.* **27**, 234-43.
- Iozzo, R.V. and Clark, C.C.** (1986). Biosynthesis of proteoglycans by rat embryo parietal yolk sacs in organ culture. *J Biol Chem.* **261**, 6658-6669.
- Iozzo, R.V. and Clark, C.C.** (1987) Chondroitin sulfate proteoglycan is a constituent of the basement membrane in the rat embryo parietal yolk sac. *Histochemistry.* **88**, 23-29.
- Iozzo, R.V.** (1994). Perlecan: a germ of proteoglycan. *Matrix Biol.* **14**, 203-208.
- Iozzo, R.V.** (1998). Matrix proteoglycans: From molecular design to cellular function. *Annu Rev Biochem.* **67**, 609-652.
- Iozzo, R.V. and Murdoch, A.D.** (1996). Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function. *FASEB J.* **10**, 598-614.
- Iozzo, R.V. and San Antonio, J.D.** (2001). Heparan sulfate proteoglycans: heavy hitters in the angiogenesis arena. *J Clin Invest.* **108**, 349-355.
- Jakobsson, L., Kreuger, J., Holmborn, K., Lundin, L., Eriksson, I., Kjellen, L. and Claesson-Welsh, L.** (2006). Heparan sulfate in trans potentiates VEGFR-mediated angiogenesis. *Dev Cell.* **10**, 625-634.
- Jennings, M.P., Virji, M., Evans, D., Forster, V., Shrikanta, Y.N., Steeghs, L., Van Der Ley, P. and Moxon, E.R.** (1998). Identification of a novel gene involved in pilin glycosylation in *Neisseria meningitidis*. *Mol Microbiol.* **29**, 975-984.
- Jepsen, K.J., Wu, F., Peragallo, J.H., Paul, J., Roberts, L., Ezura, Y., Oldberg, A., Birk, D.E. and Chakravarti, S.** (2002). A syndrome of joint laxity and impaired joint integrity in lumican- and fibromodulin-deficient mice. *J Biol Chem.* **277**, 35532-35540.
- Jianing, L. and Guosheng, J.** (2006). CD44 and Hematologic Malignancies. *Cellular & Molecular Immunology.* **3**, 359-365.
- Kallunki, P. and Tryggvason, K.** (1992). Human basement membrane heparan sulfate proteoglycan core protein: a 467-Kda protein containing multiple domains resembling elements of the low density lipoprotein receptor, laminin, neural cell adhesion molecules, and epidermal growth factor. *J Cell Biol.* **116**, 559-571.

- Kikuchi, N., Kwon, Y.D., Gotoh, M., and Narimatsu, H.** (2003). Comparison of glycosyltransferase families using the profile hidden Markov model. *Biochem Biophys Res Commun.* **310**, 574-579.
- Kisilevsky, R., Szarek, W.A., Ancsin, J., Vohra, R., Li, Z. and Marone, S.** (2004). Novel glycosaminoglycan precursors as anti-amyloid agents: part IV. *J Mol Neurosci.* **24**, 167-172.
- Kitagawa, H., Tanaka, Y., Tsuchida, K., Goto, F., Ogawa, T., Lidholt, K., Lindahl, U. and Sugahara, K.** (1995). *N*-acetylgalactosamine (GalNAc) transfer to the common carbohydrate-protein linkage region of sulfated glycosaminoglycans. *J Biol Chem.* **270**, 22190-22195.
- Kitayama, K., Hayashida, Y., Nishida, K. and Akama, T. O.** (2007). Enzymes Responsible for Synthesis of Corneal Keratan Sulfate Glycosaminoglycans *J Biol Chem.* **282**, 30085-30096.
- Kjellén, L. and Lindahl, U.** (1991). Proteoglycans: structures and interactions. *Annu Rev Biochem.* **60**, 443-475.
- Klerk, C.P., Smorenburg, S.M., Otten, H.M., Lenseing, A.W.A., Prins, M.H., Piovella, F., prandoni, A., Bos, M.M.E.M., Richel, D.J., Tienhoven G.V. and Buller, H.R.** (2005). The effect of low molecular weight heparin on survival in patients with advanced malignancy. *J Clin Oncol.* **23**, 2130-2135.
- Kobayashi, M., Sugumaran, G., Liu, J., Shworak, N.W., Silbert, J.E. and Rosenberg, R.D.** (1999). Molecular cloning and characterization of a human uronyl 2-sulfotransferase that sulfates iduronyl and glucuronyl residues in dermatan/chondroitin sulfate. *J Biol Chem.* **274**, 10474-10480.
- Kojima, T., Shworak, N.W. and Rosenberg, R.D.** (1992b). Molecular cloning and expression of two distinct cDNA-encoding heparan sulfate proteoglycan core proteins from a rat endothelial cell line. *J Biol Chem.* **267**, 4870-4877.
- Kolkman, M.A.B., Wakarchuk, W., Nuijten, P.J.M. and van der Zeijst, B.A.M.** (1997). Capsular polysaccharide synthesis in *Streptococcus pneumoniae* serotype 14: molecular analysis of the complete *cps* locus and identification of genes encoding glycosyltransferases required for the biosynthesis of the tetrasaccharide subunit. *Mol Microbiol.* **26**, 197-208.
- Kramer, K.L. and Yost, H.J.** (2003). Heparan sulfate core proteins in cell-cell signaling. *Annu Rev Genet.* **37**, 461-484.
- Kreuger, J., Matsumoto, T., Vanwildemeersch, M., Sasaki, T., Timpl, R., Claesson-Welsh, L., Spillmann, D. and Lindahl, U.** (2002). Role of heparan sulfate domain organization in endostatin inhibition of endothelial cell function. *EMBO J.* **21**, 6303-6311.

**Kröger, S. and Schröder, J.E.** (2002). Agrin in the developing CNS: new roles for a synapse organizer. *News Physiol Sci.* **17**, 207-12.

**Krusius, T.** (1986). Identification of an *O*-glycosidic mannose-linked sialylated tetrasaccharide and keratan sulfate oligosaccharides in the chondroitin sulfate proteoglycan of brain. *J Biol Chem.* **261**, 8237-8242.

**Kusche-Gullberg, M. and Kjellen, L.** (2003). Sulfotransferases in glycosaminoglycan biosynthesis. *Curr Opin Struct Biol.* **13**, 605-611.

**Laemmli, U.K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **227**, 680-185.

**Lai, J.P., Sandhu, D.S., Shire, A.M. and Roberts, L.R.** (2008). The tumor suppressor function of human sulfatase 1 (SULF1) in carcinogenesis. *J Gastrointest Cancer.* **39**, 149-58.

**Lai, J.P., Osein, A.M., Moser, C.D., Yu, C., Elswa, S.F., Hu, C., Nakamura, I., Han, T., Aderca, I., Isomoto, H., Garrity-Park, M.M., Shire, A.M., Li, J., Sanderson, S.O., Adjei, A.A., Fernandez-Zapico, M.E. and Roberts, L.R.** (2010). The oncogenic effect of sulfatase 2 in human hepatocellular carcinoma is mediated in part by glypican 3-dependent Wnt activation. *Hepatology.* **52**, 1680-9.

**Lairson, L.L., Henrissat, B., Davies, G.J. and Withers, S.G.** (2008). Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem.* **77**, 521-55.

**Lassalle, P., Molet, S., Janin, A., Heyden, J.V., Tavernier, J., Fiers, W., Devos, R. and Tonnel, A.B.** (1996). ESM-1 is a novel human endothelial cell-specific molecule expressed in lung and regulated by cytokines. *J Biol Chem.* **271**, 20458-20464.

**Legras, S., Günthert, U., Stauder, R., Curt, F., Oliferenko, S., Kluin-Nelemans, H.C., Marie, J.P., Proctor, S., Jasmin, C. and Smadja-Joffe, F.** (1998). A strong expression of CD44-6v correlates with shorter survival of patients with acute myeloid leukemia. *Blood.* **9**, 3401-3413.

**Li, D., Clark, C.C. and Myers, J.C.** (2000). Basement membrane zone type XV collagen is a disulfide-bonded chondroitin sulfate proteoglycan in human tissues and cultured cells. *J Biol Chem.* **275**, 22339-22347.

**Li, Y., Rahmanian, M., Widstrom, C., Lepperdinger, G., Frost, G.I. and Heldin, P.** (2000). Irradiation induced expression of hyaluronan (HA) synthase 2 and hyaluronidase 2 genes in rat lung tissue accompanies active turnover of HA and induction of types I and III collagen gene expression. *Am J Resp Cell Mol Biol.* **23**, 411-418.

- Lidholt, K., Fjelstad, M., Lindahl, U., Goto, F., Ogawa, T., Kitagawa, H. and Sugahara, K.** (1997). Assessment of glycosaminoglycan-protein linkage tetrasaccharides as acceptors for GalNAc- and GlcNAc transferases from mouse mastocytoma. *Glycoconj J.* 1997, **14**, 737-742.
- Linhardt, R.J. and Hileman, R.E.** (1995). Dermatan Sulfate as a Potential Therapeutic Agent. *Gen Pharmacol.* **26**, 443-451.
- Liu, J. and Mushegian, A.** (2003). Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci.* **12**, 1418-1431.
- Lyon, M., Deakin, J.A. and Gallagher, J.T.** (1994). Liver heparan sulfate structure. A novel molecular design. *J Biol Chem.* **269**, 11208-11215.
- Malmstrom, A. and Fransson, L.A.** (1975). Biosynthesis of dermatan sulfate. I. Formation of L-iduronic acid residues. *J Biol Chem.* **250**, 3419-3425.
- Malmstrom, A.** (1984). Biosynthesis of dermatan sulfate. II. Substrate specificity of the C-5 uronosyl epimerase. *J Biol Chem.* **259**, 161-165.
- Maccarana, M., Sakura, Y., Tawada, A., Yoshida, K. and Lindahl, U.** (1996). Domain structure of heparan sulfates from bovine organs. *J Biol Chem.* **271**, 17804-17810.
- Mani, K., Havsmark, B., Persson, S., Kaneda, Y., Yamamoto, H., Sakurai, K., Ashikari, S., Habuchi, H., Suzuki, S., Kimata, K., Malmstrom, A., Westergren-Thorsson, G. and Fransson, L.A.** (1998). Tumor attenuation by 2(6-hydroxynaphthyl)- $\beta$ -D-xylopyranoside requires priming of heparan sulfate and nuclear targeting of the products. *Cancer Res.* **58**, 1099-1104.
- Manten-Horst, E., Danen, E.H., Smit, L., Snoek, M., Le Poole, I.C., Van Muijen, G.N., Pals, S.T., Ruiter, D.J.** (1995). Expression of CD44 splice variants in human cutaneous melanoma and melanoma cell lines is related to tumor progression and metastatic potential. *Int J Cancer.* **64**, 182-188.
- Martin, N.B., Masson, P., Sepulchre, C., Theveniaux, J., Millet, J. and Bellamy, F.** (1996). *Semin Thromb Hemost.* **22**, 247-254.
- Marynen, P., Zhang, J., Cassiman, J.J., Van der Berghe, H. and David, G.** (1989). Partial primary structure of the 48- and 90-kilodalton core proteins of cell surface-associated heparan proteoglycans of lung fibroblasts. *J Biol Chem.* **264**, 7017-7024.
- McLean, J.** (1916). The thromboplastic action of cephalin. *Am J Physiol.* **41**, 250-7.
- McMahan, U.J.** (1990). The agrin hypothesis. *Cold Spring Harbor Symp Quant Biol.* **55**, 407-418.

- Merry, C.L. and Wilson, V.A.** (2002). Role of heparan sulfate-2-O-sulfotransferase in the mouse. *Biochim Biophys Acta.* **1573**, 319-327.
- Meyer, K. and Palmer, J.** (1934). The polysaccharide of the vitreous humor. *J Biol Chem.* **107**, 629-634.
- Meyer, K.** (1953). The mucopolysaccharides of the cornea. *J Biol Chem.* **205**, 611-616.
- Miège, C., Maréchal, E., Shimojima, M., Awai, K., Block, M.A., Ohta, H., Takamiya, K., Douce, R. and Joyard, J.** (1999). Biochemical and topological properties of type A MGDG synthase, a spinach chloroplast envelope enzyme catalyzing the synthesis of both prokaryotic and eukaryotic MGDG. *Eur J Biochem.* **265**, 990-1001.
- Migliore, A. and Granata, M.** (2008). Intra-articular use of hyaluronic acid in the treatment of osteoarthritis. *Clin Interv Aging.* **3**, 365-369.
- Miller, K.D., Guyon, V., Evans, J.N., Shuttleworth, W.A. and Taylor, L.P.** (1999). Purification, cloning, and heterologous expression of a catalytically efficient flavonol 3-O-galactosyltransferase expressed in the male gametophyte of *Petunia hybrida*. *J Biol Chem.* **274**, 34011-34019.
- Miserocchi, G., Negrini, D., Passi, A. and De Luca, G.** (2001). Development of lung edema: interstitial fluid dynamics and molecular structure. *News Physiol Sci.* **16**, 66-71.
- Miwa, T., Watanabe, A., Yamada, Y., Shino, Y., Yamada, T., Yamashita, J., Matsuda, M. and Nakano, H.** (1996). Progression in gastric carcinoma relative to the ratio of CD44 epithelial variant transcript to CD44 hematopoietic variant transcript. *Cancer.* **77**, 25-29.
- Mochizuki, H., Yoshida, K., Gotoh, M., Sugioka, S., Kikuchi, N., Kwon, Y.D., Tawada, A., Maeyama, K., Inaba, N., Hiruma, T., Kimata, K. and Narimatsu, H.** (2003). Characterization of a heparan sulfate 3-O-sulfotransferase-5, an enzyme synthesizing a tetrasulfated disaccharide. *J Biol Chem.* **278**, 26780-26787.
- Monzon, M.E., Casalino-Matsuda, S.M. and Forteza, R.M.** (2006). Identification of glycosaminoglycans in human airway secretions. *Am J Respir Cell Mol Biol.* **34**, 135-141.
- Morrison, J.F. and Ebner, K.E.** (1971). Studies of galactosyltransferase. Kinetic effects of lactalbumin with N-acetylglucosamine and glucose as galactosyl group acceptors. *J Biol Chem.* **246**, 3992-3998.
- Moses, J., Oldberg, A., Cheng, F. and Fransson, L.A.** (1997). Biosynthesis of the proteoglycan decorin. Transient 2-phosphorylation of xylose during formation of the trisaccharide linkage region. *Eur J Biochem.* **248**, 521-526.

**Moses, J., Oldberg, Å. and Fransson, L.Å.** (1999). Initiation of galactosaminoglycan biosynthesis. Separate galactosylation and dephosphorylation pathways for phosphoxylosylated decorin protein and exogenous xyloside. *Eur J Biochem.* **260**, 879-884.

**Multhaupt, H.A., Yoneda, A., Whiteford, J.R., Oh, E.S., Lee, W. and Couchman, J.R.** (2009). Syndecan signaling: when, where and why?  
*HTMLCONTROL Forms.HTML*

---

*:Hidden.1 J Physiol Pharmacol.* 60, Suppl 4, 31-8.

**Murdoch, A.D., Dodge, G.R., Cohen, I., Tuan, R.S. and Iozzo, R.V.** (1992). Primary structure of the human heparan sulfate proteoglycan for basement membrane (HSPG2/perlecan). A chimeric molecule with multi domains homologous to the low density lipoprotein receptor, laminin, neural cell adhesion molecules and epidermal growth factor. *J Biol Chem.* **267**, 8544-8557.

**Murphy, K.J., Merry, C.L.R., Lyon, M., Thompson, J.E., Roberts, I.S. and Gallagher, J.T.** (2004). A New Model for the Domain Structure of Heparan Sulfate Based on the Novel Specificity of K5 Lyase. *J Biol Chem.* **279**, 27239-27245.

**Nader, H.B., Takahashi, H.K., Straus, A.H. and Dietrich, C.P.** (1980). Selective distribution of the heparin in mammals: conspicuous presence of heparin in lymphoid tissues. *Biochim Biophys Acta.* **627**, 40-48.

**Nader, H.B., Chavante, S.R., Dos-santos, E.A., Oliveira, F.W., de-Paive, J.F., Jerônimo, S.M.B., Medeiros, G.F., de-Abreu, L.R.D., Leite, E.L., de-Sousa-Filho, J.F., Castro, R.A.B., Toma, L., Tersatiol, I.L.S., Porcionatto, E.A. and Dietrich, C.P.** (1999a). Heparan sulfates and heparins: similar compounds performing the same functions in vertebrates and invertebrates? *Braz J Med Biol Res.* **32**, 529-538.

**Naor, D. and Nedvetzki, S.** (2003). CD44 in rheumatoid arthritis. *Arthritis Res Ther.* **5**, 105-115.

**Narimatsu, H., Sinha, S., Brew, K., Okayam, H. and Qasba, P.K.** (1986). Cloning and sequencing of cDNA of bovine N-acetylglucosamine (beta 1-4) galactosyltransferase. *Proc Natl Acad Sci USA.* **83**, 4720-4724.

**Nieduszynski, I.A.** (1990). There are two major types of skeletal keratan sulfates. *Biochem J.* **271**, 243-245.

**Nile, C.J., Read, R.C., Akil, M., Duff, G.W. and Wilson, A.G.** (2008). Methylation status of a single CpG site in the IL6 promoter is related to IL6 messenger RNA levels and rheumatoid arthritis. *Arthritis Rheum.* **58**, 2686-2693.



- Noonan, D.M., Fulle, A., Vallente, P., Cai, S., Horigan, E., Sasaki, M., Yamada, Y. and Hassell, J.R.** (1991). The complete sequence of perlecan, a basement membrane heparan sulfate proteoglycan, reveals extensive similarity with laminin A chain, low density lipoprotein-receptor and neural cell adhesion molecule. *J Biol Chem.* **226**, 22939-22947.
- Nomura, T., Takizawa, M., Imokawa, G., Dohmae, N., Takio, K., Hattori, M. and Matsuo, N.** (1998). Purification cDNA cloning, and expression of the UDP-Gal: glucosylceramide  $\beta$ -1,4-galactosyltransferase from rat brain. *J Biol Chem.* **273**, 13570-13577.
- O'Donnell, C., Tiwari, V., Oh, M.J., Valyi-Nagy, T. and Shukla, D.** (2005). A role for 3-O-sulfotransferase isoform-2 in assisting HSV-1 entry and spread. *Virology.* **346**, 452-459.
- Oegma, T.R. Tr., Kraft, E.L., Jourdian, G.W. and van Valen, T.R.** (1984). Phosphorylation of chondroitin sulfate in proteoglycans from the swarm rat chondrosarcoma. *J Biol Chem.* **259**, 1720-1726.
- Oguro, K.,T. Kazama, M. Isemura, T. Nakamura, S. and Y. Sato.** (1991). Immunohistochemical alterations in basement membrane components of squamous cell carcinoma. *J Invest Dermatol.* **96**, 250-254.
- Oh, S.P., Kamagata, Y., Muragaki, Y., Timmons, S., Ooshima, A. and Olsen, B.R.** (1994). Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xaa-Yaa repeats identify a distinct family of collagenous proteins. *Proc Natl Acad Sci USA.* **91**, 4229-4233.
- Ohkawara, B., Yamamoto, T. S., Tada, M. and Ueno, N.** (2003). Role of glypican 4 in the regulation of convergent extension movements during gastrulation in *Xenopus laevis*. *Development.* **130**, 2129-2138.
- Okajima, T., Yoshida, K., Kondo, T. and Furukawa, K.** (1999a). Human homolog of *Caenorhabditis* sqv-3 gene is galactosyltransferase I involved in biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycan. *J Biol Chem.* **274**, 22915-22918.
- Okajima, T., Fukumoto, S., Furukawa, K. and Urano, T.** (1999b). Molecular basis for the progeroid variant of Ehlers–Danlos syndrome. Identification and characterization of two mutations in galactosyltransferase I gene. *J Biol Chem.* **274**, 28841-28844.
- Oldberg, A., Antonsson, P., Lindblom, K. and Heinegard, D.** (1989). A collagen-binding 59-kd protein (fibromodulin) is structurally related to the small interstitial proteoglycans PG-S1 and PG-S2 (decorin). *Embo J.* **8**, 2601-2604.
- Onnerfjord, P., Heathfield, T.F. and Heinegard, D.** (2004). Identification of tyrosine sulfation in extracellular leucine-rich repeat proteins using mass spectrometry. *J Biol Chem.* **279**, 26-33.

- Ortega, N. and Werb, Z.** (2002). New functional roles for non-collagenous domains of basement membrane collagens. *J Cell Sci.* **115**, 4201-14.
- Page, C.P.** (1997). Proteoglycans: the 'Teflon' of the airways? *Thorax.* **52**, 924-925.
- Park, P.W., Reizes, O. and Bernfield, M.** (2000b). Cell surface heparan sulfate proteoglycans: selective regulators of ligand-receptor encounters. *J Biol Chem.* **275**, 29923-29926.
- Pedersen, L.C., Dong, J., Taniguchi, F., Kitagawa, H., Krahn, J.M., Pedersen, L.G., Sugahara, K. and Negishi, M.** (2003). Crystal structure of an  $\alpha$ 1,4-N-acetylhexosaminyltransferase (EXTL2), a member of the exostosin gene family involved in heparan sulfate biosynthesis. *J Biol Chem.* **278**, 14420-14428.
- Petitou, M., Casu, B. and Lindahl, U.** (2003). 1976-1983, a critical period in the history of heparin: the discovery of the antithrombin binding site. *Biochimie.* **85**, 83-89.
- Plaas, A.H., Neame, P.J., Nivens, C.M. and Reiss, L.** (1990) Identification of the keratan sulfate attachment sites on bovine fibromodulin. *J Biol Chem.* **265**, 20634-20640.
- Poole, A.R.** (1986). Proteoglycans in health and disease: structures and functions. *Biochem J.* **236**, 1-14.
- Prandoni, P., Falanga, A. and Piccioli, A.** (2005). Cancer and venous thromboembolism. *Lancet Oncol.* **6**, 401-10.
- Prydz, K. and Dalen, K.T.** (2000). Synthesis and storing of proteoglycans. *J Cell Sci.* **113**, 193-205.
- Pusch, C.M., Zeitz, C., Brandau, O., Pesch, K., Achatz, H., Feil, S., Scharfe, C., Maurer, J., Jacobi, F.K., Pinckers, A., Andreasson, S., Hardcastle, A., Wissinger, B., Berger, W. and Meindl, A.** (2000). The complete form of X-linked congenital stationary night blindness is caused by mutations in a gene encoding a leucine-rich repeat protein. *Nature Genet.*, **26**, 324-327.
- Qasba, P.K., Ramakrishnan, B. and Boeggeman, E.** (2005). Substrate-induced conformational changes in glycosyltransferases. *Trends Biochem Sci.* **30**, 53-62.
- Qasba, P.K., Boopathy Ramakrishnan, B. and Boeggeman, E.** (2008). Structure and Function of  $\beta$ -1,4-Galactosyltransferase. *Curr Drug Targets.* **9**, 292-309.
- Quentin, E., Gladen, A., Roden, L. and Kresse, H.** (1990). A genetic defect in the biosynthesis of dermatan sulfate proteoglycan: galactosyltransferase I deficiency in fibroblasts from a patient with a progeroid syndrome. *Proc Nat Acad Sci.* **87**, 1342-1346.

- Rademacher, T.W., Parekh, R.B. and Dwek, R.A.** (1988). Glycobiology. *Annu Rev Biochem.* **57**, 785-838.
- Ramakrishnan, B. and Qasba, P.K.** (2001). Crystal structure of lactose synthase reveals a large conformational change in its catalytic component, the  $\beta$ 1,4-galactosyltransferase-I. *J Mol Biol.* **310**, 205-218.
- Ramakrishnan, B., and Qasba, P.K.** (2003) Comparison of the closed conformation of the  $\beta$ 1,4-galactosyltransferase-1 ( $\beta$ 4Gal-T1) in the presence and absence of R-lactalbumin (LA), *J Biomol Struct Dyn.* **21**, 1-8.
- Ramakrishnan, B., Balaji, P.V. and Qasba, P.K.** (2002). Crystal structure of beta1,4-galactosyltransferase complex with UDP-Gal reveals an oligosaccharide acceptor binding site. *J. Mol. Biol.* **318**, 491-502.
- Ramakrishnan, B., Ramasamy, V. and Qasba, P.K.** (2006). Structural snapshots of beta-1,4-galactosyltransferase-I along the kinetic pathway. *J Mol Biol.* **357**, 1619-1633.
- Ramakrishnan, B. and Qasba, P.** (2010). Crystal Structure of the Catalytic Domain of Drosophila  $\beta$ 1,4-Galactosyltransferase-7. *J Biol Chem.* **285**, 15619-15626.
- Raman, R., Sasisekharan, V. and Sasisekharan, R.** (2005). Structural insights into biological roles of protein-glycosaminoglycan interactions. *Chem. Biol.* **12**, 267-277.
- Ramasamy, V., Ramakrishnan, B., Boeggeman, E., and Qasba, P.K.** (2003). The role of tryptophan 314 in the conformational changes of  $\beta$ 1,4-galactosyltransferase-I. *J Mol Biol.* **331**, 1065-1076.
- Ramasamy, V., Ramakrishnan, B., Boeggeman, E., Ratner, D.M., Seeberger, P.H. and Qasba, P.K.** (2005). Oligosaccharide preferences of beta1,4-galactosyltransferase-I: crystal structures of Met340His mutant of human beta1,4-galactosyltransferase-I with a pentasaccharide and trisaccharides of the N-glycan moiety. *Mol Biol.* **353**, 53-67.
- Ramchandran, R., Dhanabal, M., Volk, R., Waterman, M.J.F., Segal, M., Lu, H., Knebelmann, B. and Sukhatme, V.P.** (1999). Antiangiogenic activity of Restin, NC10 domain of human collagen XV: comparison to endostatin. *Biochem Biophys Res Commun.* **255**, 735-739.
- Rapraeger, A. and Ott, V.L.** (1998). Molecular interactions of the syndecan core proteins. *Curr Opin Cell Biol.* **10**, 620-628.
- Rapraeger, A.C.** (2000). Syndecan-regulated Receptor Signaling. *J Cell Biol.* **149**, 1143-56.
- Reed, C.C., Gauldie, J. and Iozzo, R.V.** (2002). Suppression of tumorigenicity by adenovirus-mediated gene transfer of decorin. *Oncogene.* **21**, 3688-3695.

- Reed, C.C., Waterhouse, A., Kirby, S., Kay, P., Owens, R.A., McQuillan, D.J. and Iozzo, R.V.** (2005). Decorin prevents metastatic spreading of breast cancer. *Oncogene*. **24**, 1104-1110.
- Rehn, M. and Pihlajaniemi, T.** (1994). Alpha 1(XVIII), a collagen chain with frequent interruptions in the collagenous sequence, a distinct tissue distribution, and homology with type XV collagen. *Proc Natl Acad Sci USA*. **91**, 4234-4238.
- Rehn, M. and Pihlajaniemi, T.** (1995). Identification of three N-terminal ends of type XVIII collagen chains and tissue-specific differences in the expression of the corresponding transcripts. *J Biol Chem*. **270**, 4705-4711.
- Reichsman, F., Smith, L. and Cumberledge, S.** (1996). Glycosaminoglycans can modulate extracellular localization of the wingless protein and promote signal transduction. *J Cell Biol*. **135**, 819-827.
- Richmond, M.E., DeLuca, S. and Silbert, J.E.** (1973). Biosynthesis of chondroitin sulfate. Microsomal acceptors of sulfate, glucuronic acid and N-acetylgalactosamine. *Biochemistry* **12**, 3898-3903.
- Ridley, R.C., Xiao, H. Q., Hata, H., Woodliff, J., Epstein, J. and Sanderson, R.D.** (1993). Expression of syndecan regulates human myeloma plasma cell adhesion to type I collagen. *Blood*. **81**, 767-774.
- Robinson, J., Viti, M. and Hook, M.** (1984). Structure and properties of an undersulfated heparan sulfate proteoglycan synthesized by a rat hepatoma cell line. *J Cell Biol*. **98**, 946-953.
- Rodgers, K.D., Sasaki, T., Aszodi, A. and Jacenko, O.** (2007). Reduced perlecan in mice results in chondrodysplasia resembling Schwartz-Jampel syndrome. *Hum Mol Genet*. **16**, 515-528.
- Rong, J., Habuchi, H., Kimata, K., Lindahl, U. and Kusche-Gullberg M.** (2001). Substrate specificity of the heparan sulfate hexuronic acid 2-O sulfotransferase. *Biochemistry*. **40**, 5548-5555.
- Ropero, S., Setien, F., Espada, J., Fraga, M.F., Herranz, M., Asp, J., Benassi, M.S., Franchi, A., Patino, A., Ward, L. S., Bovee, J., Cigudosa, J.C., Wim, W. and Esteller, M.** (2004). Epigenetic loss of the familial tumor-suppressor gene exostosin-1 (EXT1) disrupts heparan sulfate synthesis in cancer cells. *Hum Mol Genet*. **13**, 2753-2765.
- Rossmann, M.G., Moras, D. and Olsen, K.W.** (1974). Chemical and biological evolution of a nucleotide-binding protein. *Nature*. **250**, 194-199.
- Roughley, P.J.** (2006). The structure and function of cartilage proteoglycans. *Eur Cell Mater*. **30**, 92-101.
- Rudd, P.M., Elliott, T., Cresswell, P., Wilson, I.A. and Dwek, R.A.** (2001). Glycosylation and the immune system. *Science*. **291**, 2370-2376.

- Ruoslahti, E.** (1988). Structure and biology of proteoglycans. *Annu. Rev. Cell. Biol.* **4**, 229-255.
- Ruoss, S.J., Gold, W.M., Caughey, G.H.** (1991). Mast cell exocytosis: evidence that granule proteoglycan processing is not coupled to degranulation. *Biochem Biophys Res Commun.* **179**, 140-146.
- Saarela, J., Ylikarppa, R., Rehn, M., Purmonen, S. and Pihlajaniemi, T.** (1998). Complete primary structure of two variant forms of human type XVIII collagen and tissue-specific differences in the expression of the corresponding transcripts. *Matrix Biol.* **16**, 319-328.
- Sanderson, R.D.** (2001). Heparan sulfate proteoglycans in invasion and metastasis. *Semin. Cell. Dev. Biol.* **12**, 89-98.
- Sasaki, T., Larsson, H., Tisi, D., Claesson-Welsh, L., Hohenester, E. and Timpl, R.** (2000). Endostatins derived from collagens XV and XVIII differ in structural and binding properties, tissue distribution and anti-angiogenic activity. *J Mol Biol.* **301**, 1179-1190.
- Sasisekharan, R., Moses, M.A., Nugent, M.A., Cooney, C.L. and Langer, R.** (1994). Heparinase inhibits neovascularization. *Proc Natl Acad Sci USA.* **91**, 1524-1528.
- Sasisekharan, R., Shriver, Z., Venkataraman, G. and Narayanasami, U.** (2002). Roles of heparan-sulphate glycosaminoglycans in cancer. *Nat Rev Cancer.* **2**, 521-528.
- Sato, T., Furukawa, K., Bakker, H., Van Den Eijnden, D.H. and Van Die, I.** (1998). Molecular cloning of a human cDNA encoding a novel  $\beta$ -1,4-galactosyltransferase with 37% identity to the mammalian UDP-Gal:GlcNAc  $\beta$ -1,4-galactosyltransferase, *Proc Natl Acad Sci USA.* **95**, 472-477.
- Saunders, S., Jalkanen, M., O'Farrel, S. and Bernfield, M.** (1989). Molecular cloning of syndecan, an integral membrane proteoglycan. *J Cell Biol.* **108**, 1547-1565.
- Scherpereel, A., Gentina, T., Grigoriu, B., Senechal, S., Janin, A., Tsicopoulos, A., Plenat, F., Bechard, D., Tonnel, A.B., Lassalle, P.** (2003). Overexpression of endocan induces tumor formation. *Cancer Res.* **63**, 6084-6089.
- Schick, B.P., Gradowski, J.F. and San Antonio, J.D.** (2001). Synthesis, secretion, and subcellular localization of serglycin proteoglycan in human endothelial cells. *Blood.* **97**, 449-458.
- Schlessinger, J. Lax, I. and Lemmon, M.** (1995). Regulation of growth factor activation by proteoglycans: what is the role of the low affinity receptors? *Cell.* **83**, 357-360.
- Schönherr, E., Witsch-Prehm, P., Harrach, B., Robenek, H. and Ranteberg, J.** (1995). Interaction of biglycan with type I collagen. *J Biol Chem.* **270**, 2776-2783.

- Schuksz, M., Fuster, M.M., Brown, J.R., Crawford, B.E., Ditto, D.P., Lawrence, R., Glass, C. A., Wang, L., Tor, Y. and Esko, J.D.** (2008). Surfen, a small molecule antagonist of heparan sulfate. *Proc Natl Acad Sci USA*. **105**, 13075-13080.
- Scott, J.E.** (1993). Dermatan Sulphate and Dermatan Sulphate Proteoglycans: Chemistry, Biology, *Chemical Pathology*. (Scott, J. E., ed) pp. 165-181, Portland Press, London.
- Scott, P.G., Nakano, T. and Dodd, C.M.** (1997). Isolation and characterization of small proteoglycans from different regions of the porcine knee meniscus. *Biochim Biophys Acta*. **1336**, 254-262.
- Scott, P.G., Dodd, C.M., Bergmann, E.M., Sheehan, J.K. and Bishop, P.N.** (2006). Crystal structure of the biglycan dimer and evidence that dimerization is essential for folding and stability of class I small leucine-rich repeat proteoglycans. *J Biol Chem*. **28**, 13324-32.
- Seidler, D.G., Faiyaz-Ul-Haque, M., Hansen, U., Yip, G.W., Zaidi, S.H., Teebi, A.S., Kiesel, L. and Götte, M.** (2006). Defective glycosylation of decorin and biglycan, altered collagen structure, and abnormal phenotype of the skin fibroblasts of an Ehlers-Danlos syndrome patient carrying the novel Arg270Cys substitution in galactosyltransferase I ( $\beta$ 4GalT-7). *J Mol Med*. **84**, 583-594.
- Sertie, A.L., Sossi, V., Camargo, A.A., Zatz, M., Brahe, C. and Passos- Bueno, M.R.** (2000). Collagen XVIII, containing an endogenous inhibitor of angiogenesis and tumor growth, plays a critical role in the maintenance of retinal structure and in neural tube closure (Knobloch syndrome). *Hum Mol Genet*. **9**, 2051-2058.
- Shridhar, V., Sen, A., Chien, J., Staub, J., Avula, R., Kovats, S., Lee, J., Lillie, J. and Smith, D.I.** (2002). Identification of underexpressed genes in early- and late-stage primary ovarian tumors by suppression subtraction hybridization. *Cancer research*. **62**, 262-270.
- Shukla, D., Liu, J., Blaiklock, P., Shworak, N.W., Bai, X., Esko, J.D., Cohen, G.H., Eisenberg, R.J., Rosenberg and R.D. Spear, P.G.** (1999). A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell*. **99**, 13-22.
- Shworak, N.W., Liu, J., Petros, L.M., Zhang, L., Kobayashi, M., Copeland, N.G., Jenkins, N.A. and Rosenberg, R.D.** (1999). Multiple Isoforms of heparan sulfate D-glucosaminyl 3-O-sulfotransferase. Isolation, characterization, and expression of human cDNAs and identification of distinct genomic loci. *J Biol Chem*. **274**, 5170-5184.
- Silbert, J.E. and Sugumaran, G.** (1995). Intracellular membranes in the synthesis, transport and metabolism of proteoglycans. *Biochim Biophys Acta*. **1241**, 371-384.
- Sjoberg, A., Onnerfjord, P., Morgelin, M., Heinegard, D., and Blom, A.M.** (2005). The Extracellular Matrix and Inflammation: Fibromodulin activates the classical pathway of complement by directly binding C1q. *J Biol Chem*. **280**, 32301-32308.

- Skold, C.M., Blaschke, E. and Eklund, A.** (1996). Transient increases in albumin and hyaluronan in broncoalveolar lavage fluid after quitting smoking: possible signs of reparative mechanisms. *Respir Med.* **90**, 523-529.
- Smeds, E., Habuchi, H., Do, A-T., Hjertson, E., Grundberg, H., Kimata, K., Lindahl, U. and Kusche-Gullberg, M.** (2003). Substrate specificities of mouse heparan sulphate glucosaminyl 6-O-sulphotransferases. *Biochem J.* **372**, 371-380.
- Spiro, R.C., Freeze, H.H., Sampath, D. and Garcia, J.A.** (1991). Uncoupling of chondroitin sulfate glycosaminoglycan synthesis by brefeldin A. *J Cell Biol.* **115**, 1463-1473.
- Stein, T., Keller, R., Stuhlsatz, H.W., Greiling, H., Ohst, E., Muller, E. and Scharf, H.D.** (1982). Structure of the linkage-region between polysaccharide chain and core protein in bovine corneal proteokeratan sulfate. *Hoppe Seylers Z Physiol Chem.* **363**, 825-833.
- Straus, A.H., Nader, H.B., Takahashi, H.K. and Dietrich, C.P.** (1982). Ontogeny of heparin in mammals: a correlation with the appearance of mast cells in tissues. *An Acad Bras Cienc.* **54**, 439-448.
- Sugahara, K., Yamashina, I., de Waard, P., Van Halbeek, H. and Vliegthart, J.F.** (1988). Structural studies on sulfated glycopeptides from the carbohydrate-protein linkage region of chondroitin 4-sulfate proteoglycans of swarm rat chondrosarcoma. Demonstration of the structure Gal(4-O-sulfate)beta 1-3Gal beta 1-4XYL beta 1-O-Ser. *J. Biol. Chem.* **263**, 10168-10174.
- Sugahara, K., Ohkita, Y., Shibata, Y., Yoshida, K. and Ikegami, A.** (1995a). Structural studies on the hexasaccharide alditols isolated from the carbohydrate-protein linkage region of dermatan sulfate proteoglycans of bovine aorta. Demonstration of iduronic acid-containing components. *J Biol Chem.* **270**, 7204-7212.
- Sugahara, K., Tsuda, H., Yoshida, K., Yamada, S., de, B.T. and Vliegthart, J.F.** (1995b). Structure determination of the octa- and decasaccharide sequences isolated from the carbohydrate-protein linkage region of porcine intestinal heparin. *J Biol Chem.* **270**, 22914-22923.
- Sugumaran, G., Katsman, M., Sunthankar, P. and Drake, R.R.** (1997). Biosynthesis of chondroitin sulfate. Purification of glucuronosyl transferase II and use of photoaffinity labeling for characterization of the enzyme as an 80-kDa protein. *J Biol Chem.* **272**, 14399-14403.
- Sugumaran, G., Katsman, M. and Silbert, J.E.** (1998). Subcellular colocalization and potential interaction of glucuronosyltransferases with nascent proteochondroitin sulphate at Golgi sites of chondroitin synthesis. *Biochem J.* **329**, 203-208.
- Superti-Furga, A., Rossi, A., Steinmann, B. and Gitzelmann, R.** (1996). A chondrodysplasia family produced by mutations in the diastrophic dysplasia sulfate transporter gene: Genotype/phenotype correlations. *Am J Med Genet.* **63**, 144-147.

- Suzuki, M.** (1939). Prosthetic group of cornea mucoid. *J. Biochem.* **30**, 185-191.
- Takagaki, K., Munakata, H., Kakizaki, I., Iwafune, M., Itabashi, T. and Endo, M.** (2002). Domain structure of chondroitin sulfate E octasaccharides binding to type V collagen. *J Biol Chem.* **277**, 8882-8889.
- Takami N, Osawa K, Miura Y, Komai K, Taniguchi M, Shiraishi M, Sato K, Iguchi T, Shiozawa K, Hashiramoto A, Shiozawa S.** (2006). Hypermethylated promoter region of DR3, the death receptor 3 gene, in rheumatoid arthritis synovial cells. *Arthritis Rheum.* **54**, 779-87.
- Tammi, M.I., Day, A.J. and Turley, E.A.** (2002). Hyaluronan and homeostasis: a balancing act. *J Biol Chem.* **277**, 4581-4584.
- Tarbouriech, N., Charnock, S.J. and Davies, G.J.** (2001). Three-dimensional structures of the Mn dTDP complexes of the family GT-2 glycosyltransferase SpsA: a comparison with related NDP-sugar glycosyltransferases. *J Mol Biol.* **314**, 655-661.
- Tkachenko, E., Rhodes, J.M. and Simons, M.** (2005). Syndecans: new kids on the signaling block. *Circ Res.* **96**, 488-500
- Toole, B.P.** (1990). Hyaluronan and its binding proteins, the hyaladherins. *Curr Opin Cell Biol.* **2**, 839-844.
- Topczewsky, J., Sepich, D.S., Myers, D.C., Walker, C., Amores, A., Lele, Z., Hammerschmidt, M., Postlethwait, J. and Solnica-Krezel, L.** (2001). The zebrafish glypican Knypek controls cell polarity during gastrulation movements of convergent extension. *Dev Cell.* **1**, 251-264.
- Tralhao, J.G., Schaefer, L., Micegova, M., Evaristo, C., Schönherr, E., Kayal, S., Veiga-Fernandes, H., Danel, C., Iozzo, R.V., Kresse, H. and Lemarchand, P.** (2003). In vivo selective and distant killing of cancer cells using adenovirusmediated decorin gene transfer. *FASEB J.* **17**, 464-466.
- Troup, S., Njue, C., Cliewer, E.V., Parisien, M., Roskelly, C., Chakravarti, S., Roughley, P.J., Murphy, L.C. and Watson, P.H.** (2003). Reduced expression of the small leucine-rich proteoglycans, lumican and decorin is associated with poor outcome in node-negative invasive breast cancer. *Clin Cancer Res.* **9**, 207-214.
- Tsen, G., Halfter W., Kröger S. and Cole, G.J.** (1995). Agrin is a heparan sulfate proteoglycan. *J Biol Chem.* **270**, 3392-3399.
- Tsuchiya, T., Osanai, T., Ogoe, A., Tamura, G., Chano, T., Kaneko, Y., Ishikawa, A., Orui, H., Wada, T., Ikeda, T., Namba, M., Takigawa, M., Kawashima, H., Hotta, T., Tsuchiya, A., Ogino, T. and Motoyama, T.** (2005). Methylation status of EXT1 and EXT2 promoters and two mutations of EXT2 in chondrosarcoma. *Cancer Genet Cytogenet.* **158**, 148-155.



- Turino, G.M. and Cantor, J.O.** (2003). Hyaluronan in respiratory injury and repair. *Am J Respir Crit Care Med.* **167**, 1169-1175.
- Turnbull, J.E. and Gallagher, J.T.** (1991). Distribution of iduronate 2-sulphate residues in heparan sulphate. Evidence for an ordered polymeric structure. *Biochem J.* **273**, 553-559.
- Tyrell, D.J., Kilfeather, S. and Page, C.P.** (1995). Therapeutic uses of heparin beyond its traditional role as an anticoagulant. *Trends Pharmacol Sci.* **16**, 198-204.
- Uchimura, K., Muramatsu, H., Kadomatsu, K., Fan, Q.W., Kurosawa, N., Mitsuoka, C., Kannagi, R., Habuchi, O. and Muramatsu, T.** (1998). Molecular cloning and characterization of an N-acetylglucosamine-6-Osulfotransferase. *J Biol Chem.* **273**, 22577-83.
- Vadaie, N., Hulinsky, R.S. and Jarvis, D.L.** (2002). Identification and characterization of a *Drosophila melanogaster* ortholog of human beta1,4-galactosyltransferase VII. *Glycobiology.* **12**, 589-97.
- Van Die, I., Van Tetering, A., Schiphorst, W.E., Sato, T., Furukawa, K. and Van Den Eijnden, D.H.** (1999). The acceptor substrate specificity of human  $\beta$ 4-galactosyltransferase V indicates its potential function in O-Glycosylation. *FEBS Lett.* **450**, 52-56.
- Van't Veer, L.J., Dai, H., van de Vijver, M.J., He, Y.D., Hart, A. A., Mao, M., Peterse, H.L., van der Kooy, K., Marton, M.J., Witteveen, A.T., Schreiber, G.J., Kerkhoven, R.M., Roberts, C., Linsley, P.S., Bernards, R., Friend, S.H.** (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature.* **415**, 530-536.
- Veugelers, M., De Cat, B., Ceulemans, H., Bruystens, A.M., Coomans, C., Durr, J., Vermeesch, J., Marynen, P. and David G.** (1999). Glypican-6, a new member of the glypican family of cell surface proteoglycans. *J Biol Chem.* **274**, 26968-26977.
- Wei, Z., Swiedler, S.J., Ishihara, M., Orellana, A. and Hirschberg, C.B.** (1993). A single protein catalyzes both N-deacetylation and N-sulfation during the biosynthesis of heparan sulfate. *Proc Nat Acad Sci USA* **90**, 3885-3888.
- Wells, L., Vosseller, K. and Hart, G.W.** (2001). Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc. *Science.* **291**, 2376-2378.
- Whitelock, J.M. and Iozzo, R.V.** (2005). Heparan sulfate: a complex polymer charged with biological activity. *Chem. Rev.* **105**, 2745-2764.
- Wiggins, C.A.R. and Munro, S.** (1998). Activity of the yeast *MNN1*  $\alpha$ -1,3 mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. *Proc Natl Acad Sci USA.* **95**, 7945-50.
- Wijdenes, J., Vooijs, W.C., Clement, C., Post, J., Morard, F., Vita, N., Laurent, P., Sun, R.X., Klein, B. and Dore, J.M.** (1996). A plasmocyte selective monoclonal antibody (B-B4) recognizes syndecan-1. *Br J Haematol.* **94**, 318-323.

- Wu, R.R. and Couchman, J.R.** (1997). cDNA cloning of the basement membrane chondroitin sulfate proteoglycan core protein, bamacan: a five domain structure including coiled-coil motifs. *J Cell Biol.* **136**, 433-44.
- Xia, G., Chen, J., Tiwari, V., Ju, W., Li, J.P., Malmstrom, A., Shukla, D. and Li, J.** (2002). Heparan sulfate 3-O-sulfotransferase isoform 5 generates both an antithrombin-binding site and an entry receptor for herpes simplex virus, type 1. *J Biol Chem.* **277**, 37912-37919.
- Xu, T., Bianco, P., Fisher, L.W., Longenecker, G., Smith, E., Goldstein, S., Bonadio, J., Boskey, A., Heegaard, A. M., Sommer, B., Satomura, K., Dominguez, P., Zhao, C., Kulkarni, A.B., Robey, P.G. and Young, M.F.** (1998). Targeted disruption of the biglycan gene leads to an osteoporosis-like phenotype in mice. *Nat Genet.* **20**, 78-82.
- Yamaguchi, A., Saito, M., Goi, T., Iida, A., Takeuchi, K. and Hirose, K.** (1995). Expression of CD44 variant exons 8-10 in gastric cancer. *Cancer Res.* **86**, 1166-71.
- Yamauchi, S., Hirahara, Y., Usui, H., Takeda, Y., Hoshino, M., Fukuta, M., Kimura, J.H. and Habuchi, O.** (1999). Purification and characterization of chondroitin 4-sulfotransferase from the culture medium of a rat chondrosarcoma cell line. *J Biol Chem.* **274**, 2456-2463.
- Yang, Y., MacLeod, V., Dai, Y., Khotskaya-Sample, Y., Shriver, Z., Venkataraman, G., Sasisekharan, R., Naggi, A., Torri, G., Casu, B., Vlodaysky, I., Suva, L.J., Epstein, J., Yaccoby, S., Shaughnessy, J.D.J., Barlogie, B. and Sanderson, R.D.** (2007). The syndecan-1 heparan sulfate proteoglycan is a viable target for myeloma therapy. *Blood.* **110**, 2041-2048.
- Yayon, A., Klagsbrun, M., Esko, J.D., Leder, P. and Ornitz, D.M.** (1991). Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell.* **64**, 841-848.
- Yip, G.W., Smollich, M. and Götte, M.** (2006). Therapeutic value of glycosaminoglycans in cancer. *Mol Cancer Ther.* **5**, 2139-48.
- Zafiroopoulos, A., Nikitovic, D., Katonis, P., Tsatsakis, A., Karamanos, N.K. and Tzanakakis, G.N.** (2008). Decorin-Induced Growth Inhibition Is Overcome through Protracted Expression and Activation of Epidermal Growth Factor Receptors in Osteosarcoma Cells. *Mol Cancer Res.* **6**, 785-794.
- Zhang, Y., Swaminathan, G.J., Deshpande, A., Boix, E., Natesh, R., Xie, Z., Acharya, K.R. and Brew, K.** (2003). Roles of individual enzyme-substrate interactions by  $\alpha$ -1,3-galactosyltransferase in catalysis and specificity. *Biochemistry.* **42**, 13512-13521.
- Zheng, J., Luo, W. and Tanzer, M.L.** (1998). Aggrecan synthesis and secretion. A paradigm for molecular and cellular coordination of multiglobular protein folding and intracellular trafficking. *J Biol Chem.* **273**, 12999-13006.

**Zuo, L., Zhang, S. M., Hu, R.L., Zhu, H. Q., Zhou, Q., Gui, S. Y., Wu, Q., Wang, Y.** (2008). Correlation between expression and differentiation of endocan in colorectal cancer. *World J Gastroenterol.* **14**, 4562-4568.

# *Annexes*

<b>Mutant</b>	<b>Amorce</b>	<b>Séquence</b>
D163A	Sens Antisens	5'GAC TAC ATT GCC ATG CAC <u>GCA</u> GTT GCA CTG CTC CCT CTC 3' 5'GAG AGG GAG CAG TGC AAC <u>TGC</u> GTG CAT GGC AAT GTA GTC 3'
D163E	Sens Antisens	5' GAC TAC ATT GCC ATG CAC <u>GAG</u> GTT GCA CTG CTC CCT CTC 3' 5' GAG AGG GAG CAG TGC AAC <u>CTC</u> GTG CAT GGC AAT GTA GTC 3'
V164A	Sens Antisens	5' TAG ATT GCC ATG CAC GAC <u>GCT</u> GAC CTG CTC CCT CTC AAC 3' 5' GTT GAG AGG GAG CAG GTC <u>AGC</u> GTC GTG CAT GGC AAT CTA 3'
D165A	Sens Antisens	5' ATT GCC ATG CAC GAC GTT <u>GCA</u> CTG CTC CCT CTC AAC GAG 3' 5' CTC GTT GAG AGG GAG CAG <u>TGC</u> AAC GTC GTG CAT GGC AAT 3'
D165E	Sens Antisens	5' ATT GCC ATG CAC GAC GTT <u>GAG</u> CTG CTC CCT CTC AAC GAG 3' 5' CTC GTT GAG AGG GAG CAG <u>CTC</u> AAC GTC GTG CAT GGC AAT 3'
A186D	Sens Antisens	5'-GG CCC TTC CAC GTG <u>GAC</u> TCC CCG GAG CTC-3' 5'-GAG CTC CGG GGA <u>GTC</u> CAC GTG GAA GGG CC-3'
L206P	Sens Antisens	5'-GC GGC ATC CTG CTG <u>CCT</u> TCC AAG CAG CAC TAC-3' 5'-GTA GTG CTG CTT GGA <u>AGG</u> CAG CAG GAT GCC GC-3'
F221A	Sens Antisens	5' ATG TCC AAC CGC <u>GCG</u> TGG GGC TGG GGC 3' 5' GCC CCA GCC CCA <u>CGC</u> GCG GTT GGA CAT 3'
F221Y	Sens Antisens	5' ATG TCC AAC CGC <u>TAC</u> TGG GGC TGG GGC 3' 5' GCC CCA GCC CCA <u>GTA</u> GCG GTT GGA CAT 3'
W222A	Sens Antisens	5' TCC AAC CGC TTC <u>GCG</u> GGC TGG GGC CGC 3' 5' GCG GCC CCA GCC <u>CGC</u> GAA GCG GTT GGA 3'
W222F	Sens Antisens	5' TCC AAC CGC TTC <u>TTT</u> GGC TGG GGC CGC 3' 5' GCG GCC CCA GCC <u>AAA</u> GAA GCG GTT GGA 3'
G223A	Sens Antisens	5' AAC CGC TTC TGG <u>GCG</u> TGG GGC CGC GAG 3' 5' CTC GCG GCC CCA <u>CGC</u> CCA GAA GCG GTT 3'
W224A	Sens Antisens	5' CGC TTC TGG GGC <u>GCG</u> GGC CGC GAG GAC 3' 5' GTC CTC GCG GCC <u>CGC</u> GCC CCA GAA GCG 3'
W224H	Sens Antisens	5' CGC TTC TGG GGC <u>CAC</u> GGC CGC GAG GAG 3' 5' CTC CTC GCG GCC <u>GTG</u> GCC CCA GAA GCG 3'
W224F	Sens Antisens	5' CGC TTC TGG GGC <u>TTC</u> GGC CGC GAG GAG 3' 5' CTC CTC GCG GCC <u>GAA</u> GCC CCA GAA GCG 3'
G225A	Sens Antisens	5' TTC TGG GGC TGG <u>GCG</u> CGC GAG GAC GAC 3' 5' GTC GTC CTC GCG <u>CGC</u> CCA GCC CCA GAA 3'

E227A	Sens Antisens	5' GGC TGG GGC CGC <b>GCG</b> GAC GAC GAC TTC 3' 5' GAA CTC GTC GTC <b>CGC</b> GCG GCC CCA GCC 3'
E227D	Sens Antisens	5' GGC TGG GGC CGC <b>GAC</b> GAC GAC GAC TTC 3' 5' GAA CTC GTC GTC <b>GTC</b> GCG GCC CCA GCC 3'
D228A	Sens Antisens	5' TGG GGC TGG GGC CGC GAG <b>GCA</b> GAC GAG TTC TAC CGG CGC 3' 5' GCG CCG GTA GAA CTC GTC <b>TGC</b> CTC GCG GCC CCA GCC CCA 3'
D228E	Sens Antisens	5' TGG GGC TGG GGC CGC GAG <b>GAG</b> GAC GAG TTC TAC CGG CGC 3' 5' GCG CCG GTA GAA CTC GTC <b>CTC</b> CTC GCG GCC CCA GCC CCA 3'
D229A	Sens Antisens	5' GGC TGG GGC CGC GAG GAC <b>GCA</b> GAG TTC TAC CGG CGC ATT 3' 5' AAT GCG CCG GTA GAA CTC <b>TGC</b> TGC CTC GCG GCC CCA GCC 3'
D229E	Sens Antisens	5' GGC TGG GGC CGC GAG GAC <b>GAG</b> GAG TTC TAC CGG CG ATT 3' 5' AAT GCG CCG GTA GAA CTC <b>CTC</b> TGC CTC GCG GCC CCA GCC 3'
E230A	Sens Antisens	5' CGC GAG GAC GAC <b>GCG</b> TTC TAC CGG CGC 3' 5' GCG CCG GTA GAA <b>CGC</b> GTC GTC CTC GCG 3'
H257A	Sens Antisens	5' AAG ACA TTT CGC <b>GCC</b> CTG CAC GAC CCA 3' 5' TGG GTC GTG CAG <b>GGC</b> GCG AAA TGA CTT 3'
H257R	Sens Antisens	5' AAG ACA TTT CGC <b>CGC</b> CTG CAC GAC CCA 3' 5' TGG GTC GTG CAG <b>GCG</b> GCG AAA TGA CTT 3'
R270C	Sens Antisens	5'-AAG AGG GAC CAG AAG <b>TGT</b> ATC GCA GCT CAA AC-3' 5'-GTT TGA GCT GCG ATA <b>CAC</b> TTC TGG TCC CTC TT-3'

### Description des amorces utilisées pour la mutagenèse dirigée par PCR de la $\beta$ 4GalT7

Le codon muté est souligné et en gras.