

UNIVERSITE DES SCIENCES ET TECHNOLOGIES DE LILLE

Laboratoire de Glycobiologie Structurale et Fonctionnelle UMR CNRS / USTL n°8576

THESE DE DOCTORAT

Sciences de la Vie et de la Santé

Présentée par

Sylvain Lehoux

**Etude de l'expression et des mécanismes de
régulation transcriptionnelle tissu-
spécifique du gène *ST6GAL2***

Soutenance le 13 novembre 2009

Jury :

Président :	Dr. Jean-Claude Michaksi	Université de Lille 1
Rapporteurs :	Pr. Abderaman Maftah	Université de Limoges UMR INRA 1061
	Dr. Rosella Mollicone	CNRS INSERM U-602 Villejuif
Examineur :	Dr. Marie-Laure Caillet-Boudin	INSERM U-837, Lille
Directeur :	Pr. Philippe Delannoy	Université de Lille 1
Co-directeur :	Dr. Marie-Ange Krzewinski-Recchi	Université de Lille 1

Remerciements

...Je ne suis pas en avance et (comme d'habitude ?) le temps me manque un peu. Je ne vais donc pas faire une liste exhaustive de toutes les personnes que je dois remercier. La liste serait trop longue. Quelqu'un a dit : « Les paroles s'envolent, les écrits restent », sûrement quelqu'un que l'on n'a jamais remercié sincèrement en face... Je pense, bien au contraire, qu'en allant remercier personnellement chacune des personnes qui m'ont aidé et soutenu pendant ces années thèses, ces paroles marqueront plus qu'une page dans un manuscrit qui finira sur le dessus d'une pile sur une étagère à prendre la poussière. Et au moins je suis sûr de n'oublier personne !

Néanmoins, de manière officielle et non moins sincère, je tiens à remercier ici :

Le Docteur Jean-Claude Michalski, pour m'avoir ouvert les portes du C9 et de l'Unité de Glycobiologie Structurale et Fonctionnelle.

Le Professeur Philippe Delannoy, pour m'avoir enseigné une rigueur scientifique à toutes épreuves durant toutes ces années.

Le Docteur Marie-Ange Krzewinski-Recchi (prononcez Réki, pas Réchi), merci d'avoir cru en moi dès le début et de m'avoir formé à la paillasse. Merci de m'avoir toujours soutenu et encouragé encore et encore. Courage pour ton HDR !

Le Docteur Sophie Groux-Degroote, pour ta précieuse aide dans tous les domaines : expérimental et rédactionnel notamment (surtout pour la relecture scientifique et orthographique de ce manuscrit). Tes conseils ont toujours été très efficaces.

Le Docteur Marie-Laure Caillet-Boudin, pour cette collaboration fructueuse. Merci pour votre curiosité, votre intérêt et votre vision différente de mon sujet.

Le Professeur Abdéraman Maftah et le Docteur Rosella Mollicone, merci d'avoir accepté si promptement de juger ce travail.

Remerciements en vrac et au kilo (ça, c'est sur que ça va être mal pris ...) :

Merci à Aurélie Cazet, ma collègue féminine préférée (pourtant la concurrence était dure), c'était cool d'avoir un sourire (en plus des autres bien sûr) tous les jours au labo, trop de bonne humeur ne tue pas la bonne humeur. Merci au Docteur Anne Harduin-Lepers, pour les discussions pertinentes que nous avons eu. Merci à Béatrice Catieau-Teylart pour ta gentillesse et pour avoir toujours été disponible. Merci à Gillou, mon sparing partner de muscu, seul je n'aurais jamais eu la motivation. Merci à Adeline Marcant pour son Western-Blot, je lui avais promis, je pense qu'elle n'a sûrement pas oublié. Merci à Cléa, ma chérie, pour me laisser me coucher tard sans râler les dernières semaines.

Merci particulièrement à toute l'équipe du 017, l'ambiance y a toujours été très bonne ! Même avec 9 filles pour seulement 2 garçons !

Enfin, merci à tous les autres, je ne vous ai pas oublié, je suis sûrement entrain de passer vous voir d'ailleurs...

Ce travail a été réalisé au sein de l'Unité de Glycobiologie Structurale et Fonctionnelle, Unité Mixte de Recherche n°8576 du CNRS, USTL, dirigée par le Docteur Jean-Claude Michalski, directeur de recherche à l'INSERM.

Ce travail a été encadré par le Professeur Philippe Delannoy et le Docteur Marie-Ange Krzewinski-Recchi, et a bénéficié d'une allocation de recherche ministérielle.

Communications et Publications scientifiques

Publications

Lehoux S, Groux-Degroote S, Cazet A, Dhaenens C M, Maurage C A, Caillet-Boudin M L, Delannoy P, Krzewinski-Recchi M A, (2009): Transcriptional regulation of the human *ST6GAL2* gene in cerebral cortex and neuronal cells. *Glycoconj J*.

Cazet A, Groux-Degroote S, Teylaert B, Kwon K M, **Lehoux S**, Slomianny C, Kim C H, Le Bourhis X, Delannoy P, (2009): GD3 synthase overexpression enhances proliferation and migration of MDA-MB-231 breast cancer cells. *Biol Chem*. 390(7): 601-9.

Groux-Degroote S, Krzewinski-Recchi M A, Cazet A, Vincent A, **Lehoux S**, Lafitte J J, Van Seuningen I, Delannoy P, (2008): IL-6 and IL-8 increase the expression of glycosyltransferases and sulfotransferases involved in the biosynthesis of sialylated and/or sulfated Lewisx epitopes in the human bronchial mucosa. *Biochem J*. 410(1): 213-23.

Orales

1st Glycobiology Symposium Korea - France, Suwon, Corée, 15 Octobre, 2007
“*Tissue-specific transcriptional regulation of hST6Gal II gene*”

18th Joint Meeting Glycobiology, Lille, France, 2 - 4 Novembre 2007
“*Tissue-specific transcriptional regulation of hST6Gal II gene*”

19th Joint Meeting Glycobiology, Wageningen, Pays-Bas, 30 Novembre - 2 Décembre 2008
“*Brain-specific transcriptional regulation of the human ST6GAL2 gene*”

Posters

Glycodev, Lille, France, 9 - 12 Décembre 2007
“*Study of transcriptional regulation of hST6Gal II gene*”

8^{ème} Journée Andrée Verbert, Lille, France, 18 Septembre 2008
“*Étude de la régulation transcriptionnelle tissu spécifique du gène ST6GAL2 chez l’Homme*”

9th JennerGlycobiology and medicine symposium, Brussels, Belgique, 13 - 15 Septembre 2009
“*Transcriptional expression and regulation of the human ST6GAL2 gene in brain*”

SOMMAIRE :

<u>Abréviations</u>	12
<u>Introduction bibliographique</u>	16
1. La sialylation	17
1.1. La glycosylation	17
1.1.1. La <i>O</i> -glycosylation de type « mucine »	18
1.1.2. La <i>N</i> -glycosylation	20
1.1.3. Extension des antennes des <i>N</i> -glycanes de type hybride et complexe.....	24
1.2. La sialylation	25
1.2.1. Une glycosylation terminale	25
1.2.2. Les acides sialiques	26
1.3. Les sialyltransférases.....	27
1.3.1. Généralités.....	27
1.3.2. La famille des ST3Gal	29
1.3.3. La famille des ST8Sia	32
1.3.4. La famille des ST6GalNAc	32
1.3.5. La famille des ST6Gal	33
2. La sialylation en α2,6- des <i>N</i>-glycosylprotéines chez l'Homme	38
2.1. Les structures sialylées en α 2,6- des <i>N</i> -glycosylprotéines humaines.....	38
2.2. Le motif Neu5Ac α 2-6Gal β 1-4GlcNAc, le Sia ₆ LacNAc	39
2.2.1. Le motif de reconnaissance des virus Influenza humains.....	39
2.2.2. Le ligand activateur du CD22 pour la réponse immunitaire médiée par les lymphocytes B.....	39
2.2.3. Un motif surexprimé au cours de la phase aiguë de l'inflammation.....	40
2.2.4. Un épitope glycanique surexprimé dans de nombreux cancers	42
2.2.5. Un motif particulier dans le phénomène de clairance hépatique des <i>N</i> -glycosylprotéines.....	42
2.3. Le motif Sia ₆ LacdiNAc	43
2.3.1. Le Sia ₆ LacdiNAc, un motif rare chez les mammifères et chez l'Homme	43
2.3.2. La glycodéline	43
2.3.3. L'hormone lutéinisante humaine, la hLH	45
2.3.4. Les activateurs du plasminogène	46
2.3.4.1. L'urokinase, u-PA	46
2.3.4.2. L'activateur de plasminogène tissulaire des cellules de Bowes, t-PA	47
2.3.5. Un rôle spécifique pour la structure Sia ₆ LacdiNAc ?	47
2.3.6. Biosynthèse des structures Sia ₆ LacNAc et Sia ₆ LacdiNAc	48

3. La régulation de l'expression des α2,6-sialyltransférases	50
3.1. La régulation transcriptionnelle des gènes de sialyltransférases.....	50
3.2. <i>ST6GAL1</i>	52
3.2.1. Organisation génomique	52
3.2.2. Promoteur de l'expression constitutive de <i>ST6GAL1</i>	55
3.2.3. Promoteur lymphocytaire de <i>ST6GAL1</i>	56
3.2.4. Promoteur hépatique de <i>ST6GAL1</i>	57
3.3. <i>ST6GAL2</i>	58
3.3.1. Organisation génomique	58
3.3.2. Expression tissulaire du gène humain, <i>ST6GAL2</i>	63
3.3.3. Le gène murin, <i>mst6gal2</i>	65
3.3.4. Le gène bovin, <i>bst6gal2</i>	68
3.4. DSiaT, la ST6Gal de drosophile, une ST6Gal I/II	69
3.4.1. Activité enzymatique	70
3.4.2. Organisation génomique	70
3.4.3. Expression tissulaire.....	71
3.5. Objectif du travail de thèse.....	73

Résultats

Introduction..... 76

1. Expression transcriptionnelle du gène <i>ST6GAL2</i> dans différentes lignées cellulaires en culture	77
2. Identification des extrémités 5' non traduites du gène <i>ST6GAL2</i> dans la lignée cellulaire de neuroblastome SH-SY5Y	80
3. Caractérisation de l'expression relative des transcrits du gène <i>ST6GAL2</i> initiés par EX, EY ou EZ dans différentes lignées cellulaires en culture et <i>in vivo</i>	87
3.1. Expression relative des transcrits contenant EX, EY et EZ dans différentes lignées cellulaires en culture	87
3.2. Expression du gène <i>ST6GAL1</i> , du gène <i>ST6GAL2</i> et de ses transcrits dans différents échantillons de cortex cérébral, de cervelet et de l'hippocampe.....	92
3.2.1. Expression de <i>ST6GAL1</i> et <i>ST6GAL2</i> dans des échantillons de cortex cérébral, de cervelet et de l'hippocampe	92
3.2.2. Expression relative des différents transcrits contenant EX, EY ou EZ dans différents échantillons de cortex cérébral, de cervelet et d'hippocampe.....	95
4. Expression de la protéine hST6Gal II dans différentes lignées cellulaires en cultures et <i>in vivo</i>	99
4.1. Test de réactivité de l'anticorps anti-hST6Gal II par Western-Blot	99

4.2. Expression de la protéine hST6Gal II dans différentes lignées cellulaires	101
4.3. Expression de la protéine hST6Gal II dans différents échantillons de cortex cérébral, de cervelet et d'hippocampe	104
5. Caractérisation des régions promotrices et des facteurs de transcription impliqués dans la régulation des transcrits de <i>ST6GAL2</i>	108
5.1. Analyses bioinformatiques d'une région de 3,6 kb en amont et englobant la région des trois exons EX, EY et EZ.....	108
5.2. Identification des régions promotrices minimales, tests d'activité luciférase de vecteurs rapporteurs	111
5.3. Identification de sites potentiels de fixation pour des facteurs de transcription par mutagenèse dirigée.....	116
6. Expression du gène <i>ST6GAL2</i> et de la protéine hST6Gal II après différenciation neuronale des cellules NT2/D1 par l'acide rétinolique	123
6.1. Les cellules NT2/D1, un modèle cellulaire de développement neuronal.....	123
6.2. Induction de l'expression du gène <i>ST6GAL2</i> dans les NT2/D1 après différenciation	124
6.3. Induction de l'expression des transcrits initiés par EX, EY et EZ dans les NT2/D1 après différenciation.....	127
6.4. Surexpression de la protéine hST6Gal II dans les cellules hNT	129
 <u>Discussion générale et Perspectives</u>	131
Rappel des résultats obtenus	133
1. Identification des extrémités 5' non traduites du gène <i>ST6GAL2</i> dans la lignée de neuroblastome humain SH-SY5Y	133
2. Expression relative des trois transcrits du gène <i>ST6GAL2</i> dans différentes lignées cellulaires en culture et dans différents échantillons du cortex cérébral, du cervelet et de l'hippocampe	134
2.1. Lignées cellulaires culture	134
2.2. Cortex cérébral, cervelet et hippocampe	134
3. Expression de la protéine hST6Gal II dans différentes lignées cellulaires en culture et dans différents échantillons du cortex cérébral, du cervelet et de l'hippocampe.....	135
4. Identification et caractérisation des régions promotrices du gène <i>ST6GAL2</i> dans la lignée cellulaire SH-SY5Y.....	135

5. Surexpression du gène <i>ST6GAL2</i> et de la protéine hST6Gal II après différenciation neuronale des cellules NT2/D1 sous l'action de l'acide rétinoïque.....	136
Discussion générale	137
1. Organisation génomique et expression des exons EX, EY et EZ	137
2. Régulation transcriptionnelle tissu spécifique du gène <i>ST6GAL2</i>	139
3. <i>ST6GAL2</i> , un gène surexprimé lors du développement neuronal.....	142
Perspectives du travail de recherche	144
1. Poursuite la caractérisation des régions promotrices du gène <i>ST6GAL2</i> et confirmer le rôle joué par certains facteurs de transcription	144
2. Poursuite de l'étude du rôle de <i>ST6GAL2</i> au cours de la différenciation neuronale, avec le modèle NT2/D1.....	145
Conclusion générale	146
<u>Matériels et Méthodes</u>	147
Culture Cellulaire	148
1. Culture des cellules.....	148
2. Récupération de culots cellulaires en vue d'une extraction d'ARN	148
3. Récupération de culots cellulaires en vue d'une extraction de protéines.....	149
4. Différenciation par l'acide rétinoïque des NT2/D1	149
5. Différenciation par l'acide rétinoïque des cellules SH-SY5Y	150
6. Transfections transitoires et Tests d'activité luciférase.....	150
7. Transfection transitoire par les vecteurs d'expression codants hST6Gal I-3xFLAG et hST6Gal II-3xFLAG	151
8. Prélèvement des échantillons de tissu cérébral humain.....	151
Biologie moléculaire	152
1. Extractions d'ADN	152
1.1. Extraction d'ADN génomique.....	152
1.2. Extraction d'ADN plasmidique	152
2. Extractions des ARN	153

2.1. ARN de cellules en cultures	153
2.2. ARN d'échantillons tissulaires	153
3. Synthèse du premier brin d'ADNc.....	154
4. 5' RACE à partir d'ARN totaux de SH-SY5Y	154
4.1. Sélection des ARNm coiffés et rétro transcription.....	154
4.2. Premières PCR et PCR nichées	155
5. Construction des vecteurs codant les formes entières de hST6Gal I et hST6Gal II étiquettées en N-terminal du 3xFLAG	156
6. Construction de vecteurs rapporteurs pour les essais luciférase	157
7. Mutations des sites de fixations des facteurs de transcription	159
8. PCR en temps réelle SYBR Green & Duplex Taqman	160
8.1. Q-PCR avec la technologie SYBR Green®	160
8.2. Q-PCR en duplex avec la technologie Taqman.....	161
9. Analyses bioinformatiques.....	161
Protéomique	163
1. Extraction de protéines	163
1.1. Protéines extraites de cellules en cultures	163
1.2. Protéines extraites d'échantillons tissulaires	163
2. Western Blotting.....	164
3. Marquage des protéines au rouge ponceau.....	164
<u>Références bibliographiques</u>	167
<u>Annexes</u>	185
Anatomie et développement cérébral	186
Article : <i>Transcriptional regulation of the human ST6GAL2 gene in cerebral cortex and neuronal cells</i>	195
Le Nouveau Chapitre de la Thèse	211

Abréviations

ABREVIATIONS :

5' RACE : Rapid amplification of 5' cDNA ends

^{7m}G : 7 méthyl-guanine

ADNc : ADN complémentaire

AR : acide rétinoïque

ARNm : ARN messenger

ASGP : asialoglycoprotéines

ASGP-R : récepteurs d'asialoglycoprotéines

BCR : B-cell receptor

BEt : bromure d'éthidium

CAZy : Carbohydrate-Active Enzymes

Cer : échantillon provenant du cervelet

CMAH : CMP-Neu5Ac hydroxylase

CMP : cytidine monophosphate

CMP-NeuAc : acide cytidyl-monophospho-*N*-acétylneuraminique

Ct : Cycle threshold

DIG : digoxigénine

DMSO : diméthyle sulfoxyde

Dol-P : dolichol phosphate

DP : degré de polymérisation

EST : Expressed Sequence Tag

Fro : échantillon provenant du lobe frontal du cortex cérébral

FSH : Follicle Stimulating Hormone

Fuc : L-fucose

Gal : D-galactose

GalA : acide D-galacturonique

GalNAc : *N*-acétyl-D-galactosamine

GAP-43 : growth-associated protein 43

GATA1 : facteur de transcription qui se fixe sur la séquence nucléotidique du type WGATAR

GBP : Glycans Binding Proteins

GdA : glycodéline-A

GdC : glycodéline-C

GdF : glycodéline-F
GDP : guanidine diphosphate
GdS : glycodéline-S
Glc : D-glucose
GlcA : acide D-glucuronique
GlcNac : *N*-acétyl-glucosamine
GnRH : Gonadotropin-releasing Hormone
HCG : Hormone Chorionique Gonadotrope
Hip : échantillon provenant de l'hippocampe
HMG : High Mobility Group
HPRT : Hypoxantine phosphoribosyltransférase
IdoA : L-iduronique
IL-6 : interleukine-6
IL6-RE-BP : IL-6 response element binding protein
kb : kilo bases
kDa : kilo Daltons
kpb : kilo paire de bases
L : Large (Sialylmotif)
LacdiNac : *N,N'*-diacétyl-lactosédiamine (GalNAc β 1-4GlcNAc)
LacNac : *N*-acétyllactosamine (Gal β 1-4GlcNAc)
LH : hormone lutéinisante (Luteinizing Hormone)
LPS : lipopolysaccharide
Man : D-mannose
ManNac : *N*-acétylmannosamine
MAP-2 : micro-tubule-associated protein 2
MTA : Multiple tissue expression array
MYT1 : Myelin Transcription Factor 1
MZF1 : Myeloid Zinc Finger 1
NCAM : Neural Cell Adhesion Molecule
NeuAc : acide *N*-acétylneuraminique
NeuGc : acide *N*-glycolylneuraminique
NF-IL6 : nuclear factor-IL-6
NF- κ B : Nuclear Factor B
NRSF : Neuron-Restrictive Silencer Factor

Occ : échantillon provenant du lobe occipital du cortex cérébral
Olf1 : Olfactory neuron-specific binding protein
PA : activateur du plasminogène
PAEP : progesterone-associated endometrial protein
Par : échantillon provenant du lobe pariétal du cortex cérébral
pb : paire de bases
PCR : Polymerase chain reaction
PeP : phosphoénolpyruvate
Pit1 : Pituitary transcription factor 1
PMA : acétate-myristate de phorbol
PP14 : Placental Protein 14
Pur α : Purine-Rich element binding protein α
Q-PCR : Quantitative-Polymerase chain reaction
REG : réticulum endoplasmique granuleux
RPLP0 : Ribosomal protein large P0
S : Small (Sialylmotif)
SD : déviation standard
Sia₆LacdiNAc : Neu5Ac α 2-6LacdiNAc
Sia₆LacNAc : Neu5Ac α 2-6LacNAc
Sox5 : SRY box 5
SP1 : Specificity Protein 1
Tem : échantillon provenant du lobe temporal du cortex cérébral
TNF α : tumor necrosis factor α
t-PA : activateur de plasminogène tissulaire
TSH : Thyroid Stimulating Hormone
TSS : site d'initiation de la transcription (Transcriptional Start Site)
UDP : uridine diphosphate
u-PA : urokinase (activateur du plasminogène de type urinaire)
vs : comparé à (Versus)
VS : Very Small (sialylmotif)
Xyl : D-xylose

Introduction

bibliographique

1. LA SIALYLATION

1.1. La glycosylation

La glycosylation constitue l'une des modifications majeures des protéines et des lipides. Cette modification covalente consiste en l'ajout séquentiel d'un ou d'une série de monosaccharides pour former une glycoprotéine ou un glycolipide. Lorsque les glycoprotéines et les glycolipides sont ancrés dans la membrane cellulaire, ils constituent une couche dense appelée « glycocalix » formant littéralement une barrière autour de la cellule. Les glycoprotéines peuvent également être retrouvées dans les milieux biologiques tels que le plasma sanguin ou la matrice extracellulaire. L'importance de ces chaînes glycaniques n'a cessé d'être soulignée durant ces dernières décennies et leurs rôles peuvent être divisés en deux grandes catégories : (1) un rôle structural et modulateur des propriétés physico-chimiques des glycoconjugués, (2) une fonction spécifique de reconnaissance des glycanes par d'autres protéines, les Glycans Binding Proteins (GBP) qui peuvent être de natures endogène ou exogène (provenant de bactéries ou de virus). La glycosylation des glycoprotéines influence notamment leur mise en conformation lors de leur biosynthèse, leur solubilité dans l'eau, leur demi-vie quand elles sont circulantes et leur reconnaissance par des molécules de liaison spécifiques. Enfin, la glycosylation intervient également dans l'adhésion cellule-cellule ou cellule-matrice (Varki, 1993).

La large variété de fonctions des déterminants glycaniques repose en grande partie sur leur diversité structurale. En effet, les chaînes oligosaccharidiques des glycoconjugués sont constituées à partir de douze monosaccharides de base que sont : le D-galactose (Gal), la *N*-acétyl-D-galactosamine (GalNAc), le D-glucose (Glc), la *N*-acétyl-glucosamine (GlcNAc), le D-mannose (Man), le L-fucose (Fuc), l'acide L-iduronique (IdoA), les acides D-glucuronique (GlcA) et D-galacturonique (GalA), le D-xylose (Xyl) et les acides *N*-acétylneuraminique (acide sialique, NeuAc) et *N*-glycolylneuraminique (NeuGc). La longueur de ces chaînes est également un élément très variable et peut comprendre d'un à plusieurs dizaines, voire centaines, de résidus monosaccharidiques. Les monosaccharides sont transférés à partir de leur forme activée, un glycosyl-nucléotide, par l'action catalytique de glycosyltransférases spécifiques du monosaccharide transféré, du substrat accepteur et du type de liaison formée. Le Gal, Glc, GalNAc, GlcNAc, IdoA, GlcA, GalA et le Xyl sont activés par l'UDP (Uridine

diphosphate), le Man et le Fuc par le GDP (Guanidine diphosphate) et le NeuAc et NeuGc par le CMP (Cytidine monophosphate).

La liaison entre le premier glycane et la partie aglycone (portion non glucidique d'une glycoprotéine ou d'un glycolipide) définissent le type de glycosylation. Cette liaison covalente entre un monosaccharide et un résidu d'acide aminé contenu dans un peptide est le plus souvent de nature *O*- ou *N*-glycosidique. La *N*-glycosylation et la *O*-glycosylation de type « mucine » sont les deux types de glycosylation majeures rencontrées pour les glycoprotéines. La chaîne glycanique d'un *N*-glycane est attachée à un résidu d'asparagine inclus dans la séquence Asn-X-Ser/Thr (X représentant n'importe quel acide aminé à l'exception de la proline) par une liaison *N*-glycosidique entre la GlcNAc du glycane et l'asparagine. La chaîne glycanique d'un *O*-glycane de type « mucine » est quant à elle engagée par une liaison *O*-glycosidique entre la GalNAc de la chaîne glycanique et un résidu de sérine ou de thréonine. Outre ces deux types de glycosylation majeures, il existe d'autres formes de glycosylation des protéines telles que la *O*-GlcNAcylation, et plus minoritairement, la *C*-glycosylation, ou la *P*-glycosylation.

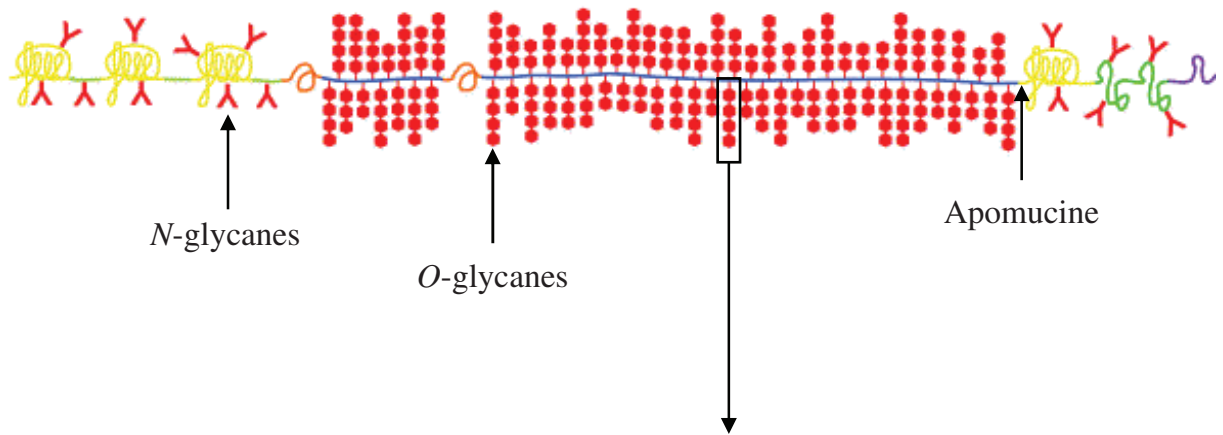
1.1.1. La *O*-glycosylation de type « mucine »

La *O*-glycosylation de type « mucine » consiste généralement en la modification covalente d'un résidu de sérine ou de thréonine par une chaîne glycanique. Elle est initiée par le transfert d'une *N*-acétylgalactosamine en anomérie α - sur le groupement OH- d'une sérine ou thréonine formant ainsi une liaison acétylique. La *O*-glycosylation débute directement au niveau de l'appareil de Golgi. Le résidu de GalNAc initial au point d'attache peut être allongé par un résidu de galactose, de *N*-acétylglucosamine ou d'acide sialique. La structure globale des chaînes de *O*-glycanes peut être divisée en trois régions (Figure 1) :

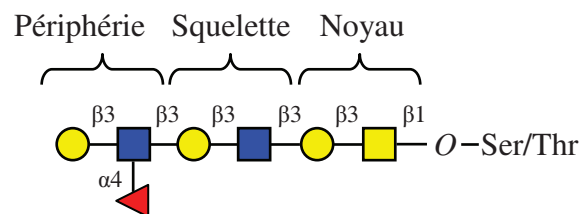
- (1), le noyau (ou core), incluant le résidu de GalNAc ancré à la partie protéique, et le ou les deux monosaccharides qui substituent cette GalNAc. Les structures oligosaccharidiques qui composent le noyau sont définies par le ou les deux monosaccharides ajoutés après la GalNAc au point d'attache. Ainsi, dix noyaux différents sont distingués (Figure 1 et Tableau 1).
- (2), le squelette, composé de l'addition alternée de résidus de Gal et de GlcNAc. Cet enchaînement forme les structures disaccharidiques de base : soit Gal β 1-3GlcNAc, appelée type 1 ou Néolacto, soit Gal β 1-4GlcNAc, appelée type 2 ou Lacto.
- (3), la partie périphérique, représentée par les derniers monosaccharides qui substituent le squelette : Neu5Ac (lié en α 2,3- ou α 2,6-), Fuc (lié en α 1,2-, α 1,3- ou α 1,4-), GalNAc (liée en

β 1,4-), GlcNAc (liée en α 1,4-), Gal (lié en α 1,3-) ou des groupements sulfates.

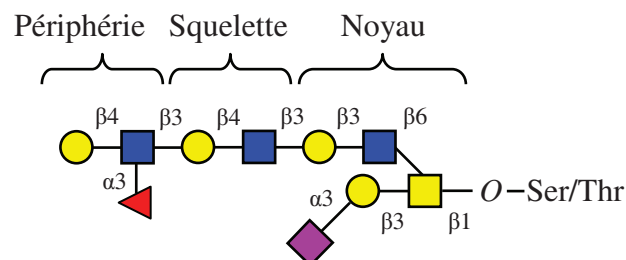
A



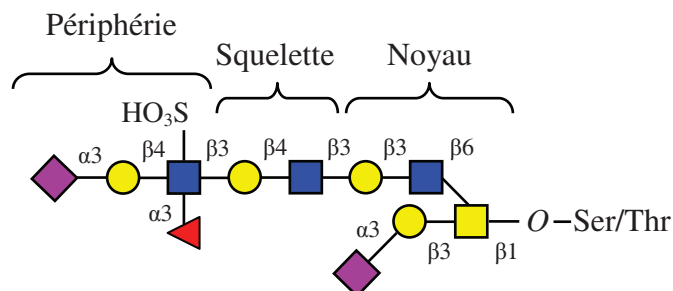
B



C



D



▲ : Fucose ◆ : NeuAc ● : Galactose ■ : GalNAc ■ : GlcNAc

Figure 1 : Schéma général de l'organisation d'une mucine de type bronchique. Exemples de O-glycanes.

A, schéma de l'organisation d'une mucine. B, C et D, exemple de O-glycanes. B, périphérie = Lewis a. C, périphérie = Lewis x. D, périphérie = 6-sulfo sialyl Lewis x.

Tableau 1 : Structures des noyaux des *O*-glycanes.

Noyau (Core)	Structure
Core 1 ou antigène T	Gal β 1-3GalNAc α -Ser/Thr
Core 2	GlcNAc β 1-6(Gal β 1-3)GalNAc α -Ser/Thr
Core 3	GlcNAc β 1-3GalNAc α -Ser/Thr
Core 4	GlcNAc β 1-6(GlcNAc β 1-3)GalNAc α -Ser/Thr
Core 5	GalNAc α 1-3GalNAc α -Ser/Thr
Core 6	GlcNAc β 1-6GalNAc α -Ser/Thr
Core 7	GalNAc α 1-6GalNAc α -Ser/Thr
Core 8	Gal α 1-3GalNAc α Ser/Thr
Antigène Tn	GalNAc α -Ser/Thr
Antigène Sialyl-Tn	Sia α 2-6GalNAc α -Ser/Thr

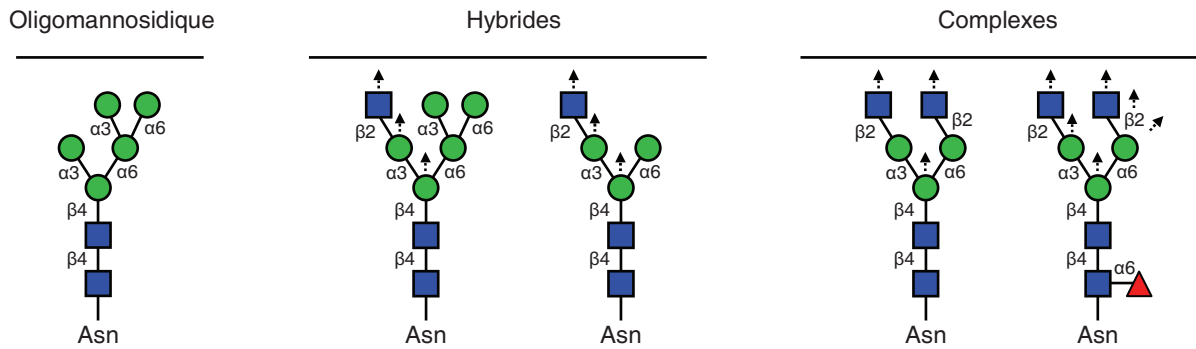
Les mucines, transmembranaires ou sécrétées dans le mucus (bronchique ou salivaire notamment) sont typiquement des exemples de glycoprotéines de haut poids moléculaires massivement *O*-glycosylées. Les *O*-glycanes des mucines représentent jusqu'à 80 % de leur masse. Les chaînes *O*-glycaniques des mucines remplissent de nombreux rôles. Elles protègent par exemple les cellules épithéliales sous-jacentes contre les agressions physiques ou contre des agents pathogènes. Elles permettent également l'adhésion cellulaire et interviennent au cours des différents processus que sont la réponse immunitaire et le développement embryonnaire (Tian *et al.*, 2009 ; Fukuda, 2002 ; Hanisch, 2001).

1.1.2. La *N*-glycosylation

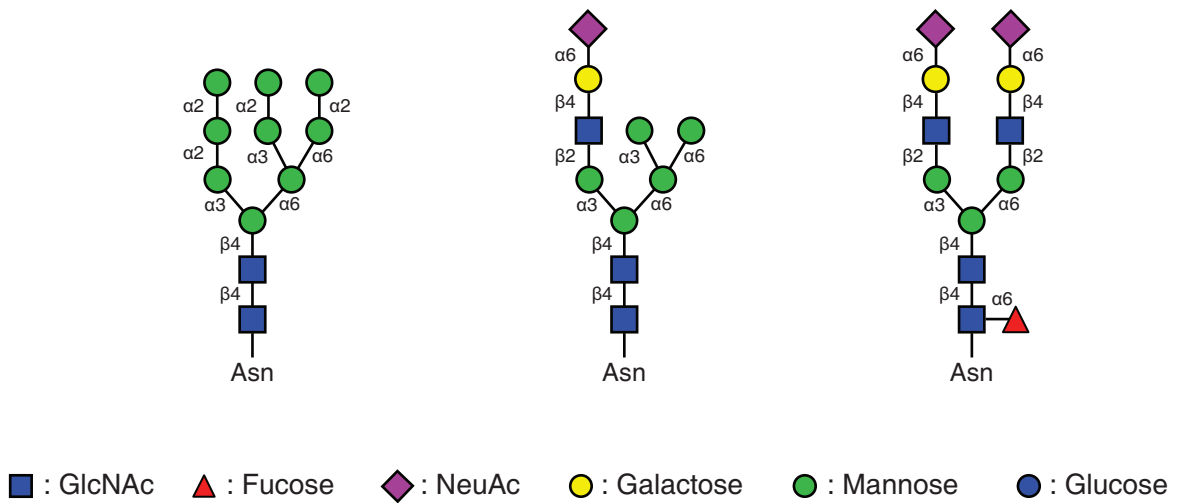
La *N*-glycosylation est un processus concomitant à la traduction protéique. Elle débute au niveau du réticulum endoplasmique granuleux (REG), pour se poursuivre dans la lumière de l'appareil de Golgi. Classiquement, les *N*-glycanes sont attachés à la protéine par une liaison *N*-glycosidique entre un résidu d'asparagine, contenu dans la séquence consensus peptidique Asn-X-Ser/Thr, et un résidu de GlcNAc. Les *N*-glycanes ont tous en commun un noyau pentasaccharidique (Man₃GlcNAc₂) à partir duquel on distingue trois classes principales de *N*-glycanes (Figure 2) :

- (1), les *N*-glycanes de type oligomannosidique
- (2), les *N*-glycanes de type hybride.
- (3), les *N*-glycanes de type complexe. Ce type de *N*-glycanes est le plus représenté chez les vertébrés.

A



B

**Figure 2 : Représentation schématique de N-glycanes.**

A, noyaux des trois classes de N-glycanes. **B**, exemples de N-glycanes appartenant aux trois classes, oligomannosidique, hybride et complexe

La biosynthèse des N-glycanes débute au niveau du REG avec l'assemblage d'un précurseur oligosaccharidique (Figure 3). Sur la face cytosolique du REG, un résidu de GlcNAc-1-P (apporté par l'UDP-GlcNAc) est ajouté au dolichol phosphate (Dol-P), pour former le GlcNAc-P-P-Dol. A partir de l'UDP-GlcNAc, un deuxième résidu de GlcNAc est transféré sur le GlcNAc-P-P-Dol pour donner le GlcNAc₂-P-P-Dol. Sur ce composé, viennent s'ajouter cinq résidus de Man, apportés par le GDP-Man, pour former la structure intermédiaire Man₅-GlcNAc₂-P-P-Dol. Après internalisation grâce à l'action d'une flippase, quatre nouveaux Man, apportés cette fois-ci par le Dol-P-Man, et trois Glc, apportés par le Dol-P-Glc sont transférés sur le composé précédent pour donner le Glc₃-Man₉-GlcNAc₂-P-P-Dol dans la lumière du REG. La structure glycanique Glc₃-Man₉-GlcNAc₂ est alors transférée

en bloc par l'oligosaccharidyltransférase sur un résidu asparaginyll d'une protéine en cours de biosynthèse (Figure 3). Les actions successives des glucosidases I et II et de la mannosidase du RER vont permettre d'obtenir la structure $\text{Man}_8\text{GlcNAc}_2\text{-(Asn)-protéine}$ transférable dès lors dans le compartiment golgien (Figure 4).

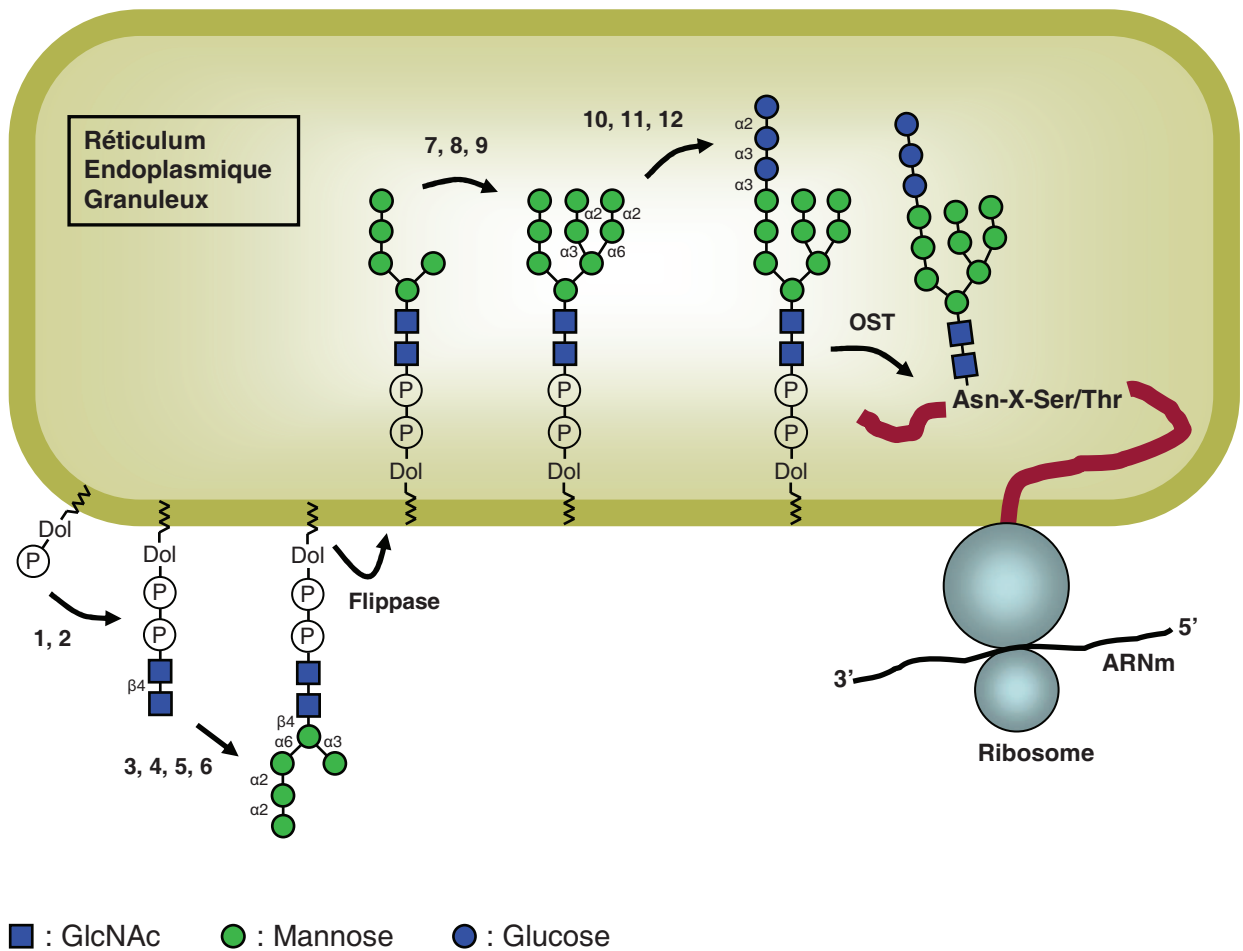


Figure 3 : Etapes réticulaires de la N-glycosylation.

Les enzymes impliquées sont : **1**, N-acétylglucosaminyll-1-P-transférase, **2**, N-acétylglucosaminyll-transférase, **3, 4, 5, 6** quatre types de mannosyltransférases, **7**, Man-P-Dol synthétase, **8, 9**, deux types de mannosyltransférases, **10**, Glc-P-Dol synthétase, **11, 12**, deux types de glucosyltransférases, **OST**, Oligosaccharidyltransférase.

Dans les premiers saccules golgiens, le *cis*-Golgi, la glycosylation se poursuit par l'action d' α -mannosidases spécifiques et de la N-acétylglucosaminylltransférase I pour donner naissance au composé $\text{GlcNAcMan}_5\text{GlcNAc}_2\text{-(Asn)-protéine}$. Les mannoses de la structure $\text{Man}_5\text{GlcNAc}_2\text{-(Asn)-protéine}$ des N-glycanes de type oligomannosidique demeurent non

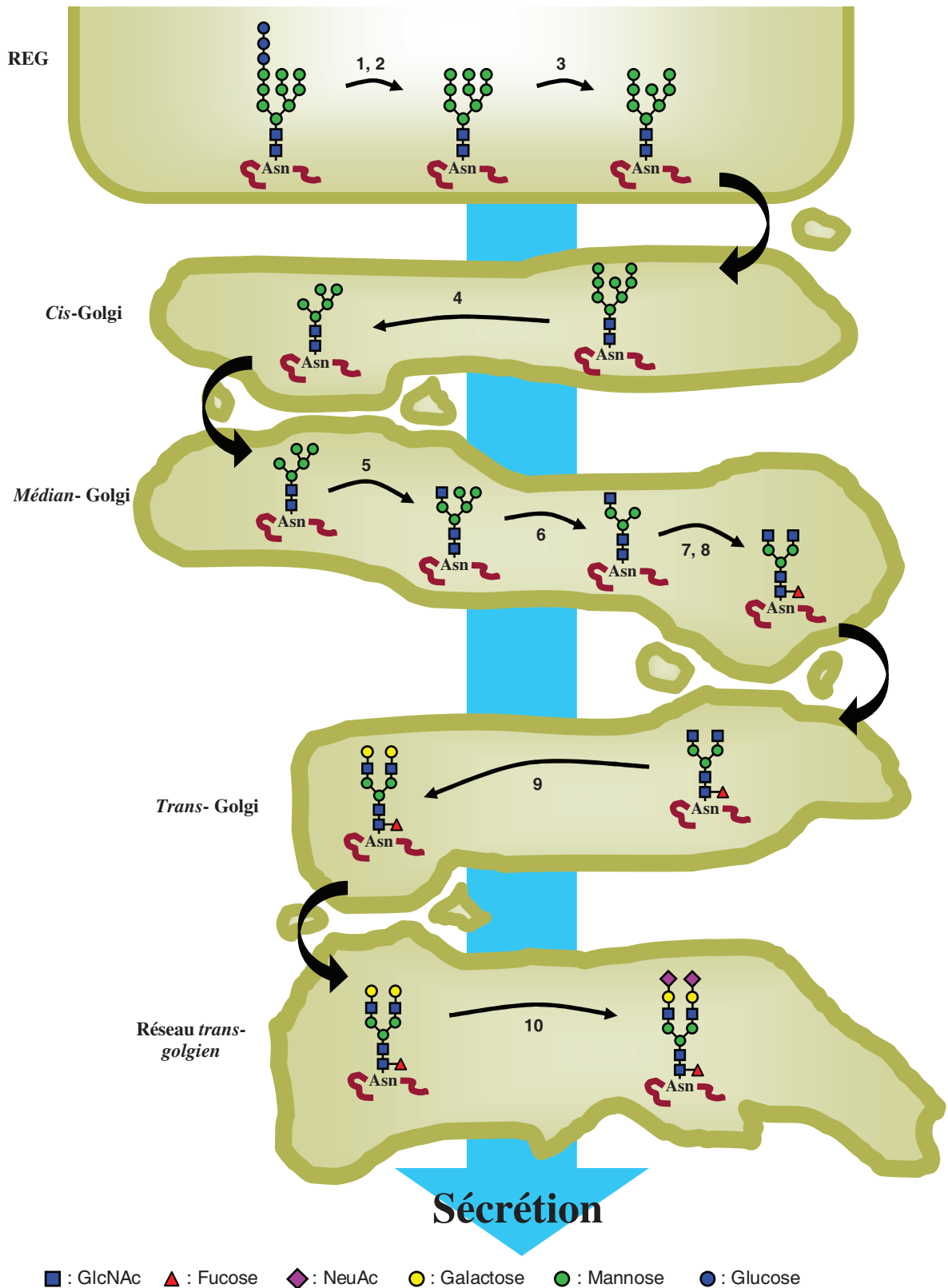


Figure 4 : Etapes golgiennes de la N-glycosylation.

Les glycosyltransférases impliquées sont : 1 et 2, α 1,2-glucosidase I et II, 3, RE α 1,2-mannosidase I, 4, α -mannosidase I, 5, β 1,2-N-acétylglucosaminyltransférase I, 6, α -mannosidase II, 7, β 1,2-N-acétylglucosaminyltransférase II, 8, α 1,6-fucosyltransférase, 9, β 1,4-galactosyltransférase, 10, α 2,3/6-sialyltransférase.

substitués. De leur côté, les *N*-glycanes de type hybride et complexe vont être soumis à l'action séquentielle de toute une série de glycosyltransférases spécifiques telles que les *N*-acétylglucosaminyl-, galactosyl-, *N*-acétylgalactosaminyl-, fucosyl- et sialyltransférases. Ces modifications vont donner naissance à la très grande diversité structurale des chaînes *N*-glycaniques retrouvées chez les eucaryotes. Les GlcNAc-transférases et mannosidases agissent dans les compartiments *cis*-Golgi et *median*-Golgi. Les Fuc-, Gal- et GalNAc-transférases sont localisées dans les citernes du *median*-Golgi et *trans*-golgi. Les sialyltransférases sont, quant à elles, actives au niveau du *trans*-Golgi et du réseau *trans*-golgien (Figure 4). Chacune de ces glycosyltransférases possède une affinité particulière pour une structure acceptrice donnée et une spécificité pour le type de liaison formée. L'expression de ces glycosyltransférases au niveau de l'appareil de Golgi est par ailleurs dépendante d'un type cellulaire et/ou de la situation physiologique, contribuant ainsi à la diversité des structures *N*-glycaniques d'un organisme à un autre. Les glycoprotéines ainsi maturées peuvent poursuivre leur chemin dans la voie de sécrétion en utilisant des vésicules afin d'être finalement exposées à la surface cellulaire ou libérées dans le milieu extracellulaire.

1.1.3. Extension des antennes des *N*-glycanes de type hybride et complexe

Les antennes des *N*-glycanes de type hybride et complexe (Figure 2B) sont majoritairement allongées par un galactose branché en β 1,4- sur la *N*-acétylglucosamine pour former la structure de type 2 Gal β 1-4GlcNAc, le LacNAc. Les antennes peuvent par ailleurs être à nouveau allongées par l'addition séquentielle de *N*-acétylglucosamine et de galactose, de laquelle résulte la formation de structures polylectosaminiques. Alternativement, un galactose peut être ajouté en β 1,3- et former le motif de type 1 Gal β 1-3GlcNAc, ou Néo-LacNAc, et une *N*-acétylgalactosamine peut être transférée en β 1,4-sur la GlcNAc pour former le LacdiNAc (GalNAc β 1-4GlcNAc) (Figure 5). L'extension de ces antennes s'effectue à partir des saccules du *trans*-Golgi par les Gal-, GalNAc- et GlcNAc-transférases et se poursuit principalement dans le réseau *trans*-golgien par la sulfatation et la sialylation.

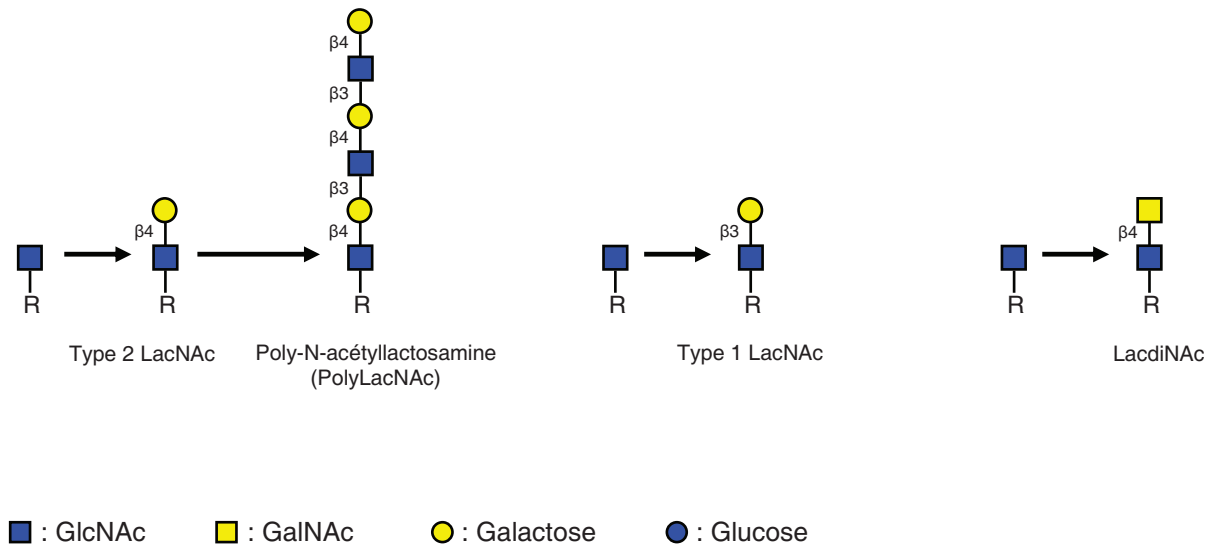


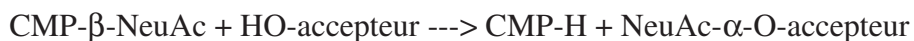
Figure 5 : Représentation schématique des structures LacNAc de type 1 et 2, PolyLacNAc et LacdiNAc.

L'élongation de la majorité des antennes des *N*-glycanes de type complexe se font par l'addition d'un résidu de Gal lié en β sur un résidu de GlcNAc. Ainsi se construit l'unité LacNAc, Gal β 1-4GlcNAc, structure de type 2, ou Gal β 1-3GlcNAc, structure de type 1. La longueur des antennes peut par ailleurs être augmentée par l'addition d'unités LacNAc et former des structures polylectosaminiques. Alternativement, une GalNAc peut être ajoutée à une GlcNAc pour former la structure GalNAc β 1-4GlcNAc, le LacdiNAc.

1.2. La sialylation

1.2.1. Une glycosylation terminale

La sialylation est l'ajout d'un résidu d'acide sialique (Figure 6) sur une chaîne glycanique préexistante. Cette glycosylation est l'une des dernières étapes de l'élongation de *N*- et *O*-glycanes des glycoprotéines. L'acide sialique est transféré à partir de sa forme activée, l'acide cytidyl-monophospho-*N*-acétylneuraminique (CMP-NeuAc), sur la partie terminale non réductrice d'un substrat accepteur glycoprotéique ou glycolipidique. Les glycosyltransférases dédiées à cette tâche sont les sialyltransférases. La réaction enzymatique peut se schématiser de la façon suivante :



La sialylation prend place au niveau des citernes du *trans*-Golgi et du réseau *trans*-golgien. Classiquement, les acides sialiques sont retrouvés à l'extrémité des chaînes glycaniques. En effet, aucun monosaccharide, à l'exception d'un autre acide sialique, ne peut être transféré sur

un résidu d'acide sialique ancré sur une chaîne glycanique. Les acides sialiques peuvent être transférés sur différents types de résidus monosaccharidiques accepteurs, sur du Gal, GalNAc, GlcNAc ou un autre NeuAc, et selon un type de liaison particulier, en α 2,3-, α 2,6- ou α 2,8-. Ces subtilités contribuent à la diversité des structures glycaniques sialylées et à couvrir ainsi une variété importante de fonctions biologiques.

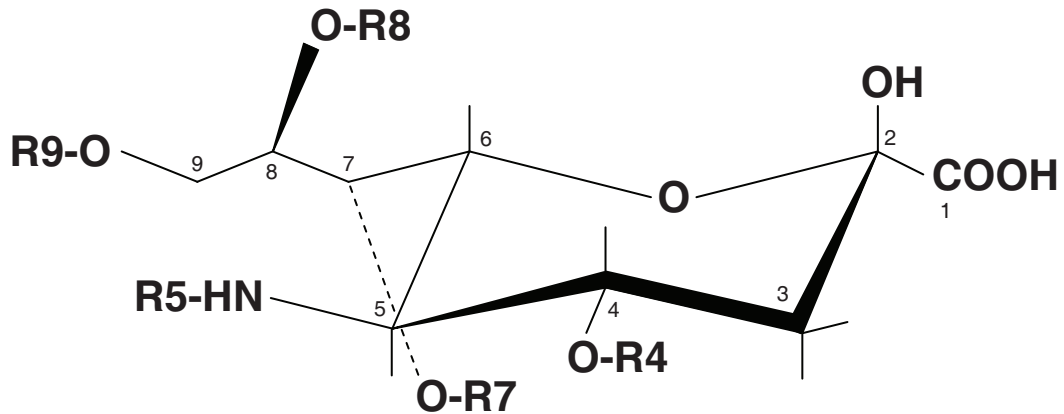


Figure 6 : Squelette de l'acide sialique et différentes substitutions possibles (NeuAc)

R5 = H, groupement acétyle, OH ou groupement glycolyl = Neu, Neu5Ac, KDN ou Neu5Gc, respectivement. R4, R7, R8 et R9 peuvent être des groupements acétyles, méthyles, sulfates ou phosphates.

1.2.2. Les acides sialiques

Les acides sialiques constituent une famille de monosaccharides acides qui ont comme squelette moléculaire commun l'acide neuraminique, un monosaccharide acide à neuf carbones : l'acide 5-amino-3,5-didéoxy-D-glycéro-D-galactononulosonique (ou acide neuraminique, Neu) (Figure 6). Communément, le carbone 5 est *N*-acétylé pour donner le l'acide 5-*N*-acétylneuraminique, le Neu5Ac. Le Neu5Ac est issu de la condensation d'un phosphoénolpyruvate (PEP) et d'une *N*-acétylmannosamine (ManNAc). Alternativement, un PEP et un Man peuvent être condensés pour former l'acide 2-kéto-3-déoxynonic, le KDN. Dans la voie de biosynthèse, le Neu5Ac, tout comme le KDN, sont ensuite activés par le CMP et forment respectivement le CMP-Neu5Ac et le CMP-KDN. Sous la forme CMP-Neu5Ac, le groupement *N*-acétyl du Neu5Ac peut être hydroxylé pour donner l'acide 5-*N*-glycolylneuraminique, le Neu5Gc. Si la découverte des acides sialiques remonte à plus de 50 ans, on dénombre aujourd'hui plus de soixante types différents d'acides sialiques retrouvés majoritairement dans le règne des deutérostomes (vertébrés, échinodermes et autres chordés),

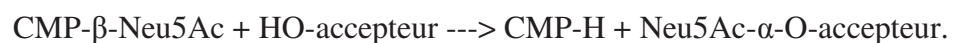
mais également chez les bactéries (Angata *et al.*, 2002). La grande diversité des acides sialiques provient essentiellement de modifications telles que la *O*-acétylation, la *O*-méthylation, la *O*-sulfatation, l'addition de groupements phosphates, mais aussi la lactonisation et la lactamisation (Figure 6) (Angata *et al.*, 2002).

Le Neu5Ac et le Neu5Gc sont les types d'acides sialiques prédominants chez les mammifères, alors que les échinodermes sont généralement plus fournis en acides sialiques *O*-substitués. Chez les protostomiens (arthropodes et mollusques), la présence d'acides sialiques n'a été que récemment mise en évidence et reste cependant rare. Par ailleurs, les acides sialiques sont, chez les vertébrés, les seuls monosaccharides dont la forme activée est couplée à un nucléotide, en l'occurrence une cytidine, monophosphate. Chez l'Homme, le Neu5Ac est retrouvé de manière quasi-exclusive, étant donnée l'absence de l'enzyme de conversion du CMP-Neu5Ac en CMP-Neu5Gc, la CMP-Neu5Ac hydroxylase (CMAH) (Chou *et al.*, 1998). Enfin, la fonction carboxyle du Neu5Ac est ionisée à pH physiologique, lui conférant ainsi une charge négative impliquée lors de phénomènes d'interaction / répulsion notamment à la surface des cellules (Kelm *et al.*, 1997).

1.3. Les sialyltransférases

1.3.1. Généralités

Les sialyltransférases forment une super famille de glycosyltransférases appartenant à la classe 29 de la classification CAZy (Carbohydrate-Active Enzymes), regroupant l'ensemble des sialyltransférases eucaryotes identifiées à ce jour (Coutinho *et al.*, 2003). Toutes les enzymes de cette super famille sont capables de transférer un acide sialique sur un substrat accepteur oligosaccharidique selon la réaction :



Les sialyltransférases sont dites invertives : la liaison s'effectue toujours en anomérie α sur les carbones 3 ou 6 d'un hexose ou d'une hexosamine ou sur le carbone 8 d'un autre résidu d'acide sialique lors de synthèse de motifs polysialylés. La super famille des sialyltransférases comporte vingt membres chez les mammifères, repartis en quatre familles. Les quatre familles se distinguent en fonction du substrat accepteur préférentiel et du type de liaison formée avec le monosaccharide accepteur. Ainsi, on distingue la famille des ST3Gal, des ST8Sia, des ST6GalNAc et des ST6Gal, qui transfèrent respectivement un acide sialique en α 2,3- sur un résidu de galactose, en α 2,8- sur un autre résidu d'acide sialique, en α 2,6- sur un résidu de *N*-

acétylgalactosamine et en $\alpha 2,6$ - sur un résidu de galactose. Six sialyltransférases composent la famille des ST3Gal, de même que la famille des ST8Sia et des ST6GalNAc, alors que deux enzymes seulement constituent la famille des ST6Gal. Cependant, la nature du substrat accepteur dépasse le simple cadre du monosaccharide cité précédemment. En effet, le disaccharide et même la partie aglycone du substrat accepteur conditionne l'affinité et la spécificité d'une sialyltransférase donnée. Il n'existe donc qu'une redondance partielle entre toutes les sialyltransférases d'une même famille.

Les sialyltransférases sont des glycoprotéines transmembranaires de type II localisées au niveau des saccules du *trans*-Golgi et du réseau *trans*-golgien. Le domaine catalytique C-terminal est orienté vers la lumière golgienne (Figure 7). Elles possèdent toutes une organisation structurale similaire. Elles sont composées d'une courte région cytoplasmique N-terminale de 5-10 acides aminés (a.a.), d'une unique région transmembranaire, d'une région tige de taille variant de 60 à 200 a.a. et enfin d'un domaine catalytique situé dans la lumière de l'appareil de Golgi (Figure 7). La région tige peut être clivée et libère ainsi la sialyltransférase sous sa forme soluble et active. C'est à l'intérieur du domaine catalytique que l'on retrouve des séquences très conservées appelées sialylmotifs ou motifs signatures dont l'homologie de séquence en a.a. est très conservée parmi les sialyltransférases eucaryotes. Ces motifs sont appelés L (pour Large, ou I), S (pour Small, ou II), VS (pour Very Small, ou IV) et III. Le motif L est constitué de 48 ou 49 a.a. dont 5 invariants. L'implication de ce motif dans la liaison avec le CMP-Neu5Ac a été démontrée par mutation en substituant certains des acides aminés conservés (Datta *et al.*, 1995). Le motif S est composé de 23 a.a. dont 2 sont invariants. Son rôle, à la fois dans la liaison du substrat donneur et celle du substrat accepteur, a été montré également par mutation en substituant certains des acides aminés conservés au niveau de ce site (Datta *et al.*, 1998). Le motif VS ainsi que le motif III, plus récemment identifié, seraient impliqués dans l'activité de transfert des sialyltransférases. Ceci a aussi été démontré par mutation en substituant certains des acides aminés présents dans ces motifs (Kitazume-Kawaguchi *et al.*, 2001 ; Jeanneau *et al.*, 2004).

Au-delà de ces sialylmotifs conservés pour toutes les sialyltransférases, certaines séquences très courtes d'acides aminés particuliers sont conservées au sein d'une même famille, ou entre certains membres d'une sous-famille. Ces motifs contribuent à la spécificité de liaison ($\alpha 2,3$ -, $\alpha 2,6$ - et $\alpha 2,8$ -) et d'accepteur (Gal, GalNAc, et Neu5Ac) des membres d'une même famille (Patel *et al.*, 2006). Ainsi, un de ces motifs de six à dix a.a. situé à la suite du sialylmotif L, serait relativement conservé entre tous les membres d'une même famille. Par exemple, il a été suggéré que ce motif contribue à la différence de substrat accepteur des

membres des familles ST3Gal et ST8Sia notamment (Fukumoto *et al.*, 1999). Enfin, pour la famille ST6GalNAc, l'étude de trois courts motifs conservés de quelques acides aminés d'un côté (Patel *et al.*, 2006) et une approche phylogénétique d'un autre côté (Harduin-Lepers *et al.*, 2005) ont mis en évidence, pour cette famille de sialyltransférases, l'existence de deux groupes de sous familles distinctes. ST6GalNAc I et II d'une part, et ST6GalNAc III, IV, V et VI d'autre part formeraient ces deux groupes. Alors que ST6GalNAc I et II transfèrent de façon indifférente sur un motif sialylé ou non, du type Gal β 1-3GalNAc-peptide (le monosaccharide accepteur est souligné ; Kurozawa *et al.*, 1994 & 1996), ST6GalNAc III, IV, V et VI n'utilisent que des motifs sialylés du type Neu5Aca2-3Gal β 1-3GalNAc-R (R étant une glycoprotéine ou un glycolipide) comme substrat accepteur (Lee *et al.*, 1999 ; Okajima *et al.*, 1999 & 2000).

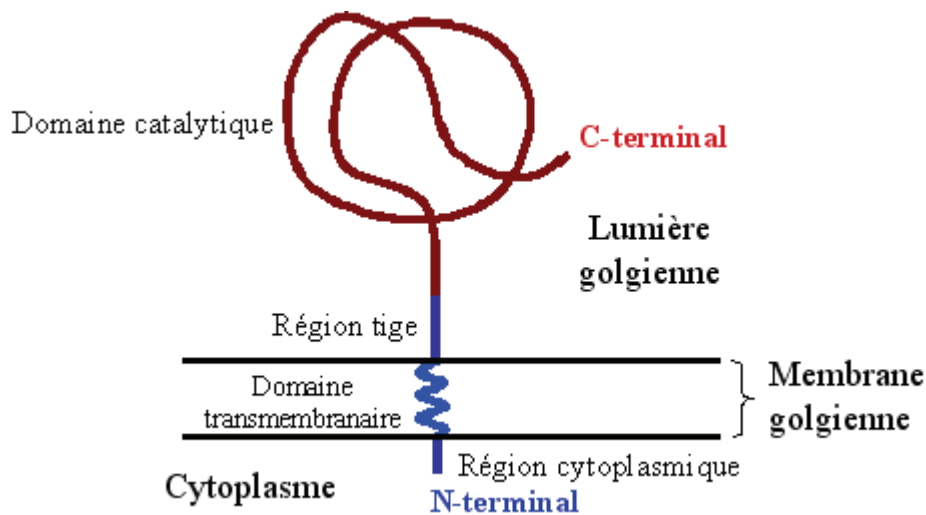


Figure 7 : Schéma de l'organisation protéique générale d'une sialyltransférase golgienne. Organisation typique d'une protéine transmembranaire golgienne de type II. A une courte région cytoplasmique N-terminale succède une unique région transmembranaire dans la membrane de l'appareil de Golgi. La région tige d'une taille variable précède, dans la lumière golgienne, la région catalytique C-terminale.

1.3.2. La famille des ST3Gal

Les sialyltransférases qui appartiennent à la famille des ST3Gal catalysent le transfert d'un acide sialique en α 2,3 sur un résidu de galactose terminal porté par des glycoprotéines ou des glycolipides. La famille des α 2,3-sialyltransférases est composée de six membres :

Tableau 2 : Récapitulatif des spécificités enzymatiques des 20 sialyltransférases humaines. Le monosaccharide accepteur est souligné.

Sialyltransférase	E.C. N°	Substrat accepteur préférentiel	Clonage ADNc humain
ST3Gal I	2.4.99.4	<u>Gal</u> β 1-3GalNAc (<i>O</i> -glycosylprotéines et glycolipides)	1995 (Chang <i>et al.</i> , 1995)
ST3Gal II	2.4.99.4	<u>Gal</u> β 1-3GalNAc (<i>O</i> -glycosylprotéines et glycolipides)	1996 (Kim <i>et al.</i> , 1996 ; Giordanengo <i>et al.</i> , 1997)
ST3Gal III	2.4.99.6	<u>Gal</u> β 1-3GlcNAc >> <u>Gal</u> β 1-4GlcNAc (glycoprotéines et glycolipides)	1994 (Kitagawa <i>et al.</i> , 1994a)
ST3Gal IV	2.4.99.-	<u>Gal</u> β 1-4GlcNAc >> <u>Gal</u> β 1-3GalNAc	1993 (Sasaki <i>et al.</i> , 1993)
ST3Gal V	2.4.99.9	<u>Gal</u> β 1-4Glc β -Cer (glycolipides) (Figure 8)	1998 (Ishii <i>et al.</i> , 1998)
ST3Gal VI	2.4.99.4	<u>Gal</u> β 1-4GlcNAc (glycoprotéines et glycolipides)	1999 (Okajima <i>et al.</i> , 1999)
ST8Sia I	2.4.99.8	<u>Neu5Ac</u> α 2-8NeuAc α 2-3Gal β 1-4Glc β -Cer (G _{D3}) (glycolipides) (Figure 8)	1994 (Nara <i>et al.</i> , 1994 ; Sasaki <i>et al.</i> , 1994 ; Haraguchi <i>et al.</i> , 1994)
ST8Sia II	2.4.99.-	(<u>Neu5Ac</u> α 2-8) _n Neu5Ac α 2-3Gal β 1-4Glc (glycoprotéines)	1997 (Angata <i>et al.</i> , 1997)
ST8Sia III	2.4.99.-	<u>Neu5Ac</u> α 2-3Gal β 1-4GlcNAc et (<u>Neu5Ac</u> α 2-8) _n Neu5Ac α 2-3Gal β 1-4GlcNAc (glycoprotéines et glycolipides)	1998 (Lee <i>et al.</i> , 1998)
ST8Sia IV	2.4.99.-	(<u>Neu5Ac</u> α 2-8) _n NeuAc α 2-3Gal β 1-4Glc (glycoprotéine)	1995 (Nakayama <i>et al.</i> , 1995)
ST8Sia V	2.4.99.-	G _{T3} , G _{T1a} , G _{D1c} et G _{Q1b} (glycolipides) (Figure 8)	1997 (Kim <i>et al.</i> , 1997)
ST8Sia VI	-	<u>Neu5Ac</u> α 2-6Gal β 1-3GalNAc (<i>O</i> -glycosylprotéines)	2005 (Teinturier-Lelièvre <i>et al.</i> , 2005)
ST6GalNAc I	2.4.99.3	(Neu5Ac α 2-3)(<u>Gal</u> β 1-3) <u>GalNAc</u> -Ser/Thr (après ou pas galactosylation, <i>O</i> -glycosylprotéines)	1999 (Ikehara <i>et al.</i> , 1999a)
ST6GalNAc II	2.4.99.3	(Neu5Ac α 2-3)Gal β 1-3 <u>GalNAc</u> -Ser/Thr (sur substrat sialylé ou non, <i>O</i> -glycosylprotéines)	2000 (Samyn-Petit <i>et al.</i> , 2000)
ST6GalNAc III	2.4.99.7	Neu5Ac α 2-3Gal β 1-3 <u>GalNAc</u> -Ser/Thr (<i>O</i> -glycosylprotéines) et G _{M1b} (glycolipides) (Figure 8)	2005 (Tsuchida <i>et al.</i> , 2005)
ST6GalNAc IV	2.4.99.7	Neu5Ac α 2-3Gal β 1-3 <u>GalNAc</u> (<i>O</i> -glycosylprotéine) et G _{D1α} (glycolipides) (Figure 8)	2000 (Harduin-Lepers <i>et al.</i> , 2000)
ST6GalNAc V	-	Neu5Ac α 2-3Gal β 1-3 <u>GalNAc</u> β 1-4Gal β 1-4Glc β 1-Cer et G _{M1b} (glycolipides) (Figure 8)	Souris 1999 (Ikehara <i>et al.</i> , 1999b)
ST6GalNAc VI	2.4.99.-	G _{M1b} , G _{D1a} et G _{T1b} (glycolipides) (Figure 8)	Souris 2000 (Okajima <i>et al.</i> , 2000)
ST6Gal I	2.4.99.1	<u>Gal</u> β 1-4GlcNAc (glycoprotéines)	1989 (Lance <i>et al.</i> , 1989)
ST6Gal II	2.4.99.1	<u>GalNAc</u> β 1-4GlcNAc (N-glycoprotéines et oligosaccharides libres)	2002 (Takashima <i>et al.</i> , 2002 ; Krzewinski-Recchi <i>et al.</i> , 2003)

1.3.3. La famille des ST8Sia

Les sialyltransférases de la famille des ST8Sia catalysent le transfert d'un acide sialique sur le carbone 8 d'un autre résidu d'acide sialique. Ces enzymes participent à la formation de structures polysialylées retrouvées aussi bien sur des glycoprotéines que sur des glycolipides. De telles structures peuvent contenir de quelques unités d'acide sialique à plus de 100 unités liées entre elles par des liaisons $\alpha 2,8$ - comme sur la glycoprotéine NCAM (Neural Cell Adhesion Molecule) (Kiss *et al.*, 1997). Bien que les six sialyltransférases qui composent cette famille transfèrent toutes un acide sialique en $\alpha 2,8$ - sur autre un acide sialique, elles présentent néanmoins des spécificités de substrats accepteurs et d'activités (Tableau 2). Les ST8Sia I, ST8Sia V et ST8Sia VI sont des monosialyltransférases et ne transfèrent qu'un seul résidu d'acide sialique, respectivement sur leur substrat accepteur préférentiel que sont le G_{M3} , le G_{D3} (Figure 8) et les *O*-glycosylprotéines se terminant par la structure Neu5Ac $\alpha 2$ -3Gal $\beta 1$ -4GlcNAc. Bien que les ST8Sia I et V soient impliquées dans l' $\alpha 2,8$ -sialylation des gangliosides, ST8Sia VI a une nette préférence pour le motif Neu5Ac $\alpha 2$ -6Gal $\beta 1$ -3GlcNAc porté par les protéines *O*-glycosylées (Nara *et al.*, 1994 ; Sasaki *et al.*, 1994 ; Kim *et al.*, 1997 ; Teinturier-Lelièvre *et al.*, 2005). Les sialyltransférases des sous familles ST8Sia II et ST8Sia IV sont, quant à elles, des polysialyltransférases qui catalysent le transfert d'acides sialiques sur la structure NeuAc $\alpha 2$ -3(6)Gal $\beta 1$ -4GlcNAc retrouvée sur les *N*-glycosylprotéines, notamment sur la NCAM (Foley *et al.*, 2009). ST8Sia III est une oligosialyltransférase capable de synthétiser des chaînes polysialylées de faible degré de polymérisation (DP) sur des glycolipides ou des glycoprotéines (Lee *et al.*, 1998 et Angata *et al.*, 2000).

1.3.4. La famille des ST6GalNAc

Les six enzymes de cette famille sont des $\alpha 2,6$ -sialyltransférases qui transfèrent un acide sialique sur un résidu de *N*-acétylgalactosamine. ST6GalNAc I et II sont capables d'utiliser comme substrat accepteur une GalNAc portée par une structure du type Gal $\beta 1$ -3GalNAc-*O*-glycoprotéine, sialylée ou non. Par ailleurs, ST6GalNAc I utilise aussi le substrat accepteur GalNAc-*O*-glycoprotéine pour former l'antigène sialyl-Tn (Neu5Ac $\alpha 2$ -6GalNAc-*O*-glycoprotéines (Ikehara *et al.*, 1999a). Par contre, ST6GalNAc III, IV, V et VI n'utilisent que le substrat sialylé Neu5Ac $\alpha 2$ -3Gal $\beta 1$ -3GlcNAc-R (R=*O*-glycoprotéine ou glycolipide). Les ST6GalNAc I, II et IV catalysent le transfert d'un acide sialique sur une GalNAc portée

par des *O*-glycosylprotéines alors que ST6GalNAc III, V et VI catalysent l'addition d'un acide sialique sur une GalNAc portée par des glycolipides. ST6GalNAc III et V sont dédiées à la synthèse du ganglioside G_{D1c} (Tsuchida *et al.*, 2005 ; Okajima *et al.*, 1999). Ces préférences de substrat accepteur pour les sous familles ST6GalNAc I et II d'une part et ST6GalNAc III, IV, V et VI d'autre part ont mis en évidence l'existence de deux groupes de sous familles parmi la famille des ST6GalNAc. De plus, une approche phylogénétique de la famille des ST6GalNAc a également corroboré cette hypothèse (Harduin-Lepers *et al.*, 2005). Enfin, il a été montré que les ST6GalNAc V et VI étaient également capables de transférer un acide sialique sur une GlcNAc, comprise dans le ganglioside NeuAc α 2-3Gal β 1-4(3)GlcNAc β 1-3Gal β 1-4Glc β 1-Cer (Tsuchida *et al.*, 2003).

1.3.5. La famille des ST6Gal

Contrairement aux trois autres familles de sialyltransférases, la famille des ST6Gal n'est composée que de 2 membres, ST6Gal I et ST6Gal II. Historiquement, ST6Gal I a été la première sialyltransférase isolée et caractérisée à partir du colostrum bovin (Paulson *et al.*, 1977), puis étudiée chez le rat (Miyagi *et al.*, 1982) et l'Homme (Lance *et al.*, 1989). Communément, ces enzymes transfèrent un résidu d'acide sialique sur un résidu de Gal retrouvé sur une structure *N*-acétyllactosaminique de type 2 : Gal β 1-4GlcNAc ou LacNAc, porté par des *N*-glycoprotéines ou des oligosaccharides libres. ST6Gal I est également supposée transférer un acide sialique en α 2,6- sur des glycolipides de la série Lacto. Cependant, ces deux sialyltransférases ont une activité catalytique quelque peu divergente vis-à-vis de certains substrats (Tableau 3). ST6Gal I a pour substrat accepteur préférentiel le LacNAc terminal de *N*-glycosylprotéines. Elle peut également catalyser le transfert d'un acide sialique sur des arylglycosides et des oligosaccharides libres se terminant par la structure GalNAc β 1-4GlcNAc (Nemansky *et al.*, 1992 ; Krzewinski-Recchi *et al.*, 2003). Ces caractéristiques enzymatiques sont conservées chez les ST6Gal I bovine (Mercier *et al.*, 1999) et murine (Takashima *et al.*, 2003). ST6Gal II, quant à elle, possède une nette préférence pour les oligosaccharides libres possédant un motif LacNAc libre et pour le LacdiNAc (GalNAc β 1-4GlcNAc) comme substrat accepteur (Krzewinski-Recchi *et al.*, 2003 ; Rohfritsch *et al.*, 2006). Chez l'Homme, le LacdiNAc est une structure beaucoup plus rare dans la composition des antennes des *N*-glycanes que le LacNAc. Par ailleurs, ces préférences de substrats sont également retrouvées pour la ST6Gal II murine vis-à-vis des oligosaccharides libres du type Gal β 1-4GlcNAc β 1-3Gal β 1-4Glc, ainsi que pour la ST6Gal II bovine vis-à-vis des

oligosaccharides libres et du motif LacdiNAc (Laporte *et al.*, 2009). Enfin, il est intéressant de noter que pour les deux ST6Gal humaines, l'efficacité de transfert de hST6Gal II est deux fois plus importante avec le LacdiNAc comme substrat accepteur par rapport au LacNAc alors que l'efficacité de hST6Gal I diminue d'un facteur quatre pour le LacdiNAc-R par rapport au LacNAc-R (Tableau 4) (Rohfritsch *et al.*, 2006), R représentant le reste de chaîne du N-glycane. Les substrats accepteurs naturels arborant du LacNAc et du LacdiNAc ainsi que ces structures sialylées en $\alpha 2,6$ - (Sia₆LacNAc et Sia₆LacdiNAc) seront plus détaillées dans le chapitre 2. **La sialylation en $\alpha 2,6$ - des N-glycosylprotéines chez l'Homme.**

Tableau 3 : Activités enzymatiques relatives comparées de hST6Gal I et hST6Gal II pour divers substrats accepteurs (adapté de Krzewinski-Recchi *et al.*, 2003 et Rohfritsch *et al.*, 2006).

Accepteur	Structures	Activité relative %	
		hST6Gal I	hST6Gal II
Glycoprotéine asialo $\alpha 1$ -acide	Gal β 1-4GlcNAc-R	100 (13,86)*	100 (0,52)*
Glycoprotéine $\alpha 1$ -acide	Neu5Ac $\alpha 2$ -6Gal β 1-4GlcNAc-R	0	0
Arylglycosides	Gal β 1-4GlcNAc β -O-octyl	72	259
	Gal β 1-3GlcNAc β -O-octyl	2,5	0
	Gal β 1-4GlcNAc β 1-2Man α 1-O-octyl	25	184
Oligosaccharides	Gal β 1-4GlcNAc	87	692
	Gal β 1-3GlcNAc	0	0
	Gal β 1-4GlcNAc β 1-3Gal β 1-4Glc	120	623
	Gal β 1-3GlcNAc β 1-3Gal β 1-4Glc	0	0

*, pmol.h⁻¹. μ l⁻¹

Tableau 4 : Activités enzymatiques relatives comparées de hST6Gal I, hST6Gal II et DSiaT pour le LacNAc et le LacdiNAc. (adapté de Krzewinski-Recchi *et al.*, 2003 ; Rohfritsch *et al.*, 2006 et Koles *et al.*, 2004).

Structures	Activité relative %		
	hST6Gal I	hST6Gal II	DSiaT
Gal β 1-4GlcNAc	100 (0,053)*	100 (0,019)*	100 (16,8)§
Gal β 1-3GlcNAc	2,5	0	36
GalNAc β 1-4GlcNAc	25	184	190

*, nmol/ml/min ; §, nmol/mg/h

D'un point de vue phylogénétique, la famille des ST6Gal est la première apparue au cours de l'évolution, avant les trois autres familles de sialyltransférases. Chez les arthropodes et en particulier chez *Drosophila melanogaster*, une seule sialyltransférase, appelé DSiaT, a été identifiée (Koles *et al.*, 2004). DSiaT appartient à la famille des ST6Gal mais n'appartient cependant ni à la sous famille ST6Gal I ni à la sous famille ST6Gal II (Figure 9). Les études menées concernant son activité enzymatique ont permis de mettre en évidence une similarité avec les propriétés enzymatiques décrites pour hST6Gal II. En effet, DSiaT est bien entendu capable de transférer un acide sialique sur une structure oligosaccharidique de type 2 : Gal β 1-4GlcNAc. Néanmoins, son substrat accepteur préférentiel *in vitro* est le LacdiNAc-R pour lequel son efficacité de transfert est deux fois supérieure à celle observée pour le LacNAc-R (R représentant le reste du *N*-glycane). Enfin, cette sialyltransférase est également active sur un motif de type 1 : Gal β 1-3GlcNAc (Koles *et al.*, 2004), à l'inverse des ST6Gal I caractérisées chez la souris et l'Homme (Tableau 4) (Takashima *et al.*, 2002 & 2003 ; Krzewinski-Recchi *et al.*, 2003). Sur le plan de l'activité enzymatique, ST6Gal II, plus qu'ST6Gal I, semble avoir conservé certaines de ces propriétés ancestrales.

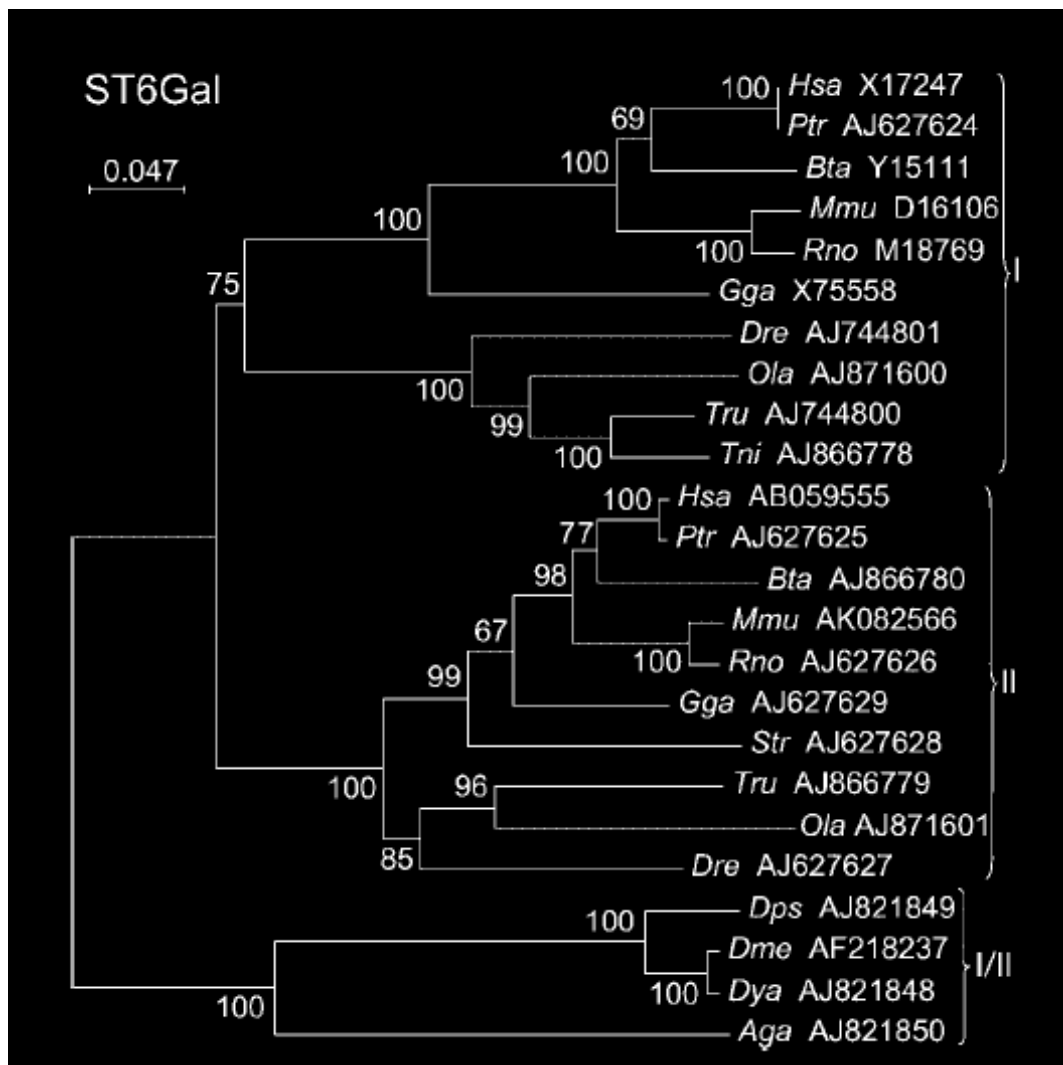


Figure 9 : Arbre phylogénétique de 24 sialyltransférases de la famille des ST6GAL, selon la méthode du Neighbor-Joining (Extrait de Harduin-Lepers *et al.*, 2005).

Aga, *Anopheles gambiae*; *Bta*, *Bos taurus*; *Dme*, *Drosophila melanogaster*; *Dps*, *Drosophila pseudoobscura*; *Dya*, *Drosophila yakuba*; *Dre*, *Danio rerio*; *Gga*, *Galus galus*; *Hsa*, *Homo sapiens*; *Mmu*, *Mus musculus*; *Ptr*, *Pan troglodytes*; *Rno*, *Rattus norvegicus*; *Str*, *Silurana tropicalis*; *Tni*, *Tetraodon nigroviridis*; *Tru*, *Takifugu rubripes*. Les quatre sialyltransférases d'insectes de *Dps*, *Dme*, *Dya* et *Aga* sont des orthologues de l'ancêtre commun (ST6Gal I/II) présent avant la séparation des deux sous familles, ST6Gal I et II, retrouvées chez les vertébrés.

Au cours de ma thèse, je me suis intéressé plus particulièrement à cette famille de sialyltransférases, les ST6Gal, chez l'Homme. hST6Gal I et hST6Gal II sont les deux seules sialyltransférases capables de synthétiser les motifs Sia₆LacNAc et Sia₆LacdiNAc terminaux rencontrés sur des *N*-glycoprotéines ou des oligosaccharides libres. Le chapitre suivant aura pour but de montrer que chez l'Homme, ces motifs remplissent des fonctions liées à la reconnaissance et l'adhésion cellulaire, à l'inflammation, la clairance hépatique et sont importants pour la demi-vie des glycoprotéines circulantes. Nous verrons également que la biosynthèse de ces motifs est dépendante d'un tissu ou d'un type cellulaire. Il sera par ailleurs montré que l'expression tissulaire et cellulaire de ces déterminants glycaniques est liée à l'expression des sialyltransférases impliquées dans leur biosynthèse. De nombreux exemples ont été décrits dans la littérature et ont montré que l'expression des sialyltransférases est contrôlée essentiellement au niveau transcriptionnel. Cette régulation s'effectue par l'utilisation de promoteurs tissulaires spécifiques régulant l'expression des gènes de sialyltransférases. L'étude de l'expression et de la régulation transcriptionnelle de l'un de ces gènes, le gène humain *ST6GAL2*, codant hST6Gal II, a fait l'objet du travail présenté ici.

2. LA SIALYLATION EN α 2,6- DES *N*-GLYCOSYLPROTEINES CHEZ L'HOMME

2.1. Les structures sialylées en α 2,6- des *N*-glycosylprotéines humaines

Chez l'Homme, les *N*-glycosylprotéines transmembranaires exposées à la surface des cellules ainsi que les *N*-glycosylprotéines circulantes sont très largement sialylées. Les motifs terminaux Neu5Ac α 2-3Gal β 1-4GlcNAc et Neu5Ac α 2-6Gal β 1-4GlcNAc sont les structures prépondérantes retrouvées sur les antennes des ces glycoprotéines sialylées. L'expression de ces structures est néanmoins souvent corrélée à un type de tissu ou un type de cellule donné mais peut également correspondre à certaines phases du développement embryonnaire ou être augmentée en réponse à un stimulus particulier. Enfin, lors des phénomènes accompagnant la différenciation tumorale d'une cellule, certains épitopes glycaniques peuvent aussi être ré-exprimés ou surexprimés.

Au cours de mes travaux, je me suis plus particulièrement intéressé au motif Neu5Ac α 2-6Gal β 1-4GlcNAc, ou Sia₆LacNAc, qui est abondamment exprimé chez l'Homme. Son rôle essentiel a été démontré dans de nombreux mécanismes d'adhésion et de reconnaissance cellulaire. Comme il sera détaillé dans le chapitre suivant, le Sia₆LacNAc est un motif indispensable à l'activation des lymphocytes B via la Siglec-2 ou CD22 (Powell LD *et al.*, 1993). Ce motif constitue aussi le ligand préférentiel des virus Influenza humains, les virus de la grippe (Shinya *et al.*, 2006). Par ailleurs, c'est également une structure glycanique déterminante dans la régulation de la demi-vie des glycoprotéines circulantes et son rôle est primordial dans la clairance hépatique des glycoprotéines sériques telles que les hormones par les récepteurs d'Ashwell (Ashwell *et al.*, 1982).

Par contraste avec le motif Sia₆LacNAc omniprésent, le motif Neu5Ac α 2-6GalNAc β 1-4GlcNAc, le Sia₆LacdiNAc, est beaucoup plus rarement exprimé au niveau des *N*-glycosylprotéines humaines. Seules quelques glycoprotéines sécrétées telles que l'hormone humaine hypophysaire LH ou la glycodéline, une glycoprotéine sécrétée par le placenta, arborent cette structure particulière (Weissharr *et al.*, 1991 ; Dell *et al.*, 1995). Le rôle de cette structure particulière est pour le moment apparenté à celui du Sia₆LacNAc. Des différences de propriétés fonctionnelles entre le Sia₆LacNAc et le Sia₆LacdiNAc n'ont actuellement pas encore été clairement démontrées.

2.2. Le motif Neu5Ac α 2-6Gal β 1-4GlcNAc, le Sia₆LacNAc

2.2.1. Le motif de reconnaissance des virus Influenza humains

De par leur position terminale et leur expression ubiquiste, les acides sialiques constituent une cible privilégiée et facilement accessible à de nombreux pathogènes et / ou à leurs toxines (Varki, 2008). Les cellules épithéliales de nos voies respiratoires supérieures expriment majoritairement des *N*-glycosylprotéines dont la périphérie est décorée par le motif Sia₆LacNAc, à l'inverse de celles de nos plus proches parents d'un point de vue évolutionniste, les grands singes, décorées par la structure Neu5Ac α 2-3Gal β 1-4GlcNAc (Gagneux *et al.*, 2003). Le Sia₆LacNAc est le ligand préférentiel reconnu par les hémagglutinines de virus « humain-spécifiques » tels que les virus Influenza de souches A et B (Shinya *et al.*, 2006 ; Stevens *et al.*, 2006). Par opposition, les virus aviaires du type Influenza ont un tropisme spécifique pour le motif Neu5Ac α 2-3Gal- prépondérant dans les épithéliums des voies aériennes supérieures des oiseaux (Suzuki, 2005 ; Yao *et al.*, 2008). Par ailleurs, les acides sialiques constituent chez l'Homme un motif privilégié de reconnaissance pour de nombreux autres virus tels que les réovirus, les polyomavirus ou les adénovirus (Forrest *et al.*, 2003 ; Gee *et al.*, 2006 ; Numéro spécial Glycoconjugate Journal 2006 Fev 1-2 : Sialic acid and Viruses).

2.2.2. Le ligand activateur du CD22 pour la réponse immunitaire médiée par les lymphocytes B

Dans le cadre de la réponse immunitaire médiée par les lymphocytes B, le motif Sia₆LacNAc joue un rôle déterminant dans l'activation du CD22. La Siglec-2, ou CD22, est une lectine appartenant à la superfamille des immunoglobulines transmembranaires (Cyster *et al.*, 1997). De par sa fonction lectinique, le CD22 possède une forte affinité pour les acides sialiques et plus spécifiquement les acides sialiques liés en α 2,6- impliqués dans la structure du Sia₆LacNAc (Powell *et al.*, 1993). Le CD22 est en outre une glycoprotéine transmembranaire retrouvée à la surface des lymphocytes B matures. En effet, il agit comme un co-récepteur des récepteurs de cellules B (B-cell receptor ou BCR) tels que les récepteurs antigéniques de surface (Sato *et al.*, 1996). Le CD22 a un rôle inhibiteur dans la transduction du signal émis par les récepteurs antigéniques et est ainsi capable de moduler les effets de la réponse immunitaire médiée par les lymphocytes B. Une fois associé aux BCR et son ligand

fixé, le CD22 est capable d'autophosphoryler ses résidus de tyrosine situés à la face intracellulaire déclenchant ainsi une cascade de signaux biochimiques permettant de limiter la réponse immunitaire des lymphocytes B lors de la réponse humorale (Law *et al.*, 1994 ; Nitschke, 2005). L'interaction du CD22 avec des *N*-glycosylprotéines circulantes ou membranaires porteuses du motif Sia₆LacNAc module ainsi l'adhérence, la migration et l'activation des lymphocytes B. Des glycoprotéines parmi lesquelles les IgM, le CD45 et le CD22 lui-même (en effet, le CD22 est porteur de *N*-glycanes terminés par du Sia₆LacNAc pouvant interagir avec un autre CD22 par reconnaissance homotypique) ont été décrites comme des ligands potentiels du CD22 (Hanasaki *et al.*, 1994 ; Stamenkovic *et al.*, 1991 ; Braesch-Andersen *et al.*, 1994 ; Walker *et al.*, 2008). De façon surprenante, il a été suggéré que la sialylation en α 2,6- des *N*-glycanes du CD22 puisse, dans le cadre d'une reconnaissance homotypique, moduler, voire abroger, l'activation des lymphocytes B en prévenant des interactions non désirées avec l'environnement moléculaire du CD22 (Braesch-Andersen *et al.*, 1994 ; Sgroi *et al.*, 1996). L'absence de l'un des deux composants du couple ligand / récepteur : Sia₆LacNAc / CD22 perturbe de façon sévère la réponse immunitaire médiée par les lymphocytes B matures. Ces propriétés ont été soulignées par la production de souris déficientes pour l'expression du CD22 (O'Keefe *et al.*, 1996 ; Otipoby *et al.*, 1996 ; Sato *et al.*, 1996) et pour l'expression de motifs Sia₆LacNAc par invalidation du gène codant ST6Gal I (Hennet *et al.*, 1998). Les souris dont le gène *mst6gal1* (le gène codant ST6Gal I) a été invalidé ne présentent aucune altération du développement et de la croissance mais présentent cependant une sévère immunodéficience. En effet, l'absence de motifs Sia₆LacNAc au niveau de la rate et des lymphocytes en particulier, liée à un défaut de l'activité de ST6Gal I, entraîne une réduction du taux d'IgM sériques, une perte de l'adhésion et de l'activation des lymphocytes B ainsi qu'une atténuation de la production d'anticorps dépendante et indépendante des antigènes T (Hennet *et al.*, 1998). Un phénotype similaire, avec toutefois une sévérité moindre, a été observé pour les souris déficientes en CD22 (O'Keefe *et al.*, 1996 ; Otipoby *et al.*, 1996 ; Sato *et al.*, 1996). Il est par ailleurs pertinent de noter qu'aucune autre fonction biologique ne semble être affectée par l'invalidation de ST6Gal I malgré une très large diversité de *N*-glycosylprotéines porteuses de motifs Sia₆LacNAc.

2.2.3. Un motif surexprimé au cours de la phase aiguë de l'inflammation

L'inflammation est une réaction locale des tissus vasculaires épithéliaux en réponse à une agression extérieure ou à un stress physique ou chimique (Gabay *et al.*, 1999). Au cours de ce processus, la phase aiguë, l'une des étapes précoces de l'inflammation, conduit à la modification quantitative et qualitative de l'expression de certaines protéines circulantes ou certaines protéines membranaires. Les protéines de la phase aiguë de l'inflammation sont produites par stimulation de cytokines pro-inflammatoires telles que l'interleukine-1 (IL1), l'IL6 ou le tumor necrosis factor α (TNF α).

Il a été rapporté voilà plus de vingt ans que la surexpression de ST6Gal I au niveau du foie et sa sécrétion sous forme soluble (tronquée au niveau de la région tige) constitue un élément associé à la phase aiguë de l'inflammation (Kaplan *et al.*, 1983 ; Dalziel *et al.*, 1999). Dans ce contexte inflammatoire, il a même été suggéré que les sialyltransférases, dont ST6Gal I, sous leur forme soluble dans le milieu puissent agir comme des lectines et se lier aux acides sialiques associés aux glycoconjugués présents à la surface des cellules empêchant ainsi la liaison à des pathogènes (Angata *et al.*, 2002). Plusieurs travaux ont également montré une augmentation de la sialylation en α 2,6- des glycoprotéines du plasma, également couplée à une défucosylation (Chavan *et al.*, 2005 ; Yasukawa *et al.*, 2005). Cette augmentation de l' α 2,6- sialylation des protéines est souvent associée à un rôle anti-inflammatoire. Par exemple, l'augmentation de la sialylation en α 2,6- des structures *N*-acétyllactosaminiques des *N*-glycanes des IgG module leur rôle lors de la réponse inflammatoire. (Kaneko *et al.*, 2006). Sialylée de cette manière, la partie Fc des IgG leur confère une activité anti-inflammatoire utilisée notamment dans le cadre du traitement par injection intraveineuse de patients souffrant de maladies auto-immunes (Nimmerjahn *et al.*, 2008). Ces propriétés anti-inflammatoires sont directement dépendantes de la sialylation en α 2,6- des épitopes LacNAc. Les IgG porteuses de Sia₆LacNAc sont dès lors capable de fixer une lectine spécifique (SIGN-R1) présente à la surface de macrophages spléniques. Cette interaction permet d'induire la sécrétion de médiateurs anti-inflammatoires qui ciblent les macrophages effecteurs situés sur le site de l'inflammation (Anthony *et al.*, 2008a & 2008b).

Enfin, un effet direct des cytokines pro-inflammatoires comme le TNF α et l'IL-1, et du lipopolysaccharide (LPS), une endotoxine bactérienne, a été montré sur des cellules endothéliales humaines en culture. Ces trois composés induisent une surexpression de l'enzyme hST6Gal I. Cette augmentation de l'expression de ST6Gal I est concomitante avec l'augmentation de l' α 2,6- sialylation des *N*-glycosylprotéines de surface de ces cellules,

témoignant à nouveau du rôle essentiel joué par l' α 2,6- sialylation et ST6Gal I au cours de l'inflammation (Hanasaki *et al.*, 1994).

2.2.4. Un épitope glycanique surexprimé dans de nombreux cancers

Les variations de la glycosylation et de la sialylation des chaînes oligosaccharidiques sont des modifications fréquemment observées dans de nombreux cas de cancers (Dall'Olio, 1996). L'augmentation d'expression de la sialylation en α 2,6- sous la forme du motif Sia₆LacNAc est souvent corrélée avec la progression tumorale, la diffusion métastatique et un mauvais pronostic de la maladie. La surexpression du Sia₆LacNAc et ses conséquences ont été rapportées dans les cancers du colon et du col de l'utérus, au cours de la leucémie myéloïde aigue, certaines tumeurs du cerveau et dans des carcinomes du sein (Sata *et al.*, 1991 ; Vierbuchen, 1995 ; Wang, 2001 ; Gessner, 1993 ; Skacel, 1991 ; Kaneko, 1996). Ces changements de sialylation ont été particulièrement étudiés dans les tissus cancéreux du colon où le déterminant Sia₆LacNAc (Sata *et al.*, 1991 ; Dall'Olio *et al.*, 1993 ; Yamashita *et al.*, 1995) et la sialyltransférase ST6Gal I (Gesner *et al.*, 1993 ; Dall'Olio *et al.*, 1989) sont fortement surexprimés et associés à un phénotype plus agressif de la maladie (Dall'Olio *et al.*, 1989 & 1993 ; Sata *et al.*, 1991). Il a également été montré que l'augmentation de la sialylation en α 2,6- du motif LacNAc des sous-unités β 1 des intégrines favoriserait l'adhésion sur le collagène de type 1 et la laminine, la mobilité et l'invasion des cellules cancéreuses (Gu *et al.*, 2004 ; Seales *et al.*, 2005). D'un autre coté, une hypo- α 2,6-sialylation des sous-unités β 1 des intégrines entraînerait une augmentation de l'adhésion à la fibronectine (Semel *et al.*, 2002) et favoriserait la prolifération et la différenciation tumorale (Chammas *et al.*, 1993 ; Gu *et al.*, 2004). De plus, la sialylation en α 2,6- de *N*-glycosylprotéines membranaires module les propriétés d'adhésion de cellules cancéreuses mammaires (Lin *et al.*, 2002). Enfin, Recchi et collaborateurs ont montré que hST6Gal I était surexprimée dans différents cas de tumeurs du sein. Cette surexpression étant corrélée avec un mauvais pronostic (Recchi *et al.*, 1998).

2.2.5. Un motif particulier dans le phénomène de clairance hépatique des *N*-glycosylprotéines

Au niveau du foie, les récepteurs identifiés par Ashwell et collaborateurs ont été initialement décrits comme des récepteurs d'asialoglycoprotéines (ASGP-R) ayant la particularité de fixer puis d'internaliser les glycoprotéines circulantes dont les glycanes se

terminent par un Gal(NAc) lié en β 1-4 (Ashwell *et al.*, 1982). Ainsi, le galactose terminal de ces glycoconjugués est un facteur déterminant pour leur demi-vie dans la circulation sanguine. Il a ainsi été suggéré que la sialylation en α 2,6- des galactoses terminaux a pour conséquence directe d'empêcher la reconnaissance et l'internalisation des glycoconjugués du sérum ainsi modifiés. La sialylation en α 2,6- des *N*-glycanes des glycoprotéines du sérum, sous forme de Sia₆LacNAc, joue ici un rôle « protecteur » et régulateur de la concentration relative des glycoprotéines sériques telles que les hormones.

2.3. Le motif Sia₆LacdiNAc

2.3.1. Le Sia₆LacdiNAc, un motif rare chez les mammifères et chez l'Homme

Cette structure constitue également une extension terminale de la périphérie des antennes des *N*-glycanes de type complexe. Chez l'Homme, les *N*-glycosylprotéines identifiées porteuses de la structure Sia₆LacdiNAc terminale sont l'urokinase, un activateur du plasminogène produit par le rein (Bergwerff *et al.*, 1995), la glycodéline, une hormone sécrétée par le placenta (Dell *et al.*, 1995) et les hormones hypophysaires (Weissharr *et al.*, 1991). Par ailleurs, le motif Sia₆LacdiNAc a également été identifié sur un activateur tissulaire du plasminogène produit par les cellules de mélanome de Bowes (Chan *et al.*, 1991) et sur la protéine C humaine recombinante produite par les cellules d'origine rénale HEK293 (Yan *et al.*, 1993). Le rôle de ce motif peu répandu s'apparente souvent à celui du Sia₆LacNAc en termes d'interactions protéiques et cellulaires. Pour autant, même si fonctionnellement la structure NeuAc α 2-6GalNAc- terminale ne semble pas se distinguer de la structure NeuAc α 2-6Gal, certaines données laissent supposer un rôle plus spécifique pour le motif Sia₆LacdiNAc.

Il apparaît cependant que plusieurs des protéines portant une structure Sia₆LacdiNAc terminale sont des *N*-glycosylprotéines libres dans le milieu extracellulaire et jouant un rôle lors de la reproduction et la fécondation.

2.3.2. La glycodéline

La glycodéline, aussi désignée par placental protein 14 (PP14) ou progesterone-associated endometrial protein (PAEP), est une glycoprotéine humaine sécrétée principalement par l'endomètre et retrouvée dans le système reproductif des primates. Quatre

isoformes de glycodéline ont été identifiées et classées en fonction de leur origine tissulaire et de leur glycosylation. La glycodéline-A (GdA) est sécrétée par l'endomètre et retrouvée dans le liquide amniotique (Seppälä *et al.*, 2002), la glycodéline-S (GdS) est synthétisée par les vésicules séminales et s'accumule dans le plasma séminal (Seppälä *et al.*, 1985), la glycodéline-F (GdF) a été identifiée dans le liquide folliculaire (Chiu *et al.*, 2003) et la glycodéline-C (GdC) est retrouvée dans la matrice extracellulaire du *cumulus oophorus* (tissu de soutien de l'ovocyte dans la cavité folliculaire) (Yeung *et al.*, 2006). Ces quatre glycoformes, dont l'expression est hautement régulée au cours du cycle menstruel (Julkunen *et al.*, 1986), possèdent toutes la même séquence en acides aminés mais ne diffèrent que par la nature de leur *N*-glycanes. La glycodéline possède trois sites de *N*-glycosylation, qui correspondent aux résidus d'asparagine 28, 63 et 85, et seules les Asn28 et Asn83 sont *N*-glycosylées. Parmi les quatre glycoformes, seule la GdA possède de façon majoritaire des *N*-glycanes terminés par du Sia₆LacdiNAc (Figure 10) (Dell *et al.*, 1995). *A contrario*, les *N*-glycanes de la GdS, qui constitue avec la GdA les deux formes majoritaires de glycodéline, sont dépourvus d'acides sialiques et sont plus largement fucosylés ou décorés par les *N*-glycanes de type oligomannosidique (Figure 10). La GdA est en outre la seule des quatre glycoformes de la glycodéline capable d'inhiber la liaison des spermatozoïdes à la zone pellucide (zona pellucida) et ceci de façon dépendante de la glycosylation et notamment de la sialylation (Oehninger *et al.*, 1995 ; Morris *et al.*, 1996). Enfin, la GdA est également décrite pour jouer un rôle dans la modulation de propriétés immunosuppressives et immunorégulatrices au cours de la fécondation / gestation en atténuant la réponse immunitaire déclenchée par l'implantation d'un embryon. Il a par ailleurs été montré que cette glycoforme inhibe la prolifération, induit l'apoptose des cellules T (Seppälä *et al.*, 2002) et réprime également l'activation des cellules Natural Killer (Okamoto *et al.*, 1991) et des cellules B (Yaniv *et al.*, 2003). L'induction de l'apoptose des cellules T, médiée par le récepteur CD45 (Rachmilewitz *et al.*, 1999), ainsi que l'inhibition de l'adhérence et de l'activation des cellules B (Yaniv *et al.*, 2003) sont directement imputées à la sialylation en α 2,6- présente uniquement sur l'isoforme GdA.

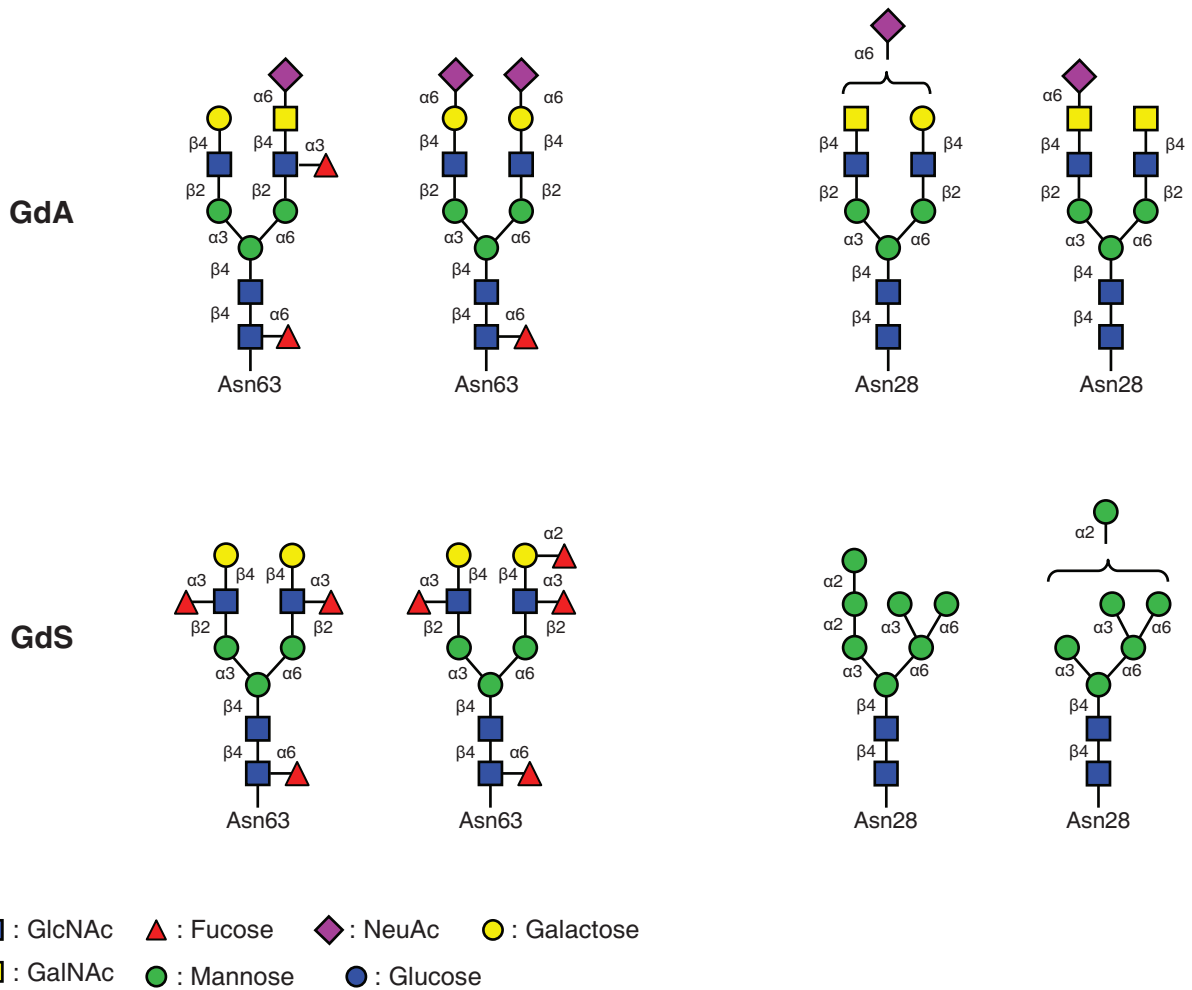


Figure 10 : N-glycanes majeurs de la glycodéline-A et de la glycodéline-S.

Les N-glycanes de la GdA sont majoritairement sialylés (Sia₆LacNAc ou Sia₆LacdiNAc), alors que les N-glycanes de la GdS sont majoritairement fucosylés ou oligomannosidiques.

2.3.3. L'hormone lutéinisante humaine, la hLH

L'hormone lutéinisante (LH) humaine, ou hLH, est un composé essentiel aux processus de reproduction. Cette hormone est produite par les glandes gonadotropes du lobe antérieur de l'hypophyse sous stimulation de la GnRH (Gonadotropin-releasing Hormone). La hLH est constituée de deux sous unités α et β . La sous unité α est une N-glycosylprotéine de 92 a.a. dont deux sites de N-glycosylation, les résidus Asn52 et Asn78, sont occupés. Par ailleurs, la sous unité α est commune aux autres hormones que sont la FSH (Follicle Stimulating Hormone), la TSH (Tyroid Stimulating Hormone) et la HCG (Hormone Chorionique Gonadotrope) (Pierce *et al.*, 1981). La sous unité β , propre à la hLH, est quant à elle une N-glycosylprotéine de 112 a. a. n'ayant qu'un site de N-glycosylation occupé. La

hLH est naturellement présente sous plusieurs glycoformes dans les milieux extracellulaires. Cependant, le site Asn78 de l' α -hLH est majoritairement occupé par des *N*-glycanes dont les antennes sont terminées par les motifs NeuAc α 2-6GalNAc β 1-4GlcNAc (Sia₆LacdiNAc) et NeuAc α 2-6Gal β 1-4GlcNAc (Sia₆LacNAc) (Weissharr *et al.*, 1991), alors que le site Asn52 de l' α -hLH et le site de la sous unité β sont habituellement occupés par des *N*-glycanes terminés par la séquence SO₄⁻4GalNAc β 1-4GlcNAc-. Il a été démontré que la présence d'acides sialiques sur les glycanes de la hLH conditionne sa demi-vie dans la circulation sanguine en jouant sur sa clairance au niveau du foie. Ces acides sialiques conditionnent également l'activité biologique de la hLH en modulant l'interaction ligand / récepteur (Burgon *et al.*, 1996). Il est par ailleurs intéressant de noter que le ratio des différentes glycoformes de hLH, plus ou moins riches en acide sialiques, varie au cours du cycle menstruel (Anobile *et al.*, 1998), postulant ainsi pour une régulation temporelle de la sialylation de cette hormone. Enfin, une étude récente a permis de mettre en évidence que la balance entre le niveau de SO₄⁻4GalNAc et d'acides sialiques présents sur les *N*-glycanes de la hLH modulerait la vitesse de clairance de l'hormone dans le flux sanguin (Wide *et al.*, 2009). Les auteurs suggèrent une compétition entre la GalNAc transférase et la Gal transférase au niveau du point d'attache sur la GlcNAc. Une compétition entre l' α 2,6-sialyltransférase et la sulfotransférase au niveau du point d'attache de la GalNAc n'est pour autant pas exclue.

2.3.4. Les activateurs du plasminogène

Les activateurs du plasminogène (PA) sont des glycoprotéines à activité protéolytique convertissant le plasminogène, une proenzyme inactive, en une forme active : la plasmine. La plasmine joue un rôle clé dans la fibrinolyse du caillot sanguin (Collen *et al.*, 1992). Par ailleurs, les activateurs du plasminogène sont exprimés dans de nombreux tissus et seraient d'une façon plus générale également impliqués dans la protéolyse de la matrice extracellulaire lors du développement et de la régénération tissulaire, comme par exemple lors de la migration cellulaire et le développement du système nerveux central (Friedman *et al.*, 1994).

2.3.4.1. L'urokinase, u-PA

L'urokinase est un activateur de plasminogène de type urinaire sécrété par les reins. L'étude de la glycosylation de l'unique chaîne *N*-glycanique a mis en évidence la présence de Sia₆LacdiNAc parmi les glycanes terminaux rencontrés. Très peu d'études ont été menées

concernant le rôle de la sialylation et du motif Sia₆LacdiNAc de l'urokinase. Seules des modifications de la clairance ont été observées en parallèle de modifications de sialylation de l'urokinase (Henkin *et al.*, 1991).

2.3.4.2. L'activateur de plasminogène tissulaire des cellules de Bowes, t-PA

L'activateur de plasminogène tissulaire (t-PA) est largement exprimé au niveau du tissu neuronal, à la fois dans le tissu neuronal en développement et au cours de la régénération du tissu neuronal (Seeds *et al.*, 1990 ; Qian *et al.*, 1993). Au niveau du système nerveux central adulte, le t-PA jouerait un rôle lors de la croissance axonale et de la migration cellulaire (Sappino *et al.*, 1993). Par ailleurs, le t-PA a été isolé à partir de plusieurs autres sources tels que le placenta humain (Rijken *et al.*, 1979), les cellules épithéliales murines C127 (Reddy *et al.*, 1987) et les cellules de Bowes, une lignée de mélanome humain d'origine neuroectodermale (Wallen *et al.*, 1983). C'est dans la lignée de Bowes que les études concernant les *N*-glycanes ont été réalisées et ont permis de mettre en évidence la présence de structures Sia₆LacdiNAc, majoritairement associées à l'un des trois sites de *N*-glycosylation occupés du t-PA (Chan *et al.*, 1991). Le rôle de ces motifs Sia₆LacNAc n'a cependant pas été étudié.

2.3.5. Un rôle spécifique pour la structure Sia₆LacdiNAc ?

Actuellement, aucun rôle particulier n'a clairement été démontré pour le motif Sia₆LacdiNAc des *N*-glycosylprotéines. Les propriétés décrites sont apparemment les mêmes que pour la structure Sia₆LacNAc et concernent surtout la modulation des propriétés de clairance, comme c'est notamment le cas pour l'hormone hypophysaire hLH (Burgon *et al.*, 1996) et pour l'urokinase (Henkin *et al.*, 1991). Dans le cas de l'urokinase, seule la présence de l'acide sialique, indépendamment de la structure sialylée terminale même, semble mise en cause (Henkin *et al.*, 1991). Concernant les propriétés immunosuppressives de la glycodéline GdA, elles sont imputées à la présence d'acide sialique liée en α 2,6- sur du galactose avant tout. La *N*-acétylgalactosamine plutôt que le galactose ne semble en rien jouer une fonction particulière.

Dans le cas de la hLH, il apparaît que la présence du motif SO₄⁻4GalNAc β 1-4GlcNAc entre en compétition avec le motif NeuAc α 2-6GalNAc β 1-4GlcNAc. Au niveau du foie, des récepteurs spécifiques du motif SO₄⁻4GalNAc β 1-4GlcNAc porté par la hLH ont été

caractérisés (Fiete *et al.*, 1991). La présence d'un acide sialique en lieu et place de la sulfatation pourrait ainsi jouer un rôle de modulateur de la clairance de la LH en inhibant l'interaction de l'épitope $\text{SO}_4\text{-}4\text{GalNAc}\beta\text{1-}4\text{GlcNAc}$ avec son récepteur spécifique des cellules hépatiques. Cette balance sulfatation / sialylation peut être considérée comme un élément de contrôle ultime, puisque situé à l'extrémité de la chaîne glycanique, vis-à-vis de propriétés de reconnaissance et d'internalisation par des récepteurs particuliers.

Jusqu'à présent, les structures $\text{Sia}_6\text{LacNAc}$ et $\text{Sia}_6\text{LacdiNAc}$ étaient considérées comme similaires vis-à-vis de leur influence sur la clairance des *N*-glycosylprotéines qui les portent au niveau des ASGP-R (récepteurs d'asialoglycoprotéines) hépatiques notamment. Cependant, il a été récemment décrit que les ASGP-R situés au niveau des hépatocytes de rat possédaient, de manière surprenante, une affinité élevée pour des glycoprotéines se terminant par la structure $\text{Sia}_6\text{LacdiNAc}$. Cette affinité pour le $\text{Sia}_6\text{LacdiNAc}$ retrouvée uniquement pour ces ASGP-R est comparable avec l'affinité observée pour les glycoprotéines se terminant par un Gal ou une GalNAc. Néanmoins, l'affinité de ces récepteurs pour des glycoprotéines porteuses de $\text{Sia}_6\text{LacdiNAc}$ serait 100 fois supérieure à celle de glycoprotéines porteuses de $\text{Sia}_6\text{LacNAc}$ terminale. Les auteurs concluent à un rôle différentiel joué par les ASGP-R dans la reconnaissance et l'internalisation de glycoprotéines porteuses de $\text{Sia}_6\text{LacdiNAc}$ ou de $\text{Sia}_6\text{LacNAc}$ terminal. (Park *et al.*, 2003 ; Park *et al.*, 2005 ; Steirer *et al.*, 2009). Les glycoprotéines circulantes décorées par du $\text{Sia}_6\text{LacdiNAc}$ pourraient être reconnues sélectivement parmi une masse d'épitopes $\text{Sia}_6\text{LacNAc}$ beaucoup plus abondante.

2.3.6. Biosynthèse des structures $\text{Sia}_6\text{LacNAc}$ et $\text{Sia}_6\text{LacdiNAc}$

Comme décrit dans ce chapitre, les structures $\text{Sia}_6\text{LacNAc}$ et $\text{Sia}_6\text{LacdiNAc}$ sont impliquées dans des processus aussi variés que l'inflammation, la clairance hépatique, l'adhésion cellulaire, la régulation de la demi-vie de glycoprotéines circulantes ou la modulation de la réponse immunitaire. L'expression de ces déterminants sialylés est quasi exclusivement liée à un type cellulaire ou un tissu donné, comme l'expression de $\text{Sia}_6\text{LacNAc}$ sur les *N*-glycanes du CD22 des lymphocytes B matures uniquement ou le $\text{Sia}_6\text{LacdiNAc}$ porté uniquement par les *N*-glycanes de la glycodéline-A produite par l'endomètre. Pour un tissu donné, l'expression de structures $\alpha\text{2,6}$ -sialylées peut également varier au cours du temps ou en réponse à un stimulus, comme la surexpression de motif $\text{Sia}_6\text{LacNAc}$ à la surface de cellules cancéreuses du colon par rapport au tissu sain ou une sialylation massive en $\alpha\text{2,6}$ - des glycoprotéines sériques due à une surexpression de hST6Gal I au cours de l'inflammation.

Ces exemples illustrent le fait que la synthèse de structures sialylées, et de motifs glycaniques plus généralement, résulte de mécanismes finement régulés.

La formation des structures Sia₆LacNAc et Sia₆LacdiNAc est directement corrélée à l'expression des glycosyltransférases impliquées dans leur biosynthèse. Chez l'Homme, deux sialyltransférases, hST6Gal I et hST6Gal II, sont capables de synthétiser ces motifs. Nous verrons dans le prochain chapitre que l'expression des glycosyltransférases et plus particulièrement celle des sialyltransférases est principalement régulée au niveau transcriptionnel, par des mécanismes impliquant l'utilisation de promoteurs alternatifs tissu-spécifiques. Durant ma thèse, je me suis focalisé sur l'étude de l'expression et de la régulation transcriptionnelle du gène *ST6GAL2*, le gène codant hST6Gal II.

3. LA REGULATION DE L'EXPRESSION DES α 2,6-SIALYLTRANSFERASES

3.1. La régulation transcriptionnelle des gènes de sialyltransférases

Les glycosyltransférases, et plus spécifiquement les sialyltransférases, n'ont généralement pas le profil d'expression tissulaire de gènes de ménage et présentent le plus souvent une expression tissu-spécifique. L'activité des sialyltransférases conditionne en effet le patron général de sialylation des glycoprotéines membranaires ou circulantes provenant d'un tissu donné. Les sialyltransférases ont communément un profil spécifique d'expression tissulaire. Ainsi, chacune des sialyltransférases caractérisées jusqu'à présent, chez l'Homme ou chez d'autres mammifères, présente une spécificité d'activité enzymatique mais aussi une spécificité d'expression tissulaire (Kitagawa *et al.*, 1994b ; Harduin-Lepers *et al.*, 2001 ; Taniguchi, 2008). De cette façon, certaines sialyltransférases sont typiquement exprimées pendant une période définie, comme ST8Sia II exprimée principalement au niveau du cerveau embryonnaire de souris (Angata *et al.*, 1997 ; Yoshida *et al.*, 1996) ou ST6Gal I surexprimée au niveau du foie durant la phase aiguë de l'inflammation (Kaplan *et al.*, 1983 ; Dalziel *et al.*, 1999), ou dans un tissu en particulier telle que hST6Gal I exprimée au niveau du foie et des lymphocytes B matures (Kitagawa *et al.*, 1994b).

Les premières études réalisées par Northern-Blotting sur l'expression des gènes codant des sialyltransférases ont mis en évidence une expression tissu-spécifique des ARN messagers (ARNm) codant ces enzymes, suggérant que l'expression transcriptionnelle de ces gènes serait régulée de manière tissu-spécifique (Kitagawa *et al.*, 1994a & 1994b ; Kim *et al.*, 1996 ; Sasaki *et al.*, 1993). La régulation transcriptionnelle est le mode de régulation majeur de l'expression des sialyltransférases. Généralement, les exons qui composent la séquence codante d'une sialyltransférase donnée sont identiques et la protéine codée toujours la même. Seuls les exons non codants situés dans la région 5' de l'ARNm sont différents. Ces exons en 5' non traduits témoignent de l'utilisation de promoteurs tissu-spécifiques ou qui répondent à une situation physiologique particulière comme lors de l'inflammation (Figure 11). Ainsi, plusieurs formes d'ARNm ne différant que pour leur extrémité 5' non codante ont été identifiées pour la majorité des gènes de sialyltransférase comme : hST3Gal V (Kim *et al.*, 2001), hST3Gal VI (Taniguchi *et al.*, 2001), hST8Sia I (Kang *et al.*, 2007), mST6Gal I (Dalziel *et al.*, 2001) et rST6Gal I (Wen *et al.*, 1992). Néanmoins, ce modèle ne s'applique pas à tous les gènes de sialyltransférases. En effet, les six formes d'ARNm isolées pour le

gène codant hST3Gal IV peuvent être classées en trois types de transcrits différents en fonction de leur site d'initiation de la traduction (ATG) (Kitagawa *et al.*, 1996 ; Taniguchi *et al.*, 1998 & 1999). Ces différences au niveau de la séquence codante sont toutefois sans conséquence apparente sur l'activité enzymatique puisqu'elles ne se traduisent que par une différence de quelques acides aminés (11 a.a. entre la forme la plus longue et la forme la plus courte) dans la région cytoplasmique et transmembranaire de hST3Gal IV. La signification biologique de ces différences n'a pas été réellement étudiée.

ST6Gal I est sans nul doute la sialyltransférase la plus étudiée, autant sur le plan de son activité enzymatique et que sur le plan de la régulation transcriptionnelle du gène qui la code. La régulation du gène humain *ST6GAL1*, codant hST6Gal I, constitue un paradigme de la régulation transcriptionnelle des gènes de sialyltransférases. Dans les paragraphes suivants, ses différents transcrits ainsi que les promoteurs qui leurs sont associés seront décrits. Je m'intéresserai ensuite au gène *ST6GAL2*, dont l'expression tissulaire est plus restreinte.

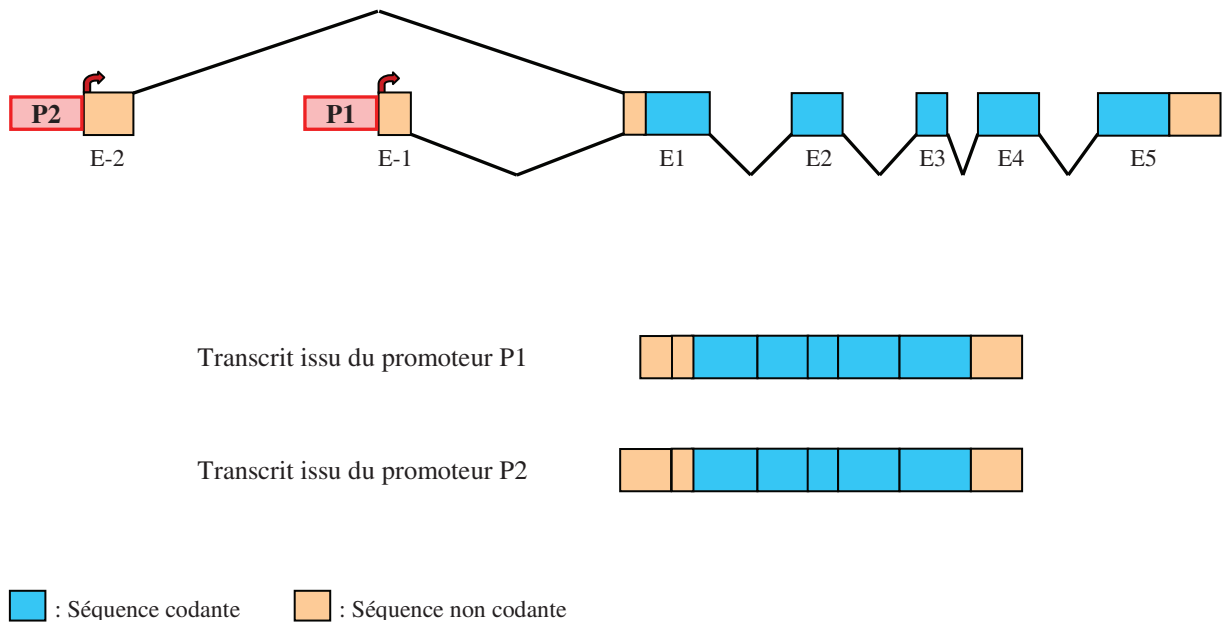


Figure 11 : Schéma général de la régulation transcriptionnelle des gènes de sialyltransférases.

Sous l'influence de promoteurs tissulaires spécifiques différentes formes d'ARNm sont transcrites ne différant que pour leurs extrémités 5' non traduites. La sialyltransférase issue de la traduction est identique. Schéma non à l'échelle.

3.2. *ST6GAL1*

Historiquement, ST6Gal I est la première sialyltransférase à avoir été isolée et identifiée, initialement à partir du colostrum bovin (Paulson *et al.*, 1977), puis à partir de foie de rat (Weinstein *et al.*, 1982) et chez l'Homme en 1989 (Lance *et al.*, 1989 ; Grundmann *et al.*, 1990). Cette enzyme a très longtemps été considérée comme l'unique sialyltransférase capable de transférer un résidu d'acide sialique en position C6 du galactose des motifs de type *N*-acétyllactosaminique (Gal β 1-4GlcNAc) et de former le motif Sia₆LacNAc. L'enzyme ST6Gal I ainsi que le gène codant cette sialyltransférase ont été intensivement étudiés notamment chez la souris, le rat et l'Homme. La structure Sia₆LacNAc, et plus généralement la sialylation en α 2,6- des *N*-glycosylprotéines, est omniprésente chez les mammifères. Il n'est donc pas surprenant de retrouver le gène codant ST6Gal I exprimé dans quasiment tous les tissus. Néanmoins, chez l'Homme, *ST6GAL1*, n'est pas ou pratiquement pas exprimé au niveau des testicules, du colon et du cerveau, et une expression importante est retrouvée au niveau des lymphocytes B et du foie (Kitagawa *et al.*, 1994b ; Kaneko *et al.*, 1995). Trois promoteurs ont été décrits et conduisent à une expression ubiquiste, spécifique du foie, et spécifique des lymphocytes B matures (Wang *et al.*, 1993 ; Aas-Eng *et al.*, 1995 ; Lo *et al.*, 1996 & 1999).

3.2.1. Organisation génomique

Les gènes codant ST6Gal I chez la souris, le rat, le bovin et l'Homme ont une organisation génomique très similaire. Tous les transcrits des gènes codants ST6Gal I chez ces espèces ont en commun cinq exons codants et un sixième exon non codant situé en amont du premier exon codant (Figure 12) (Kalcheva *et al.*, 1997 ; Wen *et al.*, 1992 ; Mercier *et al.*, 1999 ; Wang *et al.*, 1993). L'organisation génomique totale est représentée en détails sur la figure 13 pour le gène humain, *ST6GAL1*. Pour ce gène, les trois autres exons non codants, EX, EY et EZ situés en 5' sont transcrits en fonction des promoteurs P1, P2 et P3 utilisés. L'ensemble de ces neuf exons sont, chez l'Homme, répartis sur plus de 144 kpb sur le chromosome 3 (3q21-q28). Les gènes de sialyltransférases sont habituellement répartis sur des distances génomiques relativement grandes, 80 kbp pour le gène codant rST6Gal I (ST6Gal I de rat) et bST6Gal I (ST6Gal I bovine). En outre, la région non traduite en 3' du dernier exon est très longue et s'étend sur plus de 2 kpb.

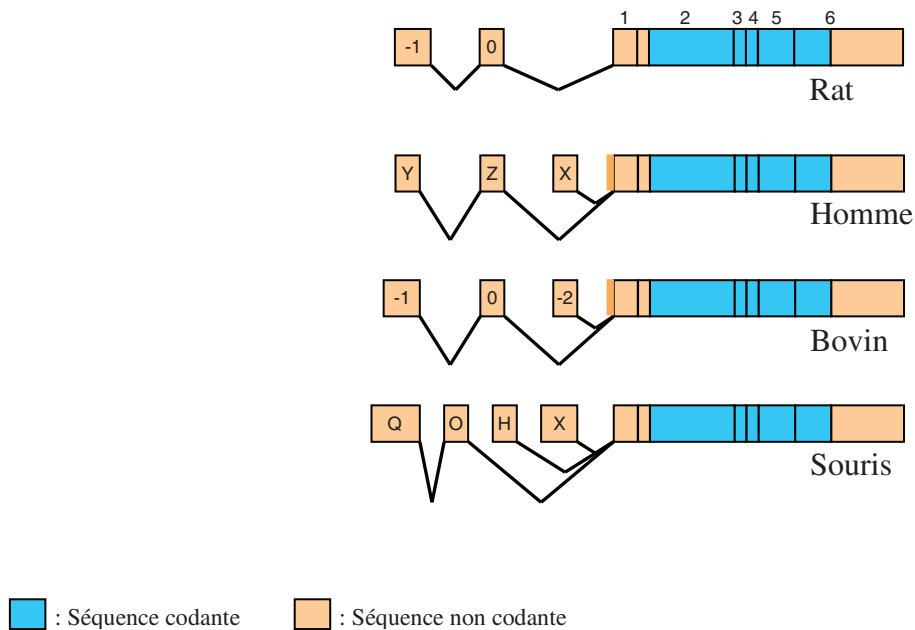


Figure 12 : Schéma de l'organisation en 5' des gènes codant ST6Gal I chez le rat, l'Homme, le bovin et la souris (Extrait et adapté de Dall'Olio, 2000).

Représentation non à l'échelle. Les exons 1, 2, 3, 4, 5 et 6 correspondent aux six exons composant la partie commune des transcrits de chaque espèce. Les exons -1 et 0, Y, Z et Y, -1, 0 et -2, Q, O, H et X correspondent aux différents exons non traduits en 5' de chaque espèce.

Les trois exons constituant la région 5' non traduite du gène humain, *ST6GALI*, sont localisés entre 3,7 kb et plus de 110 kb en amont du premier exon codant, E1, (Figure 13). Les exons EY et EZ composent les transcrits ubiquistes et l'exon EX initie les transcrits lymphocytaires (Figure 14C). Enfin, une courte séquence de 24 pb, appelée E1', située en 5' de l'exon E1 n'est retrouvée que dans les transcrits hépatiques de *ST6GALI* (Figures 13 et 14B). L'extrémité 5' de la séquence de l'exon E1 est ainsi le site accepteur d'épissage des transcrits initiés par EX ou EY (Figures 13 et 14A). A l'exception de transcrits de 2,2, 2,3 et 4,3 kb observés au niveau du foie, les transcrits de *ST6GALI* ont tous une taille proche de 4,7 kb (Kitagawa *et al.*, 1994b). Il est également intéressant de remarquer que chez le rat un transcrit de 3,6 kb est spécifiquement exprimé au niveau des reins (Paulson *et al.*, 1989 ; O'Hanlon *et al.*, 1989), ce transcrit n'ayant pas son homologue chez l'Homme. Les particularités des trois types de transcrits, constitutif, hépatique et lymphocytaire, des régions promotrices et des différents facteurs intervenant dans l'expression transcriptionnelle tissu-spécifique de *ST6GALI* sont détaillés dans les paragraphes qui suivent.

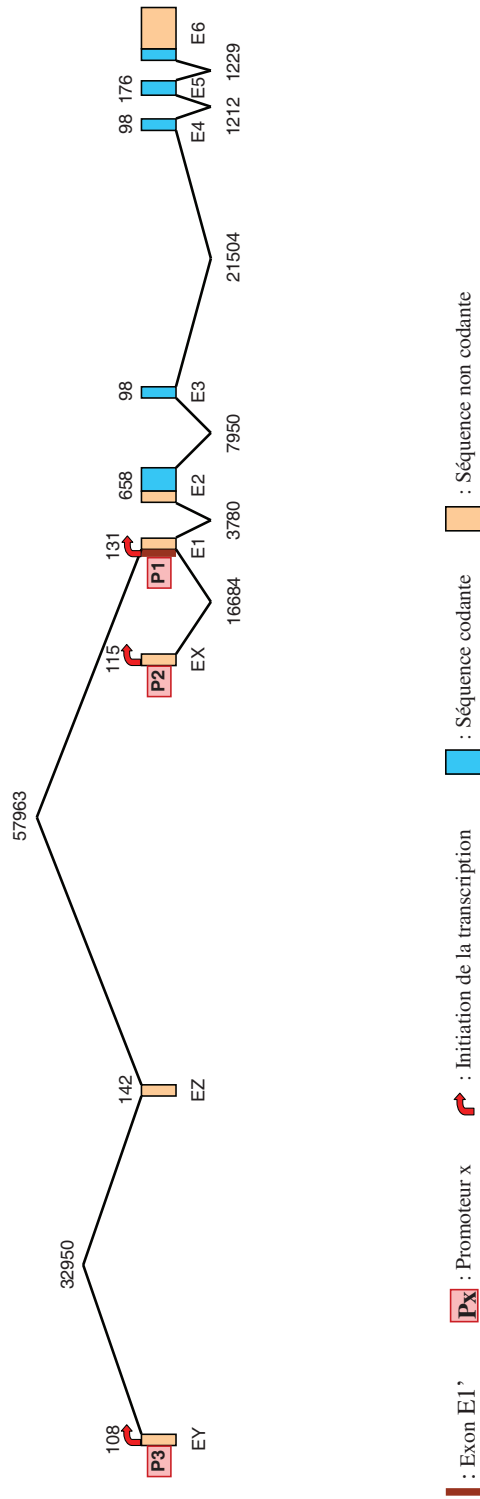


Figure 13 : Schéma de l'organisation génomique du gène *ST6GALI*. Organisation génomique du gène *ST6GALI* sur le chromosome 3. Les rectangles pleins correspondent aux exons et E1, E2, E3, etc... représentent leur dénomination. Les chiffres au dessus des exons représentent leur taille en pb. Les traits pleins noirs correspondent aux régions introniques, leur taille en pb est indiquée au dessus ou en dessous

3.2.2. Promoteur de l'expression constitutive de *ST6GALI*

Le promoteur P3 (Figures 13 et 14D) conduit à l'expression basale de *ST6GALI* dans tous les tissus. Il a été identifié et caractérisé par Wang et collaborateurs en 1993 (Wang *et al.*, 1993). Les transcrits de 4,7 kb issus de ce promoteur comportent les exons non-codants EY et EZ situés respectivement, à 90 kpb et 57 kpb, de E1 le premier exon commun à tous les transcrits (Figure 14D). Dans les cellules hépatiques et les lymphocytes B matures un promoteur spécifique (P1 et P2 respectivement) est actif et est responsable de l'expression de la majorité des transcrits exprimée dans chacun de ces tissus. Pour autant, les transcrits issus du promoteur constitutif sont également exprimés mais constituent une part minoritaire de l'expression de *ST6GALI* dans ces types cellulaires. La région promotrice en amont de EY est comparable aux régions promotrices de gènes de ménages, elle ne contient pas de boîte TATA ni de boîte CCAAT mais comporte plusieurs sites de fixation pour le facteur de transcription général SP1 (Wang, 1993). Une étude menée par Taniguchi et collaborateurs a montré qu'au cours de la différenciation des cellules HL-60 par du diméthyle sulfoxyde (DMSO), de l'acide rétinoïque (all *trans*-retinoic acid, ATRA) ou de l'acétate-myristate de phorbol (PMA), l'expression des transcrits contenant EY et EZ, seul type de transcrit exprimé dans cette lignée, était fortement diminuée (Taniguchi *et al.*, 1998). Peu d'informations sont néanmoins disponibles sur ce promoteur.

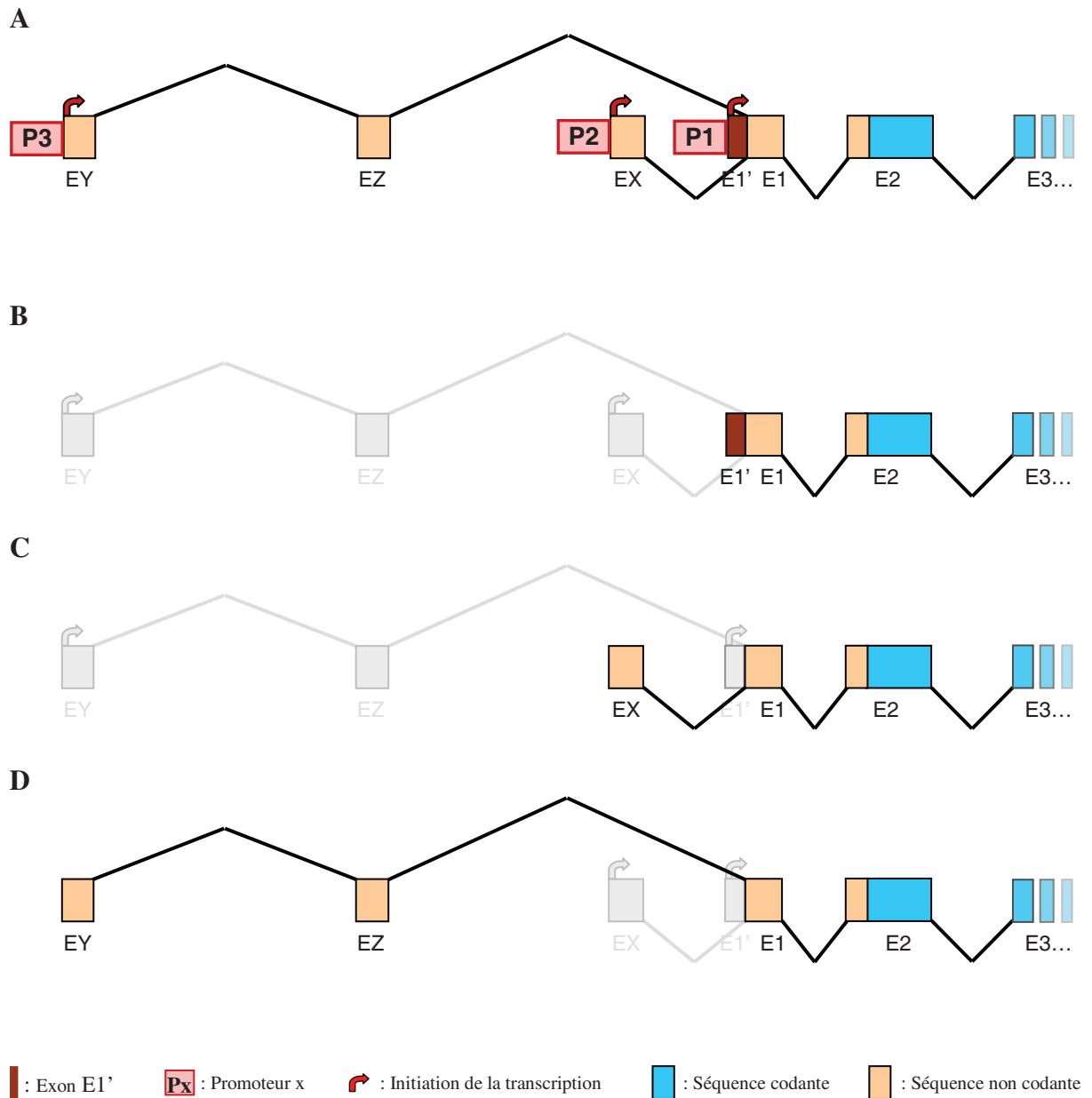


Figure 14 : Schéma des transcrits alternatifs en 5' du gène *ST6GAL1*.

A, organisation génomique en 5' du gène *ST6GAL1*. **B**, transcrit issu du promoteur hépatique P1 : E1', E1, E2... **C**, transcrit issu du promoteur lymphocytaire P2 : EX, E1, E2... **D**, transcrit issu du promoteur constitutif P3 : EY, EZ, E1, E2...

3.2.3. Promoteur lymphocytaire de *ST6GAL1*

ST6GAL1 est exprimé de manière spécifique dans les lymphocytes B matures sous la forme de transcrits comportant à leur extrémité 5' l'exon EX. Cet exon est situé à plus de 16 kpb de l'exon E1 (Figure 14B). Ce type de transcrit, d'une longueur de 4,7 kb, a uniquement

été retrouvé dans les cellules de la lignée B et le promoteur P2 a été décrit comme permettant la transcription de EX dans cette lignée cellulaire (Lo *et al.*, 1996 & 1999). L'étude de la région promotrice de P2 a permis de mettre en lumière plusieurs éléments impliqués dans l'activité de ce promoteur. Les travaux réalisés ont montré qu'une boîte TATA, localisée à l'intérieur d'une courte séquence de 68 pb située immédiatement en 5' du site d'initiation de la transcription (TSS), est suffisante pour une activité promotrice minimale. Cependant, plusieurs sites de fixation pour des éléments de réponses activateurs situés dans les séquences proches en 5' et en 3' du site d'initiation de la transcription de EX ont été décrits comme nécessaires à la transcription dans les cellules matures de la lignée B. Ainsi, les sites de fixation pour les facteurs de transcription AP2 et NF- κ B, situés en amont du TSS de EX, ont été décrits comme des éléments activateurs de la transcription. Les éléments de régulation GATA et LyF-1, situés à l'intérieur de la séquence de EX, seraient quant à eux des éléments activateurs spécifiques des stades de développement terminaux de la lignée B et actifs uniquement dans les cellules B matures (Lo *et al.*, 1996 & 1999).

3.2.4. Promoteur hépatique de *ST6GAL1*

Au niveau du foie, plusieurs tailles de transcrits ont été rapportées. En plus d'une taille de transcrit de 4,7 kb retrouvée dans tous les tissus, des ARNm de *ST6GAL1* de 2,2, 2,3 et 4,3 kb ont été révélés par Northern-Blot (Kitagawa *et al.*, 1994b). L'extrémité 5' de la forme majeure du transcrit hépatique, de 4,7 kb, démarre sa transcription au niveau de l'exon E1', une courte séquence de 24 pb située immédiatement en 5' de l'exon E1 commun à tous les types de transcrits (Figure 14B). La jonction E1' / E1 est le site accepteur d'épissage pour les transcrits constitutifs et spécifiques des lymphocytes B (Figure 14A). Les transcrits hépatiques sont sous le contrôle du promoteur P1. Cette région promotrice ne contient pas de boîte TATA ni de boîte CCAAT ni d'élément GC. Néanmoins, elle contient deux séquences nucléotidiques appelées séquences initiatrices (Inr) (Aas-Eng *et al.*, 1995 ; Lo *et al.*, 1996), éléments retrouvés pour la transcription de gènes sans boîte TATA (Smale *et al.*, 1989). De plus, cette région contient des sites de fixation pour des facteurs de transcription hépatiques tels que les Hepatocyte Nuclear Factor : HNF-1, HNF-4 et HNF-5, ainsi que trois sites de fixation pour des éléments de réponse à l'interleukine-6 (IL-6) : deux sites IL6-RE-BP (IL-6 response element binding protein) et un site NF-IL6 (nuclear factor-IL-6) (Aas-Eng *et al.*, 1995 ; Lo *et al.*, 1996). Occupés, ces sites induisent l'augmentation de la transcription de *ST6GAL1* en réponse à l'IL-6 au niveau du foie au cours des premières étapes de

l'inflammation (Gauldie *et al.*, 1987). Par ailleurs, des travaux ont montré l'implication du promoteur hépatique dans la surexpression de *ST6GAL1* dans les cas de cancers colorectaux. Il a été montré que le facteur de transcription HNF1 était exprimé, non seulement au niveau du foie, mais aussi au niveau des tissus sains ou cancéreux du colon (Eskinazi *et al.*, 1999). Cependant, la présence de DCoH/PCD, le co-facteur de dimérisation de HNF1, facilite sa fixation à l'ADN et son activité promotrice (Rhee *et al.*, 1997). DCoH/PCD est retrouvée de façon abondante dans les échantillons de tissus cancéreux de colon et non dans les tissus sains (Eskinazi *et al.*, 1999). DCoH/PCD et HNF-1 seraient ainsi impliqués dans la surexpression du transcrit hépatique de *ST6GAL1* dans les cas de cancer colorectaux.

3.3. *ST6GAL2*

Contrastant avec ST6Gal I isolée et caractérisée il y a plus de trente ans, ST6Gal II n'a été identifiée que beaucoup plus récemment. Une recherche systématique effectuée sur le génome humain par homologie de séquences avec les différents sialylmotifs a permis de mettre en évidence le gène *ST6GAL2* (Takashima *et al.*, 2002 ; Krzewinski-Recchi *et al.*, 2003). Depuis, cette seconde α 2,6-sialyltransférase a également été identifiée et caractérisée chez la souris (Takashima *et al.*, 2003) et le bovin (Laporte *et al.*, 2009). La séquence du gène codant ST6Gal II a par ailleurs été reportée chez plusieurs autres espèces telles que le rat ou le poisson zèbre *Danio rerio* (Harduin-Lepers *et al.*, 2005).

Chez l'Homme, les protéines hST6Gal I et hST6Gal II ont une homologie de séquence en acides aminés de plus de 48 % au niveau de la région catalytique. Cependant, hST6Gal II se distingue de hST6Gal I par : (1) sa spécificité de substrat, hST6Gal II préférant les oligosaccharides libres et la structure LacdiNAc comme substrats accepteurs (Tableau 3) (Krzewinski-Recchi *et al.*, 2003 ; Rohfritsch *et al.*, 2006), et (2) son expression tissulaire, *ST6GAL2* étant exprimée de façon beaucoup plus restreinte, principalement au niveau du cerveau adulte et embryonnaire (Krzewinski-Recchi *et al.*, 2003 ; Takashima *et al.*, 2002).

3.3.1. Organisation génomique

La région codante du gène humain codant ST6Gal II, *ST6GAL2*, est également composée de cinq exons, répartis sur plus de 41 kpb sur le chromosome 2 (q11.2-q12.1). L'organisation génomique de la région codante est similaire à celle du gène *ST6GAL1*, avec néanmoins quelques différences (Figure 15). L'exon E1, le premier exon codant de *ST6GAL2*,

est plus grand que celui de *ST6GAL1* (E2) (915 pb contre 606 pb), il code la région tige de l'enzyme qui est plus longue de 103 acides aminés pour hST6Gal II. De plus, la région 3' non traduite de l'exon E5 fait plus de 5 kb de long (Figure 15B).

Les extrémités 5' non traduites du gène *ST6GAL2* n'étaient pas encore connues lorsque cette thèse a débuté en 2006, seuls quelques EST (Expressed Sequence Tag) retrouvées dans les banques de données et les séquences publiées dans les travaux de Takashima et collaborateurs, et Krzewinski-Recchi et collaborateurs (Takashima *et al.*, 2002 ; Krzewinski-Recchi *et al.*, 2003) laissaient présager de régions génomiques non codantes situées à plus de 42 kbp de E1 (Figures 15B et 16). Les différentes EST appartenant à des transcrits de *ST6GAL2* ont été retrouvées dans l'hippocampe (AB 058780), dans deux adénocarcinomes ovariens (BC 008680 et BE 612797), dans l'amygdale (DA 176100) et dans un neuroblastome (BX 353569). Les extrémités 5' de ces séquences sont toutes localisées dans une région de 2 kbp, à plus de 42 kbp de l'exon E1, et regroupées en deux groupes : BX 353569 et AB 058780 d'une part, et BC 008680, BE 612797 et DA 176100 d'autre part (Figure 16). Cette répartition indiquerait que le gène *ST6GAL2* pourrait être transcrit sous la forme d'au moins deux types de d'ARNm comportant des extrémités 5' non traduites différentes. Ces séquences suggèrent également l'utilisation de deux promoteurs potentiels (Krzewinski-Recchi *et al.*, 2003).

Chez la souris et le bovin, les séquences codantes de *mst6gal2* et *bst6gal2* sont aussi composées de cinq exons (Figure 17) (Takashima *et al.*, 2003 ; Laporte *et al.*, 2009). Les travaux de Takashima et collaborateurs sur *mst6gal2* n'ont révélé la présence que d'une séquence de 420 pb constituant un exon unique non traduit en 5' localisé à 36 kbp du premier exon codant. De même, pour *bst6gal2*, Laporte et collaborateurs n'ont rapporté que la présence d'un exon en 5' non traduit d'une longueur de 132 pb et distant de 27 kbp de E1 (Figure 17).

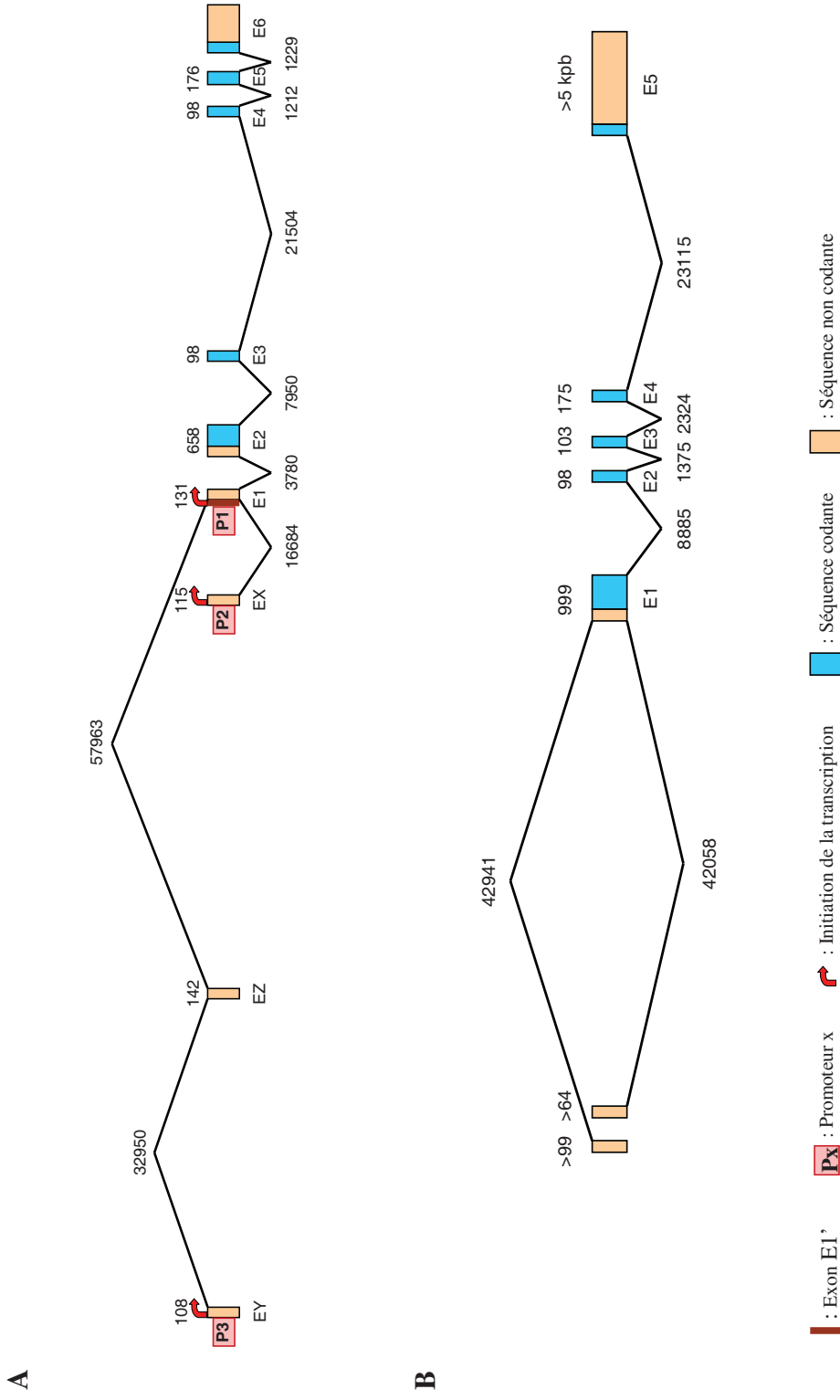


Figure 15 : Schéma comparatif de l'organisation génomique des gènes humains *ST6GAL1* et *ST6GAL2*.

A, organisation génomique du gène *ST6GAL1* sur le chromosome 3. **B**, organisation génomique du gène *ST6GAL2* sur le chromosome 2. Les rectangles pleins représentent les exons, E1, E2, E3, E4, E5 et E6 représentent leur dénomination. Les chiffres au dessus des exons représentent leur taille en pb. Les traits noirs correspondent aux régions introniques, leur taille en pb est indiquée au dessus ou en dessous.

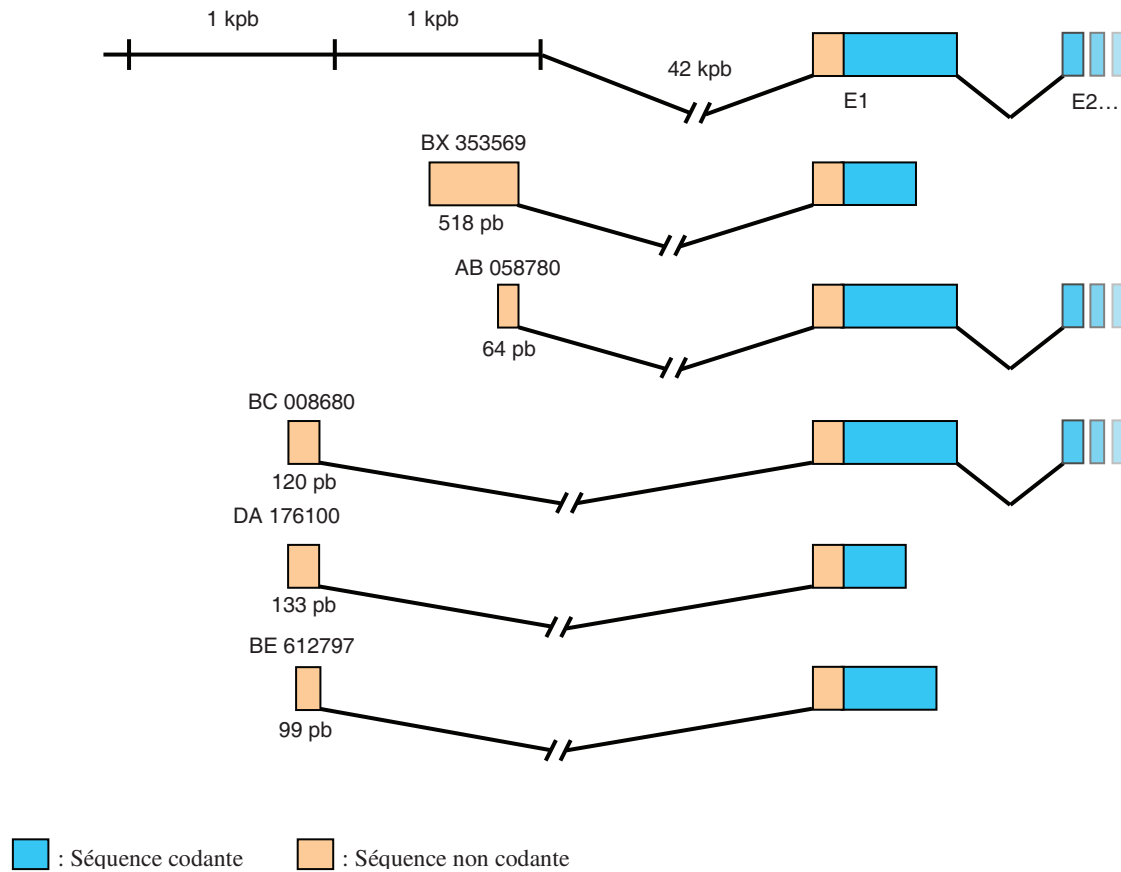


Figure 16 : Schéma des EST identifiées comme appartenant à des transcrits du gène *ST6GAL2*.
Représentation schématique de cinq différentes EST dont les extrémités 5' non traduites sont situées à plus de 42 kpb de E1

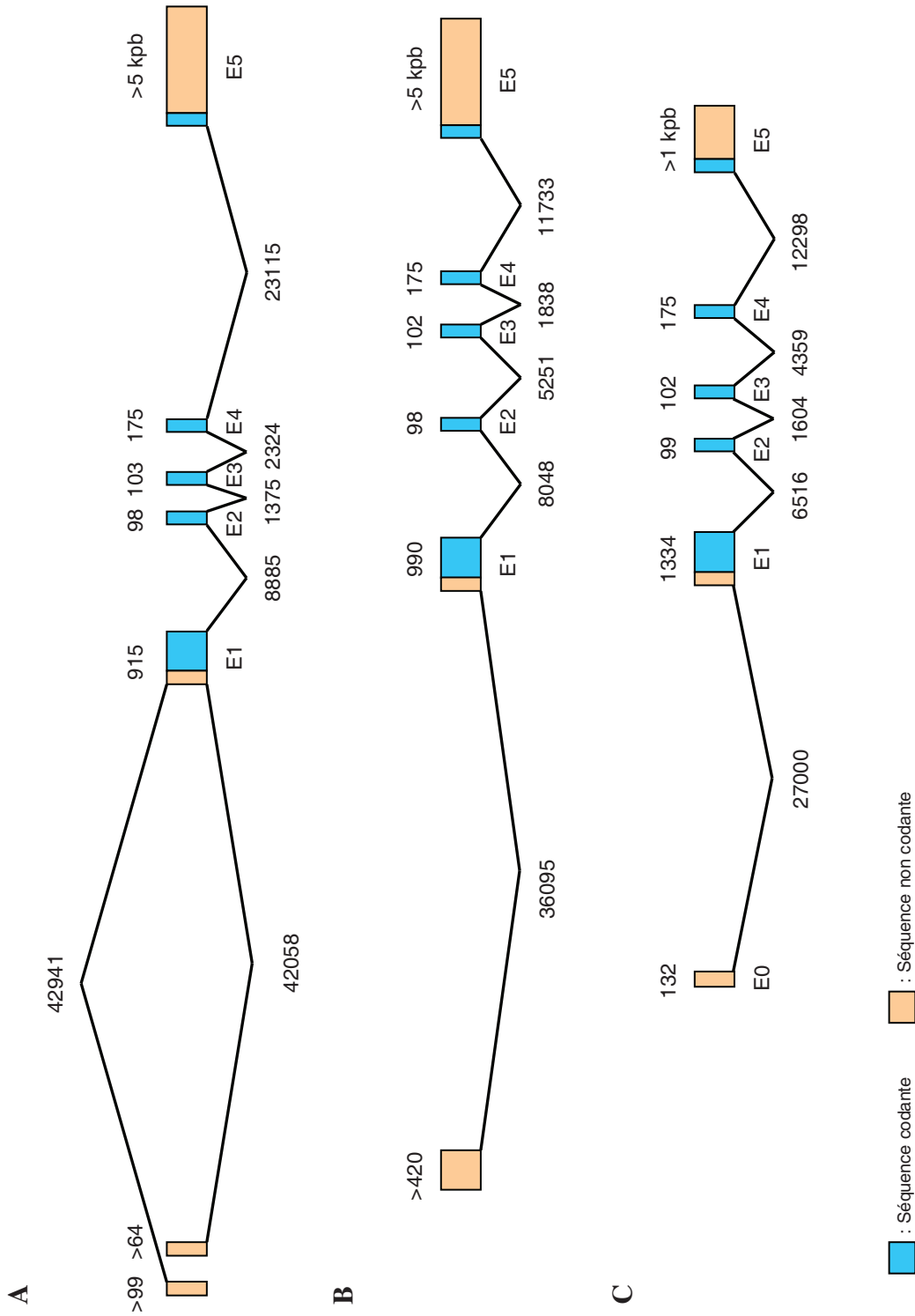


Figure 17 : Schéma de l'organisation génomique du gène humain (*ST6GAL2*), murin (*mst6gal2*) et bovin (*bst6gal2*) codant *ST6Gal II*. A, le gène humain : *ST6GAL2*. B, le gène murin : *mst6gal2*. C, le gène bovin : *bst6gal2*.

3.3.2. Expression tissulaire du gène humain, *ST6GAL2*

Comme l'ont montré les nombreuses études réalisées sur *ST6GAL1* et ses homologues murins ces trente dernières années, ce gène est à la fois largement transcrit dans quasiment tous les tissus et à la fois spécifiquement exprimé dans certains tissus comme le foie, les lymphocytes B matures et le rein (Wang *et al.*, 1993 ; Aas-Eng *et al.*, 1995 ; Lo *et al.*, 1996 & 1999 ; Paulson *et al.*, 1989). Par opposition avec le gène codant ST6Gal I, les études menées sur le gène codant ST6Gal II chez l'Homme, la souris et le bovin ont révélé un profil d'expression tissulaire plus restreint (Krzewinski-Recchi *et al.*, 2003 ; Takashima *et al.*, 2002 & 2003 ; Laporte 2009). Pour chacun des gènes étudiés, le tissu cérébral semble être le lieu privilégié de l'expression des gènes codant ST6Gal II (Figures 18 et 19). Chez l'Homme, *ST6GAL2* est presque exclusivement exprimé au niveau du cerveau adulte et embryonnaire. Le Northern-Blot réalisé par Krzewinski-Recchi et collaborateurs n'a révélé la présence d'un transcrit de *ST6GAL2* d'environ 8 kb uniquement au niveau du cerveau (Krzewinski-Recchi *et al.*, 2003). *ST6GAL2* est, dans une moindre mesure, également exprimé au niveau des testicules, du colon, de l'intestin, du poumon, de la glande thyroïde et de certains tissus fœtaux comme le cerveau, le rein, le foie et le thymus (Figure 18) (Takashima *et al.*, 2002 ; Krzewinski-Recchi *et al.*, 2003). Cette expression discrète de *ST6GAL2* pourrait être interprétée comme une compensation de l'expression ubiquiste mais néanmoins très faible de *ST6GAL1* dans certains tissus tels que le tissu cérébral ou les testicules (Kitagawa *et al.*, 1994b). Par ailleurs, la plupart des EST correspondant à des transcrits de *ST6GAL2* témoignent également d'une expression prépondérante au niveau du cerveau (AB 058780, DA 176100 et BX 353569).

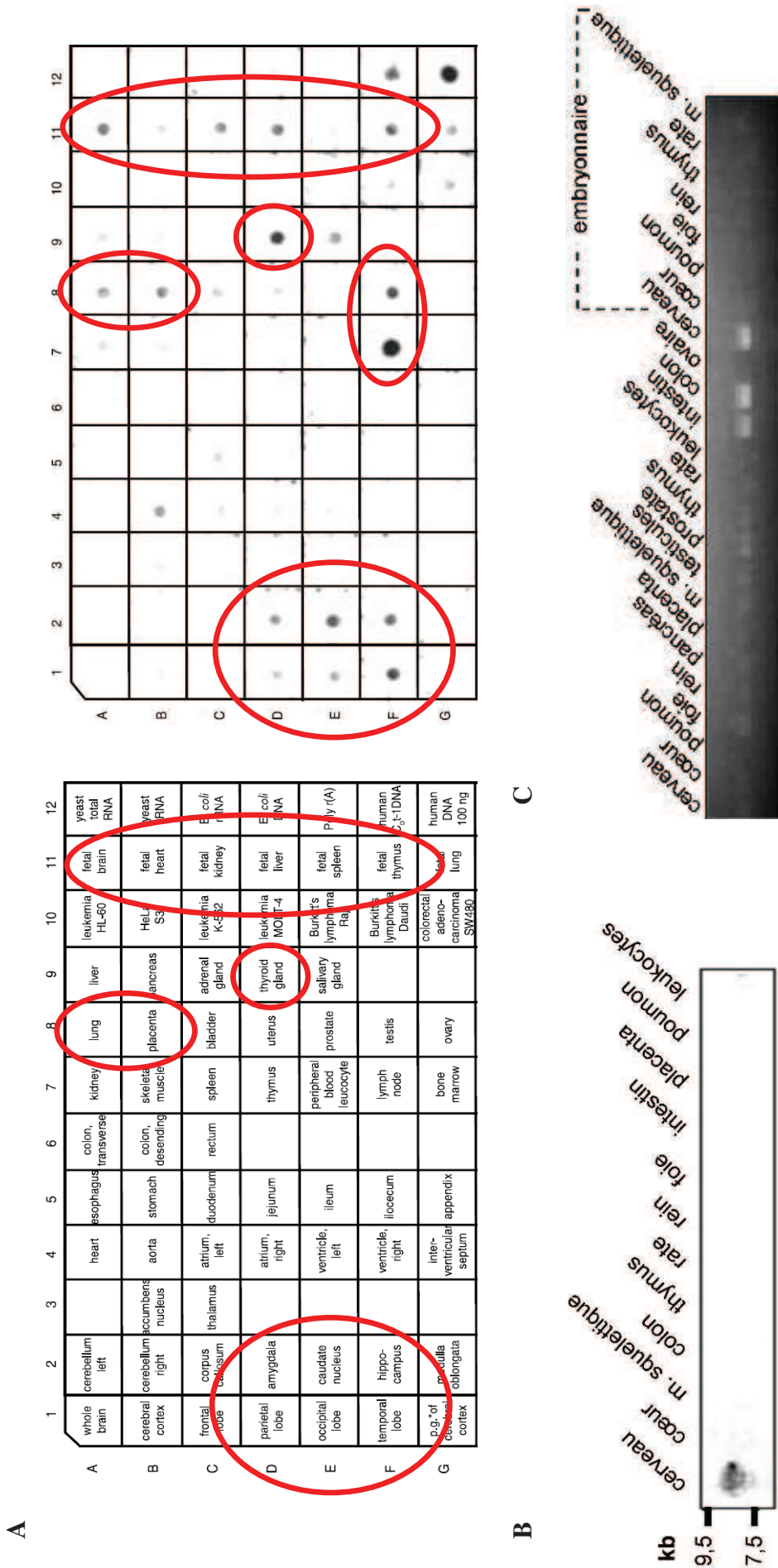


Figure 18 : Expression tissulaire du gène ST6GAL2

A, analyse d'un panneau multi tissus pour l'expression de ST6GAL2. Préparation commerciale (Clontech) d'ARN et d'ADN humain criblée avec une séquence d'ADNc radio marqué de ST6GAL2. B, analyse en Northern-Blot de l'expression de ST6GAL2. Préparation commerciale (Clontech) d'ARN de 12 tissus humains criblée avec un ADNc radio marqué de ST6GAL2. A et B extraits et adaptés de Krzewinski-Recchi *et al.*, 2003. C, analyse en PCR semi-quantitative de l'expression de ST6GAL2 de divers tissus humains. Extrait et adapté de Takashima *et al.*, 2002. m. squelettique, muscle squelettique ; *, paracental gyrus

3.3.3. Le gène murin, *mst6gal2*

Chez la souris, un Northern-Blot réalisé par Takashima et collaborateurs a montré que *mst6gal2* est exclusivement exprimé au niveau du cerveau. Par PCR semi quantitative, les auteurs ont affiné leur résultat et montré que le gène codant ST6Gal II est non seulement fortement exprimé au niveau du cerveau adulte mais également au niveau des tissus embryonnaires totaux d'embryons de 11, 15 et 17 jours. Une très faible expression a aussi été détectée dans la rate, le poumon et le muscle squelettique (Figures 19A et 19B) (Takashima *et al.*, 2003). De plus, une étude portant sur l'analyse de l'expression de plus de 140 glycosyltransférases lors du développement du cerveau embryonnaire et adulte a été réalisée plus récemment. Les auteurs montrent que *mst6gal2*, à l'opposé de la majorité des autres gènes codant des sialyltransférases, dont *mst6gal1*, est surexprimé au cours du développement cérébral (Ishii *et al.*, 2007). Dans cette étude, l'expression de *mst6gal2* est comparable à celle de *mst8sia2*, le gène codant mST8Sia II, une α 2,8-sialyltransférase connue pour être spécifiquement exprimée au cours du développement embryonnaire du cerveau (Yoshida *et al.*, 1995). Hormis une séquence de 420 pb située à 36 kbp de l'exon E1, aucune étude n'a pour le moment été publiée sur l'identification des extrémités 5' non traduites du gène *mst6gal2* et la caractérisation de sa (ou de ses) régions promotrices.

Le Allen Institute for Brain Science est un organisme dont les recherches sont focalisées sur l'étude du système nerveux central chez l'Homme et la souris. Leur site Internet (<http://www.brain-map.org>) propose en accès libre de nombreuses images d'hybridation *in-situ* révélant l'expression de plusieurs milliers de gènes sur des coupes sagittales de cerveau adulte murin. Parmi les gènes que les chercheurs ont étudiés, figurent *mst6gal1* et *mst6gal2*. L'hybridation *in-situ* a été réalisée sur la lignée murine C57BL/6J sur des souris adultes de 56 jours. Les hybridations sont réalisées à partir de sondes nucléotidiques marquées à la digoxigénine (DIG), « un matériels et méthodes » complet est disponible à <http://mouse.brain-map.org/pdf/ABADDataProductionProcesses.pdf>. Les images présentées figure 20 montrent nettement que le gène *mst6gal2* est préférentiellement exprimé au niveau du tissu cérébral par rapport à *mst6gal1*.

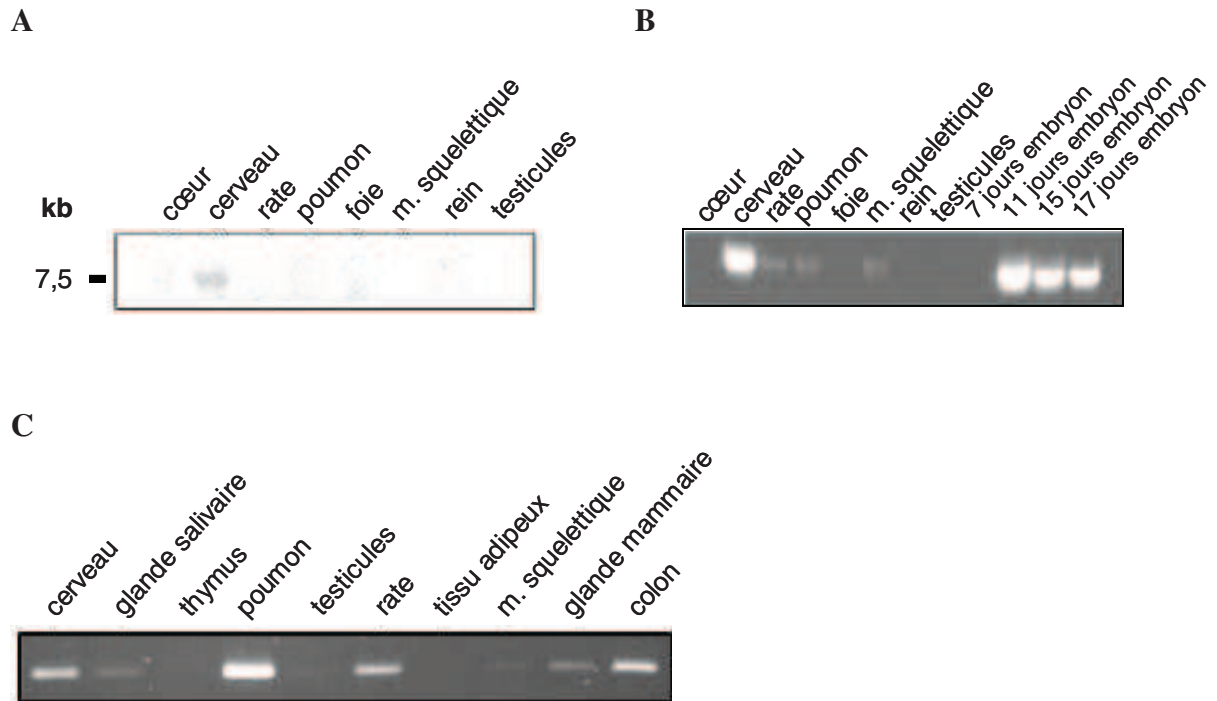


Figure 19 : Expressions tissulaires des gènes *mst6gal2* et *bst6gal2*.

A, analyse en Northern-Blot de l'expression de *mst6gal2*. Préparation commerciale (Clontech) d'ARN de 8 tissus murins criblée avec un ADNc radio marqué de *mst6gal2*. **B**, analyse en PCR semi-quantitative de l'expression de *mst6gal2* de divers tissus adultes et d'embryons de souris. **A** et **B**, extraits et adaptés de Takashima *et al.*, 2003. **C**, analyse en PCR semi-quantitative de l'expression de *bst6gal2* de divers tissus bovins. Extrait et adapté de Laporte *et al.*, 2009. m. squelettique, muscle squelettique.

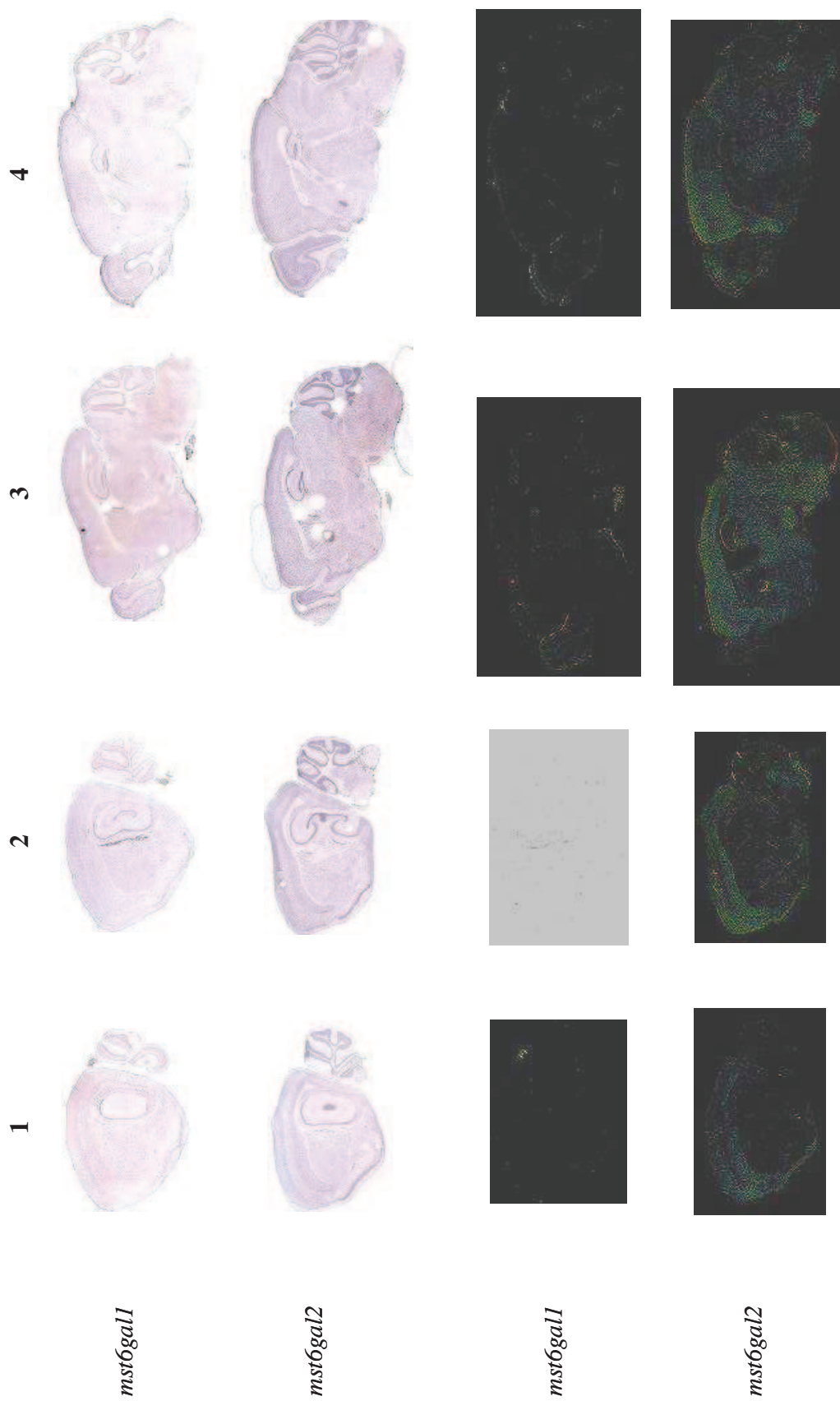


Figure 20 : Expression des gènes *mst6gal1* et *mst6gal2* dans le cerveau murin par hybridation *in-situ*. Les images ont été extraites du site Internet du Allen Institut for Brain Science (<http://www.brain-map.org>). Les chiffres 1, 2, 3 et 4 correspondent à des coupes sagittales distantes respectivement d'environ 3, 4 mm, 3 mm, 1,5 mm et 0,6 mm de l'axe sagittal du cerveau.

3.3.4. Le gène bovin, *bst6gal2*

Très récemment, l'expression tissulaire du gène bovin codant ST6Gal II, *bst6gal2*, a été rapportée par Laporte et collaborateurs. Leurs travaux ont montré que *bst6gal2* est exprimé au niveau du poumon, du cerveau, du colon, de la rate, de la glande salivaire et de la glande mammaire (Figure 19C) (Laporte *et al.*, 2009). Par ailleurs, trois types de transcrits, différents pour leurs extrémités 5' non traduites, ont été identifiés et appelés a_1 , a_2 et b (Figure 21). Deux de ces transcrits, a_1 et b, sont exprimés exclusivement au niveau du cerveau alors que le transcrit a_2 a été retrouvé dans tous les tissus dans lesquels *bst6gal2* est exprimé (la glande salivaire, le poumon, la rate, la glande mammaire et le colon) (Laporte *et al.*, 2009). Les transcrits a_1 et a_2 sont issus de deux séquences de 132 et 25 pb, respectivement, et d'un épissage alternatif de l'exon E1 excluant ses 430 premières bases (Figure 21). Ces deux transcrits sont sous le contrôle du promoteur P_A . Le transcrit b est quant à lui composé dans sa partie 5' uniquement de l'exon E1 sous sa forme entière. Le promoteur P_B est responsable de l'initiation du transcrit de type b (Figure 21). Les études préliminaires réalisées par Laporte et collaborateurs sur les promoteurs P_A et P_B rapportent la présence de sites de fixation potentiels pour les facteurs de transcription généraux SP1, AP-1 et NF-1, et pour CREB et SMAD des facteurs de transcription impliqués dans des fonctions neuronales (Laporte *et al.*, 2009).

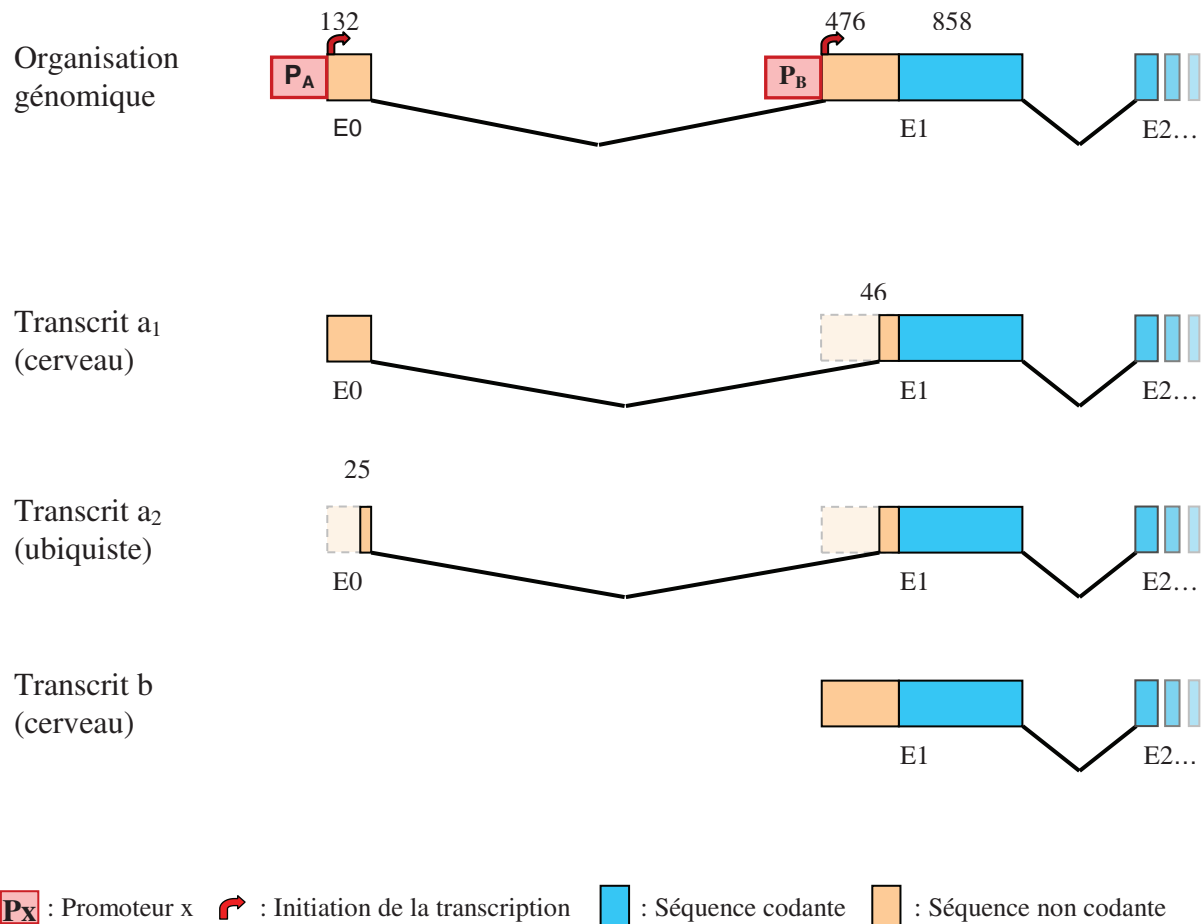


Figure 21 : Organisation génomique de l'extrémité 5' de *bst6gal2* (adapté de Laporte *et al.*, 2009). Organisation génomique en 5' de *bst6gal2*. Les transcrits a₁ et a₂ sont sous le contrôle du promoteur P_A et le transcrit b sous le contrôle du promoteur P_B. Les transcrits a₁ et b sont exprimés uniquement dans le cerveau et le transcrit a₂ dans les poumons, la glande salivaire, la glande mammaire, la rate et le colon (Laporte *et al.*, 2009). Les chiffres situés au dessus des exons représentent leur longueur respective en pb.

3.4. DSiaT, la ST6Gal de drosophile, une ST6Gal I/II

Chez les diptères, et plus particulièrement chez *Drosophila melanogaster*, une unique sialyltransférase a été identifiée, appelée DSiaT (Koles *et al.*, 2004). Les tests d'activités *in vitro* et les homologies de séquences ont montré que cette sialyltransférase appartient à la famille des α 2,6-sialyltransférases (ST6Gal). Le gène codant cette sialyltransférase est appelé *DSiaT*. L'arbre phylogénétique de la famille des ST6Gal montre que les deux sous familles ST6Gal I et ST6Gal II sont présentes chez tous les vertébrés, mais que chez les arthropodes les ST6Gal présentes n'appartiennent ni à la sous famille I ni à la II (Figure 9). Les ST6Gal identifiées chez les différentes espèces de diptères sont les orthologues du gène ancestral

existant avant l'événement de duplication qui a donné les deux sous-familles chez tous les vertébrés (Harduin-Lepers *et al.*, 2005).

3.4.1. Activité enzymatique

Les tests d'activité enzymatiques réalisés *in vitro* avec DSiaT montrent que l' α 2,6-sialyltransférase de drosophile est active sur une gamme de substrats accepteurs plus étendue que ses homologues vertébrés ST6Gal I ou ST6Gal II. DSiaT transfère sur des structures *N*-acétyllactosaminiques de type 2 (Gal β 1-4GlcNAc) mais également, avec une activité moindre, sur des motifs *N*-acétyllactosaminiques de type 1 (Gal β 1-3GlcNAc). *In vivo*, il a également été montré que DSiaT était capable d'autosialyler ses *N*-glycanes. De manière intéressante, la sialyltransférase de drosophile possède une activité enzymatique similaire à hST6Gal II vis-à-vis du LacdiNAc comme substrat accepteur préférentiel (Tableau 4). L'efficacité de transfert de DSiaT est deux fois plus importante pour le LacdiNAc par rapport au LacNAc (Koles *et al.*, 2004). De plus, DSiaT n'a aucune activité détectable envers les glycolipides comme substrat accepteur.

3.4.2. Organisation génomique

L'organisation génomique de *DSiaT* est comparable à celle des gènes codant ST6Gal II chez les vertébrés. Tout comme *ST6GAL2*, *mst6gal2* et *bst6gal2*, *DSiaT* est composé de cinq exons codants (Harduin-Lepers *et al.*, 2005). L'organisation en 5' de ce gène n'a pas encore été caractérisée (Figure 22).

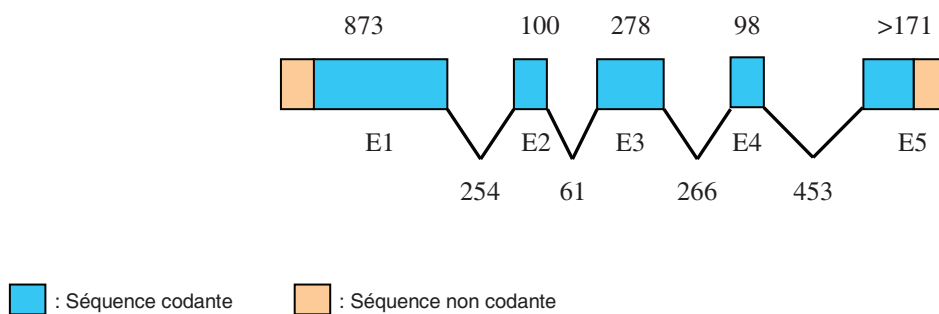


Figure 22 : Schéma de l'organisation génomique de *DSiaT*.
Les cinq exons codants du gène *DSiaT* sont répartis sur 2,5 kpb.

3.4.3. Expression tissulaire

Koles et collaborateurs ont montré que le profil d'expression de *DSiaT* était particulièrement discret. En effet, seule une expression au niveau des crêtes neurales au cours de stades tardifs du développement larvaire ainsi qu'une expression localisée au niveau des lobes oculaires à l'état adulte ont pu être mis en évidence (Figure 23) (Koles *et al.*, 2004 & 2008). A nouveau, ce profil d'expression restreint se rapproche d'avantage du profil d'expression génique observé pour les gènes *ST6GAL2*, *mst6gal2* et *bst6gal*, par opposition à de celui retrouvé pour *ST6GAL1* et ses homologues murins (souris et rat). Bien que de récentes études aient démontré chez *D. melanogaster* la présence de β 4GalNAc transférases fonctionnelles (Haines *et al.*, 2005) ainsi que d'une sialyltransférase, DSiaT, dont le substrat accepteur préférentiel est le LacdiNAc (Koles *et al.*, 2004), l'existence et la nature de structures sialylées demeurent controversées. Le profil d'expression de *DSiaT* plaide cependant pour une fonction précise au cours du développement embryonnaire du système nerveux central de la drosophile.

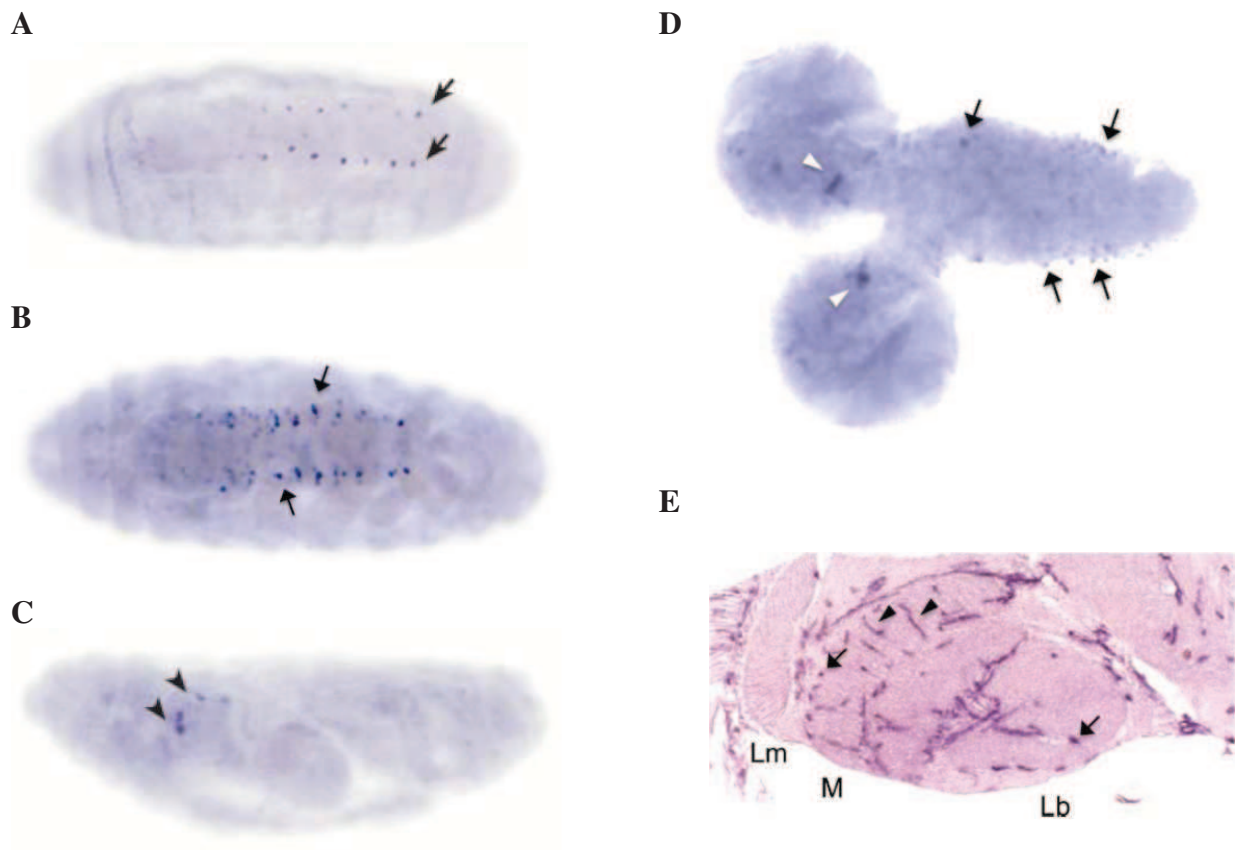


Figure 23 : Profil d'expression embryonnaire de *DSiaT*, hybridation *in situ* (Extraits de Koles *et al.*, 2004 & 2008).

A, stade 16 du développement embryonnaire, vue ventrale. Expression de *DSiaT* dans des chaînes de cellules à l'intérieur du système nerveux central (flèches). **B**, stade 17 du développement embryonnaire, vue ventrale. Davantage de cellules du système nerveux central expriment *DSiaT*. **C**, stade 17 du développement embryonnaire, vue latérale. Expression de *DSiaT* au niveau du cerveau embryonnaire (têtes de flèches). **D**, cerveau du 3^{ème} instar larvaire. Les têtes de flèches blanches indiquent l'expression de *DSiaT* dans les hémisphères cérébraux. **E**, section horizontale au cryotome d'une tête adulte. Les transcrits de *DSiaT* sont détectés dans les corps cellulaires (flèches) de même qu'au niveau des prolongements de ces cellules, présumés être des axones ou des dendrites (têtes de flèches). Les régions du lobe optique sont dénommées : Lm, lamina ; M, medulla ; Lb, lobula. **A**, **B**, **C** et **D**, la face antérieure est à gauche, **E** la face antérieure est sur le haut.

3.5. Objectif du travail de thèse

L' α 2,6-sialylation de structures *N*-acétyllactosaminiques est une modification post-traductionnelle commune observée pour de nombreuses *N*-glycoprotéines ou oligosaccharides libres. Toutefois, de tels motifs sont rares, pour ne pas dire inexistant, parmi les glycoconjugués appartenant au tissu cérébral. A l'exception d'une glycoforme de l'hormone hypophysaire LH qui arbore une structure Sia₆LacdiNAc, la grande majorité des sialoglycoconjugués associés au tissu cérébral portent des acides sialiques liés en α 2,3- et/ou en α 2,8-. La poly α 2,8-sialylation de la NCAM notamment est l'une des formes de sialylation les mieux documentées et associée à la neurogenèse, à la plasticité et au développement neuronal (Durbec *et al.*, 2001 ; Bonfanti, 2006). La présence de glycoconjugués porteur de motifs Sia₆LacNAc ou Sia₆LacdiNAc et leur(s) rôle(s) dans le développement neuronal restent encore à démontrer. Malgré un patron d'expression tissulaire étendu, *ST6GAL1* n'est que très faiblement exprimée au niveau du cerveau (Kitagawa *et al.*, 1994b ; Takashima *et al.*, 2002) ou par des cellules humaines neuronales et gliales en culture (Kaneko *et al.*, 1995 & 1996). Cependant, le profil d'expression des gènes codants ST6Gal II chez la souris, le bovin et chez l'Homme révèle clairement une expression préférentielle au niveau du système nerveux central adulte et fœtal (Takashima *et al.*, 2002 & 2003 ; Laporte *et al.*, 2009 ; Krzewinski-Recchi *et al.*, 2003), témoignant d'un rôle privilégié et d'un contrôle fin de la régulation de ces gènes au niveau du cerveau. Les travaux d'Ishii et collaborateurs ont également montré que le gène murin *mst6gal2* est l'un des trois gènes de sialyltransférases, avec *mst8sia2* et *mst6galnac5* surexprimé au cours du développement embryonnaire du tissu cérébral (Ishii *et al.*, 2007). Cette augmentation d'expression comparable à celle du gène codant ST8Sia II, *mst8sia2*, décrit pour être largement exprimé lors des étapes précoces du développement du cerveau chez la souris (Yoshida *et al.*, 1995).

De plus, hST6Gal II semble avoir conservé certaines propriétés « ancestrales » retrouvées pour l'unique sialyltransférase de drosophile en termes de spécificité enzymatique et d'expression tissulaire. Si la fonction de ST6Gal II demeure encore inconnue, il apparaît néanmoins que cette fonction semble être conservée au cours de l'évolution depuis les arthropodes jusqu'aux mammifères supérieurs.

Les travaux réalisés au laboratoire ont permis l'identification et la caractérisation de la seconde α 2,6-sialyltransférase humaine, hST6Gal II. Les données publiées dans la littérature

sur les gènes humain, murin et bovin codant ST6Gal II et sur le gène de drosophile codant une ST6Gal, aux propriétés proches de ST6Gal II, indiquent clairement une expression tissulaire restreinte principalement détectée au niveau du tissu cérébral adulte et embryonnaire (Takashima *et al.*, 2002 & 2003 ; Laporte *et al.*, 2009 Krzewinski-Recchi *et al.*, 2003 ; Koles *et al.*, 2004). Les travaux réalisés au cours de cette thèse avaient pour objectif d'étudier l'expression et la régulation transcriptionnelle du gène humain *ST6GAL2* dans un contexte « cérébral ». A travers un modèle cellulaire en culture de neuroblastome, la lignée SH-SY5Y, mon travail a permis d'identifier les extrémités 5' non traduites et de caractériser les régions promotrices associées à la transcription de ce gène. J'ai également pu étudier *in vivo* l'expression transcriptionnelle et protéique de *ST6GAL2* sur plusieurs échantillons du cortex cérébral humain. Enfin, mes travaux ont montré pour la première fois que le gène *ST6GAL2*, ses différents transcrits et la protéine hST6Gal II étaient fortement surexprimés lors de la différenciation en neurones de cellules en culture.

Résultats

INTRODUCTION

En 2003, une seconde β -galactoside- α 2,6-sialyltransférase humaine a été identifiée. Cette sialyltransférase, appelée hST6Gal II, a été enzymatiquement caractérisée *in vitro*. Le gène *ST6GAL2* est composé de cinq exons codants répartis sur le chromosome 2 (q11.2-q12.1) (Figure 15B, Introduction bibliographique). Les travaux réalisés par Krzewinski-Recchi et collaborateurs ainsi que Takashima et collaborateurs, ont montré que le profil d'expression de ce gène était assez restreint et qu'il était préférentiellement exprimé au niveau du système nerveux central (Krzewinski-Recchi *et al.*, 2003). De plus, aucune des lignées cellulaires testées (les lignées de carcinome mammaire GI-101, ovarien GI-102, pulmonaires LX-1 et GI-117 et d'adénocarcinome pancréatique GI-103, prostatique PC3 et colique CX-1 et GI-112) par Takashima et collaborateurs n'exprimaient *ST6GAL2* (Takashima *et al.*, 2002). Deux séquences rapportées par Takashima et collaborateurs ainsi que quelques EST (Expressed Sequence Tag) présentes dans les banques de données (AB 058780, BC 008680, BE 612797, DA 176100 et BX 353569) faisaient mention de séquences 5' non traduites appartenant au gène *ST6GAL2* et localisées à plus de 42 kpb de l'exon E1 sur le chromosome 2 (Figure 15B, Introduction bibliographique). Comme nous l'avons vu dans l'introduction, les extrémités 5' non traduites des gènes de sialyltransférases reflètent l'utilisation de promoteurs tissulaires spécifiques. L'identification expérimentale et l'étude des séquences non traduites en 5' du gène *ST6GAL2* constituent une première étape dans l'appréhension des mécanismes régulant la transcription de ce gène et dans la compréhension du rôle physiologique de cette sialyltransférase

Une partie des résultats présentés dans ce manuscrit a fait l'objet d'une publication dans *Glycoconjugate Journal* (voir Annexes).

1. EXPRESSION TRANSCRIPTIONNELLE DU GÈNE *ST6GAL2* DANS DIFFÉRENTES LIGNÉES CELLULAIRES EN CULTURE

Les données de la littérature décrivent le gène *ST6GAL2* et son homologue murin, comme étant préférentiellement exprimés au niveau du cerveau (Krewinski-Recchi *et al.*, 2003 ; Takashima *et al.*, 2003). Afin d'identifier le meilleur modèle cellulaire pour l'étude de l'expression et de la régulation du gène *ST6GAL2*, plusieurs lignées cellulaires humaines en culture au laboratoire et d'origines tissulaires différentes, et en particulier certaines lignées issues du système nerveux central, ont été criblées pour leur expression du gène *ST6GAL2*. Nous avons testé les lignées de neuroblastome SH-SY5Y et Kelly, la lignée de glioblastome T98G, la lignée de glioblastome / astrocytome U118, la lignée pro-neuronale de carcinome embryonnaire NT2/D1, les cellules rénales HEK, les cellules monocytaires THP-1, la lignée d'adénocarcinome du col de l'utérus HeLa, la lignée d'adénocarcinome colorectal Caco-2, la lignée de cellules épithéliales de cancer du sein MDA-MB-231, la lignée de cancer du sein BT20 et les cellules de mélanome SK-Mel28. L'expression globale des transcrits de *ST6GAL2* a été mesurée par PCR en temps réel, avec la chimie SYBR Green®, à l'aide d'amorces situées dans la séquence 3' non traduite. Dans toutes ces lignées cellulaires, l'expression de *ST6GAL2* est en général faible, voire inexistante. Le signal correspondant à la détection de l'expression du gène *ST6GAL2* est généralement visible à des valeurs de Ct (Cycle threshold) élevées lors de mesures par Q-PCR. Le Ct est le nombre de cycles de PCR pour lequel le signal fluorescent, correspondant à l'amplification du gène cible, atteint une valeur seuil (cette valeur est déterminée dans la phase exponentielle d'amplification du signal) détectable par rapport au bruit de fond. Plus la quantité d'ADNc du gène cible est faible dans le mélange réactionnel de départ plus le nombre de cycles pour atteindre la valeur seuil de fluorescence est élevé. Un Ct élevé signifie que le gène cible est faiblement exprimé. Nous avons choisi ici de normaliser l'expression de *ST6GAL2* par rapport à l'expression du gène *HPRT* (Hypoxantine phosphoribosyltransférase). Le gène *HPRT* est l'un des gènes de référence les plus faiblement exprimés (Zhang *et al.*, 2005 ; Valente *et al.*, 2009), et la détection de son expression en Q-PCR s'effectue, tout comme pour celle de *ST6GAL2*, à des valeurs de Ct élevées. Des valeurs de Ct proches entre le gène d'intérêt et son gène de référence sont nécessaires pour la précision des calculs lors de l'analyse des résultats. En effet, une trop grande différence de Ct entre les deux gènes augmente les erreurs et l'approximation des calculs finaux. Le gène de référence *RPLP0* (Ribosomal protein large P0) avait été utilisé

lors des premières mises au point de PCR en temps réel avec la chimie SYBR Green® (Groux-Degroote *et al.*, 2008). Cependant les résultats montraient une différence de Ct de dix cycles entre l'expression de *ST6GAL2* et de *RPLP0*, différence que nous avons considérée trop importante et qui nous a fait changer de gène de référence. L'expression transcriptionnelle du gène *ST6GAL1* a également été quantifiée pour chacune de ces lignées.

Les analyses réalisées sur l'expression du gène *ST6GAL2* par Northern-Blot, MTA (Multiple tissue expression array) et PCR semi-quantitatives par Krzewinski et collaborateurs (Krewinski-Recchi *et al.*, 2003), et de Takashima et collaborateurs (Takashima *et al.*, 2002) (Figure 18 Introduction bibliographique) montrent que ce gène est préférentiellement exprimé dans le système nerveux central ou embryonnaire. *ST6GAL2* n'est pas exprimé dans la plupart des autres tissus ni dans aucune des lignées cellulaires testées par Takashima et collaborateurs (Takashima *et al.*, 2002). Nos résultats, présentés en figure 24, montrent également que le gène *ST6GAL2* est exprimé dans un nombre très restreint de lignées cellulaires. Parmi ces lignées, celles qui expriment ce gène sont les lignées cellulaires neuronales SH-SY5Y et Kelly, et la lignée pro-neuronale à caractère embryonnaire NT2/D1. Les lignées cellulaires SH-SY5Y et NT2/D1 sont celles qui expriment le plus *ST6GAL2*. De plus, *ST6GAL2* est plus faiblement exprimé dans la lignée de neuroblastome Kelly et pas ou quasiment pas exprimé dans toutes les autres lignées. D'autre part, *ST6GAL1* est majoritairement exprimé dans les lignées cellulaires NT2/D2, Kelly, THP-1 et SK-Mel28 (Figure 24). Nous montrons donc ici que les seules cellules qui possèdent une expression relativement forte de *ST6GAL2* et une faible expression de *ST6GAL1* parmi les lignées testées sont les cellules SH-SY5Y.

La lignée de neuroblastome SH-SY5Y a donc été choisie comme modèle cellulaire pour l'identification des extrémités 5' de *ST6GAL2* par la technique de 5' RACE sur ARNm coiffés.

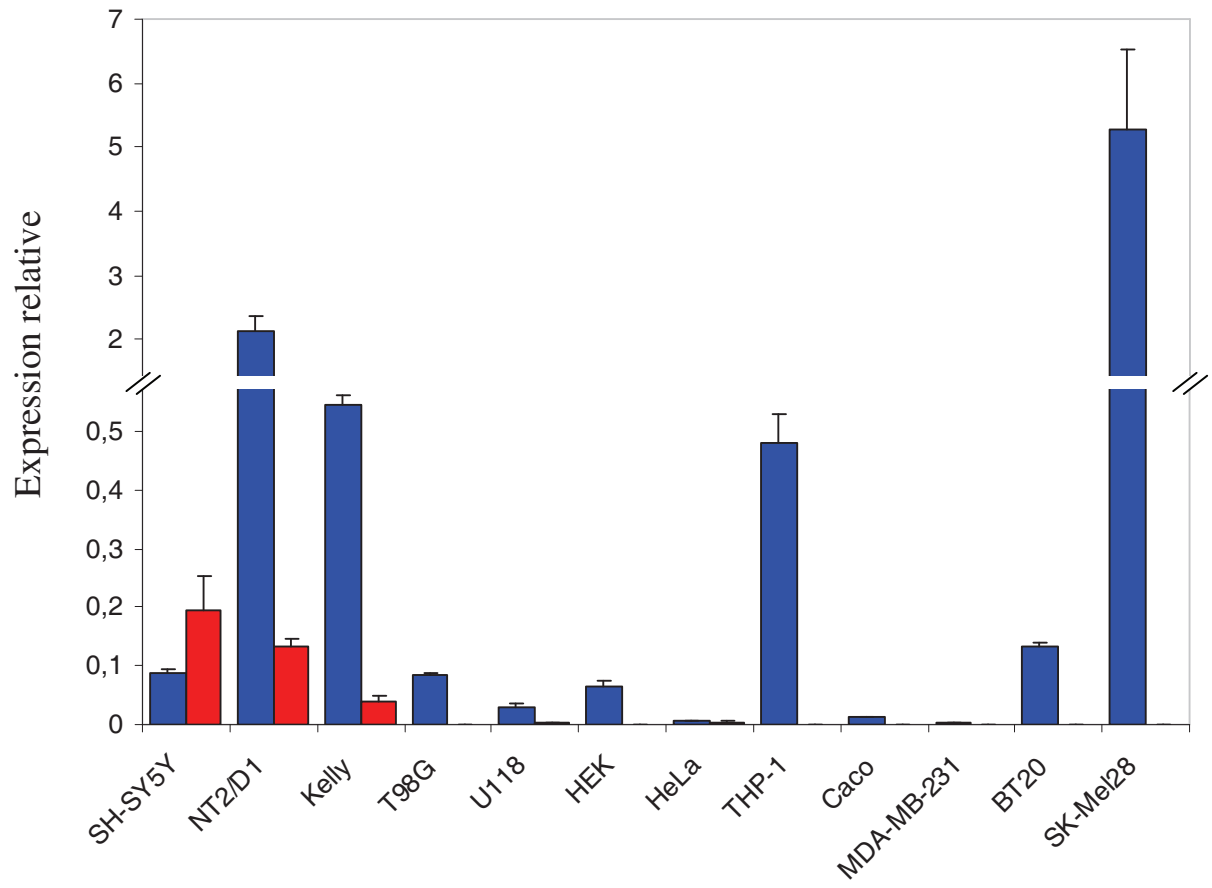


Figure 24 : Analyse par Q-PCR de l'expression des gènes *ST6GAL1* et *ST6GAL2* dans différentes lignées cellulaires.

Les valeurs relatives des expressions des gènes *ST6GAL1* et *ST6GAL2* sont exprimées par rapport à l'expression du gène de référence *HPRT*. L'expression des gènes *ST6GAL1*, *ST6GAL2* et *HPRT* a été mesurée par Q-PCR en utilisant la chimie SYBR Green®. L'expression relative du gène *ST6GAL1* est représentée en bleu, celle du gène *ST6GAL2* est représentée en rouge. Les valeurs représentent la moyenne de $n \geq 3$ expériences indépendantes \pm SD (déviation standard).

2. IDENTIFICATION DES EXTREMITES 5' NON TRADUITES DU GENE *ST6GAL2* DANS LA LIGNEE CELLULAIRE DE NEUROBLASTOME SH-SY5Y

Les cinq exons de la séquence codante du gène *ST6GAL2* s'étendent sur le chromosome 2 sur plus de 85 kpb (Figure 15B, Introduction bibliographique). Après maturation, les ARNm issus de la transcription du gène sont longs de plus de 8 kb et sont constitués d'une séquence non traduite en 3' de plus de 5 kb. Les séquences non traduites situées en 5' des transcrits de *ST6GAL2* ont été déterminées par la technique de 5' RACE sur ARNm coiffés à partir d'ARN totaux de SH-SY5Y.

La technique de 5' RACE sur ARNm coiffés permet de sélectionner les ARNm possédant à leur extrémité 5' une coiffe 7 méthyl-guanine (^{7m}G) (Figure 25, voir Matériels et Méthodes) en écartant les ARNr, ARNt et surtout les ARNm tronqués ne possédant pas de coiffe en 5'. Malgré les précautions employées lors de la manipulation des ARN et à cause de la longueur des ARNm du gène *ST6GAL2*, nous ne pouvons exclure que certaines ARNm ne se soient cassés lors des manipulations. Néanmoins, afin d'obtenir les extrémités 5' lors de la synthèse du premier brin d'ADNc, des amorces de décimères de séquences nucléotidiques aléatoires ont été utilisées pour la transcription inverse. A la place d'amorces oligo dT, les décimères permettent ainsi d'éviter la synthèse incomplète de premiers brins due à d'éventuelles cassures de l'ARNm. Par la suite, une première PCR a été effectuée avec les amorces Outer Primer et GSP1 (Figure 25 et Tableau 5, Matériels et Méthodes) s'hybridant respectivement dans la séquence de l'adaptateur oligonucléotidique du kit de 5' RACE et dans la séquence codante de l'exon E1 du gène *ST6GAL2* (Figure 25). Afin d'augmenter la quantité d'amplifications spécifiques (correspondant aux extrémités 5' du gène *ST6GAL2*) le produit de cette première amplification a été utilisé comme matrice pour un second cycle de PCR nichées avec les couples d'amorces Inner Primer / GSP2.1 ou Inner Primer / GSP2.2 (Figure 25 et Tableau 5 Matériels et Méthodes) s'hybridant à l'intérieur des amplicons générés par le premier cycle de PCR. Les amorces GSP2.1 et GPS2.2 ne sont distantes que de 95 pb sur la séquence de l'exon E1. Ces deux couples d'amorces ont été utilisés afin de prévenir un éventuel défaut d'hybridation de l'une ou l'autre des amorces GSP2.1 ou GSP2.2 qui aurait perturbé l'amplification par PCR d'une ou plusieurs des extrémités 5'. Les conditions de PCR, notamment les températures d'hybridation des cycles (de 51 à 72°C), ont été optimisées pour la première PCR et les PCR nichées. Les amplifications des PCR nichées ont été déposées et séparées sur gel d'agarose en présence de BEt.

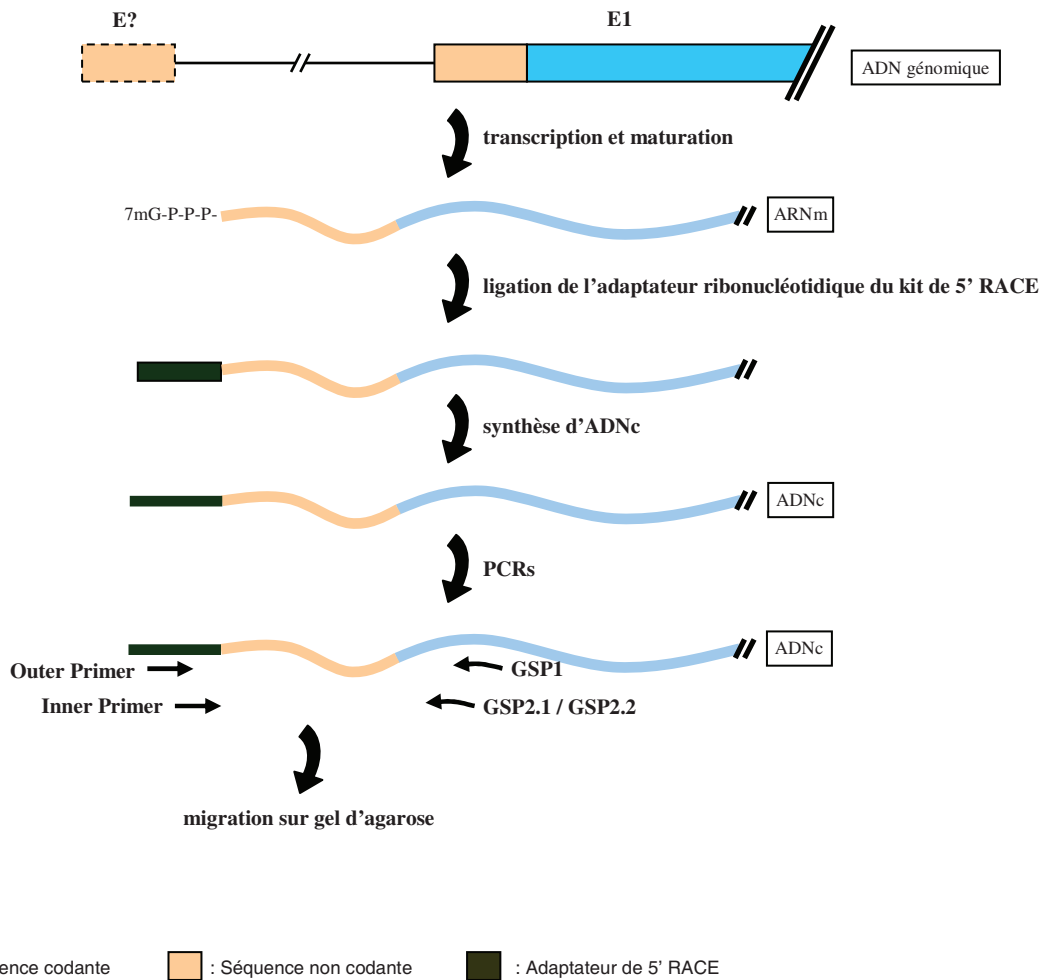


Figure 25 : Schéma de la technique de 5'RACE utilisée pour identifier les extrémités 5' non traduites des ARNm de *ST6GAL2*.

Les flèches noires orientées annotées (Outer Primer, Inner Primer, GSP1 et GSP2.1 / GSP2.2) représentent les amorces utilisées pour les différentes PCR. La transcription inverse a été réalisée avec des décimères de séquences nucléotidiques aléatoires comme amorces. Les amorces GSP1, GSP2.1 et GSP2.2 sont spécifiques de la séquence connue de l'exon E1 de *ST6GAL2*.

La séparation sur gel d'agarose présentée en figure 26A est représentative des résultats ($n > 3$) obtenus sur deux lots d'ARN extraits des cellules SH-SY5Y et après optimisation des conditions de PCR avec le couple d'amorces Inner Primer / GSP2.1. Après migration dans un gel d'agarose (2 %) contenant du bromure d'éthidium (BET), quatre amplifications, nommées A, B, C et D de 750, 550, 450 et 350 pb respectivement, ont été révélées sous lumière UV (Figure 26A). Ces amplifications ont été excisées du gel d'agarose, sous-clonées et séquencées. Des résultats strictement similaires ont été obtenus avec le couple d'amorces Inner Primer / GSP2.2, révélant sur gel un profil semblable de quatre amplifications. Les amplifications issues du couple Inner Primer / GSP2.2 sont 95 pb plus longues que celles

obtenues avec le couple Inner Primer / GSP2.1. Les séquences déduites des quatre amplifications A, B, C et D ont été alignées (<http://www.ncbi.nlm.nih.gov/BLAST/>) et comparées avec les séquences du génome humain présentes dans les bases de données. Les séquences de trois exons, que nous avons appelés EX, EY et EZ, ont pu être reconstituées et attribuées à des extrémités 5' du gène *ST6GAL2* (Figure 26B). Ces trois exons, d'une taille respective de 393, 524 et 140 pb sont localisés à plus de 42 kpb de E1 et tous situés dans un intervalle de 1,3 kpb sur le chromosome 2 (GenBank accession numéros FJ433883, FJ433884, FJ433885). 124 pb seulement séparent l'exon EX de l'exon EY, lui-même situé 87 pb en amont de l'exon EZ. De plus, ces trois exons sont mutuellement exclusifs parmi les transcrits de *ST6GAL2* que nous avons identifiés (Figure 26B). L'amplification A contient la séquence de l'exon EY, les amplifications B et C contiennent toutes les deux la séquence de l'exon EX, avec une séquence plus courte d'environ 80 pb pour le fragment C, et l'amplification D contient la séquence de l'exon EZ. Par ailleurs, les extrémités 3' des exons EX, EY et EZ, et l'extrémité 5' de l'exon E1 sont conformes à la règle GT-AG d'épissage. Enfin, EX, EY et EZ sont mutuellement exclusifs parmi les transcrits identifiés de *ST6GAL2*. La comparaison des séquences des exons EX, EY et EZ avec les séquences des EST tirées des bases de données est représentée sur la figure 27. L'extrémité 5' de toutes les EST retrouvées correspondent, avec plus ou moins de précision, aux séquences des exons EX et EY, suggérant que les transcrits contenant les exons EX et EY soient majoritairement exprimés chez l'Homme. L'extrémité 3' des EST BX 353569 et AB 058780 concorde avec l'extrémité de la séquence de l'exon EY. Inversement, l'extrémité 3' des EST BC 008680, DA 176100 et BE 612797 se situe à l'intérieur de la séquence de l'exon EX. L'extrémité 3' de ces trois dernières EST est compatible avec un site donneur d'épissage, suggérant qu'un épissage alternatif en 3' soit possible pour l'exon EX (Figure 27).

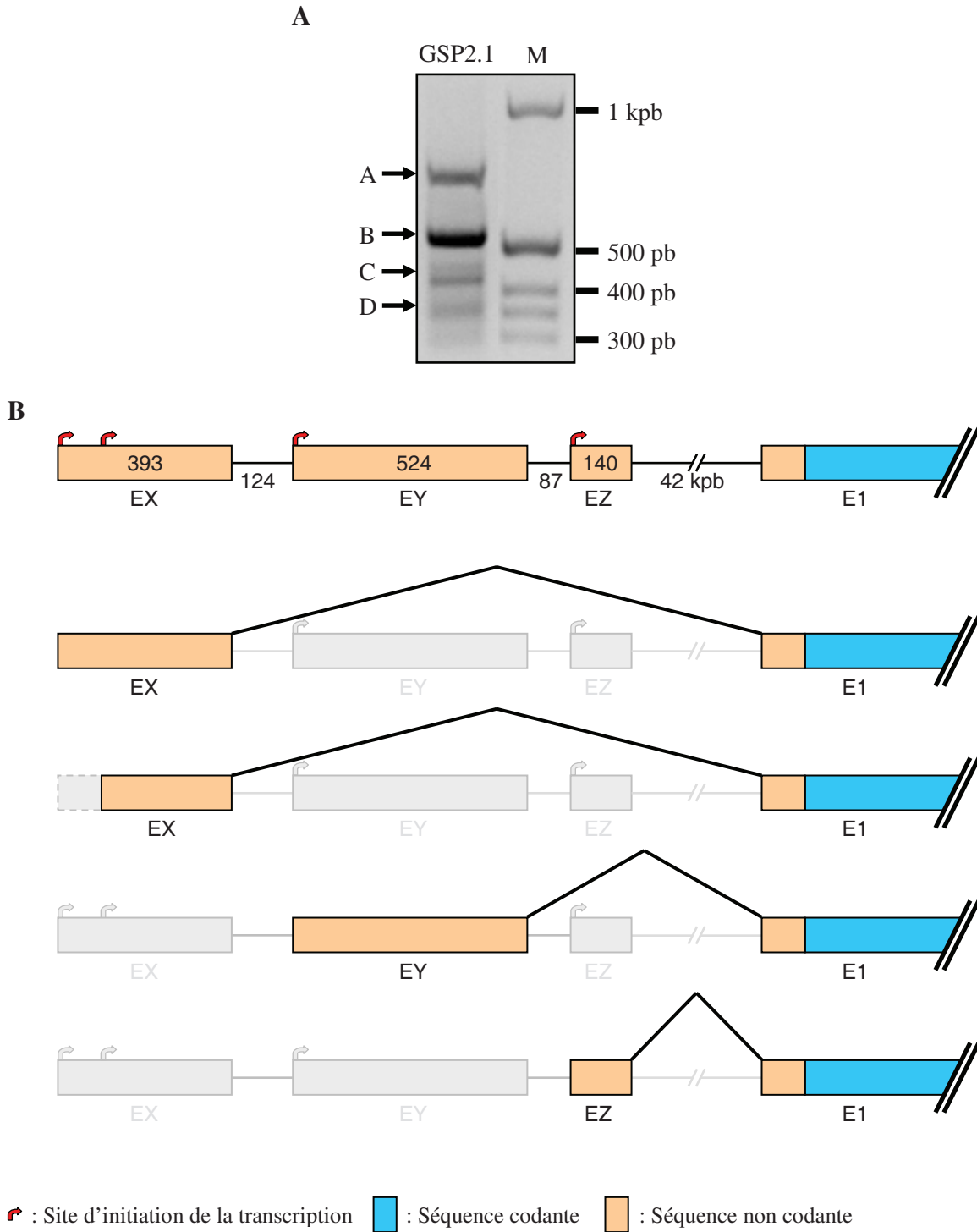


Figure 26 : Identification des exons en 5' du gène *ST6GAL2*.

A, visualisation sur gel d'agarose (2%) sous lumière UV en présence de BEt des produits (amplifications notées A, B, C et D) de la PCR nichée avec le couple d'amorce Inner Primer / GSP2.1 (piste GSP2.1). **M**, marqueur de poids moléculaire (1 kb, Invitrogen). **B**, séquence génomique reconstituée à partir du séquençage et comparaison avec le génome humain des quatre amplifications de la PCR nichée de 5'RACE. Les épissages alternatifs des exons EX, EY et EZ sont représentés. Les chiffres à l'intérieur des exons représentent leur taille en pb, les chiffres en dessous des introns représentent leur taille en pb.

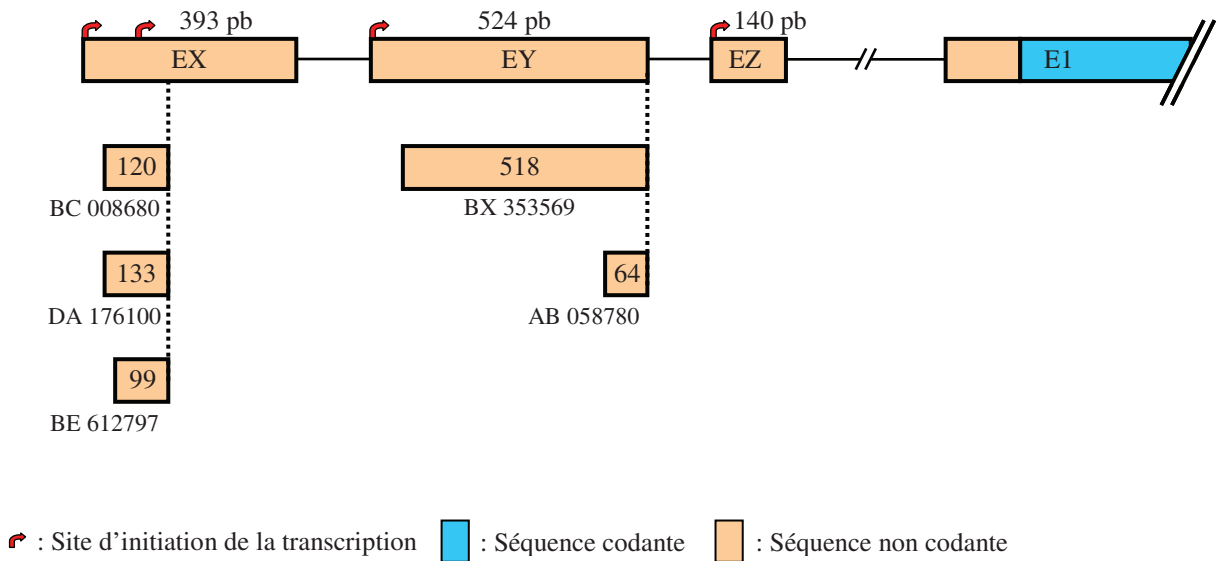


Figure 27 : Schéma comparatif des séquences des exons EX, EY et EZ identifiés expérimentalement par 5' RACE et des séquences d'EST des banques de données. Les chiffres à l'intérieur des séquences des EST représentent leur longueur en pb.

D'autres expériences de 5' RACE avec des amorces dont les séquences s'hybrident dans la région 3' de chacun des trois exons EX, EY et EZ (Figure 28B, les amorces spécifiques sont soulignées) ont été réalisées dans le but d'identifier précisément les sites d'initiation de la transcription (TSS) de chacun des exons. De manière analogue, les amplifications des PCR nichées avec les couples d'amorces Inner Primer / GSP2.X, GSP2.Y ou GSP2.Z ont été déposées, séparées sur gel d'agarose contenant du BEt et visualisées sous lumière UV. Les résultats de la figure 28A montrent, comme attendu, deux produits d'amplification concernant l'exon EX et un seul fragment pour les amplifications correspondant aux exons EY et EZ. Ces produits de PCR ont été sous-clonés, séquencés et les séquences obtenues comparées avec les séquences du génome humain. Pas moins de vingt-neuf clones appartenant aux deux amplifications de la piste EX (Figure 28A) ont été séquencés et ont révélé une hétérogénéité importante du TSS de l'exon EX. Nous avons identifié sept TSS répartis sur 76 pb et définissant deux groupes de séquences pour l'exon EX. Un premier groupe, avec des tailles de 315, 318, 320 et 326 pb, et un deuxième groupe avec des tailles de 380, 381 et 393 pb. Concernant l'exon EY, les huit clones séquencés n'ont montré qu'une variation d'un seul nucléotide du TSS, avec des tailles de 523 et 524 pb pour cet exon. Enfin aucune variation du TSS de l'exon EZ n'a été mise en évidence après le séquençage de six clones (Figure 28B).

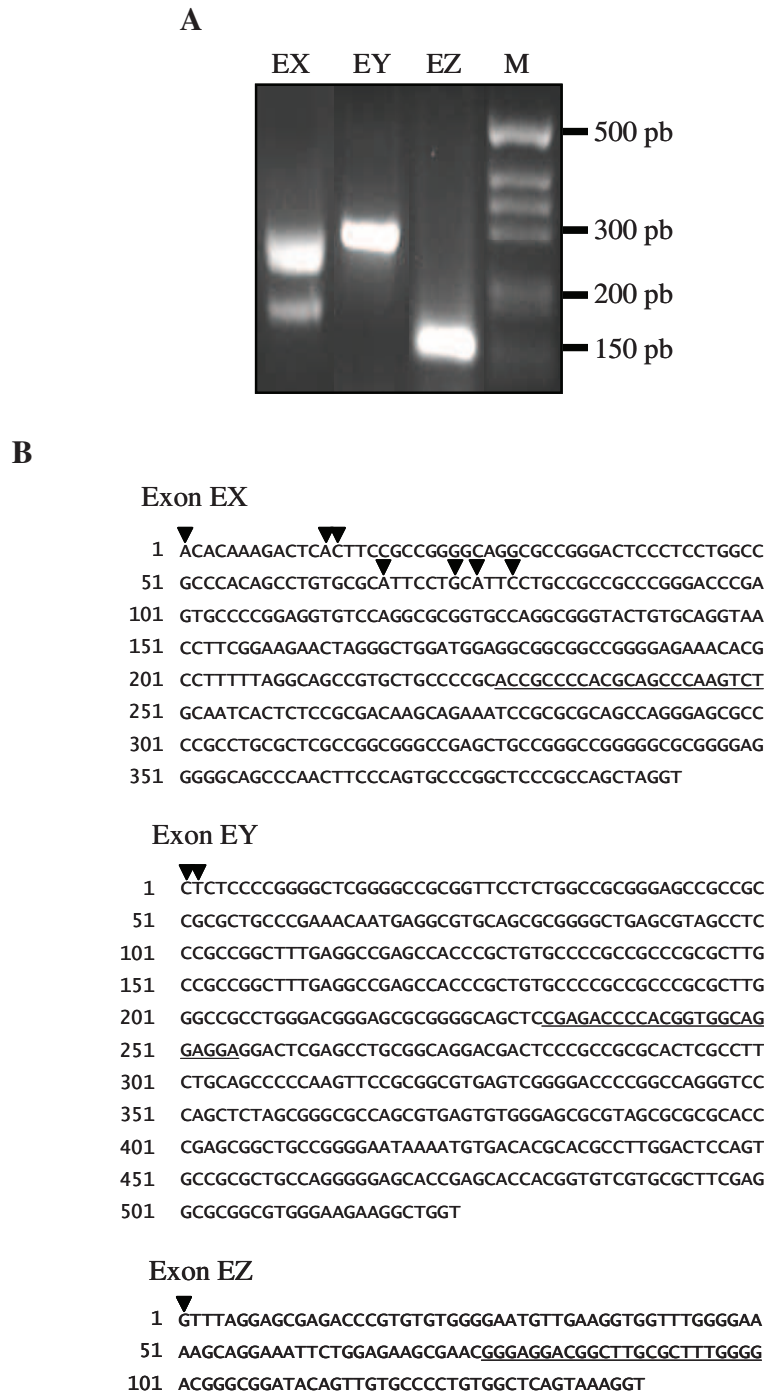


Figure 28 : Identification précise des sites d’initiation de la transcription des trois exons EX, EY et EZ.

Résultats obtenus après les expériences de 5’ RACE et PCR nichées avec des amorces spécifiques de la séquence des exons EX, EY et EZ. **A**, visualisation sous lumière UV de la migration sur gel d’agarose (2%) en présence de BEt des produits de PCR obtenus avec les couples d’amorces Inner Primer / GSP2.X (piste EX), Inner Primer / GSP2.Y (piste EY) et Inner Primer / GSP2.Z (piste EZ). M, marqueur de poids moléculaire (1kb Invitrogen). **B**, les séquences déduites des exons EX, EY et EZ après séquençage et comparaison avec le génome humain sont représentées. Les têtes de flèches représentent les différents TSS déduits du séquençage, les amorces GSP2.X, GSP2.Y et GSP2.Z sont soulignées.

En conclusion de ces expériences de 5' RACE, nous avons mis en évidence trois extrémités 5' non traduites différentes transcrites pour le gène *ST6GAL2* dans la lignée SH-SY5Y, comportant chacune de manière exclusive l'un des trois exons EX, EY ou EZ. De plus, ces trois exons sont inhabituellement proches puisque respectivement séparés par 124 et 87 pb. Enfin, la transcription de l'exon EX débute par au moins sept TSS distincts répartis sur 76 pb, l'exon EY ne présente une variation que d'un seul nucléotide de son TSS alors qu'aucune variation du TSS de l'exon EZ n'a pu être mise en évidence par cette technique.

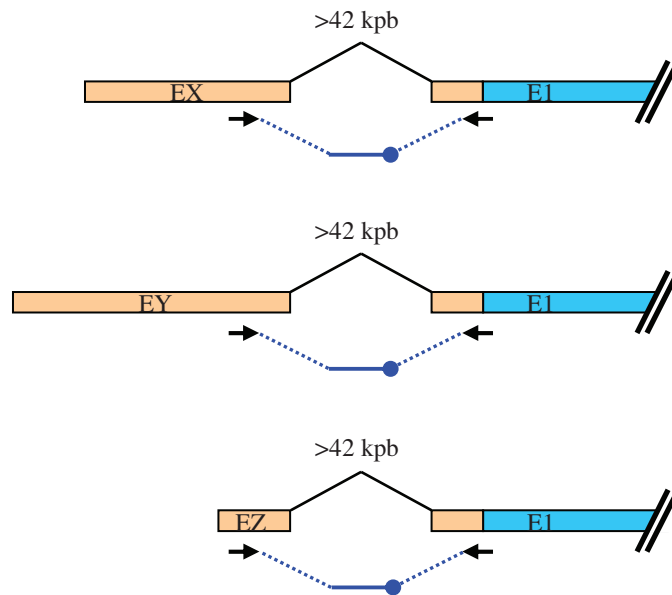
Les PCR réalisées dans le cadre de la 5' RACE et la visualisation de leur résultat sur gel d'agarose n'ont cependant qu'une valeur qualitative, particulièrement après une PCR nichée. Afin de quantifier précisément l'abondance relative de chacun des transcrits contenant les exons EX, EY ou EZ nous avons donc mis au point une méthode de quantification par Q-PCR en duplex utilisant la technologie Taqman.

3. CARACTERISATION DE L'EXPRESSION RELATIVE DES TRANSCRITS DU GENE *ST6GAL2* INITIÉS PAR EX, EY OU EZ DANS DIFFÉRENTES LIGNÉES CELLULAIRES EN CULTURE ET *IN VIVO*

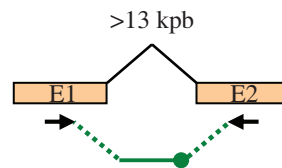
3.1. Expression relative des transcrits contenant EX, EY et EZ dans différentes lignées cellulaires en culture

L'expression des trois transcrits de *ST6GAL2* identifiés par 5' RACE a été mesurée par PCR en temps réel avec la technologie Taqman et en duplex avec le gène de référence *HPRT*. La technologie Taqman offre la possibilité de discriminer l'expression de deux gènes cibles dans un même mélange réactionnel par l'utilisation de sondes fluorescentes séquence-spécifiques. Par rapport à la chimie SYBR Green® par exemple, la spécificité de l'amplification par PCR augmente grâce à l'utilisation d'une sonde spécifique en plus du couple d'amorces de base. De plus, l'utilisation en duplex nous permet d'accroître la précision des mesures (gène d'intérêt et gène de référence) et donc la précision de l'expression relative finalement calculée. En effet, la mesure s'effectue dans un même mélange réactionnel, à partir d'une quantité strictement égale d'ADNc de départ, évitant de cette manière les aléas inhérents à la manipulation tels que les erreurs de pipetées. Ainsi, dans un même mélange réactionnel sont mesurées l'expression de l'un des transcrits de *ST6GAL2*, avec une sonde porteuse du fluorophore FAM émettant à une longueur d'onde de 520 nm, et l'expression de *HPRT*, avec une sonde porteuse du fluorophore VIC émettant à une longueur d'onde de 610 nm. Nous avons dessiné des couples amorces / sondes spécifiques de chacun des trois transcrits de *ST6GAL2* et du gène *HPRT*. Pour cela, nous avons utilisé une amorce anti-sens située dans la séquence de l'exon E1, commune à tous les transcrits, et une amorce spécifique s'hybridant dans la région 3' de chacun des trois exons EX, EY ou EZ. De plus, la séquence de chaque sonde fluorescente Taqman a été dessinée afin de s'hybrider à la jonction EX / E1, EY / E1 ou EZ / E1, renforçant ainsi la spécificité d'amplification en écartant des amplifications issues de possibles contaminations d'ADN génomique. Pour les mêmes raisons, pour le gène *HPRT*, les amorces ont été dessinées pour s'hybrider à l'intérieur des séquences des exons E1 et E2 et la sonde sur la séquence de jonction E1 / E2 (Figure 29 et Tableau 7, Matériels et Méthodes). Nous avons réalisé la synthèse d'ADNc des différents échantillons avec un mélange d'héxamères nucléotidiques de séquences aléatoires (pd(N)₆) et d'oligo dT dans les proportions 3:1. Empiriquement, les valeurs de Ct obtenues avec des ADNc

A



B



: sondes Taqman → ← : amorces : Séquence codante : Séquence non codante

Figure 29 : Schéma d’hybridation des amorces et des sondes Taqman pour la Q-PCR en duplex, pour les trois transcrits du gène *ST6GAL2* et le gène *HPRT*.

A, *ST6GAL2*, une amorce commune s’hybride dans la séquence de E1 et une amorce spécifique s’hybride dans l’exon EX, EY ou EZ. Les sondes Taqman s’hybrident spécifiquement dans les séquences à la jonction EX / E1, EY /E1 ou EZ/ E1. **B**, *HPRT*, les amorces s’hybrident dans les séquences des exons E1 et E2, la sonde Taqman s’hybride sur la séquence à la jonction de E1 / E2. Les sondes Taqman spécifiques des transcrits du gène *ST6GAL2* (en bleu) sont marquées avec le fluorophore FAM qui émet à la longueur d’onde de 520 nm. La sonde spécifique du gène *HPRT* (en vert) est marquée avec le fluorophore VIC qui émet à la longueur d’onde de 610 nm.

synthétisés avec le mélange pd(N)₆:oligo dT étaient meilleures (Ct plus faibles) comparativement à celles obtenues en utilisant des ADNc synthétisés uniquement avec des amorces du type pd(N)₆ ou du type d’oligo dT. Des valeurs de Ct plus faibles pour le gène *HPRT* et les transcrits du gène *ST6GAL2* augmentent encore un peu plus la précision des calculs finaux. Plusieurs Q-PCR de mise au point ont été préalablement effectuées afin de déterminer les quantités d’amorces et de sondes nécessaires à l’obtention d’efficacités

d'amplification proches de 100 % pour les deux gènes cibles d'intérêt (*HPRT* et l'un des transcrits de *ST6GAL2*) (Figure 30). Les droites d'efficacités ont été obtenues avec des dilutions sériées au 1/2 avec de l'ADNc de cellules SH-SY5Y (dilution initiale de 1/5). Après ces mises au point, les valeurs moyennes d'efficacités issues de plusieurs expériences et utilisées pour les calculs ultérieurs sont de 99,8 et 98,7 % pour le couple EX / *HPRT*, de 100,4 et 100,5 % pour le couple EY / *HPRT* et de 100,2 et 98,0 % pour le couple EZ / *HPRT*.

Les lignées cellulaires SH-SY5Y, NT2/D1, T98G, U118, HEK, HeLa, THP-1, Caco-2, MDA-MB-231, BT20 et SK-Mel28 précédemment testées pour leur expression totale de *ST6GAL2* ont aussi été testées pour leur expression des trois types de transcrits de *ST6GAL2*. En accord avec les résultats de PCR en temps réel sur l'expression totale de *ST6GAL2*, les SH-SY5Y et les NT2/D1 sont les lignées qui expriment le plus les trois transcrits contenant EX, EY ou EZ. La lignée Kelly exprime faiblement les trois types de transcrits et les autres lignées cellulaires n'expriment pas ou quasiment pas ces transcrits (Figure 31). Dans toutes les lignées où ces transcrits sont exprimés, les transcrits contenant l'exon EX sont plus exprimés que ceux contenant l'exon EY, les transcrits contenant l'exon EZ sont les moins exprimés. Les résultats obtenus par duplex Taqman et SYBR Green® Q-PCR (Figures 24 et 31) suggèrent qu'aucun autre type de transcrit n'est exprimé dans les lignées cellulaires SH-SY5Y, NT2/D1 et Kelly, puisque que les niveaux relatifs de l'expression globale de *ST6GAL2* et de l'expression de ses trois transcrits sont comparables. Ces données laissent également supposer une régulation au moins partiellement commune de l'initiation de la transcription des trois exons EX, EY et EZ, le profil d'expression de ces trois exons étant toujours le même quel que soit leur niveau global d'expression dans les lignées SH-SY5Y, NT2/D1 et Kelly. Il est cependant probable que la région promotrice contrôlant l'expression de l'exon EX soit plus forte que celle gouvernant l'expression de l'exon EY et que l'exon EZ soit sous le contrôle d'éléments activateurs plus faibles encore.

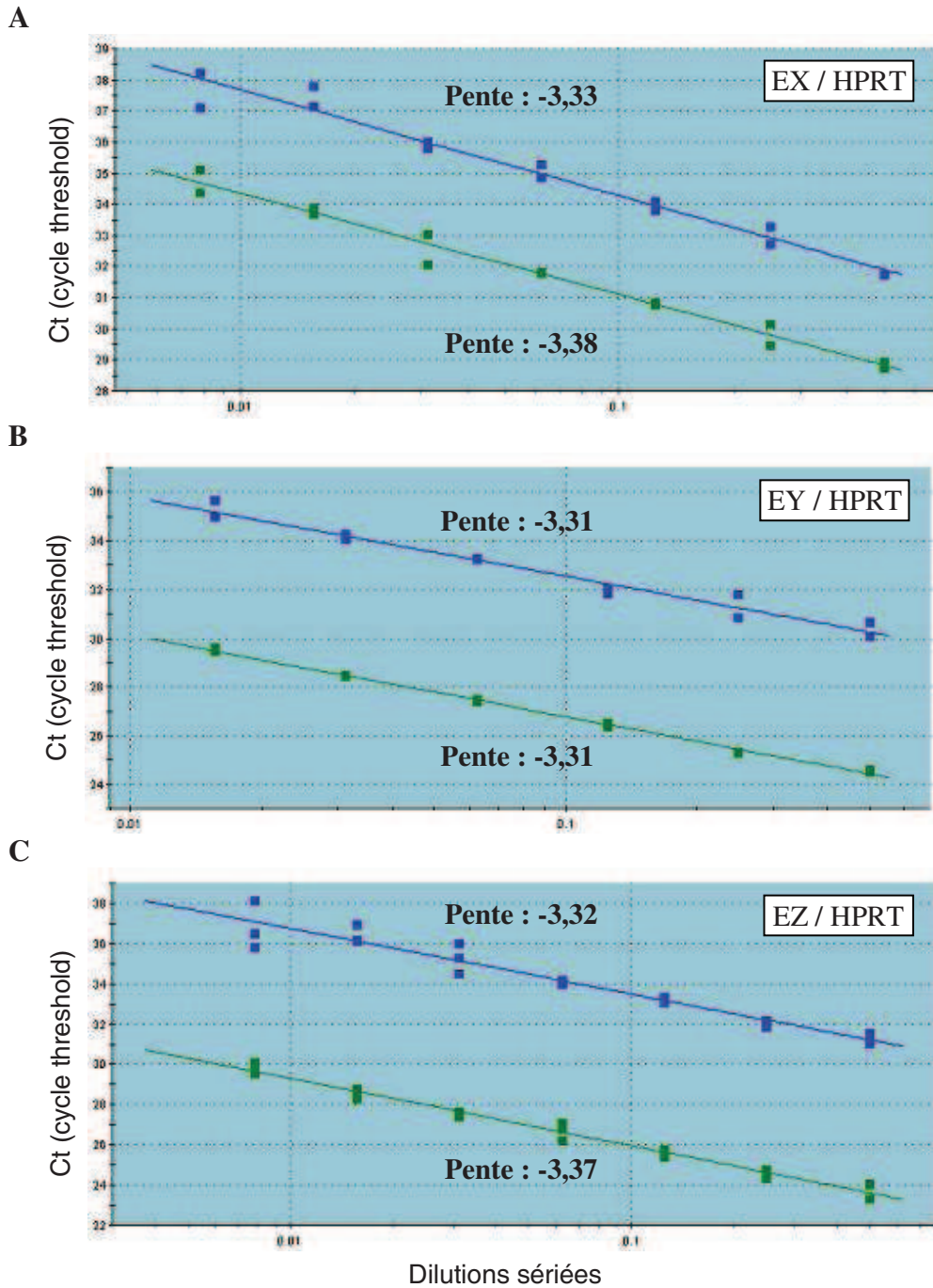


Figure 30 : Exemple de droites de calibration obtenues apr s mise au point de la Q-PCR Taqman en duplex pour le g ne *HPRT* et les diff rents transcrits du g ne *ST6GAL2*.

Les droites en bleu repr sentent l'expression de l'un des diff rents transcrits du g ne *ST6GAL2*, les droites en vert repr sentent l'expression du g ne *HPRT*. Ces droites de calibration ont  t  r alis es avec des dilutions s riees au 1/2. La formule de l' quation des pentes est $Y = \text{pente} \times \text{Log}(X) + B$, o  Y repr sente le Ct, X la quantit  d'ADN et B l'interception   l'ordonn e. L'efficacit  est  gale   : $10^{-1/\text{pente}} - 1$ (en % et multipli  par 100). **A**, droites de calibration pour le couple EX / HPRT. Les efficacit s moyennes utilis es pour ce couple sont respectivement de 99,8 et 98,7 %. **B**, droites de calibration pour le couple EY / HPRT. Les efficacit s moyennes utilis es pour ce couple sont respectivement de 100,4 et 100,5 %. **C**, droites de calibration pour le couple EZ / HPRT. Les efficacit s moyennes utilis es pour ce couple sont respectivement de 100,2 et 98,0 %.

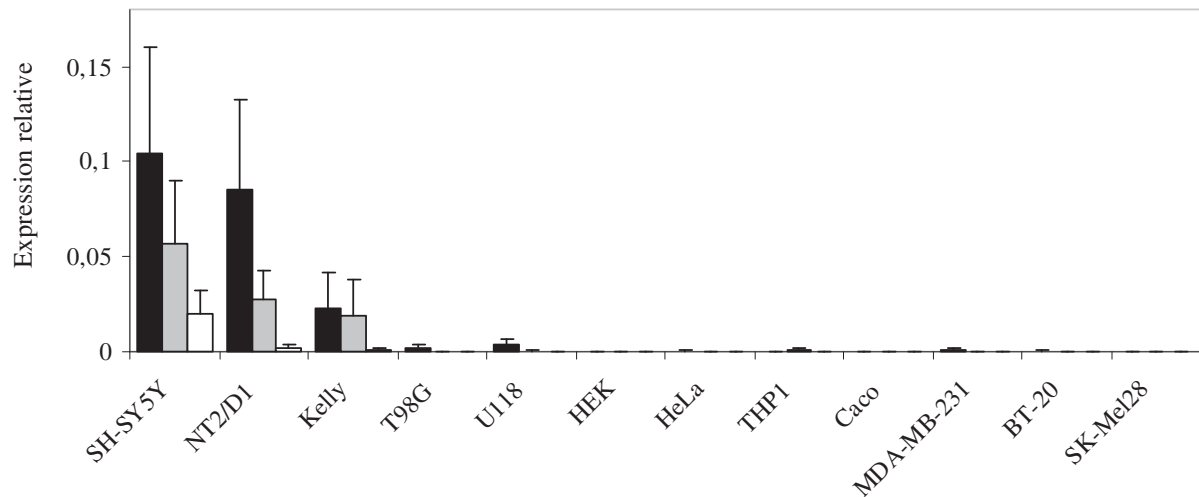


Figure 31 : Analyse par Q-PCR de l'expression des différents transcrits du gène *ST6GAL2* dans différentes lignées cellulaires.

Les valeurs relatives des expressions des transcrits du gène *ST6GAL2* sont exprimées par rapport à l'expression du gène de référence *HPRT*. En noir est représentée l'expression des transcrits initiés par l'exon EX, en gris est représentée l'expression des transcrits initiés par l'exon EY et en blanc est représentée l'expression des transcrits initiés par l'exon EZ. L'expression de chacun des transcrits ainsi que du gène de référence *HPRT* a été mesurée par la technologie Taqman et en duplex (un des transcrits + *HPRT*). Les valeurs représentent la moyenne de $n \geq 3$ expériences indépendantes \pm SD.

3.2. Expression du gène *ST6GAL1*, du gène *ST6GAL2* et de ses transcrits dans différents échantillons de cortex cérébral, de cervelet et de l'hippocampe

3.2.1. Expression de *ST6GAL1* et *ST6GAL2* dans des échantillons de cortex cérébral, de cervelet et de l'hippocampe

Notre collaboration avec l'équipe du docteur Luc Buée, du docteur Marie-Laure Caillet-Boudin et du docteur Claude-Alain Maurage (INSERM, U-837, Lille), nous a permis d'obtenir plusieurs échantillons d'ARN totaux des quatre lobes du cortex cérébral humain (lobe frontal, occipital, temporal et pariétal) et également des échantillons de cervelet et de l'hippocampe (voir Annexes). Ces échantillons *in vivo* nous ont ainsi permis de confronter les résultats obtenus avec des lignées cellulaires en culture, notamment concernant l'expression des transcrits identifiés par 5' RACE dans la lignée SH-SY5Y. Ainsi, nous avons pu quantifier l'expression *in vivo* du gène *ST6GAL2* ainsi que celle de ses trois transcrits alternatifs initiés par les exons EX, EY et EZ, et aussi celle du gène *ST6GAL1*, ce dernier étant décrit comme faiblement exprimé dans le cerveau et également dans certains types cellulaires « cérébraux » (neurones, astrocytes et cellules gliales) (Kitagawa *et al.*, 1994b ; Kaneko *et al.*, 1995 ; Takashima *et al.*, 2002 ; Krzewinski-Recchi *et al.*, 2003).

La figure 32 montre les résultats de l'expression de *ST6GAL1* et *ST6GAL2* dans les différents échantillons. Les données obtenues indiquent que pour ces vingt-cinq échantillons de différents tissus cérébraux *ST6GAL1* et *ST6GAL2* sont, de façon surprenante, tous deux exprimés à des niveaux quasiment équivalents pour un sujet donné. De plus, les niveaux d'expression de *ST6GAL1* ou de *ST6GAL2* sont relativement hétérogènes entre les différents échantillons provenant d'un même sujet. Par exemple, pour les régions frontale, occipitale, pariétale et temporale du cortex, le sujet 3 a une expression relative du gène *ST6GAL1* et du gène *ST6GAL2* de {0,65 et 0,37}, {2,57 et 1,84}, {0,99 et 0,68} et {1,19 et 0,70}, respectivement. De même, pour la région occipitale du cortex, pour le cervelet et pour l'hippocampe, le sujet 7 a par exemple une expression relative du gène *ST6GAL1* et du gène *ST6GAL2* de {0,16 et 0,17}, {0,03 et 0,03} et {6,55 et 5,68}, respectivement (Figure 32B). De plus, pour un même type de tissu provenant de différents sujets, l'expression des deux gènes est également hétérogène. Par exemple, pour le lobe occipital, les sujets 1, 2, 3, 5, 6 et 7 ont une expression relative du gène *ST6GAL1* et du gène *ST6GAL2* de {0,21 et 0,21}, {0,29 et

0,11}, {2,57 et 1,84}, {0,48 et 0,38}, {0,38 et 0,64} et {0,11 et 0,18}, respectivement. De même, les sujets 6, 7 et 8 ont pour l'hippocampe une expression relative du gène *ST6GAL1* et du gène *ST6GAL2* de {0,16 et 0,58}, {6,55 et 5,68} et {1,82 et 1,33} (Figure 32A). Enfin, les expressions de *ST6GAL1* et *ST6GAL2* sont comparables et varient similairement pour un échantillon donné, comme pour les échantillons 1-Par, 5-Occ et 7-Hip dans lesquels les expressions de *ST6GAL1* et *ST6GAL2* sont respectivement {0,13 et 0,18}, {0,48 et 0,39} et {6,55 et 5,68} (Figure 32). Il est également intéressant de noter que l'expression relative *in vivo* du gène *ST6GAL2* est en moyenne plus importante que l'expression relative retrouvée dans les cellules en culture et plus particulièrement dans notre lignée modèle, les SH-SY5Y (Figure 24). La moyenne des expressions relatives du gène *ST6GAL2* dans les échantillons de tissus cérébraux est de 0,69 contre 0,20 dans les cellules SH-SY5Y. Même si les SH-SY5Y constituent un très bon modèle pour l'étude de l'expression du gène *ST6GAL2*, le niveau d'expression de ce gène dans cette lignée reste cependant nettement inférieur à l'expression mesurée *in vivo*.

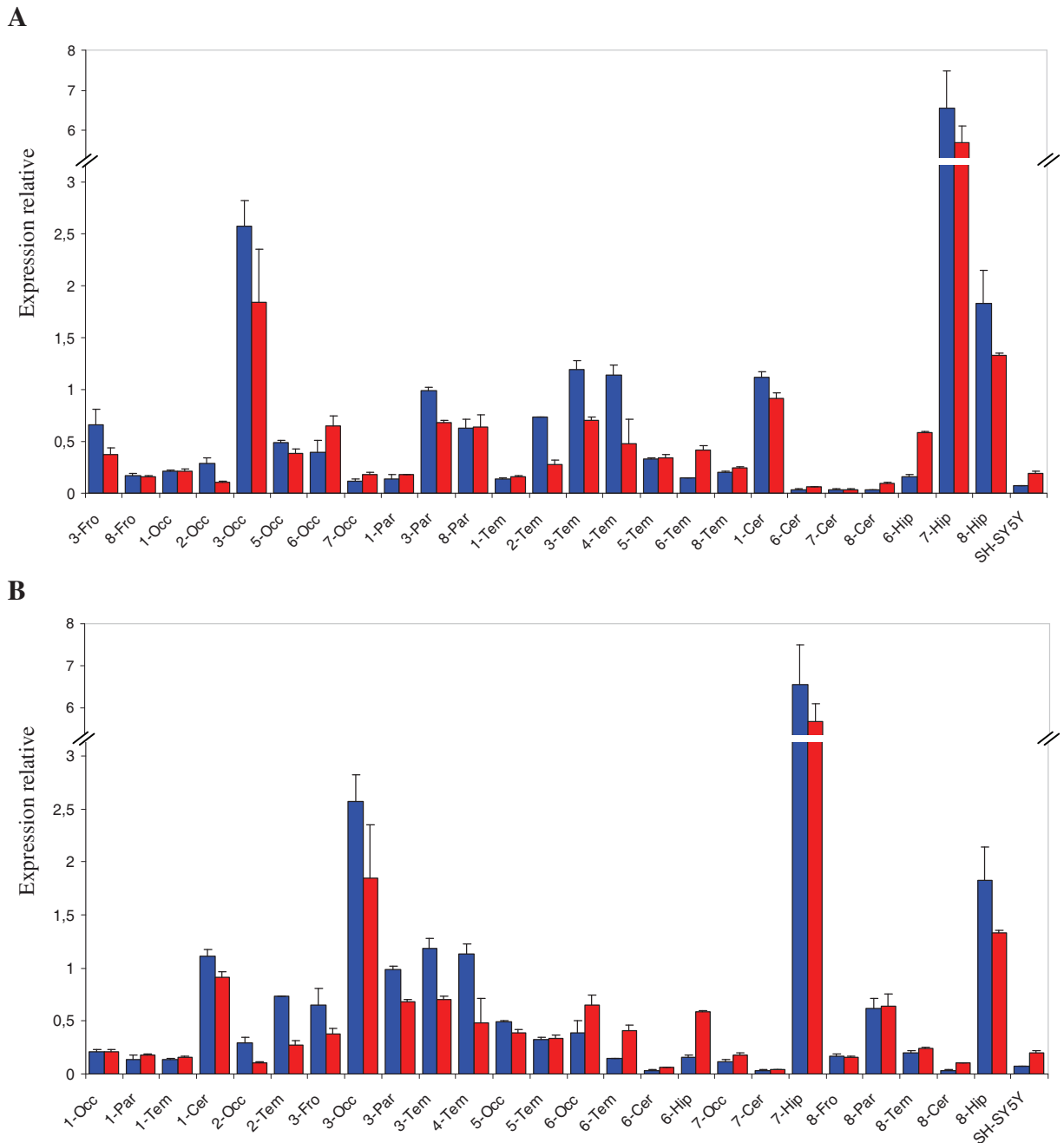


Figure 32 : Analyse par Q-PCR de l'expression des gènes *ST6GAL1* et *ST6GAL2* dans divers échantillons du tissu cérébral et dans les cellules SH-SY5Y.

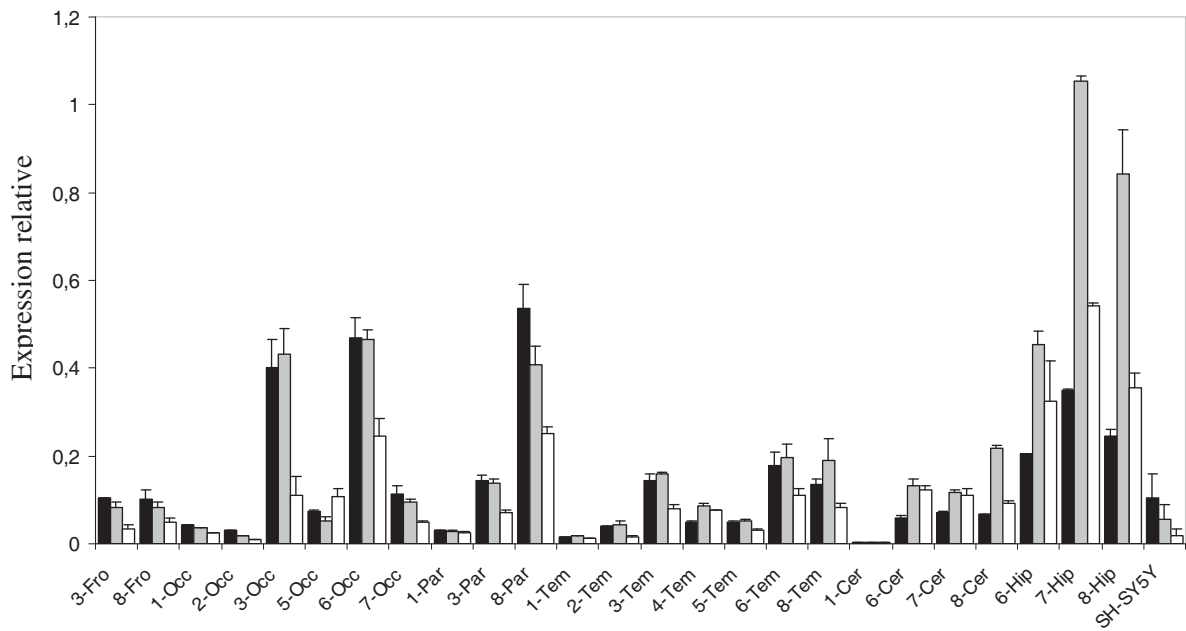
Les valeurs relatives des expressions des gènes *ST6GAL1* et *ST6GAL2* sont exprimées par rapport à l'expression du gène de référence *HPRT*. L'expression des gènes *ST6GAL1*, *ST6GAL2* et *HPRT* a été mesurée par Q-PCR en utilisant la chimie SYBR Green®. L'expression relative du gène *ST6GAL1* est représentée en bleu, celle du gène *ST6GAL2* est représentée en rouge. **A**, les échantillons sont classés par tissus : lobe frontal (Fro), occipital (Occ), pariétal (Par) et temporal (Tem), cervelet (Cer) et hippocampe (Hip). Les chiffres (1 à 8) représentent les différents sujets. **B**, les échantillons sont classés par sujets : 1 à 8. Les valeurs représentent la moyenne de n = 3 mesures ± SD

3.2.2. Expression relative des différents transcrits contenant EX, EY ou EZ dans différents échantillons de cortex cérébral, de cervelet et d'hippocampe

Les vingt-cinq échantillons précédents ont été testés afin de déterminer leur expression des trois transcrits de *ST6GAL2* contenant les exons EX, EY ou EZ. Le résultat de l'analyse par Q-PCR en duplex avec la technologie Taqman est visualisé sur la figure 33A. D'une manière générale, tous les échantillons expriment les transcrits initiés par les exons EX, EY ou EZ et l'expression de ces transcrits est cohérente avec l'expression globale de *ST6GAL2* (Figures 32 et 33A), suggérant qu'aucune autre extrémité 5' n'est utilisée au niveau du cortex cérébral, du cervelet et de l'hippocampe.

Néanmoins, trois échantillons présentent des différences entre l'expression relative globale du gène *ST6GAL2* et l'expression des différents transcrits initiés par EX, EY et EZ. L'échantillon 1-Cer est une première exception. De manière inattendue, alors que l'expression relative globale de *ST6GAL2* dans cet échantillon est de 0,91, aucun des trois transcrits contenant les exons EX, EY et EZ n'est exprimé. Cette donnée suggérerait l'existence d'un ou plusieurs autres transcrits que ceux initiés par EX, EY ou EZ. Néanmoins, les trois autres échantillons de cervelet (6-Cer, 7-Cer et 8-Cer) expriment les trois types d'ARNm de *ST6GAL2* (Figure 33A). L'analyse d'autres échantillons provenant du cervelet pourrait permettre de conclure quant à l'expression du gène *ST6GAL2* et de ses différents transcrits dans ce tissu. D'autre part, les échantillons 6-Occ et 8-Par sont également deux cas particuliers. En effet, l'expression relative globale de *ST6GAL2* pour ces deux échantillons est de 0,64. Cette valeur se situe dans la moyenne de l'expression relative observée pour les échantillons de tissu cérébral. Elle est par exemple inférieure à l'expression de l'échantillon 3-Occ (1,84) et est équivalente aux échantillons 3-Tem (0,70) ou 3-Par (0,68) (Figure 32). Cependant, l'expression relative spécifique des trois transcrits de ces mêmes échantillons, 6-Occ et 8-Par, est nettement plus importante comparativement aux autres échantillons. En effet, l'expression des trois transcrits mesurée pour les échantillons 6-Occ et 8-Par est comparable à celle de l'échantillon 3-Occ et supérieure à celle des échantillons 3-Tem ou 3-Par (Figure 33A). Ces différences pour les échantillons 6-Occ et 8-Par pourraient s'expliquer par l'existence de formes mineures de transcrits possédant une extrémité 3' alternative empêchant ainsi leur quantification par Q-PCR SYBR Green® avec les amorces que nous avons utilisées. L'EST BC 008680, par exemple, possède une extrémité 3' alternative (exon E5 alternatif) et coderait pour une protéine inactive tronquée dans sa partie C-terminale.

A



B

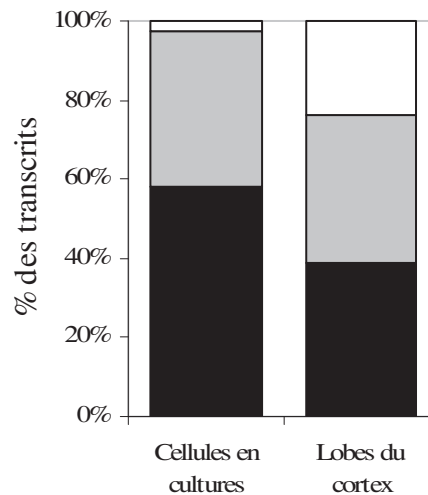


Figure 33 : Analyse par Q-PCR de l'expression des différents transcrits du gène *ST6GAL2* dans différents échantillons du tissu cérébral.

A, expression relative des différents transcrits de *ST6GAL2*. Les valeurs relatives des expressions des transcrits du gène *ST6GAL2* sont exprimées par rapport à l'expression du gène de référence *HPRT*. En noir est représentée l'expression relative des transcrits initiés par l'exon EX, en gris est représentée l'expression relative des transcrits initiés par l'exon EY et en blanc est représentée l'expression relative des transcrits initiés par l'exon EZ. L'expression de chacun des transcrits ainsi que du gène de référence *HPRT* a été mesurée par la technologie Taqman et en duplex (un des transcrits + *HPRT*). Les valeurs représentent la moyenne de n = 9 mesures ± SD. Les termes Fro, Occ, Par, Tem, Cer et Hip font respectivement références aux lobes Frontal, Occipital, Pariétal et Temporal, au Cervelet et à l'Hippocampe. Les chiffres (1 à 8) représentent les différents sujets. **B**, moyenne des expressions relatives des trois transcrits dans les lignées cellulaires en culture (SH-SY5Y, NT2/D1 et Kelly, Figure 31) et dans les différents échantillons du cortex cérébral (Figure 33A).

Une autre observation concerne le ratio des expressions relatives des différents transcrits mesurées dans les lobes du cortex cérébral (EX vs EY vs EZ). Ce ratio est différent de celui observé pour leurs expressions dans les cellules en culture (SH-SY5Y, NT2/D1 et Kelly). Pour les échantillons provenant des lobes cérébraux, les transcrits contenant les exons EX et EY sont toujours exprimés de façon prédominante par rapport aux transcrits initiés par EZ. Toutefois, les transcrits contenant EZ sont très faiblement exprimés dans les lignées en culture mais présentent une expression relative plus importante, en moyenne, dans les échantillons du cortex cérébral (Figures 31 et 33). Pour les échantillons de cervelet et d'hippocampe le ratio des expressions relatives des exons EX, EY et EZ est nettement différent. Les transcrits contenant EX sont les moins exprimés et les transcrits contenant EY sont les plus exprimés, suivis des transcrits contenant l'exon EZ (Figure 33A). L'analyse d'autres échantillons d'hippocampe serait nécessaire pour confirmer cette tendance. Néanmoins, ce profil d'expression particulier pourrait souligner une régulation transcriptionnelle légèrement différente favorisant l'expression des exons EY et EZ au détriment de l'exon EX au niveau de l'hippocampe.

Ces résultats sur l'expression spécifique de transcrits de *ST6GAL2* dans différentes régions du cerveau humain nous confortent dans la validité de notre modèle cellulaire d'étude, les cellules SH-SY5Y. Les trois extrémités 5' identifiées par 5' RACE sont exprimées *in vivo* et nos résultats ne suggèrent pas l'existence d'autres types de transcrits initiés par un ou plusieurs exons alternatifs autre que les exons EX, EY et EZ. Malgré quelques différences au niveau des ratios d'expression des trois exons EX, EY et EZ, ces données suggèrent des mécanismes de régulation transcriptionnelle du gène *ST6GAL2* similaires entre la lignée SH-SY5Y et le cortex cérébral, le cervelet et l'hippocampe. Les cellules SH-SY5Y constituent donc un modèle de choix pour l'étude *in vitro* de l'expression et de la régulation du gène *ST6GAL2*.

La variabilité inter-individuelle observée pour l'expression globale et transcrit-spécifique du gène *ST6GAL2* n'est cependant pas un phénomène inhabituel. Les études précédentes menées au laboratoire sur l'expression de plusieurs gènes codant des sialyl-, fucosyl- et sulfotransférases montrent également une variabilité inter-individuelle importante (Recchi *et al.*, 1998 ; Groux-Degroote *et al.*, 2008). De plus, les différents lobes du cortex cérébral sont différents du point de vue de leur cytoarchitecture, leur architecture chimique et moléculaire, leurs fonctions ainsi que de leur profil d'expression génique (Sur *et al.*, 2005 ;

Rash *et al.*, 2006). Ce n'est donc pas non plus une surprise d'observer des variations intra-individuelles pour les différents lobes cérébraux d'un même sujet. Néanmoins, les échantillons de cervelet ont, en moyenne, une expression relativement faible de *ST6GAL2* et les échantillons d'hippocampe une expression relativement élevée. Les autres échantillons, provenant des quatre lobes cérébraux, ont une expression de *ST6GAL2* relativement constante. Toutefois, l'analyse d'un plus grand nombre d'échantillons, des lobes frontaux et pariétaux notamment, permettrait d'obtenir des résultats plus représentatifs statistiquement.

4. EXPRESSION DE LA PROTEINE hST6Gal II DANS DIFFERENTES LIGNEES CELLULAIRES EN CULTURES ET *IN VIVO*

4.1. Test de réactivité de l'anticorps anti-hST6Gal II par Western-Blot

La protéine hST6Gal II a pu être détectée grâce à un anticorps commercial (Sigma-Aldrich) produit chez le lapin et dirigé contre une séquence peptidique de la région tige de l'enzyme (séquence dans le Matériels et Méthodes). Les séquences en acides aminés de la région tige des sialyltransférases hST6Gal I et hST6Gal II sont très différentes entre elles comparées aux séquences de leur domaine catalytique. Le domaine catalytique présente en effet 48 % d'identité en acides aminés. Ciblant ainsi une région de plus forte variabilité en acides aminés, l'anticorps anti-hST6Gal II permet d'éviter au mieux des réactions croisées et une reconnaissance aspécifique de hST6Gal I notamment. Pour vérifier la réactivité de l'anticorps anti-hST6Gal II, deux vecteurs d'expression ont été construits à partir du p3xFLAG-CMV-10. Les séquences nucléotidiques codant les formes entières de hST6Gal I et hST6Gal II ont été insérées dans le vecteur p3xFLAG-CMV-10 en phase avec 3xFLAG, générant ainsi deux vecteurs codant une protéine étiquetée en N-terminal : hST6Gal I-3xFLAG ou hST6Gal II-3xFLAG. Les cellules HeLa, qui possèdent une expression endogène très faible du gène *ST6GAL2*, ont été transfectées avec l'une ou l'autre des constructions, ou le vecteur p3xFLAG-CMV-10 vide. Après extraction des protéines totales des cellules transfectées, 40 µg de chaque extrait protéique ont été séparés dans un gel à gradient 4-20 % pour une révélation avec l'anticorps anti-hST6Gal II par la technique de Western-Blot. Une unique bande de masse moléculaire apparente d'environ 65 kDa, a été révélée dans la piste contenant les lysats protéiques de cellules HeLa transfectées avec le vecteur codant hST6Gal II-3xFLAG (Figure 34). Le Western-Blot révèle également une bande, de masse moléculaire apparente de 45 kDa, de plus faible intensité dans toutes les pistes. En accord avec l'expression endogène très faible du gène *ST6GAL2* mesurée dans les cellules HeLa (Figure 24), les extraits protéiques déposés dans les pistes HeLa + Ø-3xFLAG ou HeLa + hST6Gal I-3xFLAG n'ont révélé aucun signal à la taille attendue avec l'anticorps anti-hST6Gal II. La protéine détectée dans la piste HeLa + hST6Gal II-3xFLAG avec un poids moléculaire apparent de 65 kDa correspond à la masse prédite pour hST6Gal II, environ 60 kDa, plus le 3xFlag, environ 3 kDa. Avec un anticorps anti-FLAG, une bande unique a également été révélée dans les pistes contenant les lysats protéiques de cellules HeLa transfectées avec le

vecteur codant hST6Gal I-3xFLAG ou hST6Gal II-3xFLAG. Pour la piste HeLa + hST6Gal II-3xFLAG cette bande avait une masse moléculaire apparente d'environ 65 kDa (résultats non présentés).

A l'exception d'une révélation de faible intensité dans tous les échantillons vers 45 kDa, ces résultats suggèrent que l'anticorps commercial anti-hST6Gal II reconnaît bien la protéine hST6Gal II surexprimée dans les cellules HeLa et ne reconnaît pas la protéine hST6Gal I dans les lysats protéiques de cellules HeLa.

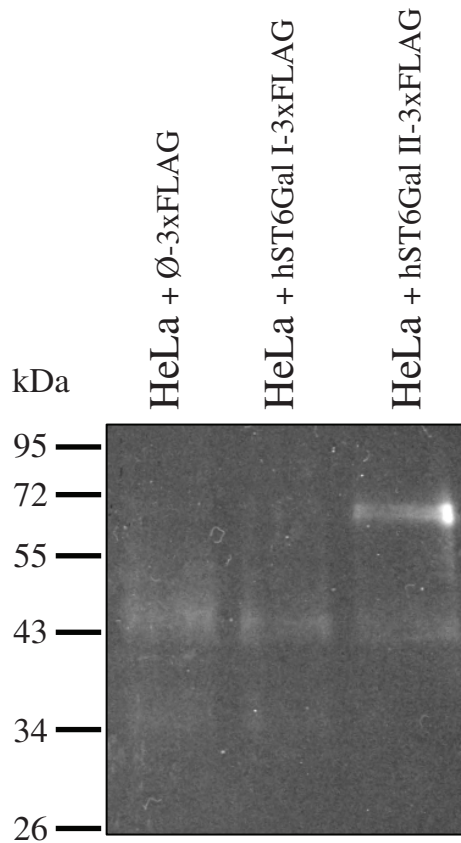


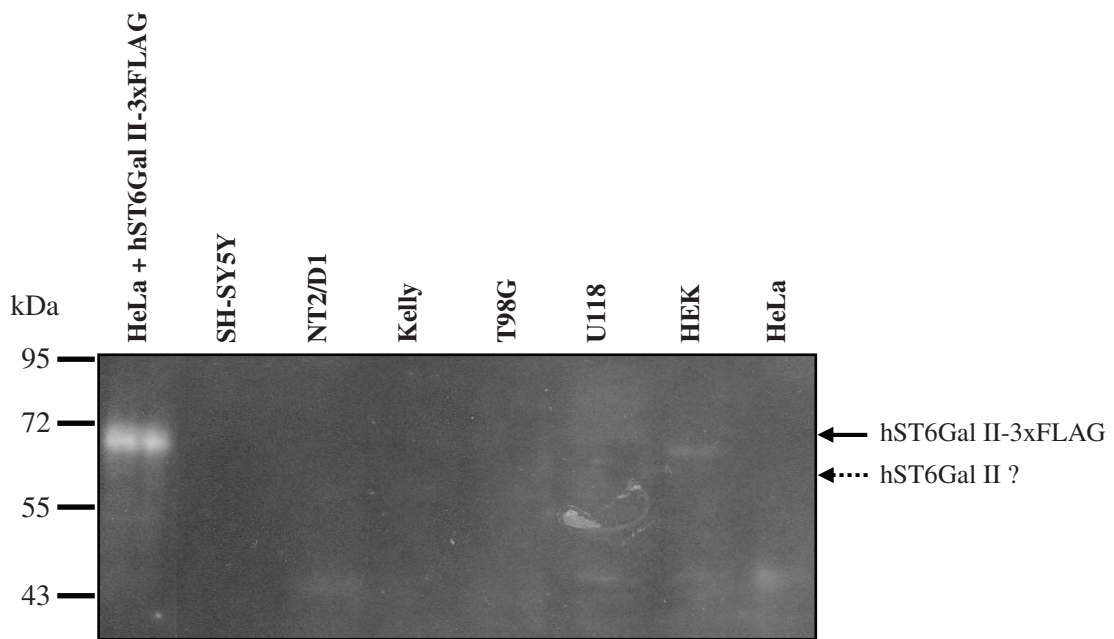
Figure 34 : Analyse par Western-Blot de la spécificité de l'anticorps anti-hST6Gal II.

40 µg d'extraits protéiques ont été déposés dans chaque piste et séparés par SDS-PAGE sur gel à gradient 4-20 %. Après transfert sur membrane de nitrocellulose, la révélation de hST6Gal II par Western-Blot est effectuée comme décrit dans le Matériels et Méthodes. La piste « HeLa + Ø-3xFLAG » correspond aux extraits cellulaires des cellules HeLa transfectées avec le vecteur p3xFLAG-CMV-10 vide. La piste « HeLa + hST6Gal I-3xFLAG » correspond aux extraits cellulaires des cellules HeLa transfectées avec le vecteur p3xFLAG-CMV-10 contenant la séquence codante la forme entière de hST6Gal I fusionnée avec le 3xFLAG. La piste « HeLa + hST6Gal II-3xFLAG » correspond aux extraits cellulaires des cellules HeLa transfectées avec le vecteur p3xFLAG-CMV-10 contenant la séquence codante la forme entière de hST6Gal II fusionnée avec le 3xFLAG.

4.2. Expression de la protéine hST6Gal II dans différentes lignées cellulaires

Après avoir vérifié la réactivité de l'anticorps anti-hST6Gal II, nous avons procédé à des analyses par Western-Blot de l'expression de la protéine hST6Gal II dans les lignées cellulaires pour lesquelles le gène *ST6GAL2* était significativement exprimé : SH-SY5Y, NT2/D1 et Kelly ; et également pour les deux lignées « cérébrales » de glioblastome T98G et U118, et pour deux lignées cellulaires pour lesquelles aucune expression de *ST6GAL2* n'a été détectée par Q-PCR : HeLa et HEK. Des lysats protéiques totaux des ces sept lignées cellulaires ont été préparés et ont été analysés par Western-Blot. 40 µg de chaque extrait protéique ainsi que 40 µg de protéines extraites des cellules HeLa transfectées par le vecteur codant hST6Gal II-3xFLAG ont été déposés par piste. De manière surprenante, aucun signal n'a été révélé par l'anticorps anti-hST6Gal II pour les différentes lignées cellulaires sur le résultat du Western-Blot présenté en figure 35A, à l'exception de la piste « témoin positif » correspondant aux extraits protéiques des cellules HeLa surexprimant hST6Gal II-3xFLAG. Un temps d'exposition plus important du film photographique révèle plusieurs bandes de masses moléculaires apparentes inférieures et supérieures à 60 kDa (la masse moléculaire théorique de hST6Gal II) en même temps qu'un faible signal à 60 kDa correspondant à la masse moléculaire de hST6Gal II (Figure 35B). Si l'on prend en considération cette bande révélée à une masse moléculaire apparente d'environ 60 kDa (cadre rouge Figure 35B), ce signal apparaît plus intense dans les pistes SH-SY5Y et NT2/D1, moins intense dans les pistes Kelly, T98G et U118, et enfin inexistant dans les pistes HEK et HeLa. L'intensité de cette bande semble en corrélation avec l'expression du gène *ST6GAL2* présenté en figure 24, ce qui suggère que cette bande pourrait correspondre à la révélation de la protéine hST6Gal II par l'anticorps. Néanmoins, comme nous l'avons mentionné plus haut, d'autres bandes dont nous ignorons l'identité sont également révélées lors de cette expérience.

A



B

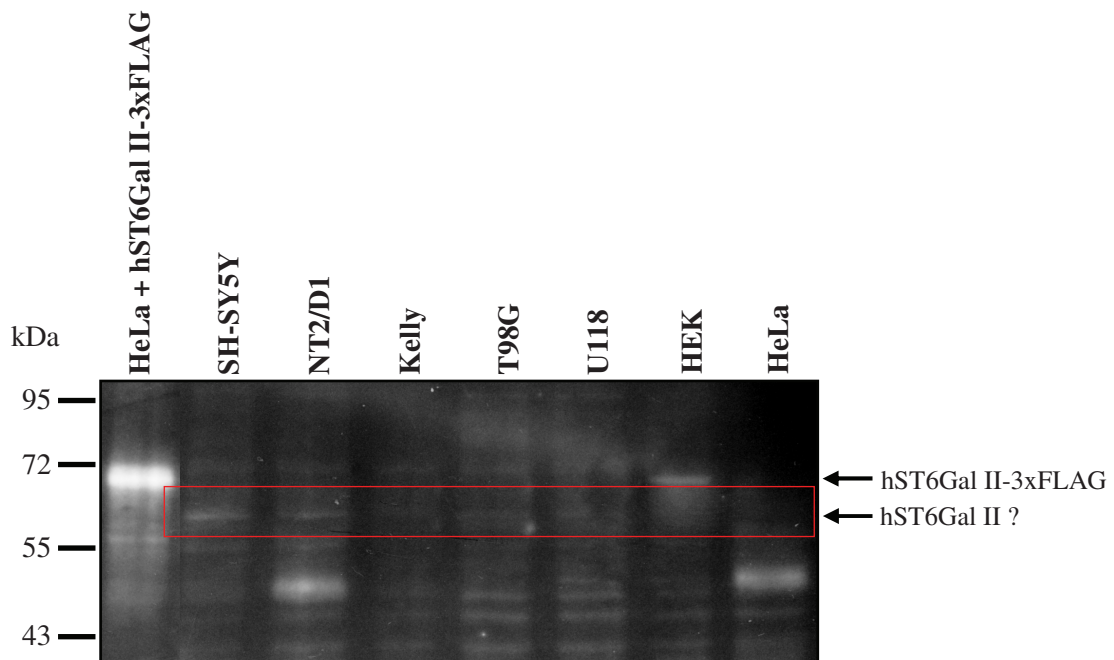


Figure 35 : Analyse par Western-Blot de l'expression de la protéine hST6Gal II dans différentes lignées cellulaires.

40 μ g d'extraits protéiques ont été déposés dans chaque piste et séparés par SDS-PAGE sur mini-gel à 8 %. Après transfert sur membrane de nitrocellulose, la révélation de hST6Gal II par Western-Blot est effectuée comme décrit dans le Matériels et Méthodes. **A**, signal dans les conditions décrites dans le matériel et méthode. **B**, temps d'exposition plus long. Les protéines extraites des cellules HeLa + hST6Gal II-3xFLAG sont utilisées comme témoin positif.

Pour obtenir une révélation plus nette, nous avons recommencé plusieurs fois l'expérience de Western-Blot en jouant sur plusieurs paramètres. Nous avons notamment modifié la quantité de protéines déposées, 20 ou 50 μg ; le temps de préchauffage des échantillons protéiques, 5 ou 12 min ; la dilution d'anticorps anti-hST6Gal II, 1:500 ou 1:2000 ; les conditions d'incubation de l'anticorps anti-hST6Gal II, en sachet plastique, en tube Falcon ou en bac plastique ainsi que les conditions d'incubation de l'anticorps anti-hST6Gal II, 1 h à température ambiante ou 1 nuit à 4°C ; et enfin le temps d'exposition du film radiographique lors de la révélation par chimioluminescence, de 3 min à 45 min. Malgré ces modifications du protocole, la visualisation de hST6Gal II exprimée dans les lignées cellulaires sélectionnées n'a pu être améliorée. Ce résultat suggère que l'expression protéique de hST6Gal II, dans la lignée cellulaire SH-SY5Y notamment, n'est pas assez importante pour être correctement détectée par Western-Blot dans nos conditions expérimentales. Même si l'expression transcriptionnelle de *ST6GAL2* dans la lignée cellulaire SH-SY5Y est de loin supérieure à celle de la majorité des autres lignées cellulaires testées (Figure 24) il semble que l'expression protéique résultante soit faible et difficilement détectable.

4.3. Expression de la protéine hST6Gal II dans différents échantillons de cortex cérébral, de cervelet et d'hippocampe

Des lysats protéiques totaux ont été préparés à partir d'échantillons de chacun des lobes du cortex cérébral (frontal, occipital, pariétal et temporal) pour trois sujets, numérotés 6, 7 et 8 (soit 12 échantillons) ainsi que deux échantillons de cervelet (sujets 6 et 7) et deux échantillons d'hippocampe (sujets 7 et 8). Pour les échantillons appelés 6-Occ, 6-Tem, 7-Occ, 8-Fro, 8-Par, 8-Tem, 7-Cer, 7-Hip et 8-Hip nous disposons de deux prélèvements différents pour effectuer à la fois une extraction d'ARN et une extraction de protéines. Ces échantillons, déjà testés pour leur expression du gène *ST6GAL2* ont également pu être testés pour leur expression de la protéine hST6Gal II. 40 µg de protéines ont été déposés dans chaque piste. Comme le montre le résultat du Western-Blot de la figure 36, une unique bande d'une masse moléculaire apparente d'environ 60 kDa correspondant à l'expression de hST6Gal II a été révélée par l'anticorps anti-hST6Gal II dans quasiment tous les échantillons du cortex cérébral, du cervelet et de l'hippocampe. Ces données suggèrent que la protéine hST6Gal II est très faiblement exprimée dans les échantillons de lobes temporaux des sujets 6, 7 et 8-Tem, l'échantillon de cervelet du sujet-6 et l'échantillon d'hippocampe des sujets 7 et 8. De plus, l'expression de la protéine est également hétérogène parmi les différents échantillons, avec néanmoins une expression visiblement plus faible pour tous les échantillons du lobe temporal. Dans les mêmes conditions expérimentales que celles décrites dans le paragraphe précédent, aucune bande n'a pu être révélée au niveau des protéines extraites des cellules SH-SY5Y (Figure 36). La protéine hST6Gal II serait donc plus exprimée dans les différents échantillons cérébraux que dans les lignées cellulaires que nous avons testées.

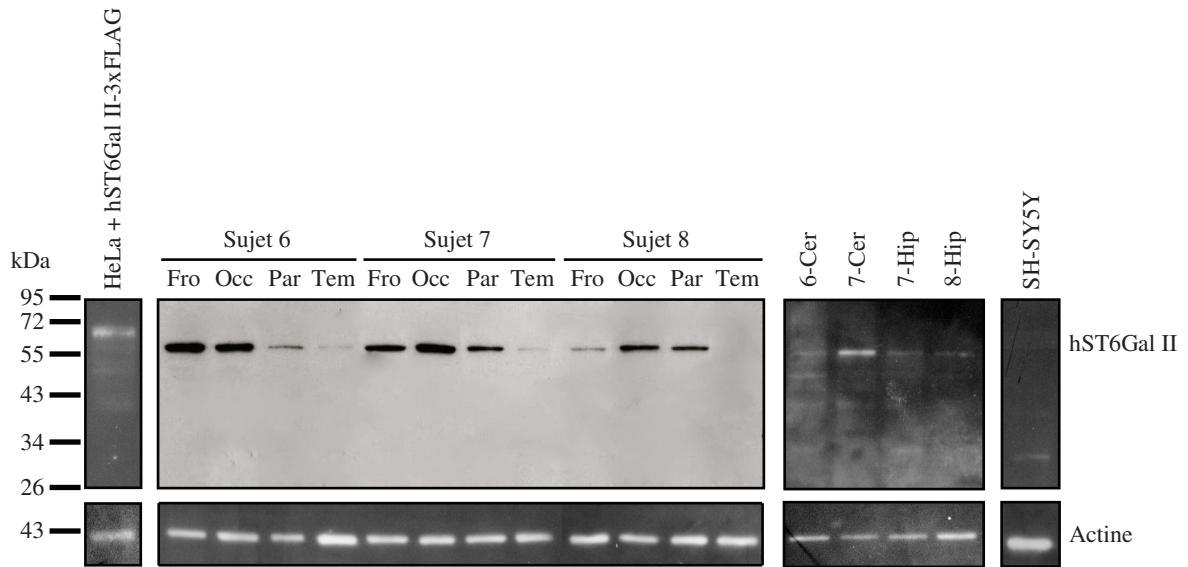


Figure 36 : Analyse par Western-Blot de l'expression de la protéine hST6Gal II dans différents échantillons du tissu cérébral et dans les cellules SH-SY5Y.

40 µg d'extraits protéiques ont été déposés dans chaque piste et séparés par SDS/PAGE sur mini-gel à 8 %. Après transfert sur membrane de nitrocellulose, la révélation de hST6Gal II et de l'actine par Western-Blot est effectuée comme décrit dans le Matériels et Méthodes. La piste HeLa + hST6Gal II-3xFLAG a été utilisée comme témoin positif. La révélation de l'actine a été utilisée comme témoin de charge. Les termes Fro, Occ, Par, Tem, Cer et Hip font références aux lobes Frontal, Occipital, Pariétal et Temporal, au Cervelet et à l'Hippocampe. Les chiffres (6 à 8) représentent les différents sujets.

Pratiquement aucune corrélation n'a été observée lorsque sont comparées l'expression transcriptionnelle du gène *ST6GAL2*, mesurée par Q-PCR, et l'expression protéique de hST6Gal II, détectée par Western-Blotting pour les échantillons 6-Occ, 6-Tem, 7-TOcc, 8-Fro, 8-Par, 8-Tem, 7-Cer, 7-Hip et 8-Hip (Figure 37). C'est particulièrement vrai pour l'échantillon du lobe occipital et de l'hippocampe du sujet 7 (7-Occ et 7-Hip) : 7-Occ est l'un des échantillons qui exprime le plus faiblement le gène *ST6GAL2* alors que la bande correspondant à l'expression de la protéine hST6Gal II est l'une de plus intense visualisée par Western-Blot ; inversement, l'échantillon 7-Hip exprime très fortement le gène *ST6GAL2* alors que l'expression protéique est l'une des plus faibles (Figure 37). De plus, les échantillons 7-Occ et 8-Tem ont une expression transcriptionnelle similaire (d'environ 0,40) malgré une différence évidente d'expression protéique : pour l'échantillon 7-Occ, une bande intense correspondant à la protéine hST6Gal II a été révélée alors que pour l'échantillon 8-Tem, aucune protéine correspondant à la masse moléculaire de hST6Gal II n'a été révélée (Figure 37). Cette « non-corrélation » entre l'expression transcriptionnelle et l'expression

protéique suppose que l'expression de la protéine hST6Gal II ne soit pas uniquement contrôlée au niveau de l'expression transcriptionnelle de son gène.

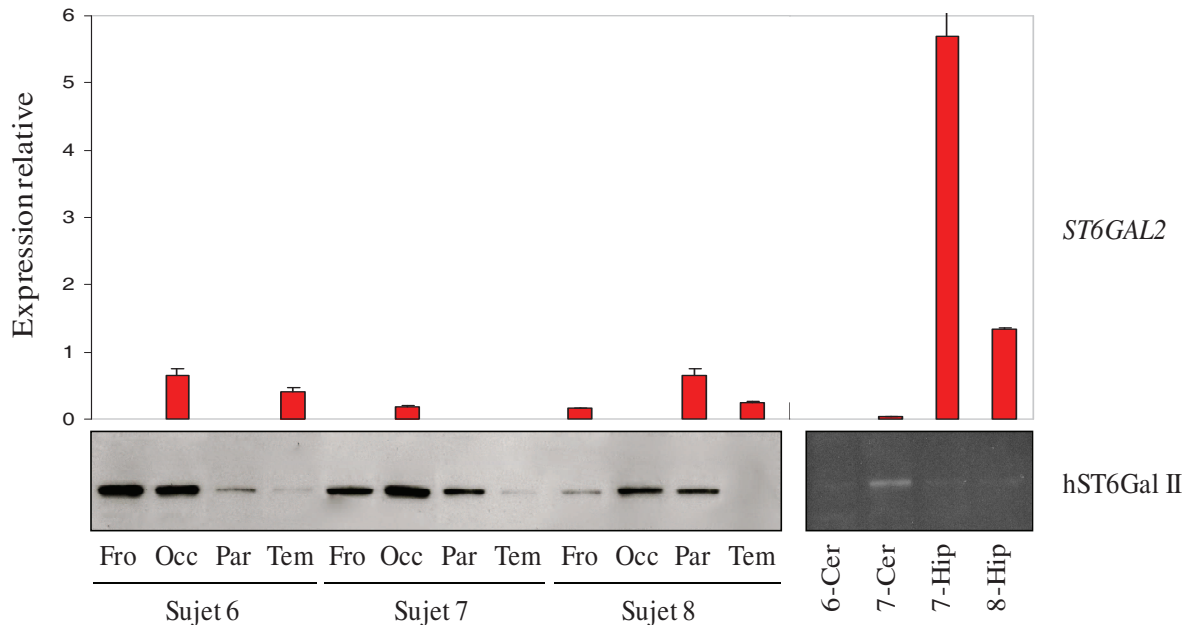


Figure 37 : Comparaison entre l'expression de la protéine hST6Gal II détectée par Western-Blot et l'expression du gène *ST6GAL2* mesurée par Q-PCR dans différents échantillons du tissu cérébral.

L'expression totale du gène *ST6GAL2* est représentée en rouge par rapport à l'expression du gène de référence *HPRT* (Figure 32). L'expression de la protéine hST6Gal II est révélée en Western-Blot par l'anticorps anti-hST6Gal II (Figure 35).

La protéine hST6Gal II est largement exprimée parmi les différents échantillons du cortex cérébral et visiblement plus faiblement exprimée dans le cervelet et l'hippocampe. Concernant ces deux derniers types de tissus, il serait nécessaire d'analyser d'autres échantillons pour définir une tendance plus représentative de l'expression de hST6Gal II dans ces régions du cerveau. Comme pour l'expression transcriptionnelle (Figures 32 et 34), il existe une variabilité inter- et intra-individuelle importante de l'expression protéique (Figure 36). La majorité des études menées sur le rôle et l'expression des sialyltransférases s'intéresse essentiellement à l'expression des gènes codant ces enzymes (par PCR semi-quantitative, PCR en temps réel ou Northern-Blot), et peu de données sont disponibles quant à leur expression protéique. L'activité des sialyltransférases et les structures qu'elles synthétisent

sont souvent appréhendées à l'aide de lectines spécifiques ou d'anticorps spécifiques dirigés contre une structure sialylée particulière. Le niveau d'expression des sialyltransférases n'est que très rarement contrôlée, par Western-Blotting notamment, du fait du peu d'anticorps disponibles reconnaissant les sialyltransférases endogènes.

L'absence de corrélation entre l'expression transcriptionnelle du gène *ST6GAL2* et l'expression protéique de la sialyltransférase hST6Gal II au niveau des différents échantillons de tissu cérébral suggère que la régulation transcriptionnelle n'est pas l'unique moyen de contrôle de l'expression de cette sialyltransférase au niveau du cerveau. La stabilité de l'ARNm et sa demi-vie pourraient jouer un rôle dans le contrôle post-transcriptionnel de l'expression du gène *ST6GAL2*. Les exemples de ce type de régulation ne sont pas très nombreux parmi les gènes de glycosyltransférases et de sialyltransférases. Des études ont montré qu'un régime supplémenté en éthanol conduisait à une déstabilisation l'ARNm de *rst6gal1* (le gène codant ST6Gal I chez le rat) et une perte de l'activité α 2,6-sialyltransférase au niveau du foie. Cette dérégulation n'affecte néanmoins pas la transcription du gène (Rao *et al.*, 1999). Ce même groupe suggère que la fixation de protéines cytosoliques stabilisantes au niveau de la longue extrémité 3' des transcrits du gène *rst6gal1* influe positivement sur l'intégrité des ARNm. Le taux de ces protéines cytosoliques diminuerait en cas de consommation d'alcool (Garige *et al.*, 2006). Des résultats similaires ont été obtenus avec des biopsies de foie humain et avec les cellules en culture HepG2 (cellules hépatiques humaines), concluant à une dérégulation de la stabilité des ARNm du gène *ST6GAL1* (Garige *et al.*, 2006 ; Gong *et al.*, 2007). La région non traduite en 3' des ARNm du gène *ST6GAL2* faisant plus de 5 kb de long, l'hypothèse d'une régulation post-transcriptionnelle tissu-spécifique n'est pas à écarter. Nous ne pouvons pas non plus écarter la possibilité que les séquences des ARNm comprenant les exons EX, EY ou EZ influent de façon différente sur la stabilité des transcrits. Cependant, aucune corrélation entre l'expression spécifique de l'un des transcrits du gène *ST6GAL2* et l'expression de la protéine hST6Gal II n'a été observée parmi les échantillons *in-vivo* que nous avons testés (Figure 37).

5. CARACTERISATION DES REGIONS PROMOTRICES ET DES FACTEURS DE TRANSCRIPTION IMPLIQUES DANS LA REGULATION DES TRANSCRITS DE *ST6GAL2*

5.1. Analyses bioinformatiques d'une région de 3,6 kb en amont et englobant la région des trois exons EX, EY et EZ

L'étape suivante dans l'étude de la régulation transcriptionnelle de *ST6GAL2* a été de caractériser les régions promotrices et les facteurs de transcription potentiels impliqués dans cette régulation. L'identification de facteurs de transcription a été en premier lieu basée sur une analyse bioinformatique des séquences juxtaposant les TSS des exons EX, EY et EZ dans le but d'identifier des sites de fixations potentiels pour des facteurs de transcription. Dans un premier temps, les analyses bioinformatiques nous avaient prédit une séquence de 600 pb en amont de l'exon EX (séquence -486 / +130, Figure 38) comme région promotrice putative du gène *ST6GAL2* (www.genomatic.de). Dans un second temps, nous avons soumis une région allant de 1,9 kb en amont de EX à 200 pb en aval de EZ, la région -1851 / +1452 (Figure 38), à une analyse détaillée des sites potentiels de fixation de facteur de transcription. Le résultat de cette analyse bioinformatique est présenté sur la figure 38. Sur cette figure, seuls les sites de fixation dont les séquences sont les plus proches du consensus et dont les facteurs associés sont des facteurs de transcription généraux ou impliqués dans la régulation de gènes à fonction « cérébrale » ont été reportés. D'après les prédictions bioinformatiques, la séquence -1851 / +1452 ne contient pas de sites consensus de boîtes TATA ou CAAT. Cependant, plusieurs sites de fixation pour des facteurs généraux de transcription tels que SP1 (Specificity Protein 1), NF- κ B (Nuclear Factor B), GATA1 (se fixe sur une séquence nucléotidique du type WGATAR) et MZF1 (Myeloid Zinc Finger 1) et pour des facteurs de transcription impliqués dans l'expression de gènes à fonction cérébrale tels que NRSF (Neuron-Restrictive Silencer Factor), Pit1 (Pituitary transcription factor 1), MYT1 (Myelin Transcription Factor 1), Pura (Purine-Rich element binding protein α), Olf1 (Olfactory neuron-specific binding protein) et Sox5 (SRY box 5) ont été identifiés et reportés (Figure 38).

A

-1851 CCCGTTTCATTTTTGTATTTTTAGTAGAGACAGGTTTCAACATTATGGCCAGGCTGGTCTCGAACTCTGACCTCGTGATCCGCCCGCCT
-1761 CGGCCTCCCAAAGTACTGGGATTACAGGTGTGAGCCACCGCCAGGCTGAAGGTAGTACTATTACATCCATTATACAGAAGAGAAACA
-1671 GAAGCCACAGACTTTGCAACCCCTCTGCGCCCTCTACTATTAAGAAGGTGAGTCTTTCTGCAGTATTTCTTTCTATCCCACCA
-1581 GCTATGGGTCTCTATCAGTTGTTTGTCTGTTTTGTTTTGGTCACTTTAACCACACTGTATGATTTTACATTTATAAAAATAAATT
-1491 TTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTGCTCTGTGCCAGGCCGGACTGCGGACTGCAGTGGCGCAAT
-1401 CTCGGCTCACTGCAAGCTCCGCTTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCAGTAGCTGGGACTACAGGCGCCTGCCACCG
-1311 CGCCCGCTAATTTTTGTATTTTTAGTAGAGACGGGTTTACCTTGTAGCCAGGATGGTCTCGATCTCTGACCTCATGATCCACCC
-1221 GCCTCGGCCTCCCAAAGTCTGGGATTACAGGCGTGAGCCACCGCCCGCCGATAATTTTTCTTTTTTAAAAATATCTGTGTAACAC
-1131 ATTGTTCAATTCAGGCAGTCCCGCAGAATTCATGTGTCAGACATAATTAATACAATTGGGAAATACTGAGTAAGCGTTTTCCGCCCA
-1041 CTTTTTGTTTGAACACAGTGCATTATACCCACTCGATTTTAGTTTAAATCCCCTGCTGTTTTACATTTCTGTAAACTCTTTGAGGGAGTTC
-951 TGGGCGCCACTGTCTAGCTCAGAGTCCAGGCTAAGGTAGGCACTAATGAATGCTGAATGAATTGGGTAGCCAGTTAATTGCTTTTTCC
-861 CCCCTAAAAAAAAGTCTAAATTAAGTAAATAAAGTATTTGCGGATTTTTAAATAATACATTTAATGGCAGATCTGTCTCCAAAACG
-771 TGTGTCCTTAGCGACTGAGGGTGTGGGTGGGGTGGGGTGGGGTACTAACCAGTCCCCTGAGATGGGCCAACCTAGCAATATTATCC
-681 ATCTACCTGGGATTCCCCTACACCTGAACAGCTCTTTGGGTCCGCGAGAGAACCAGCGGGTGGGGCACCAATGCCAGGAGTGCAGCGC
-591 GAGAGACACGCTGAAGTTCACAGGAGATTTACCAGCAAACAAGGGCGCGCGTTTGGGGTCTAGGACCAGCTGCCGCCGTCTCG
-501 AGAAAACGCAAGGGCCGCGGTGACAAGTGGACATTCCGCACGACGACCGACACTCGATGGGCACCTCCCCGGCCCCCGCCCGAG
-411 GCTCCAGTGCCAGCTCGCAGCCGCCACTGGTCCGGAGGCTGCTTCAACCACTTGTCCCCGCGCTTACCCCAAGAACTTTCCGT
-321 CCCTTTGGGGTGGGGACCGCAGTGTGGAGGGGCGGCAGTGGGGGTCTCCAGGAGCACCGTCACCGGACTTGGCGAGCTACCCCTT
-231 CCCAATGAAAGTGGCCCGCTCCCAAGTCCATCCATATCTTAGAGAAGCTCTGGGAGCGAGGGCCAGGCGGAGTGGCGAGCTGGT
-141 TGAGCGAGATGAACAAAGGGAGTAGCGAGCCGGGGAGGATGCTGCGCTCCCCTGCGCCCAACCCAGCCGGCTGCCAGGCGTGCCAGGC
-51 GCGCAGGTAGCCCGCCTTCCCGCTGGCGGGGGAGGTGTTGGACAGGGTACACAAAGACTCATTCCCGGGGGCAGGCGCGGGACT
+41 CCCTCTGCGCGCCACAGCCTGTGCGCATTCTGCATTCTGCCCGCCCGGGACCCGAGTGCCCCGAGGTTCAGGCGCGGTGCC
+131 AGCGGGTACTGTGAGGTAACCTTCGAAGAAGTGGGCTGGATGGAGCGCGCGCGGGGAGAAAACACGCCTTTTAGGCAGCCGTGC
+221 TGCCCGCACCGCCACGCAGCCCAAGTCTGCAGTCACTCTCCGCGACAAGCAGAAATCCGCGCGCAGCCAGGAGCGCCCGCCTGCG
+311 CTCGCCGCGGGCCGAGCTGCCGGCCGGGGCGCGGGAGGGGGCAGCCAACTTCCAGTGCCCGGCTCCCAGCAGTACAGTGGAGCCT
+401 GGAGCGGGAAGGATGAAAGGGATCGCGGAGAGGGAGGGGCTCTGGGGCGAGCGGAGCTCGGGGGCGGGGCGCGGCGCAGCCTCGGTGCG
+491 TGGGCGGGAACCGCGCTCTCCCGGGGCTCGGGCCGCGGTTCTCTGGCCGCGGGAGCCGCGCGCTGCCGAAACAATGAGG
+581 CGTGCAGCGCGGGGCTGAGCGTAGCCTCCCGCGGCTTTGAGGCCGAGCCACCCGCTGTGCCCCCGCGCCCGCTTGCAGCTGCCCGC
+671 GCGGGAGCAGCGCTGGCCCGGAGAGGTGATCCCGCAGGCCCTGGGACGGGAGCGGGGCGAGCTCCGAGACCCACGGTGGCAGGA
+761 GGAGGACTCGAGCTGCGGCAGGACGACTCCCGCCGCGCACTCGCTTCTGAGCCCCAAGTTCGCGCGTGAGTCCGGGACCCCGC
+851 CAGGGTCCCAGCTCTAGCGGGCGCCAGCGTGTGGAGCGCGTAGCGCGCACCCGAGCGGCTGCCGGGAATAAAATGTGACACG
+941 CACGCTTGGACTCAGTGCCGCTGCCAGGGGAGCACCAGCACACCGGTGCTGCGCTTCGAGGCGCGCGTGGGAAGAAGGCTG
+1031 GTAAGGGAGGGCACTGGGTAAAGGTTGGCGACGCCTTCTTGGTGGCGCTGCAGCTCCCGCTGCTGTGCGGGAGTTACACTCCGTGT
+1121 TTAGGAGCGAGACCCGTGTGGGAATGTTGAAGGTGTTTGGGAAAAGCAGGAAATCTGGAGAAGCGAACGGGAGGACGGCTTGGC
+1211 CTTTGGGACGGGGGATACAGTTGTCCCTGTGGCTCAGTAAAGGTATCTTTATTTGGCTCTTTGGCTCTGGAAGGAGTGAAG
+1301 ATTTTTGTCTTCTCAGCCAACCCCGTACAAAATGCACCTCTTAAAGTTAAAGTGTGTCTGTTTCTGTTCTCTGTGGCTA
+1391 TAATGTTTGGGTAGTAGGGATCTTTAGAATCCTAAATACCCGAGTGTTAATCTTATGTAG

B

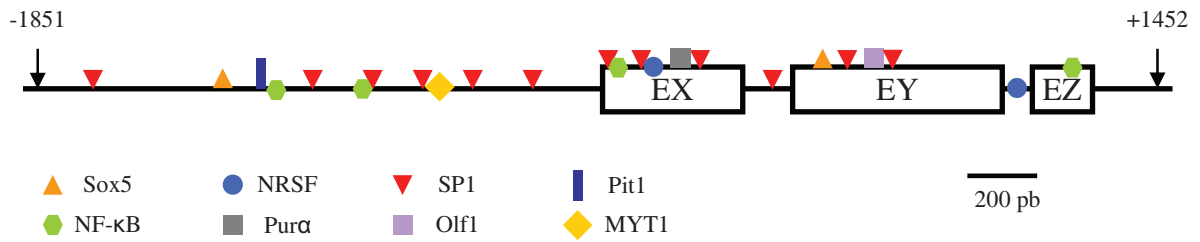


Figure 38 : Analyse bioinformatique de la séquence nucléotidique -1851 / +1452.

A, Identification de sites de fixation potentiels pour des facteurs de transcription. Les nucléotides sont numérotés avec le nucléotide +1 comme premier TSS de l'exon EX. Les têtes de flèches indiquent les TSS de l'exon EX, les triangles indiquent les deux TSS de l'exon EZ et l'étoile représente l'unique TSS de l'exon EZ. Les nucléotides en gras et en rouge représentent la séquence de 600 pb identifiée par les analyses bioinformatiques comme région promotrice putative de *ST6GAL2*. Les sites de fixation potentiels sont soulignés avec des flèches orientées vers la gauche si le site reconnu est sur le brin - et orientées vers la droite si le site reconnu est sur le brin +. Les sites de fixations reportés sont ceux présentant la meilleure homologie avec la séquence consensus pour un site et également ceux fixant des facteurs de transcription généraux ou impliqués dans des fonctions cérébrales. **B**, représentation schématique de la région -1851 / +1452 avec les différents sites potentiels de fixation identifiés en **A**.

5.2. Identification des régions promotrices minimales, tests d'activité luciférase de vecteurs rapporteurs

Pour identifier des régions promotrices gouvernant l'expression des transcrits de *ST6GAL2*, nous avons sous-cloné différentes séquences génomiques encadrant les exons EX, EY et EZ dans un vecteur rapporteur. Les différentes régions génomiques ont été insérées en amont d'un gène rapporteur dont l'activité enzymatique peut être facilement dosée et permettant ainsi de délimiter les régions promotrices minimales. Ici, le gène rapporteur utilisé est le gène codant la luciférase *Firefly*, contenu dans le vecteur pGL3-Basic. La quantité de protéine luciférase produite après transfection dans le modèle cellulaire choisi est ainsi dépendante de l'activité promotrice de la séquence insérée en amont. Après lyse des cellules transfectées et récupération des protéines totales, l'activité de la luciférase *Firefly* est mesurée par absorbance à 560 nm en présence de la luciférine, son substrat. Dans toutes les expériences, les cellules sont également toutes co-transfectées avec un vecteur normalisateur, codant la luciférase *Renilla*. La quantité de vecteur normalisateur est la même pour toutes les transfections afin de normaliser les résultats obtenus avec la luciférase *Firefly*.

Dans un premier temps nous avons isolé et amplifié la région -1851 / +263 en amont de l'exon EX (le nucléotide +1 faisant référence au TSS le plus distal de EX) par PCR à partir d'ADN génomique de cellules SH-SY5Y et à l'aide d'amorces contenant les sites de restriction *KpnI* et *HindIII* (Matériels et Méthodes). Cette séquence a ensuite été sous-clonée dans le vecteur pGL3-Basic aux sites *KpnI* / *HindIII* en amont du gène codant la luciférase *Firefly*, générant ainsi le vecteur pGL3_-1851/+263. La région -1851 / +263 a ensuite été digérée par les couples d'enzymes de restriction *KpnI* / *BglII*, *KpnI* / *XhoI*, *KpnI* / *NheI* ou *KpnI* / *SmaI*, les extrémités générées rendues bouts francs et liguées pour générer les vecteurs pGL3_-752/+263, pGL3_-498/+263, pGL3_-121/+263 et pGL3_+92/+263, respectivement (Figure 39). L'absence de mutation de tous ces vecteurs a été vérifiée par séquençage. Ces différents vecteurs ont ensuite été co-transfectés avec le vecteur normalisateur contenant le gène de la luciférase *Renilla* dans les cellules SH-SY5Y. Nos premières expériences de transfections avaient été initialement réalisées par lipofection (Lipofectamine Reagent, Invitrogen). Les valeurs d'activité luciférase mesurées pour nos différentes constructions après lipofection étaient faiblement supérieures au bruit de fond du signal et n'étaient pas pleinement exploitables. Nous avons donc changé de méthode de transfection afin d'obtenir

de meilleurs résultats. Les résultats présentés ici sont ceux obtenus par nucléofection à l'aide du système AMAXA (Matériels et Méthodes). 24 h après transfection les cellules ont été lysées, les protéines totales récupérées et l'activité luciférase mesurée. La transfection de chacun des différents vecteurs a été réalisée au moins trois fois et les mesures d'activité luciférase de chaque extrait protéique ont été réalisées en duplicata. Les résultats présentés dans la figure 39 montrent une augmentation de 7 fois de l'activité luciférase du vecteur pGL3_-121/+263 par rapport à l'activité du vecteur pGL3-Basic, alors que les constructions pGL3_-1851/+263, pGL3_-752/+263 et pGL3_-498/+263 ne montrent qu'une augmentation de 2 à 3 fois de leur activité luciférase par rapport au pGL3-Basic. L'activité du vecteur pGL3_+92/+263 est même inférieure à celle détectée pour le pGL3-Basic. Ces données suggèrent clairement l'existence d'une région promotrice minimale à l'intérieur de la séquence -121 / +92.

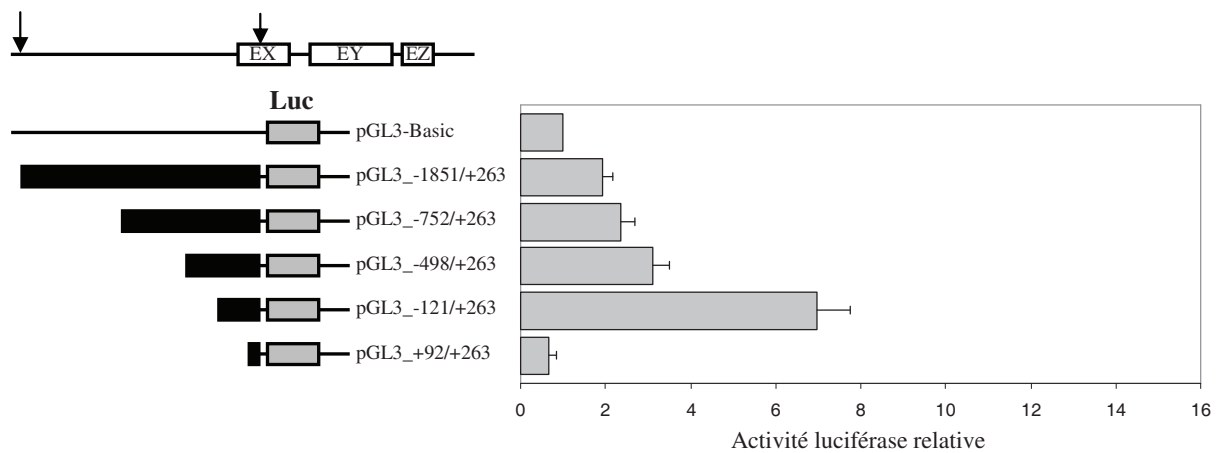


Figure 39 : Activité luciférase mesurée après transfection des différentes constructions de la région -1851 / +263 (en amont de EX) dans les cellules SH-SY5Y

Les différents vecteurs ont été co-transfectés dans les cellules SH-SY5Y avec le vecteur normalisateur codant la luciférase *Renilla*. Sur la gauche : un schéma des différentes régions insérées dans le pGL3-Basic. Sur la droite : le résultat des mesures des activités luciférase correspondantes. Les activités luciférase sont normalisées avec l'activité de la luciférase *Renilla* et exprimées par rapport à l'activité obtenue avec le vecteur pGL3-Basic. Les valeurs représentent la moyenne de $n \geq 3$ expériences indépendantes \pm SD.

L'ensemble de la région EX, EY et EZ et ses différentes délétions ont également été sous-clonées dans le but d'identifier de potentielles séquences promotrices. La séquence -121 / +1452 a été isolée et amplifiée par PCR à partir d'ADN génomique de SH-SY5Y à l'aide

d'amorces contenant le site de restriction *Bgl*II pour l'amorce anti-sens. Utilisant le site *Nhe*I naturellement présent dans la région 5' de la séquence -121 / +1452 et le site *Bgl*II incorporé dans l'amorce anti-sens, la séquence -121 / +1452 a été insérée dans le vecteur pGL3-Basic aux sites *Nhe*I / *Bgl*II en amont du gène codant la luciférase *Firefly*, générant le vecteur pGL3_-121/+1452. La région de -121 / +1452 a ensuite été traitée suivant la même stratégie de digestion par des enzymes de restriction, grâce à des sites de restriction naturellement présents dans la séquence -121 / +1452, et de ligation (bouts francs + ligation) pour générer les vecteurs pGL3_+517/+1452, pGL3_+773/+1452, pGL3_-121/+772, pGL3_+465/+772, pGL3_-121/+464 et pGL3_-121/+91. Enfin, le plasmide pGL3_+1046/1452 a été généré par PCR à partir d'ADN génomique de cellules SH-SY5Y et à l'aide d'amorces contenant les sites *Kpn*I et *Hind*III (Matériels et Méthodes). La séquence amplifiée a été insérée dans le vecteur pGL3-Basic aux sites *Kpn*I / *Hind*III. L'absence de mutation de tous ces vecteurs a été vérifiée par séquençage. L'extrémité 5' de la séquence génomique comprise dans le vecteur pGL3_+517/+1452 débute huit nucléotides en aval du premier TSS de l'exon EY. La séquence génomique comprise dans le vecteur pGL3_-121/+91 correspond à la séquence -121 / +91 se terminant douze nucléotides en aval du dernier TSS identifié par 5' RACE pour l'exon EX (Figure 28B). Ces plasmides ont ensuite été co-transfectés avec le vecteur normalisateur contenant le gène de la luciférase *Renilla* par nucléofection dans les cellules SH-SY5Y pour mesurer l'activité luciférase résultant de chacune des constructions réalisées. 24 h après transfection les cellules ont été lysées, les protéines totales récupérées et l'activité luciférase a été mesurée. La transfection de chacun des différents vecteurs a été réalisée au moins trois fois et les mesures d'activité luciférase de chaque extrait protéique ont été réalisées en duplicata. Les activités luciférase obtenues avec les plasmides pGL3_-121/+1452, pGL3_+517/+1452, pGL3_+773/+1452, pGL3_+1046/+1452 et pGL3_-121/+464 n'ont pas révélé de différences significatives avec l'activité obtenue avec le pGL3-Basic (Figure 40). D'un autre côté, les résultats obtenus avec les vecteurs pGL3_-121/+772 et pGL3_+465/+772 présentent respectivement une augmentation de l'activité luciférase de 8 et 13 fois par rapport à l'activité mesurée avec le vecteur pGL3-Basic (Figure 40). Ces données suggèrent que la séquence +465 / +772 correspond à une région promotrice. De plus, le vecteur pGL3_-121/+91, correspondant au précédent vecteur pGL3_-121/+263 délété de 181 pb en 3', présente une forte augmentation de l'activité luciférase de plus de 35 fois par rapport au pGL3-Basic, indiquant la présence d'une activité promotrice importante à l'intérieur d'une séquence de 212 pb en amont de l'exon EX (Figure 40). Enfin, toutes ces données suggèrent que la séquence +92 / +464 contienne probablement des éléments régulateurs négatifs puisque

l'activité luciférase obtenue avec plasmides pGL3_-121/+772 et pGL3_-121/+464 décroît par rapport à l'activité mesurée avec les plasmides pGL3_+465/+772 et pGL3_-121/+91 (Figure 40).

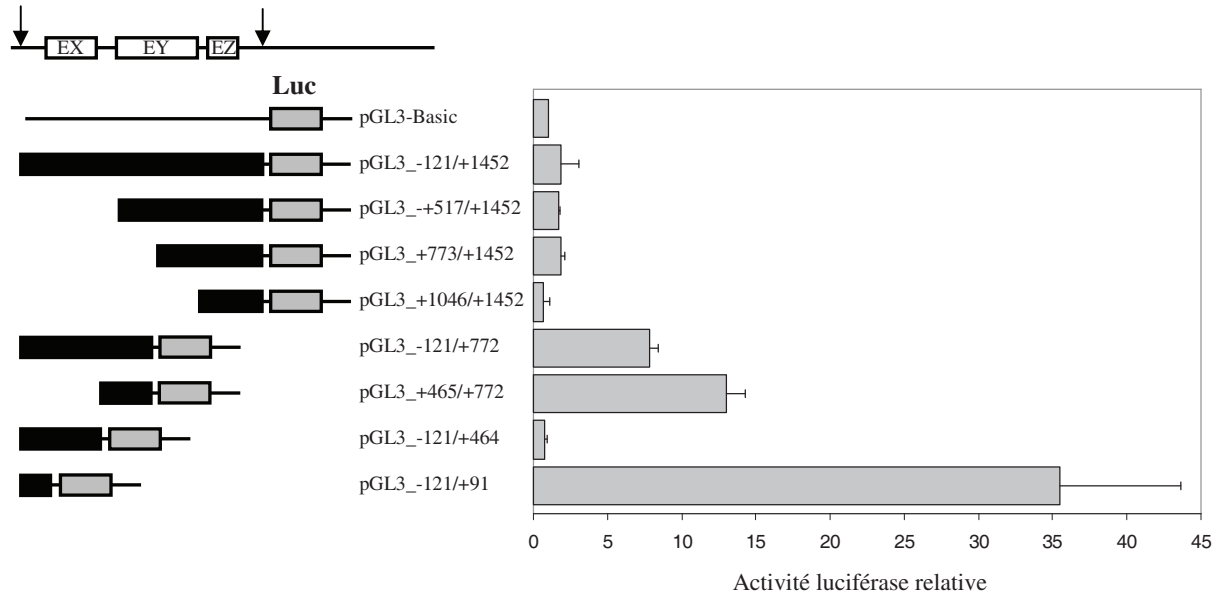


Figure 40 : Activité luciférase mesurée après transfection des différentes constructions de la région -121 / +1542 (EX/EY/EZ) dans les cellules SH-SY5Y

Les différents vecteurs ont été co-transfectés dans les cellules SH-SY5Y avec le vecteur normalisateur codant la luciférase *Renilla*. Sur la gauche : un schéma des différentes régions insérées dans le pGL3-Basic. Sur la droite : le résultat des mesures des activités luciférase correspondantes. Les activités luciférase sont normalisées avec l'activité de la luciférase *Renilla* et exprimées par rapport à l'activité luciférase obtenue avec le vecteur pGL3-Basic. Les valeurs représentent la moyenne de $n \geq 3$ expériences indépendantes \pm SD.

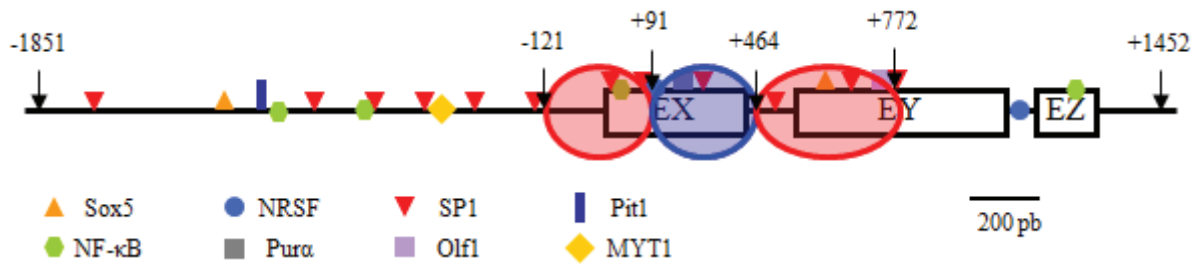


Figure 41 : Représentation schématique des régions activatrices et inhibitrices

Schéma réalisé d'après les résultats des figures 39 et 40. Les cercles rouges représentent les zones activatrices, le cercle bleu représente la zone inhibitrice.

En conclusion, les résultats présentés sur les figures 39 et 40 et résumés sur la figure 41 suggèrent qu'une série d'éléments régulateurs positifs et négatifs se succèdent à l'intérieur d'une séquence de 900 pb englobant les exons EX et EY. Cette région débute avec une forte séquence promotrice (-121 / +92) suivie d'une séquence contenant des éléments inhibant la transcription (+92 / +464) et terminée par une seconde séquence promotrice (+465 / +772) (Figure 41). Les données issues des tests d'activité luciférase n'ont cependant pas permis de mettre en évidence de séquence promotrice en amont de l'exon EZ dans la lignée SH-SY5Y. Etant donné la faible expression des transcrits initiés par EZ par rapport aux deux autres transcrits dans cette lignée cellulaire, comme en témoigne le résultat figure 31, il n'est pas surprenant que le promoteur initiant la transcription de l'exon EZ soit très faible dans le contexte des cellules SH-SY5Y.

5.3. Identification de sites potentiels de fixation pour des facteurs de transcription par mutagenèse dirigée

D'après les résultats exposés dans les figures 39 et 40, et comme résumé dans la figure 41, la région -121 / +772 contient potentiellement deux zones de régulation positives et une zone de régulation négative impliquées dans l'expression transcriptionnelle de *ST6GAL2*. Nous nous sommes donc intéressés plus particulièrement à cette séquence pour l'étude de sites potentiels de fixation de facteurs de transcription. Les analyses bioinformatiques réalisées (Figure 38) nous indiquent que cette séquence de 893 pb contient plusieurs sites potentiels de fixation pour des facteurs de transcription généraux et des facteurs de transcription associés aux fonctions neuronales. A partir des prédictions bioinformatiques nous avons sélectionné plusieurs sites potentiels de fixation. Nous avons choisi ces sites selon deux critères. D'une part en se basant sur l'homologie de leur séquence avec la séquence consensus connue pour chaque site. D'autre part, les sites de fixation potentiels ont été sélectionnés pour leur implication dans la régulation de gènes à fonctions cérébrales ou en tant que facteurs généraux de transcription. La séquence des sites de fixation des facteurs généraux NF- κ B et les quatre sites SP1 (SP1.1, SP1.2, SP1.3 et SP1.4), ainsi que des facteurs associés aux fonctions cérébrales NRSF, Olf1, Pura, et Sox5 a été mutée par substitution de base nucléotidique (Figure 42).

Ces mutations ont été effectuées par PCR. En prenant comme matrice de départ le vecteur pGL3_-121/+772 et pour chaque site, deux premières PCR ont été réalisées avec d'une part une amorce sens commune à toutes les mutations 5'-CGATAGGTACCGAGCTCTTACGCG-3' (contenant le site *KpnI*) et une amorce anti-sens spécifique du site muté, et d'autre part une amorce anti-sens commune à toutes les mutations 5'-TGCCAAGCTTACTTAGATCGCAGATCTC-3' (contenant le site *HindIII*) et une amorce sens spécifique du site muté (la séquence de l'amorce sens étant complémentaire de la séquence anti-sens pour un site muté donné). Les produits d'amplification de ces deux PCR ont été déposés sur gel d'agarose et les amplifications aux tailles attendues isolées du gel et purifiées pour servir de matrice à une troisième PCR avec le couple d'amorce commun cité ci-dessus (Figure 43). Les produits finaux ont été sous-clonés dans le vecteur pGL3-Basic aux sites *KpnI* / *HindIII*. Pour le site Olf1, une seule PCR a été réalisée avec l'amorce commune sens et l'amorce anti-sens spécifique du site Olf1 (Tableau 3, Matériels et Méthodes)

A

AGCTAGCGAGCCGGGGAGGATGCTGCGCTCCCCTGCGCCCCACCCAGCCGGCTG
 CCAGGCGCTGCCAGGCGCGCAGGTAGCCCCGGCCTTCCGCGCTGGCGGGGGGAGG
NF-κB
 TGTTGGACAGGGTACACAAAGACTCACTTCCGCCGGGGCAGGCGCCGGGACTCC
TTT A

SP1.1
CTCCTGCCGCCACAGCCTGTGCGCATTCTGCATTCTGCCGCCGCCGGGA
A AA AAA

NRSF
CCCGAGTGCCCCGAGGTGTCCAGGCGCGGTGCCAGGCGGGTACTGTGCAGGTA
GG GATT

Purα
 ACCTTCGGAAGAACTAGGGCTGGATGGAGGCGGCGGCCGGGGAGAAACACGCCT
TTACA

SP1.2
 TTTTAGGCAGCCGTGCTGCCCGCACCCGCCACGCAGCCCAAGTCTGCAGTCA
TTTT

CTCTCCGCGACAAGCAGAAATCCGCGCGCAGCCAGGGAGCGCCCCGCCTGCGCT
 CGCCGGCGGGCCGAGCTGCCGGCCGGGGGCGCGGGGAGGGGGCAGCCCACTT
 CCCAGTGCCCGGCTCCCGCCAGCTAGGTGAGCCTGGAGCGGGAAGGATGAAAGG

SP1.3
 GATCGCGGAGAGGGAGGGGCTCTGGGGCGAGCGGAGCTCGGGGGCGGGCGCG
TGTTT

GCGCAGCCTCGGTGCGTGGGCCGGGAACCGCCGCCTCTCCCCGGGGCTCGGGGC
Sox5
 CGCGGTTCTCTGGCCGCGGGAGCCGCCCGCGCTGCCGAAACAATGAGGCG
GAGT

SP1.4
TGCAGCGCGGGGCTGAGCGTAGCCTCCCGCCGGCTTTGAGGCCGAGCCACCCGC
TAAT

Olf1
 TGTGCCCGCCCGCCGCTTGAGCTGCCGCCGCGCGGAGCAGCGCTGGCCC
ATTT
 CGGAGAGGTGATCCCGCAGGCCCTGGGACGGGAGCGCGGGGCAGCTCCGAGA
 CCCACGGTGGCAGGAGGAGGACTCGAG

B

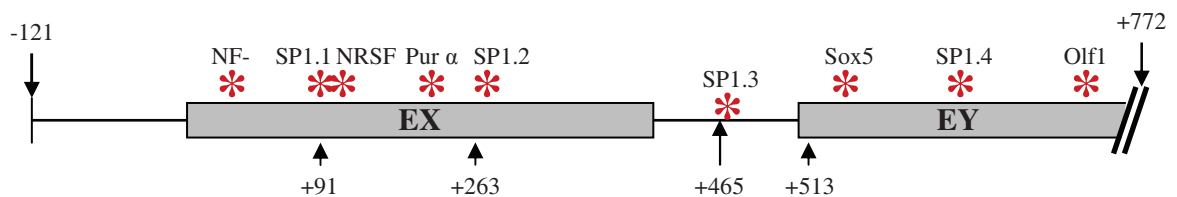


Figure 42 : Séquences et représentation des différents sites de fixation mutés sur la séquence -121 / 772

A, séquences des différents sites mutés. Pour chaque site de fixation (séquence soulignée), le nom du facteur de transcription est indiqué en bleu au dessus et les bases mutées sont en rouge. B, localisation des différents sites mutés sur la séquence génomique -121 / +772. Les nucléotides références utilisés pour les différentes constructions sont indiqués.

s'hybridant en 3' de la séquence -121 / +772. Les produits de PCR ont été sous-clonés dans le plasmide pGL3- Basic aux sites *KpnI/HindIII*. Nous avons vérifié par bioinformatique que les sites mutés ainsi générés n'étaient plus identifiés et n'avaient pas créé de nouveaux sites de fixation pour d'autres facteurs de transcription. De cette manière, les plasmides pGL3_-121/+772NF- κ B, pGL3_-121/+772NRSF, pGL3_-121/+772Olf1, pGL3_-121/+772Pur α , pGL3_-121/+772Sox5, pGL3_-121/+772SP1.1, pGL3_-121/+772SP1.2, pGL3_-121/+772SP1.3 et pGL3_-121/+772SP1.4 ont été produits. Ces vecteurs ont été transfectés dans les cellules SH-SY5Y et les protéines totales extraites 24 h après transfection. L'activité luciférase a été mesurée et comparée à l'activité obtenue avec le plasmide de référence non muté pGL3_-121/+772 (Figure 44A). La transfection de chacun des différents vecteurs a été réalisée au moins trois fois et les mesures d'activité luciférase de chaque extrait protéique ont été réalisées en duplicata. Les résultats obtenus avec les constructions pGL3_-121/+772SP1.1, pGL3_-121/+772SP1.2 et pGL3_-121/+772SP1.4 n'ont pas montré de différence d'activité significative par rapport au résultat obtenue avec le vecteur de référence (Figure 44A), suggérant que ces sites SP1 ne jouent pas de rôle dans la régulation du gène *ST6GAL2* dans les cellules SH-SY5Y. D'autre part, l'activité luciférase mesurée avec les plasmides pGL3_-121/+772NRSF et pGL3_-121/+772NF- κ B augmente respectivement de 15 % et 50 % comparativement au vecteur non muté (Figure 44A), indiquant un rôle répresseur de la transcription pour ces deux facteurs. Enfin, les activités luciférase obtenues avec les plasmides pGL3_-121/+772Pur α , pGL3_-121/+772Sox5, pGL3_-121/+772Olf1 et pGL3_-121/+772SP1.3 sont de 20 à 50 % plus faibles que l'activité observée avec le pGL3_-121/+772, suggérant que les facteurs Pur α , Sox5, Olf1 et SP1 (sur le 3^{ème} site étudié) sont des facteurs de transcription activateurs dans la régulation de ce gène (Figure 44). Prenant en compte le résultat obtenu avec le plasmide muté pGL3_-121/+772NF- κ B et le fait que le plasmide pGL3_-121/+91, a la plus forte activité luciférase et possède le site NF- κ B (Figure 40), nous avons construit le vecteur muté pGL3_-121/+91NF- κ B et avons testé son activité luciférase. Le résultat présenté dans la figure 44B révèle une forte augmentation de l'activité luciférase obtenu avec le plasmide muté comparé à celle obtenue avec le vecteur pGL3_-121/+91, confirmant que le facteur NF- κ B agirait ici comme un facteur répresseur.

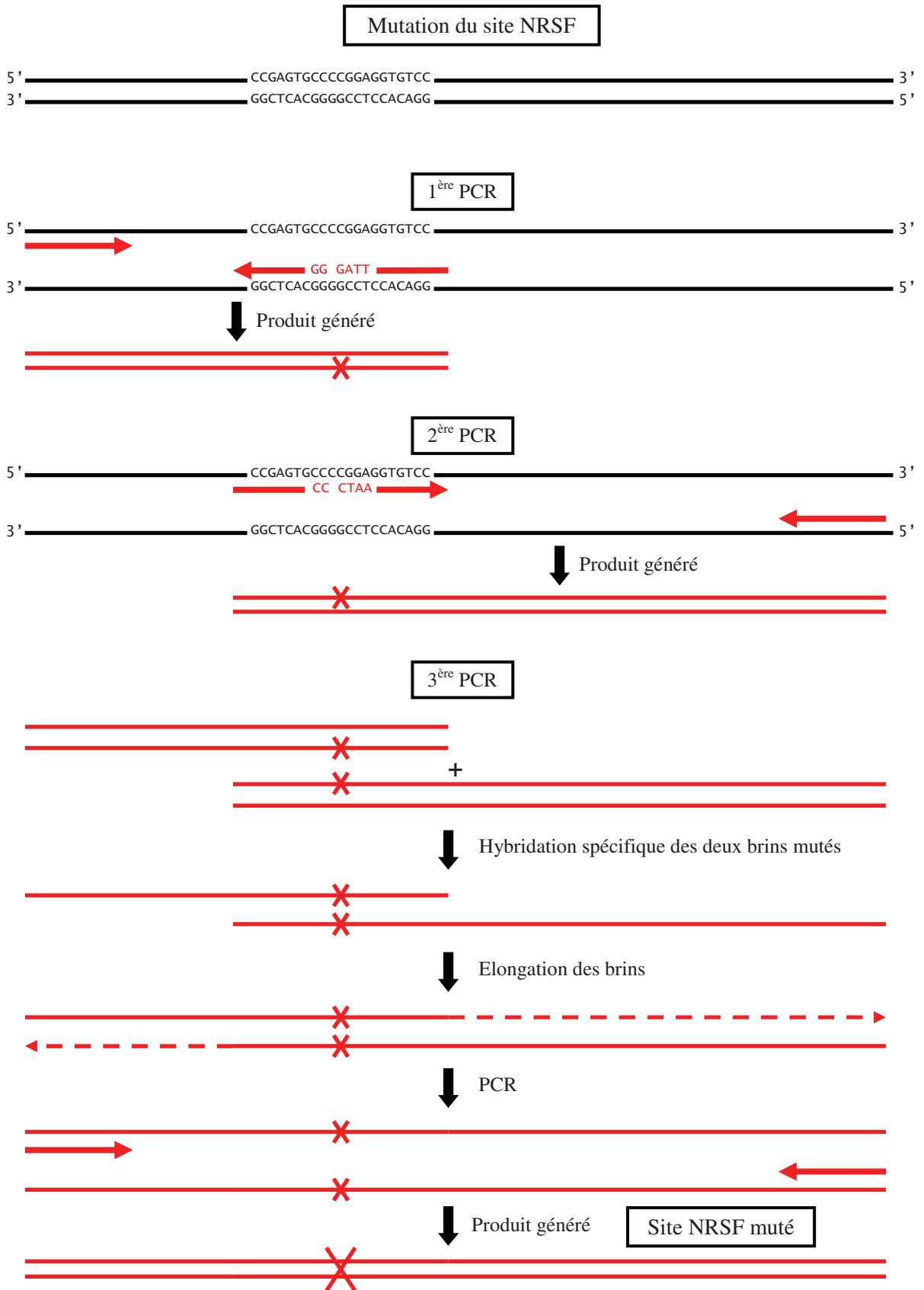


Figure 43 : Stratégie de mutation par mutagenèse dirigée de site par PCR.
 Exemple de mutation du site de fixation pour le facteur NRSF. Les flèches rouges orientées représentent les amorces utilisées. Les croix rouges représentent le site maintenant muté.

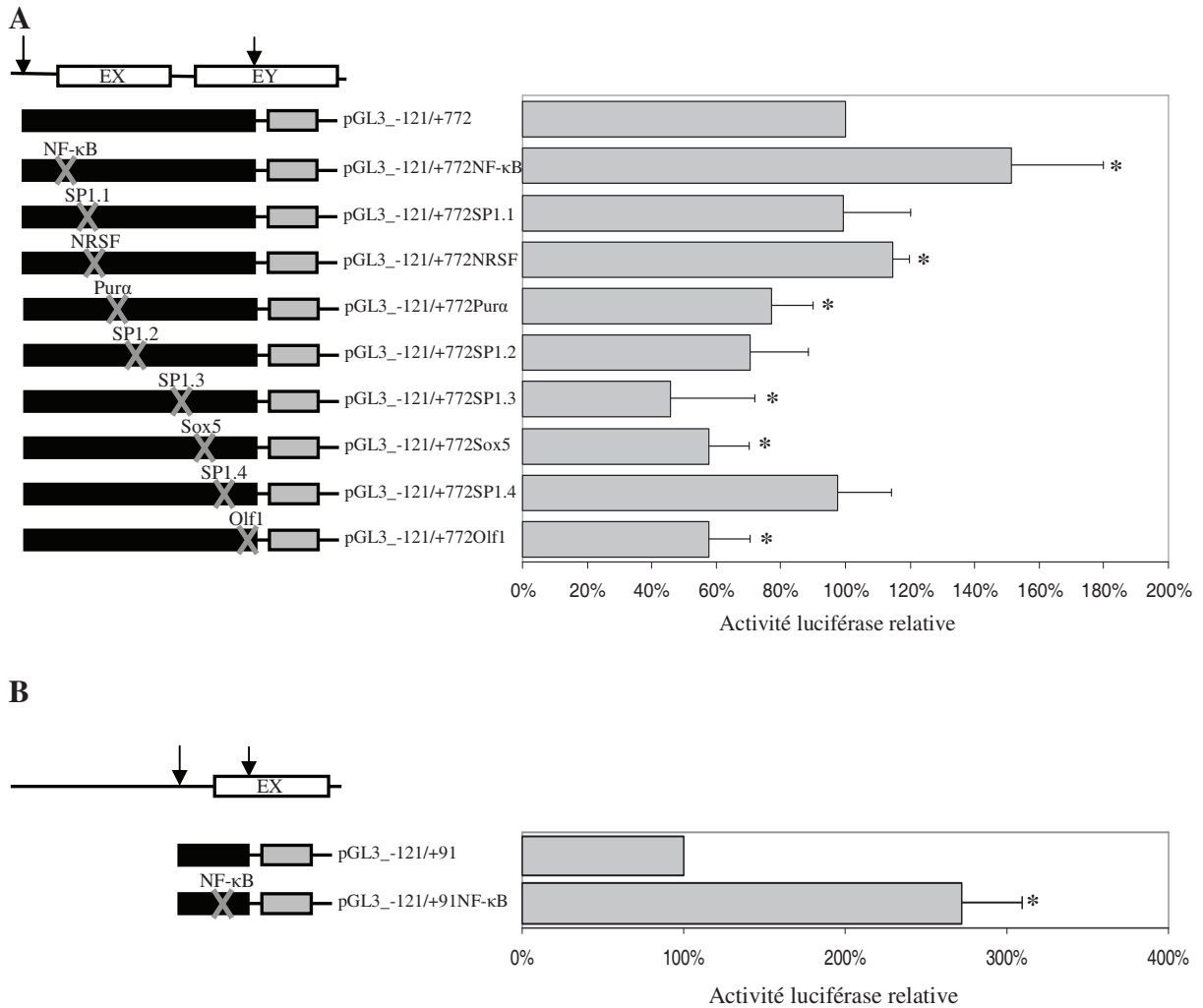


Figure 44 : Activité luciférase mesurée après transfection des différentes constructions mutées dans les cellules SH-SY5Y

Les différents vecteurs ont été co-transfectés dans les cellules SH-SY5Y avec le vecteur normalisateur codant la luciférase *Renilla*. **A**, sur la gauche : un schéma des différents plasmides mutés de la région -121 / +772. Sur la droite, le résultat de la mesure des activités luciférase. **B**, sur la gauche, un schéma des différents plasmides mutés de la région -121 / +772. Sur la droite, le résultat de la mesure des activités luciférase. Les activités luciférase sont normalisées avec l'activité de la luciférase *Renilla* et exprimées en % de l'activité du vecteur pGL3_-121/+772 (**A**) ou pGL3_-121/+91 (**B**). Les valeurs représentent la moyenne de $n \geq 3$ expériences indépendantes \pm SD. *, $p < 0,05$ vs pGL3_-121/+772 (**A**) ou pGL3_-121/+91 (**B**).

Au niveau du système nerveux central, le rôle des facteurs de transcription de la famille NF-κB a été décrit dans des processus tels que la plasticité neuronale, la formation de la myéline, la transmission synaptique et l'inflammation (Meffert *et al.*, 2005). Le site potentiel de fixation au facteur NF-κB étudié ici a été prédit par bioinformatique comme fixant l'homodimère de la protéine p50. p50 ne contient pas le domaine transactivateur et ne

peut permettre d'activer la transcription du gène. De plus l'homodimère p50/p50 agit comme un répresseur de la transcription (Kastenbauer *et al.*, 1999 ; Grunstrom *et al.*, 2004 ; Kostadinova *et al.*, 2005). Les résultats de la mutagenèse suggèrent fortement que l'homodimère p50/p50 soit impliqué dans l'inhibition de la transcription de *ST6GAL2* dans la lignée SH-SY5Y (Figure 44). Le facteur NRSF (Neuron-Restrictive Silencer Factor) inhibe principalement la transcription de gènes à fonctions neuronales dans un contexte cellulaire ou tissulaire non neuronal (Ballas *et al.*, 2005 ; Lunyak *et al.*, 2005). Nos données concernant le site putatif de fixation du facteur NRSF vont dans le sens d'un rôle répresseur de la transcription pour ce facteur et laissent supposer que l'expression de *ST6GAL2* pourrait être activée pour une ou plusieurs fonctions spécifiques des cellules neuronales. Olf1, Sox5 et Pura sont tous les trois des facteurs de transcription dont l'implication a été montrée dans la neurogenèse et, plus généralement, dans le fonctionnement neuronale (Wang *et al.*, 1997 ; Khalili *et al.*, 2003 ; Darbinian *et al.*, 2006 ; Lai *et al.*, 2008). Olf1 (Olfactory neuron-specific binding protein) est un facteur de transcription de type Hélice-boucle-hélice initialement décrit comme régulant l'expression de gènes impliqués dans les fonctions olfactives, notamment dans les neurones olfactifs (Kudrycki *et al.*, 1993 & 1998). D'une façon plus générale, Olf1 est aussi impliqué dans la différenciation et la régénération neuronale (Garel *et al.*, 1997 ; Wang *et al.*, 1997). Sox5 (SRY box 5) est un facteur possédant un domaine de fixation du type HMG (High Mobility Group) et a premièrement été rapporté comme l'un des gènes, exprimé par le chromosome Y, responsable du déterminisme masculin des gonades durant l'embryogenèse (Koopman *et al.*, 1991). Cependant, plusieurs études ont montré l'implication de ce facteur de transcription dans la genèse de certains types de neurones particuliers (neurones corticofuges) lors de la formation du cortex cérébral (Lai *et al.*, 2008 ; Leone *et al.*, 2008). Les facteurs de la famille Sox sont tout particulièrement impliqués dans l'embryogenèse et la différenciation cellulaire. Enfin, le facteur de transcription Pura (Purine-Rich element binding protein α) est un facteur plus ubiquiste qu'Olf1 et Sox5 et se fixe par ailleurs aussi bien sur l'ADN simple brin que sur l'ARN (Galia *et al.*, 2000). L'inactivation du gène codant Pura chez la souris révèle de sévères dysfonctionnements du système nerveux accompagnés d'un retard du développement neuronal dans l'hippocampe et le cervelet notamment (Khalili *et al.*, 2003).

L'implication des facteurs Olf1, Sox5 et Pura dans la régulation de la transcription du gène *ST6GAL2*, comme le suggèrent nos résultats, est en faveur d'une spécificité d'expression neuronale de ce gène. Finalement, SP1 (Specificity Protein 1) est un facteur de transcription

général activateur se liant sur des séquences riche en GC. Quatre de ses sites de fixation potentiels ont été testés dans la séquence -121 / +772. Nos données indiquent que le troisième site étudié (SP1.3) serait significativement impliqué dans l'activation de la transcription (Figure 44A).

Pour conclure sur la caractérisation des régions promotrices de *ST6GAL2*, les travaux que nous avons menés ont permis d'identifier la présence d'une région fortement promotrice en amont de l'exon EX et d'une région promotrice, plus modérée, conduisant vraisemblablement à l'initiation de la transcription de l'exon EY. Entre les deux régions promotrices réside probablement une zone de régulation négative comme l'attestent nos données. Par ailleurs, nos résultats ne nous ont pas permis de conclure quant à l'existence d'une séquence promotrice dédiée à l'expression des transcrits contenant l'exon EZ. Par mutagenèse dirigée nous avons mis en évidence un rôle potentiellement répresseur des facteurs NF- κ B et NRSF ainsi qu'un rôle activateur des facteurs SP1, Pura, Sox5 et Olf1. Néanmoins, les éléments responsables de la forte activité promotrice observée pour la séquence -121 / + 92 restent à identifier.

6. EXPRESSION DU GÈNE *ST6GAL2* ET DE LA PROTEINE HST6GAL II APRES DIFFERENCIATION NEURONALE DES CELLULES NT2/D1 PAR L'ACIDE RETINOÏQUE

L'expression du gène *ST6GAL2* semble être gouvernée par certains facteurs de transcription impliqués dans le fonctionnement et le développement cérébral. Afin d'étudier l'expression de ce gène dans un contexte de développement neuronale, nous nous sommes intéressés à un modèle de différenciation neuronale *in vitro*.

6.1. Les cellules NT2/D1, un modèle cellulaire de développement neuronal

L'acide rétinoïque (AR) est un dérivé naturel de la vitamine A qui régule de nombreux processus biologiques dont la croissance, la survie et la différenciation de plusieurs types cellulaires. Il a été montré à de nombreuses reprises que les composés de la famille des rétinoïdes, à laquelle appartiennent la vitamine A et l'AR, sont impliqués dans la régulation de l'expression de gènes contrôlant la différenciation neuronale, la croissance des neurites et l'établissement de l'axe antéropostérieur du tube neural (Maden, 2002 ; McCaffery *et al.*, 2003). Plus récemment, le rôle des rétinoïdes a été souligné dans différents aspects du fonctionnement du système nerveux central adulte (Mey *et al.*, 2004 ; Lane *et al.*, 2005).

Les cellules NT2/D1 sont une lignée cellulaire de carcinome embryonnaire pluripotente qui ressemble aux cellules embryonnaires précoces dans leur morphologie, leur profil d'expression antigénique et leur profil d'expression de gènes (Andrews, 1984 & 1998). De plus, ces cellules expriment plusieurs marqueurs typiques de cellules épithéliales neurales et de cellules pro-neurales tels que des filaments intermédiaire neuroépithéliaux de kératine et de nestine et les molécules d'adhésion cellulaires NCAM et N-cadhérine (Lee *et al.*, 1986 ; Pleasure *et al.*, 1993). En présence d'AR, les cellules NT2/D1 s'engagent de façon irréversible dans un lignage neuronal (Andrews, 1984 ; Lee *et al.*, 1986), fournissant ainsi un excellent modèle d'étude *in vitro* pour l'étude de gènes humains impliqués dans la différenciation neuronale. Après traitement à l'AR, les NT2/D1 se différencient en cellules neuronales ressemblant aux neurones du système nerveux central (Pleasure *et al.*, 1992). Les cellules neuronales ainsi obtenues possèdent des caractéristiques neuronales morphologiques, telles que l'extension de dendrites et d'axones, et des caractéristiques neuronales moléculaires,

comme l'expression de la MAP-2 (micro-tubule-associated protein 2) et de neurofilaments, ou l'expression de protéine de surface comme la NCAM et la GAP-43 (growth-associated protein 43) (Andrews, 1984 ; Pleasure *et al.*, 1992 ; Zigova *et al.*, 2000).

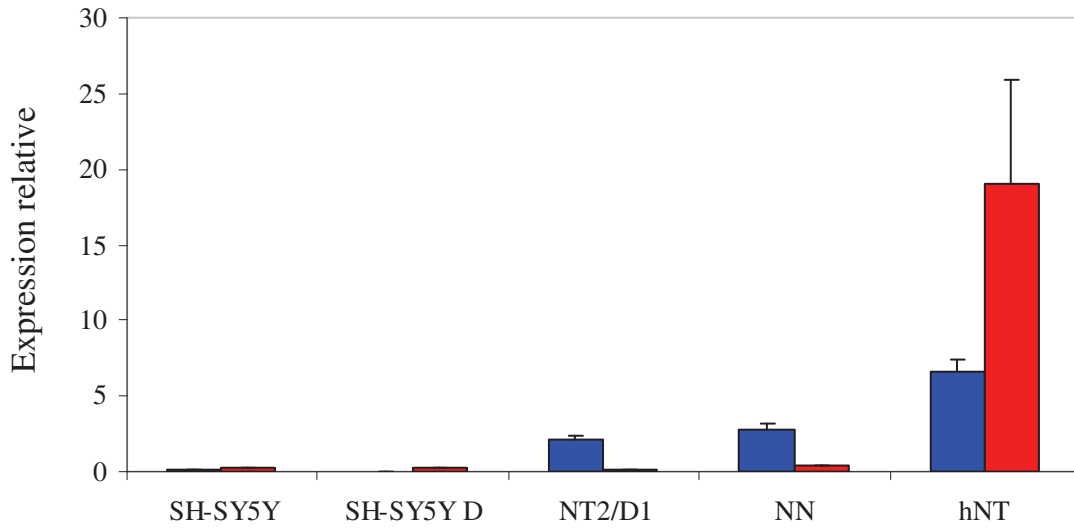
Dérivant d'un neuroblastome, les cellules SH-SY5Y peuvent également être exposées à l'AR pour acquérir davantage de propriétés neuronales, dont une croissance des neurites et un phénotype de neurones cholinergiques matures (Pahlman *et al.*, 1984). L'AR induit également les signaux de survie cellulaire et augmente la résistance des SH-SY5Y aux neurotoxines (Cheung *et al.*, 2009). Par opposition aux cellules NT2/D1, la différenciation des cellules SH-SY5Y par l'AR est un phénomène réversible dès l'arrêt du traitement et leur différenciation neuronale est moins prononcée sur le plan morphologique et biochimique.

6.2. Induction de l'expression du gène *ST6GAL2* dans les NT2/D1 après différenciation

Après exposition à l'AR et différenciation morphologique des cellules NT2/D1, les neurones post-mitotiques forment la couche supérieure et sont assemblés en neurosphères dans la boîte de culture. Les neurones, appelés hNT, sont séparés de la couche non différenciée de cellules, appelées NN (Matériels et Méthodes). Les deux types de populations cellulaires ont été récupérés pour la suite des analyses. Par ailleurs, les cellules SH-SY5Y ont également été différenciées par AR et récupérées (les SH-SY5Y forment une population cellulaire unique) après traitement (Matériels et Méthodes). Trois lots d'ARN issus de trois expériences indépendantes, pour les SH-SY5Y et pour les NT2/D1, ont été analysés. L'expression globale des gènes *ST6GAL1* et *ST6GAL2* a été analysée par Q-PCR pour toutes ces populations cellulaires : SH-SY5Y différenciées (SH-SY5Y D) ou non (SH-SY5Y), les NT2/D1 avant traitement (NT2/D1), différenciées en neurones (hNT) ou non différenciées (NN) après traitement. Le résultat de la figure 45A montre nettement une très forte augmentation de l'expression de *ST6GAL2*, d'un facteur 150 en moyenne, dans la fraction neuronale (hNT) alors que l'expression reste pratiquement inchangée dans la fraction non différenciée (NN) par comparaison aux NT2/D1. Dans le même temps, l'expression du gène *ST6GAL1* n'est augmentée que d'un facteur 3 dans les hNT et reste stable dans les NN comparée à l'expression dans les NT2/D1 (Figure 45A). En parallèle, l'expression des gènes

ST6GAL1 et *ST6GAL2* reste inchangée après différenciation des cellules SH-SY5Y par l'AR (Figure 45B).

A



B

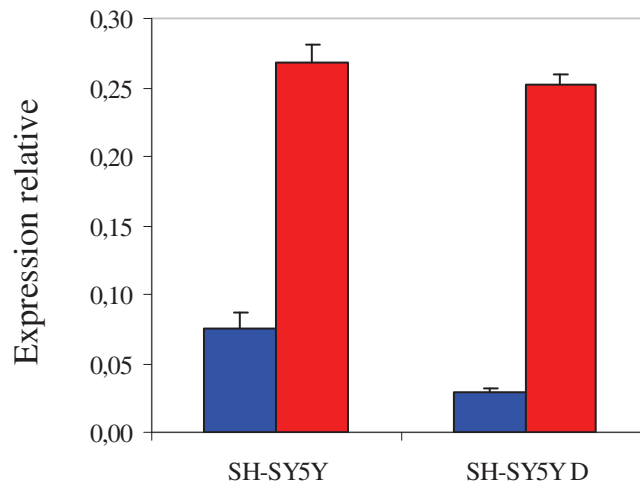


Figure 45 : Analyse par Q-PCR de l'expression des gènes *ST6GAL1* et *ST6GAL2* avant et après différenciation par l'AR des cellules SH-SY5Y et NT2/D1.

Les valeurs relatives des expressions des gènes *ST6GAL1* et *ST6GAL2* sont exprimées par rapport à l'expression du gène de référence *HPRT*. L'expression des gènes *ST6GAL1*, *ST6GAL2* et *HPRT* a été mesurée par Q-PCR en utilisant la chimie SYBR Green®. L'expression relative du gène *ST6GAL1* est représentée en bleu, celle du gène *ST6GAL2* est représentée en rouge. **A**, expression de *ST6GAL1* et *ST6GAL2* dans les cellules SH-SY5Y, SH-SY5Y D, NT2/D1, NN et hNT. **B**, expression de *ST6GAL1* et *ST6GAL2* dans les cellules SH-SY5Y et SH-SY5Y D. Les valeurs représentent la moyenne de n = 3 expériences indépendantes ± SD

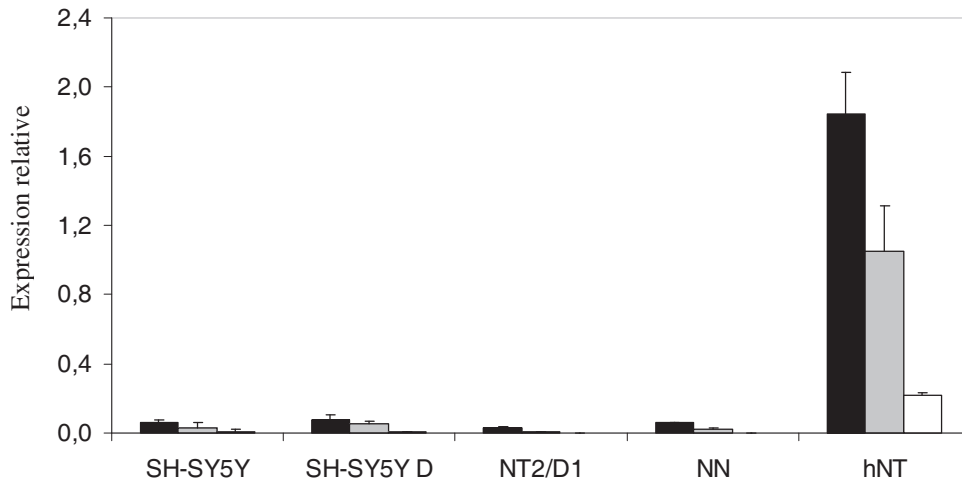
Ces données révèlent une activation transcriptionnelle importante du gène *ST6GAL2* lors de la différenciation neuronale des cellules. Bien que les deux modèles étudiés ici soient différents, la différenciation neuronale sous acide rétinoïque des cellules NT2/D1 est considérée comme plus « aboutie » et définitive par comparaison avec les cellules SH-SY5Y pour lesquelles la différenciation est moins prononcée et réversible dès l'arrêt du traitement. Néanmoins, une surexpression de plus d'un facteur 100 suggère fortement un rôle important de hST6Gal II au cours de ce processus.

6.3. Induction de l'expression des transcrits initiés par EX, EY et EZ dans les NT2/D1 après différenciation

Les trois transcrits contenant les exons EX, EY et EZ sont exprimés dans les cellules NT2/D1 avant différenciation et nos données suggèrent qu'ils soient les seuls transcrits du gène *ST6GAL2* exprimés cette lignée (Figure 31). Afin de vérifier si ces trois types de transcrits sont surexprimés lors de la différenciation neuronale nous avons quantifié leur expression dans les populations neuronale (hNT) et non-neuronale (NN). L'expression des transcrits contenant EX, EY ou EZ a été mesurée par Taqman Q-PCR en duplex avec le gène *HPRT*. Il apparaît clairement dans la figure 46 que l'expression de ces trois transcrits est fortement augmentée et d'une manière similaire (augmentation de 150 fois en moyenne) à celle de l'expression globale du gène *ST6GAL2* présentée dans la figure 45. Ce résultat suggère qu'aucun autre type de transcrite du gène *ST6GAL2* à l'exception de ceux identifiés précédemment ne soit exprimé après différenciation. Conformément au résultat de la figure 45 concernant l'expression de *ST6GAL2* dans les cellules SH-SY5Y au cours de leur différenciation, l'expression des trois types de transcrits ne varie pratiquement pas (Figure 46).

Pour autant, cette forte augmentation transcriptionnelle ne modifie pas le profil d'expression des transcrits contenant EX, EY ou EZ, il reste identique à celui observé avant différenciation : EX>EY>EZ. Cette observation laisse sous-entendre deux choses : (1), les régions promotrices impliquées dans cette surexpression sont probablement identiques à celles décrites précédemment. (2), l'ensemble de ces régions est activé puisque que les trois types de transcrits sont surexprimés avec un profil similaire à celui observé avant différenciation. Par ailleurs l'expression de *ST6GAL2* et de ses différents transcrits dans les hNT est bien supérieure à celle observée dans les différents échantillons du tissu cérébral que nous avons testé auparavant (Figures 32 et 33). L'expression relative moyenne du gène *ST6GAL2* dans les hNT est de 19,0 contre 0,69 en moyenne dans les échantillons *in vivo* que nous avons testés. Ceci suggère que le niveau d'expression du gène *ST6GAL2* pourrait être plus important au cours du développement du système nerveux central qu'à l'état adulte.

A



B

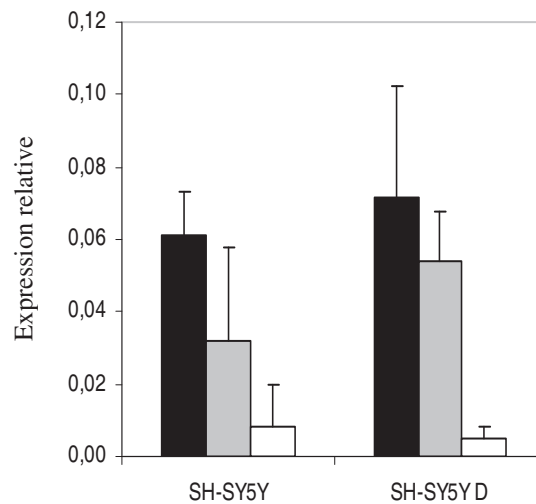


Figure 46 : Analyse par Q-PCR de l'expression des différents transcrits du gène *ST6GAL2* avant et après différenciation par l'acide rétinoïque des cellules SH-SY5Y et NT2/D1.

Les valeurs relatives des expressions des transcrits du gène *ST6GAL2* sont exprimées par rapport à l'expression du gène de référence *HPRT*. En noir est représentée l'expression des transcrits initiés par l'exon EX, en gris est représentée l'expression des transcrits initiés par l'exon EY et en blanc est représentée l'expression des transcrits initiés par l'exon EZ. L'expression de chacun des transcrits ainsi que du gène de référence *HPRT* a été mesurée par la technologie Taqman et en duplex (un des transcrits + *HPRT*). **A**, expression relative des différents transcrits de *ST6GAL2* dans les cellules SH-SY5Y, SH-SY5Y D, NT2/D1, NN et hNT. **B**, expression relative des différents transcrits de *ST6GL2* dans les cellules SH-SY5Y et SH-SY5Y D. Les valeurs représentent la moyenne de n = 3 expériences indépendantes \pm SD.

6.4. Surexpression de la protéine hST6Gal II dans les cellules hNT

Nous avons voulu vérifier si l'augmentation transcriptionnelle du gène *ST6GAL2* mesurée par Q-PCR était corrélée à une surexpression de la protéine hST6Gal II. Pour cela, les extraits protéiques des cellules NN et hNT après traitement à l'AR ainsi que des cellules NT2/D1 avant traitement ont été analysés par Western-Blot afin d'étudier leur expression de la protéine hST6Gal II. En accord avec l'augmentation de l'expression transcriptionnelle du gène *ST6GAL2* observée (Figures 45 et 46), nous avons révélé une augmentation importante de la quantité de la protéine hST6Gal II dans les cellules hNT par rapport aux cellules NT2/D1 non traitées et par rapport aux cellules de la fraction non différenciée NN (Figure 47).

A l'inverse des expressions transcriptionnelles et protéiques observées pour les échantillons cérébraux, ces données issues de la différenciation des NT2/D1 suggèrent une bonne corrélation entre l'expression transcriptionnelle et protéique de *ST6GAL2*.

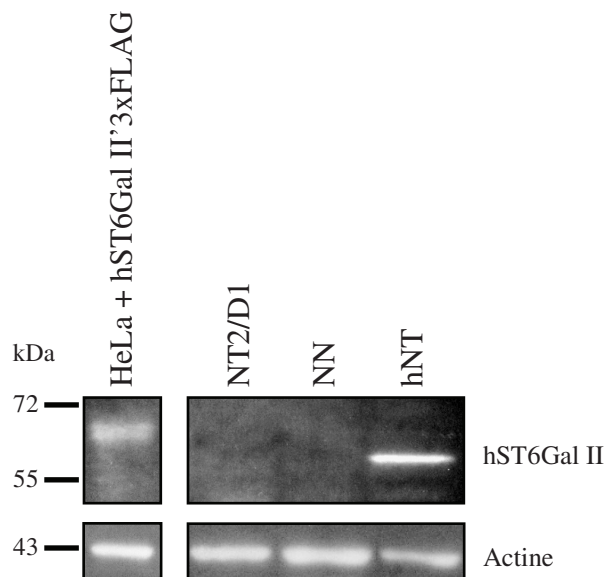


Figure 47 : Analyses par Western-Blot de l'expression de la protéine hST6Gal II avant et après différenciation par l'AR des cellules NT2/D1.

40 μ g d'extraits protéiques ont été déposés dans chaque piste et séparés par SDS/PAGE sur mini-gel à 8 %. Après transfert sur membrane de nitrocellulose, la révélation de hST6Gal II et de l'actine par Western-Blot est effectuée comme décrit dans le Matériels et Méthodes. La piste HeLa + hST6Gal II-3xFLAG a été utilisée comme témoin positif. La révélation de l'actine a été utilisée comme témoin de charge.

L'ensemble de ces résultats préliminaires suggère un rôle important de la sialyltransférase hST6Gal II au cours de la différenciation neuronale. De plus, l'ensemble des régions promotrices contrôlant l'expression des exons EX, EY et EZ semble être fortement activé durant ce processus. Dans ce cadre, la corrélation entre l'expression (et la surexpression) du gène *ST6GAL2* et de la protéine hST6Gal II ne fait aucun doute ; elle s'oppose aux observations faites dans les échantillons cérébraux précédemment testés (Figure 37). Cependant, le modèle cellulaire employé ici ne comporte qu'un type cellulaire et est plus uniforme qu'un échantillon tissulaire regroupant plusieurs types de cellules (cellules neuronales, cellules gliales...).

L'étude des structures sialylées avant et après différenciation des NT2/D1 pourrait confirmer une augmentation de l' α 2,6- sialylation globale ou l'augmentation/apparition plus spécifique de structures Sia₆LacdiNAc induites par la surexpression de hST6Gal II. Des études préliminaires en utilisant la lectine SNA, spécifique des motifs Sia₆Lac(di)NAc, sur des extraits protéiques totaux de NT2/D1, NN et hNT n'ont pas révélé de différences significatives. Il n'est cependant pas du tout exclu qu'un nombre très restreint de *N*-glycoprotéines soient les substrats de hST6Gal II, particulièrement s'il s'agit de former du Sia₆LacdiNAc, une structure, rappelons-le, rare parmi les *N*-glycosylprotéines humaines. Enfin, l'invalidation du gène *ST6GAL2* et/ou *ST6GAL1* nous renseignerait sur le rôle et l'importance de la sialylation en α 2,6- médiée par hST6Gal II au cours de la différenciation neuronale.

Discussion et Perspectives

hST6Gal II est l'une des dernières sialyltransférases à avoir été identifiée et caractérisée chez l'Homme (Takashima *et al.*, 2002 ; Krzewinski-Recchi *et al.*, 2003). Elle se distingue de hST6Gal I, l'autre membre de la famille ST6Gal, par sa spécificité de substrat et par son profil d'expression tissulaire. Contrairement à hST6Gal I, le substrat préférentiel de hST6Gal II déterminé *in vitro* est le motif LacdiNAc (GalNAc β 1-4GlcNAc) (Tableaux 3 et 4, Introduction bibliographique) (Rohfritsch *et al.*, 2006). Cependant, les structures Sia₆LacdiNAc n'ont été que rarement décrites sur les *N*-glycosylprotéines humaines. Concernant l'expression du gène *ST6GAL2*, celui-ci présente un profil d'expression très restreint. Alors que *ST6GAL1* est exprimé dans presque tous les tissus, *ST6GAL2* est principalement exprimé dans le cerveau adulte et embryonnaire (Figure 18, Introduction bibliographique) (Krzewinski-Recchi *et al.*, 2003). Les travaux réalisés sur le gène murin et bovin codant ST6Gal II (respectivement, *mst6gal2* et *bst6gal2*) révèlent également une expression préférentielle au niveau du système nerveux central (Figure 19, Introduction bibliographique) (Takashima *et al.*, 2003 ; Laporte *et al.*, 2009). De plus, l'unique α 2,6-sialyltransférase identifiée chez *D. melanogaster*, DSiaT, présente elle aussi des similitudes avec hST6Gal II en termes d'activité enzymatique et d'expression tissulaire (Tableau 4, Figure 23, Introduction bibliographique) (Koles *et al.*, 2004 & 2008). En effet, si *DSiaT* est un orthologue du gène ancestral apparu avant la duplication qui a donné les deux sous-familles ST6Gal I et II chez les vertébrés (Figure 9, Introduction bibliographique) (Harduin-Lepers *et al.*, 2005), hST6Gal II semble avoir retenu certaines caractéristiques ancestrales vis-à-vis notamment de la spécificité de substrat accepteur et de l'expression tissulaire du gène. L'expression transcriptionnelle du gène *ST6GAL2* semble donc finement régulée par des mécanismes apparemment conservés depuis la drosophile jusque la souris et l'Homme.

L'expression des sialyltransférases est principalement régulée au niveau transcriptionnel. Cette régulation s'effectue par l'utilisation de promoteurs tissulaires spécifiques, qui conduisent à l'expression de transcrits codant la même protéine mais possédant des séquences 5'-non traduites différentes. L'un des exemples les mieux décrits concerne la régulation de l'expression du gène *ST6GAL1* (Figures 13 et 14, Introduction bibliographique). Trois promoteurs ont été caractérisés, conduisant à l'expression de transcrits ubiquistes, hépatiques et lymphocytaires (Figure 14, Introduction bibliographique).

Le gène *ST6GAL2* est composé de cinq exons codants répartis sur plus de 85 kpb sur le chromosome 2 (Figure 15, Introduction bibliographique). Avant que les travaux présentés dans cette thèse ne débutent, seules deux séquences de 64 et 119 pb rapportés par Takashima et collaborateurs, et Krzewinski-Recchi et collaborateurs, ainsi que quelques EST présentes

dans les bases de données suggéraient la présence de deux exons alternatifs en 5' du gène *ST6GAL2* (Figures 15 et 16, Introduction bibliographique). Les résultats présentés dans ce manuscrit ont essentiellement contribué à l'identification et à la caractérisation des extrémités 5' non traduites du gène *ST6GAL2* chez l'Homme.

RAPPEL DES RESULTATS OBTENUS

1. IDENTIFICATION DES EXTREMITES 5' NON TRADUITES DU GENE *ST6GAL2* DANS LA LIGNEE DE NEUROBLASTOME HUMAIN SH-SY5Y

En préalable à l'étude de l'expression et de la régulation du gène *ST6GAL2*, nous avons testé par Q-PCR plusieurs lignées cellulaires afin de sélectionner notre modèle d'étude, la lignée cellulaire de neuroblastome humain SH-SY5Y. Les cellules SH-SY5Y présentent le double intérêt d'exprimer fortement *ST6GAL2*, par rapport aux autres lignées cellulaires testées, et d'exprimer faiblement le gène *ST6GAL1* (Figure 24, Résultats).

L'identification des extrémités 5' non traduites a été réalisée par 5' RACE sur ARNm coiffés à partir d'ARN totaux extraits des cellules SH-SY5Y. Les multiples expériences de 5' RACE que nous avons réalisées nous ont permis de mettre en évidence l'utilisation de trois exons alternatifs, mutuellement exclusifs, parmi les transcrits du gène *ST6GAL2*. Ces exons, que nous avons appelés EX, EY et EZ ont une taille respective de 393, 524 et 140 pb (Figure 26, Résultats). Ils sont localisés à plus de 42 kb du premier exon codant E1 et ne sont séparés que de 124 et 87 pb sur le génome.

Nous avons également montré que le site d'initiation de la transcription (TSS) de l'exon EX est très variable, pas moins de sept TSS répartis sur 76 pb ont été mis en évidence. A l'inverse, seulement deux TSS différents d'une pb ont été mis en évidence pour l'exon EY et aucune variation du TSS de EZ n'a été révélée expérimentalement (Figure 28, Résultats).

2. EXPRESSION RELATIVE DES TROIS TRANSCRITS DU GENE *ST6GAL2* DANS DIFFERENTES LIGNEES CELLULAIRES EN CULTURE ET DANS DIFFERENTS ECHANTILLONS DU CORTEX CEREBRAL, DU CERVELET ET DE L'HIPPOCAMPE

2.1. Lignées cellulaires en culture

Pour étudier l'expression relative des trois transcrits identifiés par 5' RACE, nous avons mis au point une méthode de quantification par Taqman Q-PCR en duplex avec le gène de référence *HPRT*. Cette technique nous a permis de montrer que dans les lignées cellulaires SH-SY5Y, NT2/D1 et Kelly les transcrits contenant l'exon EX sont exprimés de façon prédominante, suivis des transcrits initiés par l'exon EY alors que les transcrits contenant l'exon EZ sont beaucoup plus faiblement exprimés (Figure 31, Résultats).

2.2. Cortex cérébral, cervelet et hippocampe

A partir de 25 échantillons de tissu cérébral provenant de 8 sujets différents (deux échantillons du lobe frontal, 6 du lobe occipital, 3 du lobe pariétal, 7 du lobe temporal, 4 du cervelet et 3 de l'hippocampe), nous avons montré que le gène *ST6GAL1* et le gène *ST6GAL2* sont exprimés à des niveaux quasiment équivalents dans chacun des échantillons testés. L'expression de ces gènes est par ailleurs hétérogène entre les différents sujets et entre les différents tissus cérébraux examinés d'un même sujet (Figure 32, Résultats).

Nos résultats montrent que les trois transcrits initiés par les exons EX, EY et EZ sont exprimés dans les 25 échantillons de tissu cérébral. A l'exception de trois échantillons (6-Occ, 8-Par et 1-Cer), le niveau d'expression relatif des transcrits contenant EX, EY ou EZ est comparable au niveau relatif de l'expression globale de *ST6GAL2* pour tous les échantillons. Pour les régions du cortex cérébral et le cervelet, le profil d'expression relatif des transcrits contenant les exons EX, EY et EZ est comparable à celui observé pour les lignée SHSY5Y, NT2/D1 et Kelly (Figure 33B, Résultats), EX et EY sont les exons majoritairement exprimés. Les trois échantillons d'hippocampe ont, quant à eux, un profil d'expression des trois transcrits nettement différent. Dans ce tissu, les transcrits initiés par EY sont les plus exprimés, suivis par ceux initiés par EZ et les transcrits contenant EX sont les moins exprimés (Figure 33A, Résultats).

3. EXPRESSION DE LA PROTEINE hST6GAL II DANS DIFFERENTES LIGNEES CELLULAIRES EN CULTURE ET DANS DIFFERENTS ECHANTILLONS DU CORTEX CEREBRAL, DU CERVELET ET DE L'HIPPOCAMPE

Avec un anticorps anti-hST6Gal II, nous avons pu montrer que la protéine hST6Gal II est exprimée dans différents échantillons du cortex cérébral, du cervelet et de l'hippocampe (Figure 36, Résultats). Comme pour l'expression transcriptionnelle, l'expression protéique est relativement hétérogène. Par ailleurs, comme nous l'avons montré pour certains échantillons, l'expression du gène *ST6GAL2* n'est que faiblement corrélée à l'expression de la protéine hST6Gal II (Figure 37, Résultats). Dans des conditions expérimentales similaires, l'expression de la protéine hST6Gal II dans les lignées cellulaires SH-SY5Y et NT2/D1 n'a été que très faiblement révélée (Figure 35, Résultats).

4. IDENTIFICATION ET CARACTERISATION DES REGIONS PROMOTRICES DU GENE *ST6GAL2* DANS LA LIGNEE CELLULAIRE SH-SY5Y

L'étude de l'activité promotrice de diverses régions génomiques par des tests d'activité luciférase, nous a permis d'identifier deux séquences promotrices. La première séquence est localisée entre les nucléotides -121 et +91 en amont de l'exon EX et possède une forte activité promotrice. La seconde région, située en +465 / +772 et partiellement en amont de l'exon EY, possède une activité promotrice plus modérée (Figures 39 et 40, Résultats). Avec ces expériences, nous n'avons pu mettre en évidence de région promotrice correspondant à la transcription de l'exon EZ.

La région -121 / +772, qui regroupe des zones activatrices et inhibitrices de la transcription (Figure 41, Résultats), a été sélectionnée pour l'identification de facteurs de transcription potentiels. Par des analyses bioinformatiques et une étude par mutagenèse dirigée des sites de fixation potentiels de facteurs de transcription, nous avons montré que les facteurs NF- κ B et NRSF pourraient jouer un rôle répresseur de la transcription. Parallèlement, ces résultats indiquent également que les facteurs SP1 (sur le troisième site étudié), Pura, Sox5 et Olf1 seraient des éléments activateurs de la transcription du gène *ST6GAL2* (Figures 38 et 44, Résultats).

5. SUREXPRESSION DU GENE *ST6GAL2* ET DE LA PROTEINE hST6GAL II APRES DIFFERENCIATION NEURONALE DES CELLULES NT2/D1 SOUS L'ACTION DE L'ACIDE RETINOÏQUE

La lignée cellulaire NT2/D1 est un excellent modèle *in vitro* de différenciation neuronale. Nous avons montré que l'expression du gène *ST6GAL2* est augmentée d'environ 150 fois uniquement dans la population cellulaire différenciée en neurones. Dans le même temps, l'expression du gène *ST6GAL1* ne présente qu'une faible augmentation d'un facteur 3 (Figure 45, Résultats).

Nous avons également montré que les transcrits initiés par les exons EX, EY et EZ sont les transcrits fortement induits au cours du processus de différenciation (Figure 46, Résultats). Par ailleurs, le profil d'expression relatif des trois types de transcrits reste inchangé entre les cellules NT2/D1 et les cellules hNT (EX>EY>EZ), suggérant que l'ensemble de la région promotrice EX/EY/EZ est activée.

En parallèle, nous avons également montré par des expériences de Western-Blot que la protéine hST6Gal II est également fortement surexprimée après différenciation neuronale (Figure 47, Résultats).

DISCUSSION GENERALE

1. ORGANISATION GENOMIQUE ET EXPRESSION DES EXONS EX, EY ET EZ

Une organisation génomique singulière, quelle incidence ?

La distance intronique de plus de 42 kbp qui sépare l'exon E1 des exons EX, EY et EZ est un trait commun des gènes de sialyltransférases. Les exemples de *mst8sia2*, *ST6GAL1*, *rst6gal1* ou *ST3GAL3* soulignent à quel point les différents exons qui composent les gènes de sialyltransférases sont répartis sur de longues distances chez les mammifères (Yoshida *et al.*, 1996 ; Dall'Olio *et al.*, 2004 ; Svensson *et al.*, 1990 ; Taniguchi *et al.*, 2003). Pour *ST6GAL1*, l'exon EX est par exemple séparé de plus de 57 kbp de l'exon E1 (Figure 13, Introduction bibliographique) et plus de 28 kbp séparent l'exon E1 de l'exon E2 pour le gène *ST3GAL3* (Taniguchi *et al.*, 2003). La promiscuité des exons EX, EY et EZ est, quant à elle, bien plus surprenante. Un modèle de promoteurs alternatifs a été proposé par Kimura et collaborateurs en se basant sur l'étude de plus de 1,7 millions d'extrémités 5' d'ADNc humains non dégradés. Les résultats de cette étude postulent que des TSS séparés par un minimum de 500 pb sont théoriquement suffisant pour constituer des régions promotrices alternatives, exclusives et indépendantes. Dans ces travaux, le cerveau est même considéré comme le deuxième tissu, après les testicules, pour lequel le nombre le plus élevé de promoteurs alternatifs tissu-spécifiques a été recensé (Kimura *et al.*, 2006). Ici, le groupe de TSS de l'exon EX est séparé des deux TSS de l'exon EY par 427 pb qui sont eux-mêmes distants de 608 pb de l'extrémité 5' de EZ.

L'organisation génomique des exons en 5' de *ST6GAL2* est donc à la limite théorique du modèle proposé par Kimura et collaborateurs. Le gène bovin codant ST6Gal II, *bst6gal2*, est le seul à avoir été également étudié. Comme présenté dans la figure 20 (Introduction bibliographique), un unique exon de 132 pb, E0, a été identifié (Laporte *et al.*, 2009). Les trois transcrits alternatifs en 5' décrits par les auteurs correspondent à : (1), la séquence entière de l'exon E0 + une forme tronquée en 5' de l'exon E1, (2) une forme tronquée en 5' de l'exon E0 + l'exon E1 en entier, et (3), un transcrit initié directement par la séquence entière de l'exon E1. Ces résultats sont assez éloignés de ce que nous avons obtenu. Cette organisation des exons alternatifs en 5' du gène *ST6GAL2* est-elle une évolution particulière propre au gène humain codant ST6Gal II ? L'identification des extrémités 5' non traduites du

gène codant ST6Gal II chez d'autres espèces, comme la souris, et la comparaison des séquences pourrait nous apporter un éclairage nouveau sur les données présentées dans ce manuscrit et sur une éventuelle conservation de l'organisation génomique au cours de l'évolution.

Cette organisation a-t-elle un rôle biologique particulier ?

Contrairement aux autres études portées sur la régulation transcriptionnelle des gènes de sialyltransférases, les trois exons alternatifs en 5' que nous avons identifiés ici, ne sont pas dispersés outre mesure sur le génome. Les exons initiant les transcrits alternatifs tissu-spécifiques du gène *ST6GAL1* par exemple, sont séparés par plusieurs dizaines de kpb (Figure 13, Introduction bibliographique). Pour *ST6GAL1*, des promoteurs tissu-spécifiques clairement individualisés, conduisent à la transcription d'exons « génomiquement » isolés en 5' puis d'une partie commune d'exons codants (Figure 13, Introduction bibliographique). Nos mesures effectuées en Q-PCR sur l'expression relative des trois transcrits initiés par EX, EY, ou EZ dans différentes lignées cellulaires en culture ou *in vivo* (lobes du cortex cérébral, cervelet et hippocampe), ainsi que nos expériences concernant l'identification et la caractérisation des régions promotrices ne sont pas en faveur d'une régulation de la transcription complètement indépendante de ces trois exons. Malgré certaines différences entre le profil d'expression relatif de ces trois types de transcrits dans les cellules en culture et *in vivo*, les transcrits contenant les exons EX, EY et EZ sont toujours tous les trois exprimés. Dans le cadre de la différenciation neuronale de cellules NT2/D1, nous avons également observé une surexpression similaire des trois transcrits. Si nous ne pouvons cependant pas exclure que l'un des transcrits puisse être spécifiquement exprimé au détriment des deux autres dans un autre tissu que le tissu cérébral, nos résultats plaident toutefois en faveur d'une régulation commune, ou du moins en partie commune, de l'expression des exons EX, EY et EZ. Avec les données dont nous disposons, nous ne pouvons réellement déterminer si l'existence et la transcription de l'un des transcrits initiés par les exons EX, EY ou EZ joue un rôle particulier ou si leurs expressions sont étroitement liées. La question reste posée : quelle intérêt a la cellule d'exprimer trois transcrits différents uniquement dans leur séquence non traduite en 5' et codant, à priori, pour une protéine identique ? Cela contraste néanmoins avec le modèle de régulation transcriptionnelle établie pour les gènes de sialyltransférases : un promoteur tissulaire alternatif, une extrémité 5' non traduite alternative.

Une explication pourrait venir de l'unique différence entre ces trois transcrits alternatifs : la séquence ribonucléotidique en 5' de ces trois ARNm. La différence de séquence pourrait influencer sur la stabilité et la vitesse de dégradation de ces ARNm. En fonction du contexte tissulaire, la cellule pourrait tirer partie d'une dégradation plus ou moins rapide des ARNm de *ST6GAL2*. Toutefois, les rares travaux menés sur la stabilité de l'ARNm de gènes de sialyltransférases se sont uniquement intéressés à l'extrémité 3' des transcrits, de *rst6gal1* et *ST6GAL1* notamment (Rao *et al.*, 1999 ; Garige *et al.*, 2005 ; Garige *et al.*, 2006 ; Gong *et al.*, 2007). De plus, l'extrémité 3' du gène *ST6GAL2* étant très longue (>5 kpb), il est plus probable qu'elle soit le site d'une éventuelle régulation post-transcriptionnelle.

2. REGULATION TRANSCRIPTIONNELLE TISSU-SPECIFIQUE DU GENE *ST6GAL2*

Transcrits spécifiques du cerveau ou régulation spécifique dans le cerveau ?

Actuellement, aucune autre extrémité 5' alternative que celles identifiées par 5' RACE dans la lignée SH-SY5Y n'a été décrite. Toutes les séquences d'EST contenues dans les banques de données correspondent, dans leur extrémité 5', à la séquence de l'un des exons EX ou EY. Pour autant, les travaux que nous avons réalisés ne nous permettent pas d'affirmer que les transcrits contenant les exons EX, EY et EZ soient spécifiques du tissu cérébral. La mesure de l'expression de l'un ou des trois transcrits dans d'autres types cellulaires ou d'autres tissus pourrait confirmer ou non si l'expression de ces exons est spécifique du système nerveux central. Il ne fait cependant aucun doute que l'expression des transcrits contenant les exons EX, EY et EZ soit préférentiellement régulée dans des lignées cellulaires neuronales et dans les tissus cérébraux. Cependant, des études préliminaires que nous avons réalisées ont révélé que le gène *ST6GAL2* et les transcrits initiés par les exons EX, EY et EZ sont exprimés dans la lignée de cancer du poumon A549. Les niveaux d'expression sont équivalents à ceux observés pour la lignée SH-SY5Y. *ST6GAL2* avait été rapporté comme faiblement exprimé dans des explants bronchiques mais cette expression était très nettement augmentée sous l'influence des cytokines pro-inflammatoires IL-6 et IL-8 (Groux-Degroote *et al.*, 2008). Par ailleurs, les travaux de Laporte et collaborateurs rapportent une expression importante de *bst6gal2* dans le poumon, expression qui est augmentée sous l'influence de l'IL-6 (Laporte *et al.*, 2009). Ces données suggèrent que *ST6GAL2* puisse être également

exprimé au niveau du poumon et surexprimée en présence de cytokines pro-inflammatoires. Une étude de l'expression des éventuels différents transcrits exprimés et surexprimés au niveau des explants bronchiques pourrait confirmer ou non l'expression des exons EX, EY ou EZ.

Cependant, les résultats obtenus par les tests d'activité luciférase et de mutagenèse dirigée dans la lignée SH-SY5Y suggèrent que des facteurs de transcription « cérébraux » puissent influencer la transcription des exons EX et EY principalement. Considérant nos résultats, il est possible d'imaginer un modèle de régulation dans lequel la région génomique englobant les trois exons EX, EY et EZ soit une zone regroupant des éléments activateurs et répresseurs de la transcription. La séquence -121 / +772 semble regrouper ces propriétés (Figures 40 et 41, Résultats). A l'intérieur de cette séquence, des facteurs de transcription spécifiques de fonctions cérébrales pourraient influencer positivement ou négativement sur l'expression de *ST6GAL2*. Il est également concevable que d'autres facteurs de transcription, spécifiques d'un autre type tissulaire, que nous n'avons pu mettre en évidence à travers notre modèle puissent jouer un rôle dans un autre contexte cellulaire. L'utilisation de nos différents vecteurs rapporteurs dans d'autres environnements cellulaires pourrait nous aider à définir davantage les séquences promotrices activatrices et inhibitrices de la transcription des exons EX, EY et EZ, ainsi que les facteurs de transcription impliqués. A l'exception des trois échantillons provenant de l'hippocampe, les résultats compilés des mesures effectuées par Q-PCR et de l'étude des régions promotrices réalisée dans les SH-SY5Y suggèrent fortement une régulation transcriptionnelle similaire entre les lignées cellulaires en culture et les lobes du cortex cérébral et le cervelet. Le niveau d'expression relatif moyen observé dans les différents échantillons cérébraux demeure cependant supérieur à celui des lignées cellulaires en culture. Par ailleurs, les trois échantillons provenant de l'hippocampe montrent un profil d'expression des exons EX, EY et EZ nettement différent de celui des cellules en culture et des autres échantillons du tissu cérébral. Dans ces échantillons, l'exon EX est visiblement moins exprimé alors que l'exon EY est nettement le plus exprimé (Figure 33A, Résultats). Il est raisonnable de postuler que dans ce tissu, la balance des éléments activateurs et répresseurs soit différente de celle présentée pour les cellules SH-SY5Y, et par extension pour le cervelet et le cortex cérébral. Il n'est pas non plus exclu que d'autres facteurs de transcription soient actifs et favorisent l'expression de l'exon EY et EZ au détriment de l'exon EX. A nouveau se pose la question de la pertinence d'une telle régulation pour la cellule où les trois types de transcrits sont néanmoins toujours exprimés.

***ST6GAL1* et *ST6GAL2* exprimés au niveau du cerveau**

L'essentiel des travaux menés jusqu'à présent sur l'expression du gène codant ST6Gal I chez l'Homme, mais aussi chez le rat, la souris, et le bovin, présente ce gène comme pas ou peu exprimé au niveau du cerveau (Takashima *et al.*, 2002 & 2003 ; Kitagawa *et al.*, 1994 ; Ishii *et al.*, 2007 ; Kaneko *et al.*, 1995 ; Paulson *et al.*, 1989, Mercier *et al.*, 1999). Par opposition, *ST6GAL2* et ses orthologues murin et bovin sont présentés comme presque exclusivement exprimés dans le système nerveux central embryonnaire et adulte (Kzewinski-Recchi *et al.*, 2003 ; Takashima *et al.*, 2003 ; Ishii *et al.*, 2007 ; Laporte *et al.*, 2009). Des données sont également disponibles sur le site Internet de l'Allen Institute for Brain Science (<http://www.brain-map.org>). Ce service propose des images permettant de visualiser l'expression d'un très grand nombre de gènes par hybridation *in-situ* sur des coupes sagittales de cerveau de souris. L'expression des gènes *mst6gal1* et *mst6gal2* est visualisée dans la figure 21 (Introduction bibliographique). Sur ces images, il apparaît indiscutablement que le gène *mst6gal2* est plus exprimé que le gène *mst6gal1* au niveau du cerveau. Contrastant avec ces images, une étude réalisée par Nairn et collaborateurs sur l'expression de l'ensemble des gènes de glycosylation chez la souris rapporte cependant que *mst6gal2* est exclusivement exprimé dans le cerveau, mais que dans le même temps *mst6gal1* est exprimé à un niveau équivalent dans le tissu cérébral (Nairn *et al.*, 2008). Enfin, les travaux présentés dans cette thèse montrent que, de façon inattendue, tous les échantillons de tissu cérébral que nous avons examinés expriment les gènes *ST6GAL1* et *ST6GAL2* à des niveaux équivalents (Figure 32, Résultats). De plus, même si le niveau d'expression est relativement hétérogène entre les différents échantillons, l'expression des deux gènes est toujours de niveau égal pour un même échantillon.

Nos analyses par Q-PCR montrent que le gène *ST6GAL2* est exprimé dans les cellules de type neuronal (SH-SY5Y et Kelly), et pratiquement pas dans les cellules de type glial (U118 et T98G) (Figure 24, Résultats). Par ailleurs, le gène *ST6GAL1* est faiblement exprimé dans la lignée de neuroblastome SH-SY5Y et plus fortement exprimé dans une autre lignée de neuroblastome, les cellules Kelly. Il n'est également que faiblement exprimé dans les cellules gliales U118 et T98G (Figure 24, Résultats). Le gène *ST6GAL2* pourrait très bien n'être exprimé que dans les cellules de type neuronal à l'intérieur du système nerveux central. Dans cette configuration, il serait intéressant de pouvoir isoler les différents types cellulaires du cortex cérébral pour y mesurer l'expression du gène *ST6GAL2*.

En parallèle, nous avons visualisé l'expression de la protéine hST6Gal II dans les différents lobes du cortex cérébral ainsi que, plus faiblement, dans le cervelet et l'hippocampe. Pouvoir observer l'expression de la protéine hST6Gal II directement sur des coupes histologiques de cerveau serait une bonne opportunité de confirmer une expression tissu- et/ou cellule-spécifique de cette sialyltransférase.

3. *ST6GAL2*, UN GENE SUREXPRIME LORS DU DEVELOPPEMENT NEURONAL

La lignée cellulaire NT2/D1 est un modèle *in vitro* qui reproduit les événements précoces impliqués dans la différenciation neuronale et la neurogenèse (Couillard-Despres *et al.*, 2008). Nos travaux ont montré sans ambiguïté que le gène *ST6GAL2* et la protéine hST6Gal II sont fortement surexprimés après la différenciation neuronale des cellules NT2/D1. Dans le même temps l'expression du gène *ST6GAL1* n'augmente que très faiblement (Figure 45, Résultats), suggérant que seule hST6Gal II jouerait un rôle dans le développement neuronal. L'augmentation des transcrits initiés par EX, EY et EZ suggère aussi que les régions génomiques que nous avons décrites contiennent les éléments de réponse permettant cette augmentation de la transcription. Les éléments de réponse à l'acide rétinoïque sont les sites RAR (Retinoic acid receptor), RXR (Retinoïd X receptor) (Rastinejad, 2001) ou ROR (Retinoïd-related orphan receptor). Cependant, un seul site de fixation RXR a été prédit par les analyses informatiques dans la séquence -1851 / +1452, situé à l'intérieur de la séquence de l'exon EY. Plusieurs autres sites potentiels de fixation pour des éléments de réponse à l'acide rétinoïque sont répartis dans une région génomique allant de 2 à 8 kpb en amont de l'exon EX. Par ailleurs, plusieurs études ont montré que le facteur de transcription SP1 jouerait un rôle dans l'activation de la transcription induit par l'acide rétinoïque (Husmann *et al.*, 2000 ; Safe *et al.*, 2004 ; Chen *et al.*, 2007). De nombreux sites de fixation pour le facteur SP1 ont été prédits à l'intérieur des régions génomiques entourant les exons EX, EY et EZ (Figure 38, Résultats). Nous avons montré par mutagenèse qu'au moins un de ces sites était potentiellement activateur (Figure 44, Résultats). Des études ultérieures sur la régulation transcriptionnelle du gène *ST6GAL2* dans la population neuronale permettraient de conclure quant à une éventuelle occupation des autres sites SP1 au cours de la différenciation neuronale des NT2/D1 sous l'action de l'acide rétinoïque.

Ce résultat préliminaire des plus intéressants laisse supposer d'un rôle prépondérant de la sialyltransférase hST6Gal II au cours du développement du système nerveux central.

Certains résultats de Ishii et collaborateurs chez la souris soulignent l'augmentation de l'expression de *mst6gal2* au cours du développement cérébral (Ishii *et al.*, 2007). Takashima et collaborateurs rapportent également une forte expression de *mst6gal2* dans les tissus embryonnaires (sans plus de précision sur un tissu en particulier) (Takashima *et al.*, 2003). Enfin, l'unique sialyltransférase de drosophile, DSiaT, aux propriétés similaires à celle de hST6Gal II, voit son expression transcriptionnelle visualisée uniquement dans les crêtes neurales au cours du développement larvaire. Le gène *ST6GAL2* est vraisemblablement exprimé, voire surexprimé, pendant l'embryogenèse du système nerveux central. Pour autant, les protéines cibles de cette sialyltransférase, ou plus généralement les *N*-glycanes terminés par du Sia₆Lac(di)NAc, sont des structures peu représentées dans le cerveau. Les gènes *DSiaT* (Koles *et al.*, 2004), *mst6gal2* (Takashima *et al.*, 2003 ; Ishii *et al.*, 2007) et *ST6GAL2* (Krzewinski-Recchi *et al.*, 2002) présentent une expression au niveau du cerveau embryonnaire. De plus, l'expression du gène *mst6gal2* est augmentée au cours de développement embryonnaire du cerveau et celle du gène *DSiaT* est retrouvée uniquement dans les crêtes neurales en formation (Koles *et al.*, 2004). Enfin, nos données montrent également une augmentation importante de l'expression du gène *ST6GAL2* et de la protéine hST6Gal II au cours de la différenciation neuronale des cellules NT2/D1. Tenant compte de ces données, il est donc raisonnable de penser que l' α 2,6-sialylation potentiellement engendrée par les différentes orthologues de ST6Gal II et DSiaT puisse n'être que transitoire dans le cadre du développement cérébral. Il est possible d'imaginer que des *N*-glycosylprotéines porteuses de structures terminales Sia₆Lac(di)NAc soient uniquement exprimées, ou sialylées de cette manière, lors de la neurogenèse pour remplir une fonction particulière. Dans cette hypothèse, ces motifs Sia₆Lac(di)NAc ne seraient pas ou peu représentés dans le cerveau adulte et donc difficilement identifiables. Dans le cerveau adulte plusieurs raisons pourraient justifier de l'absence des structures Sia₆Lac(di)NAc : les *N*-glycosylprotéines portant ces structures ne seraient plus exprimées, les glycosyltransférases (dont ST6Gal II) impliquées dans la biosynthèse de ces structures ne seraient plus exprimées, des sialidases viendraient supprimer cette sialylation. Un exemple typique de dynamisme de structures sialylées est représenté par la NCAM, une *N*-glycosylprotéine polysialylée et largement présente dans le cerveau adulte et embryonnaire. La NCAM est hautement polysialylée au début du développement du système nerveux central alors que le degré de polymérisation des acides sialiques de la NCAM diminue fortement dans le cerveau adulte (Rothbard *et al.*, 1982 ; Brusés *et al.*, 2001 ; Rutishauser, 2008). Chez la souris, mST8Sia II, l'une des deux polysialyltransférases impliquées dans la polysialylation de la NCAM est

surexprimée dans le cerveau durant le développement embryonnaire pour ne plus être exprimée dès la deuxième semaine post-natale (Yoshida *et al.*, 1995). Pour autant, aucune étude n'a pour le moment rapporté de *N*-glycosylprotéines porteuses de structures terminales LacNAc et/ou LacdiNAc dans le cerveau embryonnaire ou adulte de mammifères.

PERSPECTIVES DU TRAVAIL DE RECHERCHE

Dans la perspective de la poursuite de ce travail, plusieurs orientations peuvent être envisagées :

1. Poursuivre la caractérisation des régions promotrices du gène *ST6GAL2* et confirmer le rôle joué par certains facteurs de transcription

Caractérisation des facteurs de transcription potentiels

Nos résultats suggèrent que les facteurs de transcription NF- κ B, NRSF, SP1, Puro α , Sox5 et Olf1 jouent un rôle dans la régulation transcriptionnelle du gène *ST6GAL2* en se fixant à l'intérieur de la région -121 / +772. Des études complémentaires par chromatine immuno-précipitation (DNA ChIP) dans les cellules SH-SY5Y, ainsi que des expériences de retard sur gel pourraient confirmer la spécificité de fixation suggérée par bio-informatique et mutagenèse dirigée, pour ces différents facteurs. Le rôle activateur ou répresseur de ces facteurs pourrait également être examiné par surexpression transitoire de l'un de ces facteurs dans les cellules SH-SY5Y ou dans un contexte cellulaire dans lequel le gène *ST6GAL2* n'est pas exprimé, comme les cellules HeLa.

Etude des régions promotrices des exons EX, EY et EZ dans un contexte cellulaire non-neuronal, les cellules de cancer de poumon A549

Nos résultats préliminaires indiquent que le gène *ST6GAL2* et les différents transcrits que nous avons identifiés sont exprimés dans la lignée pulmonaire A549. En dehors du contexte neuronal que nous avons étudié, il serait intéressant de transfecter nos différentes constructions des régions promotrices en vue de tests d'activité luciférase. Des différences entre les séquences activatrices et inhibitrices identifiées dans les cellules SH-SY5Y et les cellules A549 pourraient être mises en lumière. Ces résultats pourraient permettre d'identifier d'autres facteurs de transcription potentiellement impliqués dans l'expression pulmonaire du gène *ST6GAL2* suggérée par les résultats de Groux-Degroote et collaborateurs (Groux-Degroote *et al.*, 2008) et Laporte et collaborateurs pour son homologue bovin (Laporte *et al.*, 2009).

2. Poursuite de l'étude du rôle de *ST6GAL2* au cours de la différenciation neuronale, avec le modèle NT2/D1

Visualisation et caractérisation des *N*-glycanes avant et après différenciation des cellules NT2/D1

Les cellules NT2/D1 et leur différenciation en neurones représentent pour la première fois l'opportunité de visualiser le rôle ou l'un des rôles de hST6Gal II dans un contexte « semi-physiologique ». La lignée NT2/D1 est en effet considérée comme l'un des meilleurs modèles *in vitro* mimant la différenciation neuronale au cours de la neurogenèse embryonnaire. Nos résultats montrent indiscutablement que le gène *ST6GAL2* et la protéine hST6Gal II sont fortement surexprimés au cours de la différenciation neuronale des cellules NT2/D1 par l'acide rétinoïque. Afin de caractériser davantage le rôle de cette sialyltransférase dans ce processus il serait nécessaire, dans un premier temps, d'identifier les *N*-glycosylprotéines arborant du Sia₆Lac(di)NAc avant et après différenciation. Des analyses en Lectine-Blot avec la lectine SNA (qui reconnaît les structures Sia₆Lac(di)NAc) suivies d'analyses protéomiques nous permettront d'identifier des protéines substrat potentiel de hST6Gal II. Ce projet a déjà été initié en collaboration avec le Docteur Willy Morelle de notre unité.

Définition du rôle de hST6Gal II dans la différenciation neuronale des cellules NT2/D1

En parallèle, le rôle et l'impact de la surexpression du gène *ST6GAL2* peuvent être étudiés par plusieurs approches. Les cellules NT2/D1 pourront être transfectées par un shRNA afin d'invalider stablement l'expression du gène *ST6GAL2* et d'analyser les conséquences de cette invalidation au cours de la différenciation. Une stratégie similaire pourrait être employée avec le gène *ST6GAL1* pour confirmer un rôle prépondérant de hST6Gal II par rapport à hST6Gal I dans ce contexte. Une approche alternative pourrait consister à surexprimer la sialyltransférase hST6Gal II et/ou hST6Gal I dans les cellules NT2/D1 pour y en étudier d'éventuels changements de comportements de la cellule. Nous essaierons de répondre à la question, hST6Gal II induit-elle un phénotype plus « neuronal » de la cellule ?

CONCLUSION GENERALE

Nos travaux ont permis d'identifier pour le gène humain *ST6GAL2* trois types de transcrits alternatifs au niveau de leur extrémité 5' non traduite dans la lignée cellulaire de neuroblastome SH-SY5Y. Ces transcrits comportent chacun un exon 5' alternatif que nous avons appelé EX, EY et EZ. Nos résultats montrent que ces exons sont aussi exprimés dans différentes régions du cortex cérébral, dans le cervelet et dans l'hippocampe. Nous montrons aussi pour la première fois que la protéine hST6Gal II est exprimée dans le tissu cérébral. Nos travaux sur les régions promotrices du gène *ST6GAL2* rapportent la présence de deux régions promotrices, l'une située en amont de l'exon EX et l'autre en amont de l'exon EY. Plusieurs facteurs de transcription, généraux ou associés aux fonctions cérébrales, semblent jouer un rôle dans la régulation de l'expression du gène *ST6GAL2*. Les facteurs NF- κ B et NRSF réprimerait la transcription alors que les facteurs SP1, Sox5, Pur α et Olf1 l'activeraient

Pour terminer, nos travaux suggèrent fortement que la protéine hST6Gal II soit impliquée dans le développement neuronal. Nous avons en effet fait la preuve de la surexpression du gène *ST6GAL2* et de la protéine hST6Gal II dans un modèle *in vitro* de différenciation neuronale.

Les travaux réalisés au cours de cette thèse permettent une meilleure compréhension de l'expression et des mécanismes transcriptionnels de régulation du gène *ST6GAL2* dans un contexte cérébral.

Matériels et Méthodes

CULTURE CELLULAIRE

1. CULTURE DES CELLULES

Les lignées cellulaires NT2/D1, Kelly, T98G et U118 nous ont été fournies par les Drs Luc Buée et Marie-Laure Caillet-Boudin (INSERM U837, CHR, Lille). Les lignées cellulaires SH-SY5Y, NT2/D1, T98G, U118, HeLa, HEK-293, THP-1, Caco2, MDA-MB-231, BT-20, et SK-Mel 28 ont été cultivées avec du milieu DMEM (Dulbecco's modified Eagle's medium) contenant 4500 mg/l de glucose (BioWhittaker, Lonza, Belgium). Les cellules Kelly ont été cultivées avec du milieu RPMI 1640 (BioWhittaker). Les milieux ont été supplémentés avec 10 % du sérum de veau fœtal inactivé à la chaleur (Eurobio, Courtaboeuf, France), 2 mM de L-glutamine, 100 U/ml de pénicilline et 10 µg/ml de streptomycine. Les cellules sont incubées à 37°C sous une atmosphère de 5 % de CO₂.

En routine, le milieu de culture est changé au moins une fois par semaine et les cellules sont divisées à confluence au moins au 1/10^{ème}.

2. RECUPERATION DE CULOTS CELLULAIRES EN VUE D'UNE EXTRACTION D'ARN

A la confluence souhaitée, le milieu de culture est aspiré, 5 ml de PBS sont ajouté puis à nouveau aspirés dans le but ne laisser aucune trace de milieu de culture. Les cellules sont décollées par l'action de la trypsine (environ 2 ml pour une flasque de 75 cm²) 5 à 15 min, ou jusqu'à décollement des cellules, à 37°C. Les cellules ont été resuspendues avec 5 ml de milieu ou de PBS et réparties en fonction de la quantité souhaitée dans un ou plusieurs tubes Falcon de 15 ml ou tube Eppendorf de 1,5 ml. Le culot est obtenu par centrifugation à 1500 rpm pendant 5 min. Le surnageant est alors aspiré et le culot cellulaire stocké à -20°C ou -80°C pour un stockage prolongé.

3. RECUPERATION DE CULOTS CELLULAIRES EN VUE D'UNE EXTRACTION DE PROTEINES

A la confluence souhaitée, le milieu de culture est aspiré, 5 ml de PBS sont ajoutés puis à nouveau aspirés dans le but ne laisser aucune trace de milieu de culture. 5 ml de PBS contenant de l'EDTA 0,004 M sont ajoutés et laissés en contact avec les cellules pendant 20 à 30 min, jusqu'à décollement des cellules, à 37°C. Les cellules sont remises en suspension et réparties en fonction de la quantité souhaitée dans un ou plusieurs tubes Falcon de 15 ml ou tubes Eppendorf de 1,5 ml. Les cellules sont alors centrifugées à 1500 rpm pendant 5 min. Le surnageant est alors aspiré et le culot ressuspendu dans du PBS pour être lavé et débarrassé des traces d'EDTA. Les cellules sont à nouveau centrifugées à 1500 rpm pendant 5 min pour les culoter. L'opération de lavage est répétée deux fois et le culot cellulaire obtenu stocké à -20°C ou -80°C pour un stockage prolongé.

4. DIFFERENCIATION PAR L'ACIDE RETINOÏQUE DES NT2/D1

1.10⁶ cellules NT2/D1 ont étéensemencées dans une flasque de 75 cm² avec 10 ml de milieu complet contenant 10 µM d'acide rétinoïque (Sigma-Aldrich, Saint-Quentin Fallavier, France) et 10 µM d'amphotéricine B (Fungizone) (Sigma-Aldrich). La flasque a été emballée dans du papier aluminium pour protéger au maximum l'acide rétinoïque de la dégradation par la lumière. Le milieu a été changé 3 fois par semaine, les cellules restant dans la même flasque. Après cinq semaines de traitement à l'acide rétinoïque, les cellules forment un tapis très dense et multicouches de NT2/D1 et ont été divisées au ¼ puisensemencées dans 4 flasques de 75 cm² avec du milieu complet pendant trois jours. Les cellules on ensuite été traitées deux à trois semaines par ajout d'inhibiteurs mitotiques dans 10 ml de milieu de culture milieu complet (10 µM 5-fluoro-2'-deoxyuridine, 10 µM uridine, 1 µM cytosine-β-D-arabinofuranoside) (Sigma-Aldrich). A l'issue de ce traitement, la différenciation neuronale est morphologiquement achevée. Les cellules formes deux couches, une couche inférieure de NT2/D1 non différenciées, appelées NN, et une couche supérieure de NT2/D1 différenciées en neurones, appelée hNT. Les hNT s'organisent en neurosphères desquelles s'étendent des prolongements dendritiques. Ces deux types cellulaires ont été séparés mécaniquement par ajout d'1 ml de trypsine, qu'on laisse agir 1 min, avant de donner un coup sec et franc sur la flasque. Les hNT se décollent alors en un tapis surnageant. Les hNT ont été récupérées en

ajoutant 2 à 3 ml de PBS pour faciliter cette opération. Un culot cellulaire est ensuite récupéré par centrifugation 5 min à 1500 rpm. Les NN peuvent également être récupérées par action prolongée de la trypsine puis centrifugation 5 min à 1500 rpm pour obtenir un culot cellulaire.

5. DIFFERENCIATION PAR L'ACIDE RETINOÏQUE DES CELLULES SH-SY5Y

1.10⁶ cellules SH-SY5Y ont étéensemencées dans une flasque de 75 cm² avec 10 ml de milieu complet contenant 10 µM d'acide rétinoïque *all-trans* (Sigma-Aldrich). Les cellules ont été laissées 7 jours dans ce milieu de différenciation et le milieu changé tous deux jours. A l'issue de ce traitement, les cellules ont été récoltées pour réaliser l'extraction d'ARN.

6. TRANSFECTIONS TRANSITOIRES ET TESTS D'ACTIVITE LUCIFERASE

Les transfections préalables aux tests d'activité luciférase ont été réalisés avec le Cell Line Nucleofector[®] Kit V (AMAXA, Europe), en accord avec les recommandations du fabricant. Brièvement, deux millions de cellules ont été récoltées à ~70 % de confluence, centrifugées 5 min à 1000 tr/min et ressuspendues dans 100 µl de solution Nucleofector[®] Kit V. 2 µg de plasmide d'intérêt (pGL3-Basic ou constructions plasmidiques) et 40 ng du vecteur normalisateur codant la luciférase *Renilla* sont ajoutés à la suspension cellulaire. La transfection a ensuite été effectuée dans une cuvette AMAXA avec le programme G-004, selon les recommandations du kit pour les cellules SH-SY5Y. Les cellules ont alors étéensemencées en plaques six puits avec 2 ml de milieu complet et incubées à 37°C. 24 h après, le milieu est retiré, les cellules rincées avec du PBS puis lysées avec 250 µl de Passive Lysis Buffer fourni par le kit Dual Luciferase Reporter Assay System (Promega, Madison, USA) à température ambiante pendant 15 min. Les 250 µl contenant les lysats cellulaires sont repris et centrifugés à 2000 rpm ; les protéines totales, dont la luciférase, sont alors en suspension. 20 µl de cette suspension ont été utilisés pour les tests d'activité luciférase. Les activités luciférase *Firefly* et *Renilla* ont été quantifiées avec le kit Dual Luciferase Reporter Assay System (Promega) par mesure de l'absorbance à 560 nm. La luminescence a été mesurée 10 sec par point avec le Centro Luminomètre de Berthold Technologie.

7. TRANSFECTION TRANSITOIRE PAR LES VECTEURS D'EXPRESSION CODANTS HST6GAL I-3xFLAG ET HST6GAL II-3xFLAG

Les transfections ont été réalisées par lipofection avec le kit Lipofectamine Reagent (Invitrogen, Carlsbad, USA), suivant les recommandations du fabricant, dans les cellules HeLa à ~70 % de confluence. 3.10^5 cellules ont étéensemencées par puits de plaque six puits avec 2 ml de milieu complet la veille de la transfection. Pour préparer la transfection, le milieu de culture a été aspiré, les cellules ont été rincées au PBS et 800 μ l d'OptiMEM ont été ajoutés (BioWhittaker). Dans un premier tube Eppendorf sont ajoutés 1 μ g de vecteur, 5 μ l de Reagent Plus (fourni par le kit) et qsp 100 μ l d'OptiMEM. Dans un deuxième tube Eppendorf, sont ajoutés 5 μ l de Lipofectamine et 95 μ l d'OptiMEM. Chacun des tubes a été laissé 15 min à température ambiante. Le contenu des deux tubes a ensuite été mélangé et laissé 15 min à température ambiante. Les 200 μ l de solution vecteur / lipofectamine ont alors été ajoutés goutte à goutte dans le puit préparé. La plaque est ensuite placée dans l'incubateur 5 h à 37°C. Après ce temps, le milieu de transfection a été remplacé par 2 ml de milieu complet. 24 h après, la production des protéines recombinantes a été induite par l'ajout de butyrate de sodium (concentration finale 10 mM). Les cellules sont ensuite placées 16 h dans l'incubateur avant d'être récupérées pour l'extraction des protéines.

8. PRELEVEMENT DES ECHANTILLONS DE TISSU CEREBRAL HUMAIN

Les échantillons de tissu cérébral provenant du cortex cérébral, du cervelet et de l'hippocampe ont été obtenus lors de l'autopsie de 8 sujets ne présentant pas de pathologie cérébrale. Ces échantillons ont été obtenus auprès de la Neurobanque de Lille en accord avec les protocoles d'éthiques appliqués en France. Le délai post-mortem avant prélèvement n'excède pas 36 h pour tous les échantillons.

BIOLOGIE MOLECULAIRE

1. EXTRACTIONS D'ADN

1.1. Extraction d'ADN génomique

L'extraction d'ADN génomique à partir de culots cellulaires ($\sim 5 \cdot 10^6$ cellules) de SH-SY5Y a été réalisée avec le kit NucleoSpin® Tissue (Macherey-Nagel, Hoerd, France) selon les recommandations du fournisseur. Brièvement, un culot cellulaire de $5 \cdot 10^6$ cellules a été ressuspendu dans 200 μ l de tampon de lyse (fourni par le kit) auquel a été ajouté de la protéinase K ; le tout a été incubé à 70°C pendant 15 min. 210 μ l d'éthanol (96-100 %) ont été ajoutés et l'échantillon a été vortexé vigoureusement. L'ADN a alors été fixé par chromatographie d'absorption sur une membrane de silice, la colonne a été lavée avec deux tampons (fournis par le kit). L'ADN a finalement été élué par 100 μ l d'un tampon d'éluion (fourni par le kit).

1.2. Extraction d'ADN plasmidique

Les extractions d'ADN plasmidique ont été réalisées à partir de 5 ml de culture bactérienne avec le kit NucleoSpin® Plasmid Mini Prep (Macherey-Nagel) ou à partir de 50 ml de culture bactérienne avec le kit NucleoSpin® Plasmid Midi Prep (Macherey-Nagel) pour un volume et un rendement plus important. Brièvement, pour les « mini-preps », à partir de 5 ml de culture bactérienne, les bactéries ont été culottées puis resuspendues dans un tampon de lyse (fourni par le kit). L'ADN plasmidique a été dénaturé / renaturé de manière différentielle par rapport à l'ADN génomique par un tampon de lyse alcalin contenant du SDS fourni par le kit ; les débris cellulaires les protéines et l'ADN génomique ont été précipités par centrifugation. L'ADN plasmidique a ensuite été fixé par chromatographie d'absorption sur une membrane de silice. Il a été élué par 50 μ l d'eau. Pour les « midis preps », 50 ml de culture bactérienne ont été centrifugés puis les bactéries resuspendues dans un tampon de lyse (fourni par le kit). L'ADN plasmidique a été libéré par un tampon de lyse alcalin contenant du SDS (fourni par le kit), les débris cellulaires et les protéines ont été éliminés par filtration et l'ADN plasmidique a été récupéré sur une colonne échangeuse d'anions. L'ADN plasmidique a ensuite été élué par 250 μ l d'eau.

La quantité et la pureté de l'ADN extrait ont été obtenues par mesure de l'absorbance à 260 nm et du rapport de l'absorbance à 260 nm / 280 nm réalisées avec le NanoDrop® (Labtech, Paris, France).

2. EXTRACTIONS DES ARN

2.1. ARN de cellules en cultures

Les ARN totaux des cellules SH-SY5Y, Kelly, U118, T98G, HEK 293, HeLa, THP1, Caco-2, MDA-MB-231, BT-20, SK Mel-28, NT2/D1, hNT et NN ont été extraits à partir de culots cellulaires (2 à $5 \cdot 10^6$ cellules) avec le kit NucleoSpin® RNA II (Macherey-Nagel), selon les recommandations du fabricant. Brièvement, 3,5 μ l de β -mercaptoéthanol et 350 μ l d'un tampon de lyse (fourni par le kit) ont été ajoutés au culot cellulaire, le tout a été vigoureusement vortexé. Le lysat visqueux obtenu a été filtré sur une colonne par centrifugation (10000 rpm, 1 min) et l'éluât récupéré. 350 μ l d'éthanol 70 % ont ensuite été ajoutés à l'éluât, le tout a été déposé sur une colonne de silice et centrifugé (10000 rpm, 1 min) pour fixer l'ARN. La colonne a ensuite été lavée avant d'y déposer de la DNase (fournie par le kit) 15 min à température ambiante. Trois lavages ont été effectués (tampons fournis par le kit), puis l'ARN a été élué avec de 50 μ l d'eau sans RNase. Les ARN ont été stockés à -20°C ou -80°C .

2.2. ARN d'échantillons tissulaires

Les ARN totaux issus des échantillons des différents lobes du cortex cérébral, de l'hippocampe et du cervelet ont été extraits à l'aide du kit RNeasy Lipid Tissue Mini Kit (Qiagen GmbH, Germany), selon les recommandations du fournisseur. Brièvement, les échantillons ont été homogénéisés mécaniquement sur glace avec un piston et un homogénéisateur en verre d'une contenance de 5 ml. L'échantillon tissulaire a été homogénéisé dans la solution de lyse QIAzol Lysis Reagent (1 ml pour 100 mg de tissu). L'homogénat ainsi préparé a ensuite été placé à température ambiante pendant 15 min. 200 μ l de chloroforme ont été ajoutés avant d'agiter vigoureusement pendant 15 sec et laisser reposer 3 min à température ambiante. Après centrifugation à 12000 rpm pendant 15 min à 4°C , la phase supérieure a été transférée dans un nouveau tube dans lequel est ajouté un volume

équivalent d'éthanol 70 % ; le mix a alors été vigoureusement vortexé avant d'être passé sur une colonne et centrifugé 15 sec à 10000 rpm à température ambiante. La colonne, sur laquelle l'ARN s'est fixé, a été lavée par deux tampons (du kit) puis l'ARN élué avec 50 µl d'eau dépourvue de RNase. Les ARN ont été stockés à -20°C ou -80°C.

La quantité et la pureté des ARN ont été obtenues par mesure de l'absorbance à 260 nm et du rapport de l'absorbance à 260 nm / 280 nm réalisées à l'aide du NanoDrop® (Labtech, Paris, France).

3. SYNTHÈSE DU PREMIER BRIN D'ADNC

La synthèse du premier brin d'ADN complémentaire a été réalisée avec le First-Strand cDNA Synthesis Kit (Amersham Biosciences, Freiburg, Germany) selon les instructions du kit à partir d'ARN (2 à 5 µg) provenant de cellules ou de tissus. La rétro transcription a été effectuée avec pour amorces un mix d'héxanucléotides (pd(N)₆) et d'oligo dT dans les proportions 3:1. Expérimentalement, les ADNc synthétisés avec ce mix d'amorce donnent de meilleurs résultats en Q-PCR : les Ct observés sont plus faibles pour l'expression des gènes quantifiés par rapport à une synthèse d'ADNc pour laquelle seuls des pd(N)₆ ou que des oligo dT ont été utilisés comme amorces.

4. 5' RACE A PARTIR D'ARN TOTAUX DE SH-SY5Y

4.1. Sélection des ARNm coiffés et rétro transcription

Les ARN totaux des cellules SH-SY5Y ont été extraits avec le kit NucleoSpin® RNA II (Macherey-Nagel). Les expériences de 5' RACE ont été réalisées avec le kit First Choice RLM-RACE Kit (Ambion, Huntingdon, Royaume-Uni) selon les instructions du kit pour la « réaction standard » à partir de 10 µg d'ARN. Les 10 µg d'ARN ont été traités par la CIP (Calf Intestine Alkaline) à 37°C pendant 1 h puis extraits par une solution de phénol:chloroforme 5:1 et précipités par de l'éthanol 70 %. Les ARN ont ensuite été traités par la TAP (Tobacco Acid Pyrophosphatase) à 37°C pendant 1 h puis liés à l'adaptateur de 5' RACE (GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUGAAA) à 37°C pendant 1 h et la transcription inverse effectuée. L'ADN complémentaire (ADNc) a été

synthétisé à partir de l'ARN traité par la CIP et la TAP, d'amorces composées de décamères aléatoires et avec la M-MLV Reverse Transcriptase à 42°C pendant 1 h.

4.2. Premières PCR et PCR nichées

Après synthèse de l'ADNc, une première PCR a été réalisée avec les amorces Outer Primer (OP, fournie par le kit et spécifique de la séquence de l'adaptateur de 5'RACE) (Tableau 5) et les amorces GSP1, GSP1.X, GSP1.Y ou GSP1.Z (Tableau 5) dans les conditions de PCR : 94°C 2 min ; 40 cycles de 94°C 30 sec, 62°C 30 sec et 72°C 3 min ; 72°C 5 min. Les PCR nichées subséquentes ont été réalisées avec les amorces Inner Primer (IP, fournie par le kit et spécifique de la séquence de l'adaptateur) et les amorces GPS2.1, GSP2.2, GSP2.X, GSP2.Y ou GSP2.Z dans les conditions de PCR suivantes : 94°C 2 min ; 40 cycles de 94°C 30 sec, 72°C 30 sec et 72°C 2 min ; 72° 5 min. Les amplifications de PCR ont été réalisées avec la Taq polymérase Taq PCR Core Kit (Qiagen) dans un mix d'un volume final de 25 µl comprenant : 5 µl de Q-Solution 5x, 2,5 µl de tampon PCR-Buffer (Tris-Cl, KCl, (NH₄)₂SO₄ et 15 mM de MgCl₂ ; pH 8,7) 10x, 0,5 µl de dNTP (10 mM chaque), 0,5 µl de chaque amorce (10 µM chaque), 0,15 µl de Taq PCR Core Kit (0,5 U), 1 µl d'ADN matrice (issu de l'ADNc de la 5' RACE pour la première PCR ou issu de la première PCR pour les PCR nichées) et qsp 25 µl d'eau. Les produits de PCR nichées ont ensuite été séparés par électrophorèse sur gel d'agarose 2 % contenant du BEt, extraits et purifiés avec le kit MinElute (Qiagen). Pour la purification, la bande de gel découpée a été dissoute dans un tampon pour ajuster les conditions de fixations sur colonne. La solution a ensuite été passée sur une chromatographie d'absorption sur une membrane de silice pour fixer l'ADN par centrifugation à 10000 rpm pendant 1min. La colonne a été lavée par un tampon (fourni par le kit) et élué avec 10 µl de tampon d'éluion (10 mM Tris-Cl pH 8,5, fourni par le kit). Après purification, les amplicons ont été sous-clonés dans le pCR2.1 du kit TOPO-TA cloning (Invitrogen) et les bactéries TOP10 (fournies par le kit) ont été transformées par choc thermique avec ces vecteurs. Les plasmides ont été extraits, purifiés (Cf : extraction de plasmides) puis séquencés par Genoscreen (Lille, France).

5. CONSTRUCTION DES VECTEURS CODANT LES FORMES ENTIÈRES DE hST6Gal I ET hST6Gal II ETIQUETÉES EN N-TERMINAL DU 3xFLAG

Les séquences nucléotidiques codant les formes entières des protéines hST6Gal I et hST6Gal II ont été respectivement amplifiées par PCR avec les amorces sens 5'-ATCTTCAAGCTTATTCACACCAAC-3' et anti-sens 5'-AGAGTGAGGATCCTGTGCTTAGCAGTGAATG-3' (les sites *Hind*III et *Bam*HI incorporées sont soulignés), et l'amorce sens 5'-GAGACCGAATTCGAAACCACACTTGAAGCAATGG-3' et anti-sens 5'-ACATTGGATCCCAAGAAACCCTTTTAAAGAGTGTGG-3' (les sites *Eco*RI et *Bam*HI sont soulignés). Les deux séquences amplifiées ont été insérées dans le vecteur d'expression p3xFLAG-CMV-10 (Figure 48) en phase en N-terminal avec l'étiquette 3xFLAG. La séquence codant hST6Gal I a été insérée aux sites *Hind*III / *Bam*HI et la séquence codant hST6Gal II a été insérée aux sites *Eco*RI / *Bam*HI, générant ainsi deux vecteurs codant hST6Gal I-3xFLAG ou hST6Gal II-3xFLAG.

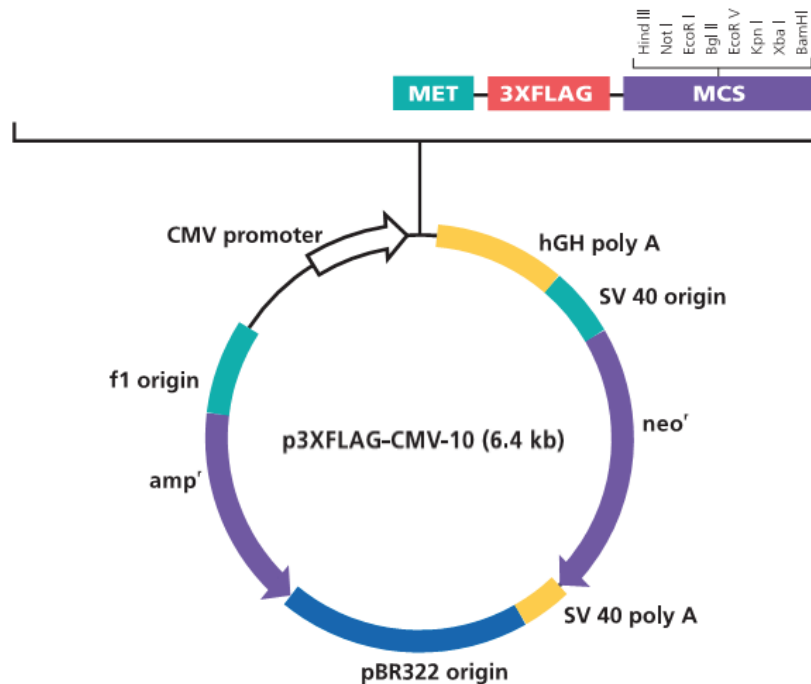


Figure 48 : Carte schématique du vecteur p3xFLAG-CMV-10

D'après Sigma-Aldrich. Les séquences codantes de hST6Gal I ou hST6Gal II ont été insérées en phase et en aval de l'étiquette 3xFLAG.

6. CONSTRUCTION DE VECTEURS RAPPORTEURS POUR LES ESSAIS LUCIFERASE

La région génomique -1851 / +253 en amont de l'exon EX a été amplifiée par PCR, à partir d'ADN génomique provenant de la lignée SH-SY5Y avec les amorces suivantes : 5'-CGCTTGCCACGGTACCCGGTCATTTTTGT-3' et 5'-AGACTTAAGCTTCGTGGGGCGGTGCGGG-3', incluant respectivement les sites de restriction *KpnI* et *HindIII* (nucléotides soulignés). Cette séquence a été insérée dans le vecteur pGL3-Basic (Promega, Figure 49) aux sites *KpnI* / *HindIII* en amont du gène rapporteur codant la luciférase *Firefly*, générant ainsi le plasmide pGL3_-1851/+263. Le plasmide pGL3_-1851/+263 a ensuite été digéré par différents couples d'enzymes de restrictions (*KpnI* / *BglII*, *KpnI* / *XhoI*, *KpnI* / *NheI* et *KpnI* / *SmaI*), en utilisant des sites de restriction naturellement présents dans la séquence -1851 / +263 (Figure 50). 2 unités d'enzymes de restriction ont été utilisées pour 2 µg de plasmide dans un volume final de 20 µl 1h à 37°C (25°C pour le couple *KpnI* / *SmaI*). Les plasmides ainsi digérés ont été purifiés avec le kit MinElute (Qiagen) et élués dans un volume final de 10 µl. Les extrémités digérées ont été rendues bouts francs par l'enzyme DNA Terminator® End Repair Kit (Euromedex, Mundolsheim, France) : 10 µl de plasmides ouverts et purifiés précédemment, 10 µl de tampon DNA Terminator® End Repair, 2 µl d'enzyme DNA Terminator® End Repair et qsp 50 µl d'eau. Ce mélange a été laissé 30 min à température ambiante puis 15 min à 70°C pour stopper la réaction. Les vecteurs ainsi traités ont été purifiés avec le kit MinElute (Qiagen) et récupérés dans un volume final de 10 µl. Enfin, ces 10 µl ont été traités avec la Quick T4 DNA Ligase (Biolabs, Ozyme, Saint Quentin Yvelines, France) pour les recirculariser : 10 µl de vecteurs ouverts avec extrémités franches sont incubés avec 10 unités de T4 DNA Ligase dans 10 µl de tampon (100 mM de Tris-HCl (pH 7,5), 20 mM de MgCl₂, 20 mM de dithiothreitol, 2 mM d'ATP et 50 µg/ml de sérum albumine bovine) pendant 5 min à température ambiante. Ainsi, les couples d'enzymes *KpnI* / *BglII*, *KpnI* / *XhoI*, *KpnI* / *NheI* ainsi que *KpnI* / *SmaI* ont été utilisées pour générer respectivement les plasmides pGL3_-752/+263, pGL3_-498/+263, pGL3_-121/+263 et pGL3_+92/+263.

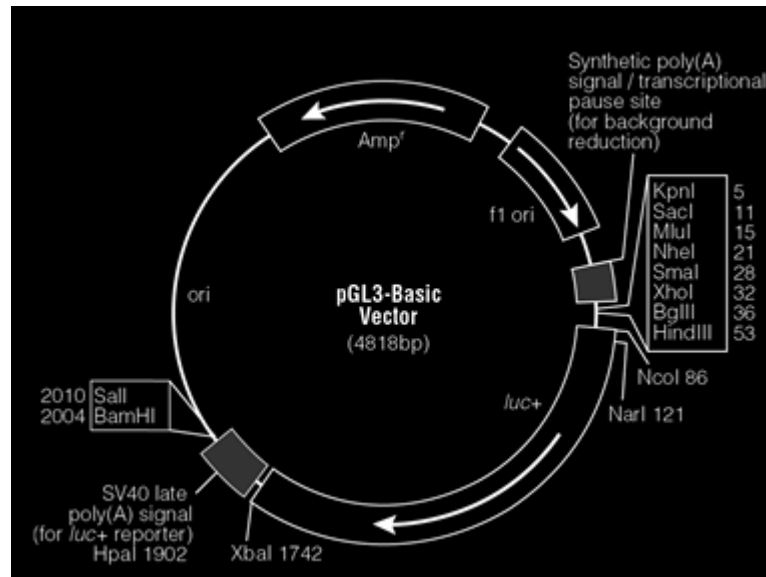


Figure 49 : Carte simplifiée du vecteur pGL3-Basic

D'après Promega. Luc+ correspond au gène codant la luciférase *Firefly*. Les différentes constructions pour les tests d'activité luciférase ont été insérées au site de clonage multiple (nucléotides 5 à 53).

En utilisant la même stratégie, la région génomique -121 / + 1451, englobant les exons EX, EY et EZ, a été isolée partir d'ADN génomique provenant de la lignée SH-SY5Y. Cette région a été amplifiée par PCR avec les amorces suivantes : 5'-TCACCGGACTTGGCGAGCTA-3' et 5'-GCCGCCTGAGATCTACATAAG-3', incluant le site de restriction *Bgl*III (nucléotides soulignés). En utilisant le site naturel *Nhe*I présent dans la séquence amplifiée et le site introduit *Bgl*III, le fragment obtenu par PCR a été inséré dans le vecteur pGL3-Basic (Promega), en amont du gène rapporteur codant la luciférase *Firefly*, aux sites *Nhe*I / *Bgl*III générant le plasmide pGL3_-121/+1452. Les plasmides pGL3_+517/+1452, pGL3_+773/+1452, pGL3_+465/+772, pGL3_-121/+772, pGL3_-121/+464 et pGL3_-121/+91 ont été générés par digestions enzymatiques, utilisation de bouts francs et ligation (Figure 50). Le plasmide pGL3_+1046/+1452 a été généré par PCR à partir d'ADN génomique de SH-SY5Y avec les amorces 5'-GGAGGGTACCGGGTAAGGGTT-3' et 5'-CCGTACCTTAATAAGCTTTCTACATA-3', contenant les sites de restrictions *Kpn*I et *Hind*III (nucléotides soulignés), respectivement. Le fragment amplifié a ensuite été inséré dans le vecteur pGL3-Basic (Promega) aux sites *Kpn*I / *Hind*III en amont du gène rapporteur codant la luciférase *Firefly*.

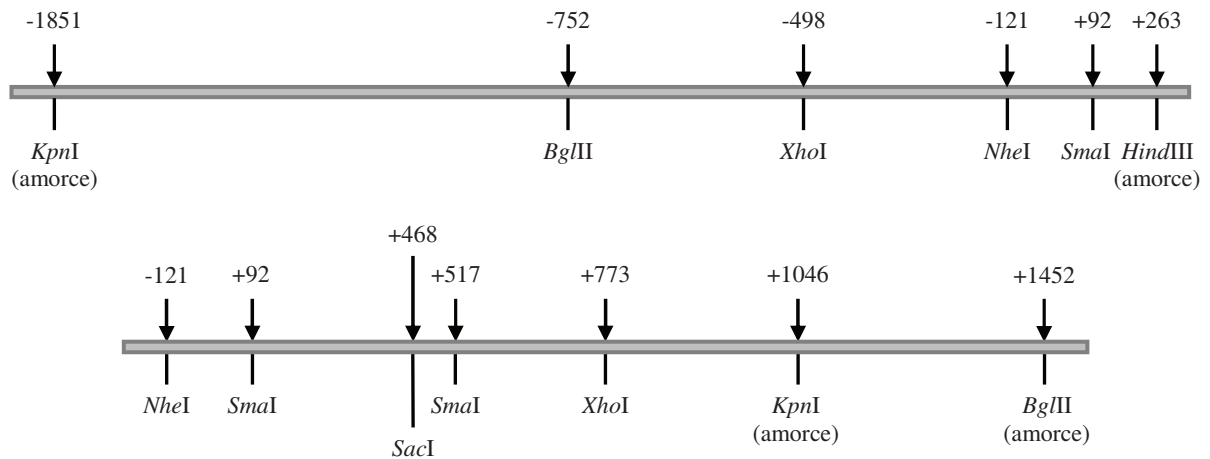


Figure 50 : Sites de restriction utilisés dans les régions -1851 / +263 et -121 / +1452 pour créer les différents vecteurs utilisés dans les tests d'activité luciférase.

Le nucléotide +1 correspond site d'initiation de la transcription de la plus longue séquence identifiée par 5' RACE pour l'exon EX. Les sites *KpnI*, *HindIII* et *BglIII* (sur la séquence -121 / +1452) ont été introduits par les amorces utilisées pour la PCR.

Les différentes PCR ont été réalisées avec la Taq polymérase HotStar HiFidelity (Qiagen) dans un mix final de 25 µl comprenant : 5 µl de Q-Solution 5x, 5 µl de PCR-Buffer 5x (1,5 mM de chaque dNTP, 7,5 mM de MgSO₄, Tris-Cl, du KCl, du (NH₄)₂SO₄, sérum albumine bovine, Triton® X-100 et Facteur SB (additif de Qiagen) ; pH 8,7), 1 µl de Taq HiFi (1,25 U), 1 µl d'ADN génomique (environ 20 ng) et qsp 25 µl.

L'ensemble de ces constructions a été séquencé afin de s'assurer de l'absence de mutation (Genoscreen).

7. MUTATIONS DES SITES DE FIXATIONS DES FACTEURS DE TRANSCRIPTION

Les mutations par substitutions de bases pour les sites de fixation prédits pour les facteurs de transcription NF-κB, NRSF, Olf1, Pura, Sox5, et les quatre sites SP1, désignés SP1.1, SP1.2, SP1.3 et SP1.4 ont été introduites par mutagenèse dirigée à l'aide de la HotStar HiFidelity Polymerase kit (Qiagen), utilisée selon les recommandations du kit (conditions d'utilisation et concentrations des différents réactifs identiques à celles du paragraphe ci-dessus). Deux premières PCR ont été réalisées avec d'une part, l'amorce sens commune 5'-CGATAGGTACCGAGCTCTTACGCG-3' (contenant le site *KpnI*) et une amorce anti-sens contenant la mutation spécifique d'un site (Tableau 6), et d'autre part, l'amorce anti-sens commune 5'-TGCCAAGCTTACTTAGATCGCAGATCTC-3' (contenant le site *HindIII*) et une amorce sens (complémentaire de l'amorce anti-sens) contenant la mutation spécifique

d'un site. Le plasmide pGL3_-121/+772 a été utilisé comme matrice pour ces deux PCR. Les PCR ont été réalisées dans les conditions suivantes : 95°C 5 min ; 30 cycles de 95°C 1 min, 53-60°C (en fonction des couples d'amorces) 1 min et 72°C 1 min ; 72°C 5 min. Les amplicons attendus ont été séparés et sélectionnés par électrophorèse sur gel d'agarose et purifiés et utilisés comme matrice pour une PCR finale avec les amorces communes sens et anti-sens dans les mêmes conditions que précédemment (Figure 43, Résultats). La mutation du site Olf1 a été générée par PCR avec l'amorce commune sens et l'amorce anti-sens Olf-3' (contenant le site *HindIII*, Tableau 6) dans les conditions de PCR suivantes : 95°C 5 min ; 30 cycles de 95°C 1 min, 53°C 1 min et 72°C for 1 min ; 72°C 5 min. Les produits finaux de PCR (contenant à leurs extrémités les sites de restrictions *KpnI* et *HindIII*) ont été insérés dans le vecteur pGL3-Basic aux sites *KpnI* / *HindIII* (Figure 49). Les vecteurs générés ont été séquencés pour s'assurer de l'absence d'autres mutations. Le plasmide pGL3_-121/+91NF-κB a été obtenu par digestion avec le couple d'enzyme *SmaI* / *BglIII* à partir du vecteur pGL3_-121/+772NF-κB, puis les extrémités digérées rendu bouts francs et liguées, selon le protocole décrit ci-dessus.

8. PCR EN TEMPS REELLE SYBR GREEN & DUPLEX TAQMAN

Les PCR en temps réel et les analyses consécutives ont été réalisées avec la machine MX4000 Multiplex Quantitative System (Stratagene, La Jolla, CA, USA) équipée du logiciel d'analyses Version 3.0.

8.1. Q-PCR avec la technologie SYBR Green®

Les amorces utilisées pour la quantification des gènes *ST6GAL1*, *ST6GAL2* et *HPRT* ont été dessinées pour s'hybrider dans la séquence de l'exon E6 (*ST6GAL1*), de l'exon E5 (*ST6GAL2*) et des exons E2 et E4 (Zhang *et al.*, 2005), respectivement (Tableau 7). Le volume réactionnel final de 25 µl était composé de 12,5 µl de Brilliant® SYBR Green® QPCR Master Mix Kit (Stratagene), du couple d'amorces considérées à la concentration déterminée (Tableau 7), de 4 µl d'ADNc cible dilué 1/40^{ème} et d'eau qsp 25 µl. Les conditions de PCR étaient : 95°C 10 min ; 40 cycles de 95°C 1 min, 51°C 1min et 72°C 30 sec ; 51°C jusqu'à 95°C en incrément de 1°C à chaque cycle pour la courbe de dissociation. Chaque point testé en Q-PCR a été réalisé en triplicata en plaque 96 puits. Les courbes standards

d'efficacité ont été réalisées pour chacun des couples d'amorces avec des dilutions sériées au 1/5^{ème} d'ADNc de SH-SY5Y. Les efficacités utilisées pour les calculs des résultats sont de 101,3 % pour *ST6GAL2*, de 100,9 % pour *ST6GAL1* et de 99,6 % pour *HPRT*. Elles représentent la moyenne d'au moins trois expériences indépendantes.

8.2. Q-PCR en duplex avec la technologie Taqman

Pour le gène *ST6GAL2* et ses différents transcrits, des amorces sens spécifiques ont été dessinées s'hybridant à l'intérieur de chacune des séquences des exons EX, EY et EZ et compatibles avec l'amorce anti-sens commune exon1-3'Taq s'hybridant dans la séquence de l'exon commun E1 (Tableau 7 et Figure 29, Résultats). Pour le gène normalisateur *HPRT*, les amorces ont été dessinées pour s'hybrider avec les séquences des exons E1 et E2 (Tableau 7 et Figure 29, Résultats). Les sondes Taqman ont été synthétisées par Applied Biosystems. Les sondes relatives aux transcrits du gène *ST6GAL2* ont été marquées avec le fluorophore FAM et la sonde relative au gène *HPRT* a été marquée avec le fluorophore VIC. Les réactions de PCR ont été réalisées avec 12,5 µl du QuantiTect Multiplex PCR Kit Mix (Qiagen), 0,4 µM de chacune des quatre amorces, 0,4 µM de chaque sonde et 4 µl d'ADNc cible dilué au 1/40^{ème} dans un volume final de 25 µl. Les conditions suivantes ont été appliquées pour les Q-PCR : 95°C 15 min ; 40 cycles de 95°C 1min et 60° C 1 min (hybridation et élongation). Chaque point analysé a été réalisé en triplicata dans une plaque 96 puits. Pour s'assurer de l'efficacité de la technique, les courbes standards d'efficacité ont été réalisées pour chaque couple *HPRT* / EX, *HPRT* / EY et *HPRT* / EZ en utilisant des dilutions sériées au 1/2 avec de l'ADNc de cellules SH-SY5Y. Les efficacités utilisées pour les calculs des résultats sont de 99.5 % et 97.8 % pour le transcrit spécifique EX et *HPRT*, 100.4 % et 100.5 % pour le transcrit spécifique EY et *HPRT* et enfin 100.2 % et 98.0 % pour le transcrit spécifique EZ et *HPRT*. Elles représentent la moyenne d'au moins trois expériences indépendantes.

9. ANALYSES BIOINFORMATIQUES

Les analyses bioinformatiques concernant la recherche de sites potentiels de fixation de facteurs de transcription ont été réalisées avec le logiciel MatInspector 2.2 à partir du portail internet de Genomatix (www.genomatix.de) en utilisant les matrices de TRANSAC 4.0 (Quandt, 1995) avec comme paramètres : "core similarity: 0.95" et "matrices similarity:

Optimized”. Ces paramètres ont été utilisés dans le but de diminuer le nombre de « faux positifs » au niveau des sites de fixation reconnus par les matrices bioinformatiques.

PROTEOMIQUE

1. EXTRACTION DE PROTEINES

La concentration des différents extraits protéiques a été déterminée avec le kit Micro BCA™ Protein Assay Reagent (Pierce, Rockford, USA) en utilisant la sérum albumine bovine (BSA) comme standard pour la gamme étalon, selon les instructions fournis par le fabricant.

1.1. Protéines extraites de cellules en cultures

Les protéines totales des lignées cellulaires SH-SY5Y, Kelly, U118, T98G, HEK 293, HeLa, THP1, Caco-2, MDA-MB-231, BT-20, SK Mel-28, NT2/D1, hNT et NN ont été extraites à partir de culots cellulaires ($\sim 5 \cdot 10^6$ cellules). Ces culots ont été repris sur la glace dans 500 μ l de tampon de lyse (150 mM de NaCl, 50 mM de Tris-HCl, 1% de Triton® X-100 ; pH 7,5) contenant un cocktail de d'inhibiteurs de phosphatases et de protéases (Roche, Meylan, France). La solution a été vigoureusement vortexée pendant 1 min, soniquée dans un bain à sonication pendant 2 min (puissance 50 %), puis placée 30 min dans la glace. Le surnageant (contenant les protéines totales) et les débris cellulaires ont été séparés par centrifugation pendant 20 min à 10 000 tr/min à 4°C.

1.2. Protéines extraites d'échantillons tissulaires

Les échantillons des différents lobes du cortex cérébral, de l'hippocampe et du cervelet ont été homogénéisés sur glace à l'aide d'un piston et d'un homogénéisateur en verre d'une capacité de 5 ml dans un tampon composé de CHAPS (0,5 %) ainsi qu'un cocktail d'inhibiteurs de protéases et de phosphatases (Roche). 100 μ l de tampon de lyse ont été ajoutés pour 10 mg de tissu. L'homogénat a ensuite été soniqué sur glace à 50 % de puissance environ 30 pulsations puis mis en agitation douce sur une roue 2 h à 4°C. Le surnageant contenant les protéines totales et les débris tissulaires ont été séparés par centrifugation pendant 20 min à 10000 tr/min à 4°C. Les protéines ont été stockées à -20°C ou -80°C.

2. WESTERN BLOTTING

L'anticorps anti-hST6Gal II a été produit chez le lapin et est dirigé contre la séquence peptidique PSSLSFLETRLLPVQGKQRAIMGAAHEPSPPGGLDARQALPRAHPAGSFH AGPGDLQKWAQSQDGFHKEFFSSQVGRKSQSFAFYPEDDDYFFAAGQPGWHSHTQG TLGFSPGEPGPREGAFPAAQ localisée dans la région tige de l'enzyme. 40 µg de protéines totales par échantillon ont été chauffés à 95° pendant 10 min avec du tampon de Laemmli réducteur (8X : 12,5 % de Tris-HCl pH=6,8 2 M, 40 % de glycérol, 20 % de SDS, 5 % de β-mercaptoéthanol, 0,01 % de bleu de bromophénol, H₂O). Les protéines ont ensuite été séparées par SDS/PAGE sur un mini gel de 8 % (Bio-Rad, Richmond, USA) ou sur un grand gel à gradient 4-20 % (Bio-Rad). Après transfert sur membrane de nitrocellulose (200 mA, 2 h), l'étape de blocage a été réalisée avec du TBS (Tris Buffer Saline) contenant 0,05 % de Tween 20 et 5 % de lait écrémé en poudre (Auchan) pendant 1 h à température ambiante (~20°C) sous agitation douce. L'incubation de l'anticorps primaire, anti-ST6Gal II (Sigma-Aldrich, St Louis, USA) ou anti-actine (Santa Cruz Biotechnology inc., Europe), a été réalisée pendant une nuit (~16 h) à 4°C sous agitation douce dans du TBS contenant 0,05 % de Tween 20 et 5 % de lait écrémé en poudre. Les anticorps ont été ajoutés à la dilution de 1:1000 pour l'anticorps anti-ST6Gal II et 1:10000 pour l'anticorps anti-actine. Après trois étapes de lavage de 10 min avec du TBS, 0,05 % Tween 20 les membranes ont été incubées 1 h à température ambiante sous agitation douce avec du TBS, 0,05 % Tween 20 et l'anticorps secondaire IgG de chèvre anti-lapin (HRP) à la dilution de 1:10000 dans du TBS contenant 0,05 % Tween 20 et 5 % de lait écrémé en poudre. Enfin, les membranes ont été lavées trois fois pendant 10 min dans du TBS, 0,05 % Tween et la détection des signaux a été effectuée par chimioluminescence avec le kit ECL+[®] advance Western blotting detection reagents (Amersham Biosciences, Little Chalfont, Bucks, U.K.).

3. REVELATION DES PROTEINES AU ROUGE PONCEAU

Après transfert sur membranes de nitrocellulose, quelques ml de rouge ponceau (Rouge ponceau 0,1 % acide acétique 5%) ont été déposés pour visualiser les protéines transférées et ainsi vérifier l'efficacité du transfert et l'équichargement des pistes en protéines sur la membrane avant de procéder à la révélation par Western-Blot. L'excédent de rouge ponceau est enlevé et la membrane rincée à l'eau.

Tableau 5 : Séquences des amorces utilisées pour la 5' RACE

Amorce	Séquence	T°C d'hybridation
Adaptateur 5' RACE	5'-GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGC UUUGAUGAAA-3	-
Outer Primer	5'-GCTGATGGCGATGAATGAACACTG-3'	61°
Inner Primer	5'-CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG-3'	72°
GSP1	5'-CCTCTTCTTCACCCGCCTC-3'	61°
GSP2.1	5'-AGCGGGGTTGCTGTCGGT-3'	72°
GSP2.2	5'-AGGCGGGGAGGGCTCATG-3'	72°
GSP1.X	5'-GGAGCCGGGCACTGGGAAGTTGGG-3'	60°
GSP2.X	5'-AGACTTGGGCTGCGTGGGGCGGTGC-3'	72°
GSP1.Y	5'-CTGGAGTCCAAGGCGTGCGTGTCA-3'	60°
GSP2.Y	5'-TCCTCCTGCCACCGTGGGGTCTCG-3'	72°
GSP1.Z	5'-TACTGAGCCACAGGGGCACAACTG-3'	60°
GSP2.Z	5'-CCCCAAAGCGCAAGCCGTCCTCCC-3'	72°

Tableau 6 : Séquences des amorces utilisées pour la mutagenèse dirigée

Amorce	Séquence	T°C d'hybridation
NFκB-5'	5'-CAGGCGCCTTTACTCAATCCTGG-3'	60°
NFκB-3'	5'-CCAGGATTGAGTAAAGGCGCCTG-3'	60°
NRSF-5'	5'-CCGAGTGGGCGATTGGTGTCCA-3'	60°
NRSF-3'	5'-TGGACACCAATCGCCCACTCGG-3'	60°
Olf1-3'	5'-GCTGCCAAGCTTTCCCGTCAAATGC CGCCTG-3'	60°
Purα-5'	5'-GGATGGAGGCTTACACCGGGG-3'	60°
Purα-3'	5'-CCCCGGTGTAAGCCTCCATCC-3'	55°
Sox5-5'	5'-GCCCGAGAGTATGAGGCGTGC-3'	60°
Sox5-3'	5'-GCACGCCTCATACTCTCGGGC-3'	60°
SP1.1-5'	5'-CTGCCGAAGAAAGGGACCCG-3'	60°
SP1.1-3'	5'-CGGGTCCCTTTCTTCGGCAG-3'	60°
SP1.2-5'	5'-CCCGCACTTTTCCACGCAG-3'	60°
SP1.2-3'	5'-CTGCGTGGAAAAGTGCGGG-3'	60°

SP1.3-5'	5'-GAGCTCGGGTGTGTTGGCGC-3'	53°
SP1.3-3'	5'-GCGCCAAACACCCGAGCTC-3'	60°
SP1.4-5'	5'-GCCCCGCTAAATGCGCTTG-3'	60°
SP1.4-3'	5'-CAAGCGCATTTAGCGGGGC-3'	60°

Tableau 7 : Séquences des amorces et des sondes utilisées en Q-PCR

Amorce SYBR Green®	Séquence	T°C d'hybridation
ST6GAL2 5'	5'-ACGCTGCTGATTGACTCTTCT-3'	51°
ST6GAL2 3'	5'-CACATACTGGCACTCATCTAA-3'	51°
ST6GAL1 5'	5'-GGGCTCCAAACTAACCATCTC-3'	51°
ST6GAL1 3'	5'-AAATCCAGGCTTTCTCACTCC-3'	51°
HPRT 5' *	5'-GACCAGTCAACAGGGGACAT-3'	51°
HPRT 3' *	5'-AACACTTCGTGGGGTCCTTTTC-3'	51°
Amorce Taqman	Séquence	T°C d'hybridation
HPRT5'Taq	5'-GCCCTGGCGTCGTGATT-3'	60°
HPRT3'Taq	5'-CTCAGCATAATGATTAGGTATGCAAAA-3'	60°
EX5'Taq	5'-GCAGCCCAACTTCCCAGTG-3'	60°
EY5'Taq	5'-GGTGTCGTGCGCTTCGA-3'	60°
EZ5'Taq	5'-ACGGGCGGATACAGTTGTG-3'	60°
Exon1-3'Taq	5'-TCAGCACCTTGTGTCTTAATGCA-3'	60°
Sonde Taqman	Séquence	T°C d'hybridation
HPRT	5'-VIC-TGATGATGAACCAGGTAT-TAMRA-3'	60°
EX	5'-FAM-CCAGCTAGGTTTCATTCT-MGB-3'	60°
EY	5'-FAM-AAGAAGGCTGGTTCATT-MGB-3'	60°
EZ	5'-FAM-CAGTAAAGGTTTCATTCTGCCA-MGB-3'	60°

* : Zhang, 2005

Références bibliographiques

REFERENCES BIBLIOGRAPHIQUES

- Aas-Eng D A, Asheim H C, Deggerdal A, Smeland E, Funderud S, (1995):** Characterization of a promoter region supporting transcription of a novel human beta-galactoside alpha-2,6-sialyltransferase transcript in HepG2 cells. *Biochim Biophys Acta*. 1261(1): 166-9.
- Andrews P W, (1984):** Retinoic acid induces neuronal differentiation of a cloned human embryonal carcinoma cell line in vitro. *Dev Biol*. 103(2): 285-93.
- Andrews P W, (1998):** Teratocarcinomas and human embryology: pluripotent human EC cell lines. Review article. *Apmis*. 106(1): 158-67; discussion 167-8.
- Angata K, Nakayama J, Fredette B, Chong K, Ranscht B, Fukuda M, (1997):** Human STX polysialyltransferase forms the embryonic form of the neural cell adhesion molecule. Tissue-specific expression, neurite outgrowth, and chromosomal localization in comparison with another polysialyltransferase, PST. *J Biol Chem*. 272(11): 7182-90.
- Angata K, Suzuki M, McAuliffe J, Ding Y, Hindsgaul O, Fukuda M, (2000):** Differential biosynthesis of polysialic acid on neural cell adhesion molecule (NCAM) and oligosaccharide acceptors by three distinct α 2,8-sialyltransferases, ST8Sia IV (PST), ST8Sia II (STX), and ST8Sia III. *J Biol Chem*. 275(24): 18594-601.
- Angata T, Varki A, (2002):** Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chem Rev*. 102(2): 439-69.
- Anobile C J, Talbot J A, McCann S J, Padmanabhan V, Robertson W R, (1998):** Glycoform composition of serum gonadotrophins through the normal menstrual cycle and in the post-menopausal state. *Mol Hum Reprod*. 4(7): 631-9.
- Anthony R M, Nimmerjahn F, Ashline D J, Reinhold V N, Paulson J C, Ravetch J V, (2008):** Recapitulation of IVIG anti-inflammatory activity with a recombinant IgG Fc. *Science*. 320(5874): 373-6.
- Anthony R M, Wermeling F, Karlsson M C, Ravetch J V, (2008):** Identification of a receptor required for the anti-inflammatory activity of IVIG. *Proc Natl Acad Sci U S A*. 105(50): 19571-8.
- Ashwell G, Harford J, (1982):** Carbohydrate-specific receptors of the liver. *Annu Rev Biochem*. 51: 531-54.
- Ballas N, Mandel G, (2005):** The many faces of REST oversee epigenetic programming of neuronal genes. *Curr Opin Neurobiol*. 15(5): 500-6.
- Bergwerff A A, Van Oostrum J, Kamerling J P, Vliegthart J F, (1995):** The major N-linked carbohydrate chains from human urokinase. The occurrence of 4-O-sulfated, (alpha 2-6)-sialylated or (alpha 1-3)-fucosylated N-acetylgalactosamine(beta 1-4)-N-acetylglucosamine elements. *Eur J Biochem*. 228(3): 1009-19.

- Bonfanti L**, (2006): PSA-NCAM in mammalian structural plasticity and neurogenesis. *Prog Neurobiol.* 80(3): 129-64.
- Braesch-Andersen S, Stamenkovic I**, (1994): Sialylation of the B lymphocyte molecule CD22 by alpha 2,6-sialyltransferase is implicated in the regulation of CD22-mediated adhesion. *J Biol Chem.* 269(16): 11783-6.
- Bruses J L, Rutishauser U**, (2001): Roles, regulation, and mechanism of polysialic acid function during neural development. *Biochimie.* 83(7): 635-43.
- Burgon P G, Stanton P G, Robertson D M**, (1996): In vivo bioactivities and clearance patterns of highly purified human luteinizing hormone isoforms. *Endocrinology.* 137(11): 4827-36.
- Chammas R, Veiga S S, Travassos L R, Brentani R R**, (1993): Functionally distinct roles for glycosylation of alpha and beta integrin chains in cell-matrix interactions. *Proc Natl Acad Sci U S A.* 90(5): 1795-9.
- Chan A L, Morris H R, Panico M, Etienne A T, Rogers M E, Gaffney P, Creighton-Kempford L, Dell A**, (1991): A novel sialylated N-acetylgalactosamine-containing oligosaccharide is the major complex-type structure present in Bowes melanoma tissue plasminogen activator. *Glycobiology.* 1(2): 173-85.
- Chang M L, Eddy R L, Shows T B, Lau J T**, (1995): Three genes that encode human beta-galactoside alpha 2,3-sialyltransferases. Structural analysis and chromosomal mapping studies. *Glycobiology.* 5(3): 319-25.
- Chavan M M, Kawle P D, Mehta N G**, (2005): Increased sialylation and defucosylation of plasma proteins are early events in the acute phase response. *Glycobiology.* 15(9): 838-48.
- Chen Y, Kundakovic M, Agis-Balboa R C, Pinna G, Grayson D R**, (2007): Induction of the reelin promoter by retinoic acid is mediated by Sp1. *J Neurochem.* 103(2): 650-65.
- Cheung Y T, Lau W K, Yu M S, Lai C S, Yeung S C, So K F, Chang R C**, (2009): Effects of all-trans-retinoic acid on human SH-SY5Y neuroblastoma as in vitro model in neurotoxicity research. *Neurotoxicology.* 30(1): 127-35.
- Chiu P C, Koistinen R, Koistinen H, Seppala M, Lee K F, Yeung W S**, (2003): Zona-binding inhibitory factor-1 from human follicular fluid is an isoform of glycodelin. *Biol Reprod.* 69(1): 365-72.
- Chou H H, Takematsu H, Diaz S, Iber J, Nickerson E, Wright K L, Muchmore E A, Nelson D L, Warren S T, Varki A**, (1998): A mutation in human CMP-sialic acid hydroxylase occurred after the Homo-Pan divergence. *Proc Natl Acad Sci U S A.* 95(20): 11751-6.
- Collen D, Lijnen H R**, (1992): Fibrin-specific fibrinolysis. *Ann N Y Acad Sci.* 667: 259-71.

- Couillard-Despres S, Quehl E, Altendorfer K, Karl C, Ploetz S, Bogdahn U, Winkler J, Aigner L, (2008):** Human in vitro reporter model of neuronal development and early differentiation processes. *BMC Neurosci.* 9: 31.
- Coutinho P M, Deleury E, Davies G J, Henrissat B, (2003):** An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol.* 328(2): 307-17.
- Cyster J G, Goodnow C C, (1997):** Tuning antigen receptor signaling by CD22: integrating cues from antigens and the microenvironment. *Immunity.* 6(5): 509-17.
- Dall'Olio F, Malagolini N, di Stefano G, Minni F, Marrano D, Serafini-Cessi F, (1989):** Increased CMP-NeuAc:Gal beta 1,4GlcNAc-R alpha 2,6 sialyltransferase activity in human colorectal cancer tissues. *Int J Cancer.* 44(3): 434-9.
- Dall'Olio F, Trere D, (1993):** Expression of alpha 2,6-sialylated sugar chains in normal and neoplastic colon tissues. Detection by digoxigenin-conjugated Sambucus nigra agglutinin. *Eur J Histochem.* 37(3): 257-65.
- Dall'olio F, (1996):** Protein glycosylation in cancer biology: an overview. *Clin Mol Pathol.* 49(3): M126-M135.
- Dall'Olio F, (2000):** The sialyl-alpha2,6-lactosaminyl-structure: biosynthesis and functional role. *Glycoconj J.* 17(10): 669-76.
- Dall'Olio F, Chiricolo M, D'Errico A, Gruppioni E, Altimari A, Fiorentino M, Grigioni W F, (2004):** Expression of beta-galactoside alpha2,6 sialyltransferase and of alpha2,6-sialylated glycoconjugates in normal human liver, hepatocarcinoma, and cirrhosis. *Glycobiology.* 14(1): 39-49.
- Dalziel M, Lemaire S, Ewing J, Kobayashi L, Lau J T, (1999):** Hepatic acute phase induction of murine beta-galactoside alpha 2,6 sialyltransferase (ST6Gal I) is IL-6 dependent and mediated by elevation of exon H-containing class of transcripts. *Glycobiology.* 9(10): 1003-8.
- Dalziel M, Huang R Y, Dall'Olio F, Morris J R, Taylor-Papadimitriou J, Lau J T, (2001):** Mouse ST6Gal sialyltransferase gene expression during mammary gland lactation. *Glycobiology.* 11(5): 407-12.
- Darbinian N, White M K, Khalili K, (2006):** Regulation of the Pur-alpha promoter by E2F-1. *J Cell Biochem.* 99(4): 1052-63.
- Datta A K, Paulson J C, (1995):** The sialyltransferase "sialylmotif" participates in binding the donor substrate CMP-NeuAc. *J Biol Chem.* 270(4): 1497-500.
- Datta A K, Sinha A, Paulson J C, (1998):** Mutation of the sialyltransferase S-sialylmotif alters the kinetics of the donor and acceptor substrates. *J Biol Chem.* 273(16): 9608-14.
- Dell A, Morris H R, Easton R L, Panico M, Patankar M, Oehniger S, Koistinen R, Koistinen H, Seppala M, Clark G F, (1995):** Structural analysis of the oligosaccharides

derived from glycodeilin, a human glycoprotein with potent immunosuppressive and contraceptive activities. *J Biol Chem.* 270(41): 24116-26.

Durbec P, Cremer H, (2001): Revisiting the function of PSA-NCAM in the nervous system. *Mol Neurobiol.* 24(1-3): 53-64.

Eskinazi R, Thony B, Svoboda M, Robberecht P, Dassel D, Heizmann C W, Van Laethem J L, Resibois A, (1999): Overexpression of pterin-4a-carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 in human colon cancer. *Am J Pathol.* 155(4): 1105-13.

Fiete D, Srivastava V, Hindsgaul O, Baenziger J U, (1991): A hepatic reticuloendothelial cell receptor specific for SO4-4GalNAc beta 1,4GlcNAc beta 1,2Man alpha that mediates rapid clearance of lutropin. *Cell.* 67(6): 1103-10.

Foley D A, Swartzentruber K G, Colley K J, (2009): Identification of sequences in the polysialyltransferases ST8Sia II and ST8Sia IV that are required for the protein-specific polysialylation of the neural cell adhesion molecule, NCAM. *J Biol Chem.* 284(23): 15505-16.

Forrest J C, Dermody T S, (2003): Reovirus receptors and pathogenesis. *J Virol.* 77(17): 9109-15.

Friedman G C, Seeds N W, (1994): Tissue plasminogen activator expression in the embryonic nervous system. *Brain Res Dev Brain Res.* 81(1): 41-9.

Fukuda M, (2002): Roles of mucin-type O-glycans in cell adhesion. *Biochim Biophys Acta.* 1573(3): 394-405.

Fukumoto S, Miyazaki H, Goto G, Urano T, Furukawa K, Furukawa K, (1999): Expression cloning of mouse cDNA of CMP-NeuAc:Lactosylceramide alpha2,3-sialyltransferase, an enzyme that initiates the synthesis of gangliosides. *J Biol Chem.* 274(14): 9271-6.

Gabay C, Kushner I, (1999): Acute-phase proteins and other systemic responses to inflammation. *N Engl J Med.* 340(6): 448-54.

Gagneux P, Cheriyan M, Hurtado-Ziola N, van der Linden E C, Anderson D, McClure H, Varki A, Varki N M, (2003): Human-specific regulation of alpha 2-6-linked sialic acids. *J Biol Chem.* 278(48): 48245-50.

Gallia G L, Johnson E M, Khalili K, (2000): Puralpha: a multifunctional single-stranded DNA- and RNA-binding protein. *Nucleic Acids Res.* 28(17): 3197-205.

Garel S, Marin F, Mattei M G, Vesque C, Vincent A, Charnay P, (1997): Family of Ebf/Olf-1-related genes potentially involved in neuronal differentiation and regional specification in the central nervous system. *Dev Dyn.* 210(3): 191-205.

Garige M, Gong M, Rao M N, Zhang Y, Lakshman M R, (2005): Mechanism of action of ethanol in the down-regulation of Gal(beta)1, 4GlcNAc alpha2,6-sialyltransferase messenger RNA in human liver cell lines. *Metabolism.* 54(6): 729-34.

Garige M, Gong M, Lakshman M R, (2006): Ethanol destabilizes liver Gal beta 1, 4GlcNAc alpha2,6-sialyltransferase, mRNA by depleting a 3'-untranslated region-specific binding protein. *J Pharmacol Exp Ther.* 318(3): 1076-82.

Gauldie J, Richards C, Harnish D, Lansdorp P, Baumann H, (1987): Interferon beta 2/B-cell stimulatory factor type 2 shares identity with monocyte-derived hepatocyte-stimulating factor and regulates the major acute phase protein response in liver cells. *Proc Natl Acad Sci U S A.* 84(20): 7251-5.

Gee G V, Dugan A S, Tsomaia N, Mierke D F, Atwood W J, (2006): The role of sialic acid in human polyomavirus infections. *Glycoconj J.* 23(1-2): 19-26.

Gessner P, Riedl S, Quentmaier A, Kemmner W, (1993): Enhanced activity of CMP-neuAc:Gal beta 1-4GlcNAc:alpha 2,6-sialyltransferase in metastasizing human colorectal tumor tissue and serum of tumor patients. *Cancer Lett.* 75(3): 143-9.

Giordanengo V, Bannwarth S, Laffont C, Van Miegem V, Harduin-Lepers A, Delannoy P, Lefebvre J C, (1997): Cloning and expression of cDNA for a human Gal(beta1-3)GalNAc alpha2,3-sialyltransferase from the CEM T-cell line. *Eur J Biochem.* 247(2): 558-66.

Gong M, Garige M, Hirsch K, Lakshman M R, (2007): Liver Galbeta1,4GlcNAc alpha2,6-sialyltransferase is down-regulated in human alcoholics: possible cause for the appearance of asialoconjugates. *Metabolism.* 56(9): 1241-7.

Groux-Degroote S, Krzewinski-Recchi M A, Cazet A, Vincent A, Lehoux S, Lafitte J J, Van Seuningen I, Delannoy P, (2008): IL-6 and IL-8 increase the expression of glycosyltransferases and sulfotransferases involved in the biosynthesis of sialylated and/or sulfated Lewisx epitopes in the human bronchial mucosa. *Biochem J.* 410(1): 213-23.

Grundmann U, Nerlich C, Rein T, Zettlmeissl G, (1990): Complete cDNA sequence encoding human beta-galactoside alpha-2,6-sialyltransferase. *Nucleic Acids Res.* 18(3): 667.

Grundstrom S, Anderson P, Scheipers P, Sundstedt A, (2004): Bcl-3 and NFkappaB p50-p50 homodimers act as transcriptional repressors in tolerant CD4+ T cells. *J Biol Chem.* 279(9): 8460-8.

Gu J, Taniguchi N, (2004): Regulation of integrin functions by N-glycans. *Glycoconj J.* 21(1-2): 9-15.

Haines N, Irvine K D, (2005): Functional analysis of Drosophila beta1,4-N-acetylgalactosaminyltransferases. *Glycobiology.* 15(4): 335-46.

Hanasaki K, Varki A, Stamenkovic I, Bevilacqua M P, (1994): Cytokine-induced beta-galactoside alpha-2,6-sialyltransferase in human endothelial cells mediates alpha 2,6-sialylation of adhesion molecules and CD22 ligands. *J Biol Chem.* 269(14): 10637-43.

Hanisch F G, (2001): O-glycosylation of the mucin type. *Biol Chem.* 382(2): 143-9.

- Haraguchi M, Yamashiro S, Yamamoto A, Furukawa K, Takamiya K, Lloyd K O, Shiku H, Furukawa K, (1994):** Isolation of GD3 synthase gene by expression cloning of GM3 alpha-2,8-sialyltransferase cDNA using anti-GD2 monoclonal antibody. *Proc Natl Acad Sci U S A*. 91(22): 10455-9.
- Harduin-Lepers A, Stokes D C, Steelant W F, Samyn-Petit B, Krzewinski-Recchi M A, Vallejo-Ruiz V, Zanetta J P, Auge C, Delannoy P, (2000):** Cloning, expression and gene organization of a human Neu5Ac alpha 2-3Gal beta 1-3GalNAc alpha 2,6-sialyltransferase: hST6GalNAcIV. *Biochem J*. 352 Pt 1: 37-48.
- Harduin-Lepers A, Vallejo-Ruiz V, Krzewinski-Recchi M A, Samyn-Petit B, Julien S, Delannoy P, (2001):** The human sialyltransferase family. *Biochimie*. 83(8): 727-37.
- Harduin-Lepers A, Mollicone R, Delannoy P, Oriol R, (2005):** The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology*. 15(8): 805-17.
- Henkin J, Dudlak D, Beebe D P, Sennello L, (1991):** igh sialic acid content slows prourokinase turnover in rabbits. *Thromb Res*. 63(2): 215-25.
- Hennet T, Chui D, Paulson J C, Marth J D, (1998):** Immune regulation by the ST6Gal sialyltransferase. *Proc Natl Acad Sci U S A*. 95(8): 4504-9.
- Husmann M, Dragneva Y, Romahn E, Jehnichen P, (2000):** Nuclear receptors modulate the interaction of Sp1 and GC-rich DNA via ternary complex formation. *Biochem J*. 352 Pt 3: 763-72.
- Ikehara Y, Kojima N, Kurosawa N, Kudo T, Kono M, Nishihara S, Issiki S, Morozumi K, Itzkowitz S, Tsuda T, Nishimura S I, Tsuji S, Narimatsu H, (1999a):** Cloning and expression of a human gene encoding an N-acetylgalactosamine-alpha2,6-sialyltransferase (ST6GalNAc I): a candidate for synthesis of cancer-associated sialyl-Tn antigens. *Glycobiology*. 9(11): 1213-24.
- Ikehara Y, Shimizu N, Kono M, Nishihara S, Nakanishi H, Kitamura T, Narimatsu H, Tsuji S, Tatematsu M, (1999b):** A novel glycosyltransferase with a polyglutamine repeat; a new candidate for GD1alpha synthase (ST6GalNAc V)(1). *FEBS Lett*. 463(1-2): 92-6.
- Ishii A, Ohta M, Watanabe Y, Matsuda K, Ishiyama K, Sakoe K, Nakamura M, Inokuchi J, Sanai Y, Saito M, (1998):** Expression cloning and functional characterization of human cDNA for ganglioside GM3 synthase. *J Biol Chem*. 273(48): 31652-5.
- Ishii A, Ikeda T, Hitoshi S, Fujimoto I, Torii T, Sakuma K, Nakakita S, Hase S, Ikenaka K, (2007):** Developmental changes in the expression of glycogenes and the content of N-glycans in the mouse cerebral cortex. *Glycobiology*. 17(3): 261-76.
- Jeanneau C, Chazalet V, Auge C, Soumpasis D M, Harduin-Lepers A, Delannoy P, Imberty A, Breton C, (2004):** Structure-function analysis of the human sialyltransferase ST3Gal I: role of n-glycosylation and a novel conserved sialylmotif. *J Biol Chem*. 279(14): 13461-8.

- Julkunen M, Apter D, Seppala M, Stenman U H, Bohn H, (1986):** Serum levels of placental protein 14 reflect ovulation in nonconceptional menstrual cycles. *Fertil Steril.* 45(1): 47-50.
- Kalcheva I, Elliott R W, Dalziel M, Lau J T, (1997):** The gene encoding beta-galactoside alpha2,6-sialyltransferase maps to mouse chromosome 16. *Mamm Genome.* 8(8): 619-20.
- Kaneko Y, Yamamoto H, Colley K J, Moskal J R, (1995):** Expression of Gal beta 1,4GlcNAc alpha 2,6-sialyltransferase and alpha 2,6-linked sialoglycoconjugates in normal human and rat tissues. *J Histochem Cytochem.* 43(9): 945-54.
- Kaneko Y, Yamamoto H, Kersey D S, Colley K J, Leestma J E, Moskal J R, (1996):** The expression of Gal beta 1,4GlcNAc alpha 2,6 sialyltransferase and alpha 2,6-linked sialoglycoconjugates in human brain tumors. *Acta Neuropathol.* 91(3): 284-92.
- Kaneko Y, Nimmerjahn F, Ravetch J V, (2006):** Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science.* 313(5787): 670-3.
- Kang N Y, Kim C H, Kim K S, Ko J H, Lee J H, Jeong Y K, Lee Y C, (2007):** Expression of the human CMP-NeuAc:GM3 alpha2,8-sialyltransferase (GD3 synthase) gene through the NF-kappaB activation in human melanoma SK-MEL-2 cells. *Biochim Biophys Acta.* 1769(11-12): 622-30.
- Kaplan H A, Woloski B M, Hellman M, Jamieson J C, (1983):** Studies on the effect of inflammation on rat liver and serum sialyltransferase. Evidence that inflammation causes release of Gal beta 1 leads to 4GlcNAc alpha 2 leads to 6 sialyltransferase from liver. *J Biol Chem.* 258(19): 11505-9.
- Kastenbauer S, Ziegler-Heitbrock H W, (1999):** NF-kappaB1 (p50) is upregulated in lipopolysaccharide tolerance and can block tumor necrosis factor gene expression. *Infect Immun.* 67(4): 1553-9.
- Kelm S, Schauer R, (1997):** Sialic acids in molecular and cellular interactions. *Int Rev Cytol.* 175: 137-240.
- Khalili K, Del Valle L, Muralidharan V, Gault W J, Darbinian N, Otte J, Meier E, Johnson E M, Daniel D C, Kinoshita Y, Amini S, Gordon J, (2003):** Puralpha is essential for postnatal brain development and developmentally coupled cellular proliferation as revealed by genetic inactivation in the mouse. *Mol Cell Biol.* 23(19): 6857-75.
- Kim K W, Kim S W, Min K S, Kim C H, Lee Y C, (2001):** Genomic structure of human GM3 synthase gene (hST3Gal V) and identification of mRNA isoforms in the 5'-untranslated region. *Gene.* 273(2): 163-71.
- Kim Y J, Kim K S, Kim S H, Kim C H, Ko J H, Choe I S, Tsuji S, Lee Y C, (1996):** Molecular cloning and expression of human Gal beta 1,3GalNAc alpha 2,3-sialyltransferase (hST3Gal II). *Biochem Biophys Res Commun.* 228(2): 324-7.

- Kim Y J, Kim K S, Do S, Kim C H, Kim S K, Lee Y C, (1997):** Molecular cloning and expression of human alpha2,8-sialyltransferase (hST8Sia V). *Biochem Biophys Res Commun.* 235(2): 327-30.
- Kimura K, Wakamatsu A, Suzuki Y, Ota T, Nishikawa T, Yamashita R, Yamamoto J, Sekine M, Tsuritani K, Wakaguri H, Ishii S, Sugiyama T, Saito K, Isono Y, Irie R, Kushida N, Yoneyama T, Otsuka R, Kanda K, Yokoi T, Kondo H, Wagatsuma M, Murakawa K, Ishida S, Ishibashi T, Takahashi-Fujii A, Tanase T, Nagai K, Kikuchi H, Nakai K, Isogai T, Sugano S, (2006):** Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res.* 16(1): 55-65.
- Kiss J Z, Rougon G, (1997):** Cell biology of polysialic acid. *Curr Opin Neurobiol.* 7(5): 640-6.
- Kitagawa H, Paulson J C, (1994):** Cloning of a novel alpha 2,3-sialyltransferase that sialylates glycoprotein and glycolipid carbohydrate groups. *J Biol Chem.* 269(2): 1394-401.
- Kitagawa H, Paulson J C, (1994):** Differential expression of five sialyltransferase genes in human tissues. *J Biol Chem.* 269(27): 17872-8.
- Kitagawa H, Mattei M G, Paulson J C, (1996):** Genomic organization and chromosomal mapping of the Gal beta 1,3GalNAc/Gal beta 1,4GlcNAc alpha 2,3-sialyltransferase. *J Biol Chem.* 271(2): 931-8.
- Kitazume-Kawaguchi S, Kabata S, Arita M, (2001):** Differential biosynthesis of polysialic or disialic acid Structure by ST8Sia II and ST8Sia IV. *J Biol Chem.* 276(19): 15696-703.
- Koles K, Irvine K D, Panin V M, (2004):** Functional characterization of Drosophila sialyltransferase. *J Biol Chem.* 279(6): 4346-57.
- Koles K, Repnikova E, Pavlova G, Korochkin L I, Panin V M, (2008):** Sialylation in protostomes: a perspective from Drosophila genetics and biochemistry. *Glycoconj J.* 26(3): 313-24.
- Koopman P, Gubbay J, Vivian N, Goodfellow P, Lovell-Badge R, (1991):** Male development of chromosomally female mice transgenic for Sry. *Nature.* 351(6322): 117-21.
- Kostadinova R M, Nawrocki A R, Frey F J, Frey B M, (2005):** Tumor necrosis factor alpha and phorbol 12-myristate-13-acetate down-regulate human 11beta-hydroxysteroid dehydrogenase type 2 through p50/p50 NF-kappaB homodimers and Egr-1. *Faseb J.* 19(6): 650-2.
- Krzewinski-Recchi M A, Julien S, Juliant S, Teinturier-Lelievre M, Samyn-Petit B, Montiel M D, Mir A M, Cerutti M, Harduin-Lepers A, Delannoy P, (2003):** Identification and functional expression of a second human beta-galactoside alpha2,6-sialyltransferase, ST6Gal II. *Eur J Biochem.* 270(5): 950-61.

- Kudrycki K, Stein-Izsak C, Behn C, Grillo M, Akeson R, Margolis F L, (1993):** Olf-1-binding site: characterization of an olfactory neuron-specific promoter motif. *Mol Cell Biol.* 13(5): 3002-14.
- Kudrycki K E, Buiakova O, Tarozzo G, Grillo M, Walters E, Margolis F L, (1998):** Effects of mutation of the Olf-1 motif on transgene expression in olfactory receptor neurons. *J Neurosci Res.* 52(2): 159-72.
- Lai T, Jabaudon D, Molyneaux B J, Azim E, Arlotta P, Menezes J R, Macklis J D, (2008):** SOX5 controls the sequential generation of distinct corticofugal neuron subtypes. *Neuron.* 57(2): 232-47.
- Lance P, Lau K M, Lau J T, (1989):** Isolation and characterization of a partial cDNA for a human sialyltransferase. *Biochem Biophys Res Commun.* 164(1): 225-32.
- Lane M A, Bailey S J, (2005):** Role of retinoid signalling in the adult brain. *Prog Neurobiol.* 75(4): 275-93.
- Laporte B, Gonzalez-Hilarion S, Maftah A, Petit J M, (2009):** The second bovine beta-galactoside-alpha2,6-sialyltransferase (ST6Gal II): genomic organization and stimulation of its in vitro expression by IL-6 in bovine mammary epithelial cells. *Glycobiology.* 19(10): 1082-93.
- Law C L, Sidorenko S P, Clark E A, (1994):** Regulation of lymphocyte activation by the cell-surface molecule CD22. *Immunol Today.* 15(9): 442-9.
- Lee V M, Andrews P W, (1986):** Differentiation of NTERA-2 clonal human embryonal carcinoma cells into neurons involves the induction of all three neurofilament proteins. *J Neurosci.* 6(2): 514-21.
- Lee Y C, Kim Y J, Lee K Y, Kim K S, Kim B U, Kim H N, Kim C H, Do S I, (1998):** Cloning and expression of cDNA for a human Sia alpha 2,3Gal beta 1, 4GlcNA:alpha 2,8-sialyltransferase (hST8Sia III). *Arch Biochem Biophys.* 360(1): 41-6.
- Lee Y C, Kaufmann M, Kitazume-Kawaguchi S, Kono M, Takashima S, Kurosawa N, Liu H, Pircher H, Tsuji S, (1999):** Molecular cloning and functional expression of two members of mouse NeuAcalpha2,3Galbeta1,3GalNAc GalNAcalpha2,6-sialyltransferase family, ST6GalNAc III and IV. *J Biol Chem.* 274(17): 11958-67.
- Leone D P, Srinivasan K, Chen B, Alcamo E, McConnell S K, (2008):** The determination of projection neuron identity in the developing cerebral cortex. *Curr Opin Neurobiol.* 18(1): 28-35.
- Lin S, Kemmer W, Grigull S, Schlag P M, (2002):** Cell surface alpha 2,6 sialylation affects adhesion of breast carcinoma cells. *Exp Cell Res.* 276(1): 101-10.
- Lo N W, Lau J T, (1996):** Transcription of the beta-galactoside alpha 2,6-sialyltransferase gene in B lymphocytes is directed by a separate and distinct promoter. *Glycobiology.* 6(3): 271-9.

- Lo N W, Lau J T, (1999):** Transcription of the beta-galactoside alpha2,6-sialyltransferase gene (SIAT1) in B-lymphocytes: cell type-specific expression correlates with presence of the divergent 5'-untranslated sequence. *Glycobiology*. 9(9): 907-14.
- Lunyak V V, Rosenfeld M G, (2005):** No rest for REST: REST/NRSF regulation of neurogenesis. *Cell*. 121(4): 499-501.
- Maden M, (2002):** Retinoic acid and limb regeneration--a personal view. *Int J Dev Biol*. 46(7): 883-6.
- McCaffery P J, Adams J, Maden M, Rosa-Molinar E, (2003):** Too much of a good thing: retinoic acid as an endogenous regulator of neural differentiation and exogenous teratogen. *Eur J Neurosci*. 18(3): 457-72.
- Meffert M K, Baltimore D, (2005):** Physiological functions for brain NF-kappaB. *Trends Neurosci*. 28(1): 37-43.
- Mercier D, Wierinckx A, Oulmouden A, Gallet P F, Palcic M M, Harduin-Lepers A, Delannoy P, Petit J M, Leveziel H, Julien R, (1999):** Molecular cloning, expression and exon/intron organization of the bovine beta-galactoside alpha2,6-sialyltransferase gene. *Glycobiology*. 9(9): 851-63.
- Mey J, McCaffery P, (2004):** Retinoic acid signaling in the nervous system of adult vertebrates. *Neuroscientist*. 10(5): 409-21.
- Miyagi T, Tsuiki S, (1982):** Purification and characterization of beta-galactoside (alpha 2 leads to 6)sialyltransferase from rat liver and hepatomas. *Eur J Biochem*. 126(2): 253-61.
- Morris H R, Dell A, Easton R L, Panico M, Koistinen H, Koistinen R, Oehninger S, Patankar M S, Seppala M, Clark G F, (1996):** Gender-specific glycosylation of human glycodelin affects its contraceptive activity. *J Biol Chem*. 271(50): 32159-67.
- Nairn A V, York W S, Harris K, Hall E M, Pierce J M, Moremen K W, (2008):** Regulation of glycan structures in animal tissues: transcript profiling of glycan-related genes. *J Biol Chem*. 283(25): 17298-313.
- Nakayama J, Fukuda M N, Fredette B, Ranscht B, Fukuda M, (1995):** Expression cloning of a human polysialyltransferase that forms the polysialylated neural cell adhesion molecule present in embryonic brain. *Proc Natl Acad Sci U S A*. 92(15): 7031-5.
- Nara K, Watanabe Y, Maruyama K, Kasahara K, Nagai Y, Sanai Y, (1994):** Expression cloning of a CMP-NeuAc:NeuAc alpha 2-3Gal beta 1-4Glc beta 1-1'Cer alpha 2,8-sialyltransferase (GD3 synthase) from human melanoma cells. *Proc Natl Acad Sci U S A*. 91(17): 7952-6.
- Nemansky M, Van den Eijnden D H, (1992):** Bovine colostrum CMP-NeuAc:Gal beta(1-->4)GlcNAc-R alpha(2-->6)-sialyltransferase is involved in the synthesis of the terminal NeuAc alpha(2-->6)GalNAc beta(1-->4)GlcNAc sequence occurring on N-linked glycans of bovine milk glycoproteins. *Biochem J*. 287 (Pt 1): 311-6.

- Nimmerjahn F, Ravetch J V**, (2008): Anti-inflammatory actions of intravenous immunoglobulin. *Annu Rev Immunol*. 26: 513-33.
- Nitschke L**, (2005): The role of CD22 and other inhibitory co-receptors in B-cell activation. *Curr Opin Immunol*. 17(3): 290-7.
- O'Hanlon T P, Lau K M, Wang X C, Lau J T**, (1989): Tissue-specific expression of beta-galactoside alpha-2,6-sialyltransferase. Transcript heterogeneity predicts a divergent polypeptide. *J Biol Chem*. 264(29): 17389-94.
- O'Keefe T L, Williams G T, Davies S L, Neuberger M S**, (1996): Hyperresponsive B cells in CD22-deficient mice. *Science*. 274(5288): 798-801.
- Oehninger S, Coddington C C, Hodgen G D, Seppala M**, (1995): Factors affecting fertilization: endometrial placental protein 14 reduces the capacity of human spermatozoa to bind to the human zona pellucida. *Fertil Steril*. 63(2): 377-83.
- Okajima T, Fukumoto S, Ito H, Kiso M, Hirabayashi Y, Urano T, Furukawa K**, (1999): Molecular cloning of brain-specific GD1alpha synthase (ST6GalNAc V) containing CAG/Glutamine repeats. *J Biol Chem*. 274(43): 30557-62.
- Okajima T, Chen H H, Ito H, Kiso M, Tai T, Furukawa K, Urano T, Furukawa K**, (2000): Molecular cloning and expression of mouse GD1alpha/GT1aalpha/GQ1balpha synthase (ST6GalNAc VI) gene. *J Biol Chem*. 275(10): 6717-23.
- Okamoto N, Uchida A, Takakura K, Kariya Y, Kanzaki H, Riittinen L, Koistinen R, Seppala M, Mori T**, (1991): Suppression by human placental protein 14 of natural killer cell activity. *Am J Reprod Immunol*. 26(4): 137-42.
- Otipoby K L, Andersson K B, Draves K E, Klaus S J, Farr A G, Kerner J D, Perlmutter R M, Law C L, Clark E A**, (1996): CD22 regulates thymus-independent responses and the lifespan of B cells. *Nature*. 384(6610): 634-7.
- Pahlman S, Ruusala A I, Abrahamsson L, Mattsson M E, Esscher T**, (1984): Retinoic acid-induced differentiation of cultured human neuroblastoma cells: a comparison with phorbol ester-induced differentiation. *Cell Differ*. 14(2): 135-44.
- Park E I, Manzella S M, Baenziger J U**, (2003): Rapid clearance of sialylated glycoproteins by the asialoglycoprotein receptor. *J Biol Chem*. 278(7): 4597-602.
- Park E I, Mi Y, Unverzagt C, Gabius H J, Baenziger J U**, (2005): The asialoglycoprotein receptor clears glycoconjugates terminating with sialic acid alpha 2,6GalNAc. *Proc Natl Acad Sci U S A*. 102(47): 17125-9.
- Patel R Y, Balaji P V**, (2006): Identification of linkage-specific sequence motifs in sialyltransferases. *Glycobiology*. 16(2): 108-16.
- Paulson J C, Rearick J I, Hill R L**, (1977): Enzymatic properties of beta-D-galactoside alpha2 leads to 6 sialyltransferase from bovine colostrum. *J Biol Chem*. 252(7): 2363-71.

- Paulson J C, Weinstein J, Schauer A, (1989):** Tissue-specific expression of sialyltransferases. *J Biol Chem.* 264(19): 10931-4.
- Pierce J G, Parsons T F, (1981):** Glycoprotein hormones: structure and function. *Annu Rev Biochem.* 50: 465-95.
- Pleasure S J, Page C, Lee V M, (1992):** Pure, postmitotic, polarized human neurons derived from Ntera 2 cells provide a system for expressing exogenous proteins in terminally differentiated neurons. *J Neurosci.* 12(5): 1802-15.
- Pleasure S J, Lee V M, (1993):** Ntera 2 cells: a human cell line which displays characteristics expected of a human committed neuronal progenitor cell. *J Neurosci Res.* 35(6): 585-602.
- Powell L D, Sgroi D, Sjoberg E R, Stamenkovic I, Varki A, (1993):** Natural ligands of the B cell adhesion molecule CD22 beta carry N-linked oligosaccharides with alpha-2,6-linked sialic acids that are required for recognition. *J Biol Chem.* 268(10): 7019-27.
- Qian Z, Gilbert M E, Colicos M A, Kandel E R, Kuhl D, (1993):** Tissue-plasminogen activator is induced as an immediate-early gene during seizure, kindling and long-term potentiation. *Nature.* 361(6411): 453-7.
- Rachmilewitz J, Riely G J, Tykocinski M L, (1999):** Placental protein 14 functions as a direct T-cell inhibitor. *Cell Immunol.* 191(1): 26-33.
- Rao M N, Lakshman M R, (1999):** Chronic ethanol consumption leads to destabilization of rat liver beta-galactoside alpha2,6-sialyltransferase mRNA. *Metabolism.* 48(6): 797-803.
- Rash B G, Grove E A, (2006):** Area and layer patterning in the developing cerebral cortex. *Curr Opin Neurobiol.* 16(1): 25-34.
- Rastinejad F, (2001):** Retinoid X receptor and its partners in the nuclear receptor family. *Curr Opin Struct Biol.* 11(1): 33-8.
- Recchi M A, Hebbar M, Hornez L, Harduin-Lepers A, Peyrat J P, Delannoy P, (1998):** Multiplex reverse transcription polymerase chain reaction assessment of sialyltransferase expression in human breast cancer. *Cancer Res.* 58(18): 4066-70.
- Reddy V B, Garramone A J, Sasak H, Wei C M, Watkins P, Galli J, Hsiung N, (1987):** Expression of human uterine tissue-type plasminogen activator in mouse cells using BPV vectors. *DNA.* 6(5): 461-72.
- Rijken D C, Wijngaards G, Zaal-de Jong M, Welbergen J, (1979):** Purification and partial characterization of plasminogen activator from human uterine tissue. *Biochim Biophys Acta.* 580(1): 140-53.
- Rohfritsch P F, Joosten J A, Krzewinski-Recchi M A, Harduin-Lepers A, Laporte B, Juliant S, Cerutti M, Delannoy P, Vliegenthart J F, Kamerling J P, (2006):** Probing the substrate specificity of four different sialyltransferases using synthetic beta-D-Galp-(1-->4)-beta-D-GlcpNAc-(1-->2)-alpha-D-Manp-(1-->O) (CH(2))7CH3 analogues general activating

- effect of replacing N-acetylglucosamine by N-propionylglucosamine. *Biochim Biophys Acta*. 1760(4): 685-92.
- Rothbard J B, Brackenbury R, Cunningham B A, Edelman G M, (1982):** Differences in the carbohydrate structures of neural cell-adhesion molecules from adult and embryonic chicken brains. *J Biol Chem*. 257(18): 11064-9.
- Rutishauser U, (2008):** Polysialic acid in the plasticity of the developing and adult vertebrate nervous system. *Nat Rev Neurosci*. 9(1): 26-35.
- Safe S, Kim K, (2004):** Nuclear receptor-mediated transactivation through interaction with Sp proteins. *Prog Nucleic Acid Res Mol Biol*. 77: 1-36.
- Samyn-Petit B, Krzewinski-Recchi M A, Steelant W F, Delannoy P, Harduin-Lepers A, (2000):** Molecular cloning and functional expression of human ST6GalNAc II. Molecular expression in various human cultured cells. *Biochim Biophys Acta*. 1474(2): 201-11.
- Sappino A P, Madani R, Huarte J, Belin D, Kiss J Z, Wohlwend A, Vassalli J D, (1993):** Extracellular proteolysis in the adult murine brain. *J Clin Invest*. 92(2): 679-85.
- Sasaki K, Watanabe E, Kawashima K, Sekine S, Dohi T, Oshima M, Hanai N, Nishi T, Hasegawa M, (1993):** Expression cloning of a novel Gal beta (1-3/1-4) GlcNAc alpha 2,3-sialyltransferase using lectin resistance selection. *J Biol Chem*. 268(30): 22782-7.
- Sasaki K, Kurata K, Kojima N, Kurosawa N, Ohta S, Hanai N, Tsuji S, Nishi T, (1994):** Expression cloning of a GM3-specific alpha-2,8-sialyltransferase (GD3 synthase). *J Biol Chem*. 269(22): 15950-6.
- Sata T, Roth J, Zuber C, Stamm B, Heitz P U, (1991):** Expression of alpha 2,6-linked sialic acid residues in neoplastic but not in normal human colonic mucosa. A lectin-gold cytochemical study with Sambucus nigra and Maackia amurensis lectins. *Am J Pathol*. 139(6): 1435-48.
- Sato S, Miller A S, Inaoki M, Bock C B, Jansen P J, Tang M L, Tedder T F, (1996):** CD22 is both a positive and negative regulator of B lymphocyte antigen receptor signal transduction: altered signaling in CD22-deficient mice. *Immunity*. 5(6): 551-62.
- Seales E C, Jurado G A, Brunson B A, Wakefield J K, Frost A R, Bellis S L, (2005):** Hypersialylation of beta1 integrins, observed in colon adenocarcinoma, may contribute to cancer progression by up-regulating cell motility. *Cancer Res*. 65(11): 4645-52.
- Seeds N W, Haffke S, Christensen K, Schoonmaker J, (1990):** Cerebellar granule cell migration involves proteolysis. *Adv Exp Med Biol*. 265: 169-78.
- Semel A C, Seales E C, Singhal A, Eklund E A, Colley K J, Bellis S L, (2002):** Hyposialylation of integrins stimulates the activity of myeloid fibronectin receptors. *J Biol Chem*. 277(36): 32830-6.

- Seppala M, Koskimies A I, Tenhunen A, Rutanen E M, Sjoberg J, Koistinen R, Julkunen M, Wahlstrom T, (1985):** Pregnancy proteins in seminal plasma, seminal vesicles, preovulatory follicular fluid, and ovary. *Ann N Y Acad Sci.* 442: 212-26.
- Seppala M, Taylor R N, Koistinen H, Koistinen R, Milgrom E, (2002):** Glycodelin: a major lipocalin protein of the reproductive axis with diverse actions in cell recognition and differentiation. *Endocr Rev.* 23(4): 401-30.
- Sgroi D, Nocks A, Stamenkovic I, (1996):** A single N-linked glycosylation site is implicated in the regulation of ligand recognition by the I-type lectins CD22 and CD33. *J Biol Chem.* 271(31): 18803-9.
- Shinya K, Ebina M, Yamada S, Ono M, Kasai N, Kawaoka Y, (2006):** Avian flu: influenza virus receptors in the human airway. *Nature.* 440(7083): 435-6.
- Skacel P O, Edwards A J, Harrison C T, Watkins W M, (1991):** Enzymic control of the expression of the X determinant (CD15) in human myeloid cells during maturation: the regulatory role of 6-sialyltransferase. *Blood.* 78(6): 1452-60.
- Smale S T, Baltimore D, (1989):** The "initiator" as a transcription control element. *Cell.* 57(1): 103-13.
- Stamenkovic I, Sgroi D, Aruffo A, Sy M S, Anderson T, (1991):** The B lymphocyte adhesion molecule CD22 interacts with leukocyte common antigen CD45RO on T cells and alpha 2-6 sialyltransferase, CD75, on B cells. *Cell.* 66(6): 1133-44.
- Steirer L M, Park E I, Townsend R R, Baenziger J U, (2009):** The asialoglycoprotein receptor regulates levels of plasma glycoproteins terminating with sialic acid alpha2,6-galactose. *J Biol Chem.* 284(6): 3777-83.
- Stevens J, Blixt O, Tumpey T M, Taubenberger J K, Paulson J C, Wilson I A, (2006):** Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science.* 312(5772): 404-10.
- Sur M, Rubenstein J L, (2005):** Patterning and plasticity of the cerebral cortex. *Science.* 310(5749): 805-10.
- Suzuki Y, (2005):** Sialobiology of influenza: molecular mechanism of host range variation of influenza viruses. *Biol Pharm Bull.* 28(3): 399-408.
- Svensson E C, Soreghan B, Paulson J C, (1990):** Organization of the beta-galactoside alpha 2,6-sialyltransferase gene. Evidence for the transcriptional regulation of terminal glycosylation. *J Biol Chem.* 265(34): 20863-8.
- Takashima S, Tsuji S, Tsujimoto M, (2002):** Characterization of the second type of human beta-galactoside alpha 2,6-sialyltransferase (ST6Gal II), which sialylates Galbeta 1,4GlcNAc structures on oligosaccharides preferentially. Genomic analysis of human sialyltransferase genes. *J Biol Chem.* 277(48): 45719-28.

- Takashima S, Tsuji S, Tsujimoto M, (2003):** Comparison of the enzymatic properties of mouse beta-galactoside alpha2,6-sialyltransferases, ST6Gal I and II. *J Biochem.* 134(2): 287-96.
- Taniguchi A, Matsumoto K, (1998):** Down-regulation of human sialyltransferase gene expression during in vitro human keratinocyte cell line differentiation. *Biochem Biophys Res Commun.* 243(1): 177-83.
- Taniguchi A, Hasegawa Y, Higai K, Matsumoto K, (1999):** Down-regulation of human Galbeta1,3GalNAc/Galbeta1,4GlcNAc alpha2,3-sialyltransferase (hST3Gal IV) gene during differentiation of the HL-60 cell line. *Mol Cell Biol Res Commun.* 2(3): 197-201.
- Taniguchi A, Kaneta R, Morishita K, Matsumoto K, (2001):** Gene structure and transcriptional regulation of human Gal beta1,4(3) GlcNAc alpha2,3-sialyltransferase VI (hST3Gal VI) gene in prostate cancer cell line. *Biochem Biophys Res Commun.* 287(5): 1148-56.
- Taniguchi A, Morishima T, Tsujita Y, Matsumoto Y, Matsumoto K, (2003):** Genomic structure, expression, and transcriptional regulation of human Gal beta 1,3 GalNAc alpha 2,3-sialyltransferase gene. *Biochem Biophys Res Commun.* 300(2): 570-6.
- Taniguchi A, (2008):** Promoter structure and transcriptional regulation of human beta-galactoside alpha2, 3-sialyltransferase genes. *Curr Drug Targets.* 9(4): 310-6.
- Teinturier-Lelievre M, Julien S, Juliant S, Guerardel Y, Duonor-Cerutti M, Delannoy P, Harduin-Lepers A, (2005):** Molecular cloning and expression of a human hST8Sia VI (alpha2,8-sialyltransferase) responsible for the synthesis of the diSia motif on O-glycosylproteins. *Biochem J.* 392(Pt 3): 665-74.
- Tian E, Ten Hagen K G, (2009):** Recent insights into the biological roles of mucin-type O-glycosylation. *Glycoconj J.* 26(3): 325-34.
- Tsuchida A, Okajima T, Furukawa K, Ando T, Ishida H, Yoshida A, Nakamura Y, Kannagi R, Kiso M, Furukawa K, (2003):** Synthesis of disialyl Lewis a (Le(a)) structure in colon cancer cell lines by a sialyltransferase, ST6GalNAc VI, responsible for the synthesis of alpha-series gangliosides. *J Biol Chem.* 278(25): 22787-94.
- Tsuchida A, Ogiso M, Nakamura Y, Kiso M, Furukawa K, Furukawa K, (2005):** Molecular cloning and expression of human ST6GalNAc III: restricted tissue distribution and substrate specificity. *J Biochem.* 138(3): 237-43.
- Valente V, Teixeira S A, Neder L, Okamoto O K, Oba-Shinjo S M, Marie S K, Scrideli C A, Paco-Larson M L, Carlotti C G, Jr., (2009):** Selection of suitable housekeeping genes for expression analysis in glioblastoma using quantitative RT-PCR. *BMC Mol Biol.* 10: 17.
- Varki A, (1993):** Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology.* 3(2): 97-130.
- Varki A, (2008):** Sialic acids in human health and disease. *Trends Mol Med.* 14(8): 351-60.

- Vierbuchen M J, Fruechtnicht W, Brackrock S, Krause K T, Zienkiewicz T J, (1995):** Quantitative lectin-histochemical and immunohistochemical studies on the occurrence of alpha(2,3)- and alpha(2,6)-linked sialic acid residues in colorectal carcinomas. Relation to clinicopathologic features. *Cancer*. 76(5): 727-35.
- Walker J A, Smith K G, (2008):** CD22: an inhibitory enigma. *Immunology*. 123(3): 314-25.
- Wallen P, Pohl G, Bergsdorf N, Ranby M, Ny T, Jornvall H, (1983):** Purification and characterization of a melanoma cell plasminogen activator. *Eur J Biochem*. 132(3): 681-6.
- Wang P H, Li Y F, Juang C M, Lee Y R, Chao H T, Tsai Y C, Yuan C C, (2001):** Altered mRNA expression of sialyltransferase in squamous cell carcinomas of the cervix. *Gynecol Oncol*. 83(1): 121-7.
- Wang S S, Tsai R Y, Reed R R, (1997):** The characterization of the Olf-1/EBF-like HLH transcription factor family: implications in olfactory gene regulation and neuronal development. *J Neurosci*. 17(11): 4149-58.
- Wang X, Vertino A, Eddy R L, Byers M G, Jani-Sait S N, Shows T B, Lau J T, (1993):** Chromosome mapping and organization of the human beta-galactoside alpha 2,6-sialyltransferase gene. Differential and cell-type specific usage of upstream exon sequences in B-lymphoblastoid cells. *J Biol Chem*. 268(6): 4355-61.
- Weinstein J, de Souza-e-Silva U, Paulson J C, (1982):** Purification of a Gal beta 1 to 4GlcNAc alpha 2 to 6 sialyltransferase and a Gal beta 1 to 3(4)GlcNAc alpha 2 to 3 sialyltransferase to homogeneity from rat liver. *J Biol Chem*. 257(22): 13835-44.
- Weisshaar G, Hiyama J, Renwick A G, Nimitz M, (1991):** NMR investigations of the N-linked oligosaccharides at individual glycosylation sites of human lutropin. *Eur J Biochem*. 195(1): 257-68.
- Wen D X, Svensson E C, Paulson J C, (1992):** Tissue-specific alternative splicing of the beta-galactoside alpha 2,6-sialyltransferase gene. *J Biol Chem*. 267(4): 2512-8.
- Wide L, Eriksson K, Sluss P M, Hall J E, (2009):** Serum half-life of pituitary gonadotropins is decreased by sulfonation and increased by sialylation in women. *J Clin Endocrinol Metab*. 94(3): 958-64.
- Yamashita K, Fukushima K, Sakiyama T, Murata F, Kuroki M, Matsuoka Y, (1995):** Expression of Sia alpha 2-->6Gal beta 1-->4GlcNAc residues on sugar chains of glycoproteins including carcinoembryonic antigens in human colon adenocarcinoma: applications of *Trichosanthes japonica* agglutinin I for early diagnosis. *Cancer Res*. 55(8): 1675-9.
- Yan S B, Chao Y B, van Halbeek H, (1993):** Novel Asn-linked oligosaccharides terminating in GalNAc beta (1-->4)[Fuc alpha (1-->3)]GlcNAc beta (1-->.) are present in recombinant human protein C expressed in human kidney 293 cells. *Glycobiology*. 3(6): 597-608.
- Yaniv E, Borovsky Z, Mishan-Eisenberg G, Rachmilewitz J, (2003):** Placental protein 14 regulates selective B cell responses. *Cell Immunol*. 222(2): 156-63.

Yao L, Korteweg C, Hsueh W, Gu J, (2008): Avian influenza receptor expression in H5N1-infected and noninfected human tissues. *FASEB J.* 22(3): 733-40.

Yasukawa Z, Sato C, Kitajima K, (2005): Inflammation-dependent changes in alpha2,3-, alpha2,6-, and alpha2,8-sialic acid glycotopes on serum glycoproteins in mice. *Glycobiology.* 15(9): 827-37.

Yeung W S, Lee K F, Koistinen R, Koistinen H, Seppala M, Ho P C, Chiu P C, (2006): Roles of glycodelin in modulating sperm function. *Mol Cell Endocrinol.* 250(1-2): 149-56.

Yoshida Y, Kojima N, Kurosawa N, Hamamoto T, Tsuji S, (1995): Molecular cloning of Sia alpha 2,3Gal beta 1,4GlcNAc alpha 2,8-sialyltransferase from mouse brain. *J Biol Chem.* 270(24): 14628-33.

Yoshida Y, Kurosawa N, Kanematsu T, Kojima N, Tsuji S, (1996): Genomic structure and promoter activity of the mouse polysialic acid synthase gene (mST8Sia II). Brain-specific expression from a TATA-less GC-rich sequence. *J Biol Chem.* 271(47): 30167-73.

Zhang X, Ding L, Sandford A J, (2005): Selection of reference genes for gene expression studies in human neutrophils by real-time PCR. *BMC Mol Biol.* 6(1): 4.

Zigova T, Barroso L F, Willing A E, Saporta S, McGrogan M P, Freeman T B, Sanberg P R, (2000): Dopaminergic phenotype of hNT cells in vitro. *Brain Res Dev Brain Res.* 122(1): 87-90

Annexes

Le cerveau, le système nerveux central

Anatomiquement, le cerveau est l'organe central du corps humain localisé à l'intérieur de la boîte crânienne. Ce terme englobe l'ensemble du système nerveux central situé à l'intérieur de la tête. Chez les invertébrés, le terme « cerveau » peut également être employé pour désigner les centres nerveux. Cet organe assure la coordination de l'ensemble de l'organisme. Chez les vertébrés, il contrôle le comportement, la plupart des mouvements et l'homéostasie des fonctions internes comme la pression artérielle, la température interne et le rythme cardiaque.

D'un point de vue neuroanatomique, le cerveau est l'ensemble des structures nerveuses formant le diencephale et le télencéphale. Le diencephale comprend essentiellement le thalamus et le l'hypothalamus. Le télencéphale est constitué des hémisphères cérébraux (cortex cérébral, substance blanche et structures sous-corticales). Communément, et de manière imprécise, le mot « cerveau » est employé pour désigner l'ensemble de l'encéphale. L'encéphale, qui est la partie du système nerveux contenu dans la boîte crânienne, est en réalité composé du cerveau (diencephale et télencéphale), du cervelet et du tronc cérébral.

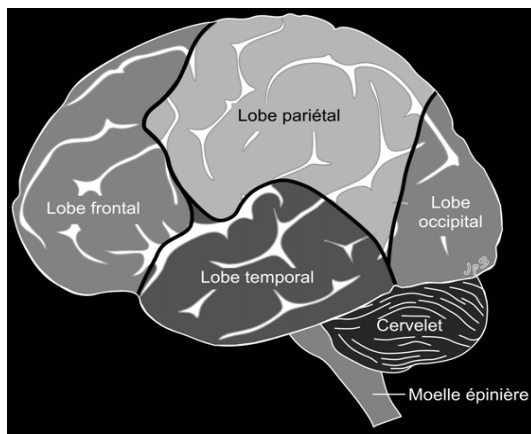
Le cortex cérébral

Le cortex cérébral désigne la substance grise entourant les deux hémisphères cérébraux. D'un point de vue de l'évolution, le terme néocortex désigne la plus récente évolution du cortex cérébral. Le néocortex correspond aux couches les plus externes du cortex cérébral. Chez l'Homme, il est divisé en six couches. Une septième couche existe de façon transitoire durant l'embryogenèse et disparaît rapidement avec la maturation cérébrale (Figure 48) :

1. La couche moléculaire. Elle contient des dendrites et des axones. Les neurones des couches internes y projettent leurs dendrites, courtes et orientées perpendiculairement à la surface du cortex. Les axones, longs, sont orientés parallèlement à la surface du cortex.
2. La couche granulaire externe. Constituée de neurones de type granulaire, elle reçoit les afférences d'autres aires du cortex, on parle de « connexions cortico-corticales afférentes ».
3. La couche pyramidale externe. Cette couche est constituée de neurones de type pyramidal et émet des connexions vers d'autres zones du cortex, on parle de « connexions cortico-corticales efférentes ».

4. La couche granulaire interne. Elle contient des neurones étoilés et pyramidaux. C'est au niveau de cette zone que les informations extérieures au cortex entrent. Elle reçoit aussi les afférences provenant de l'hémisphère cérébral opposé.
5. La couche pyramidale externe. De cette couche sont envoyés des connexions efférentes sortant du cortex. Les neurones qui innervent les motoneurones partent par exemple de cette couche.
6. La couche polymorphe. C'est la couche la plus interne à l'état adulte, elle envoie des prolongements axonaux en direction du thalamus permettant ainsi une rétroaction sur les entrées du cortex cérébral.

A



B

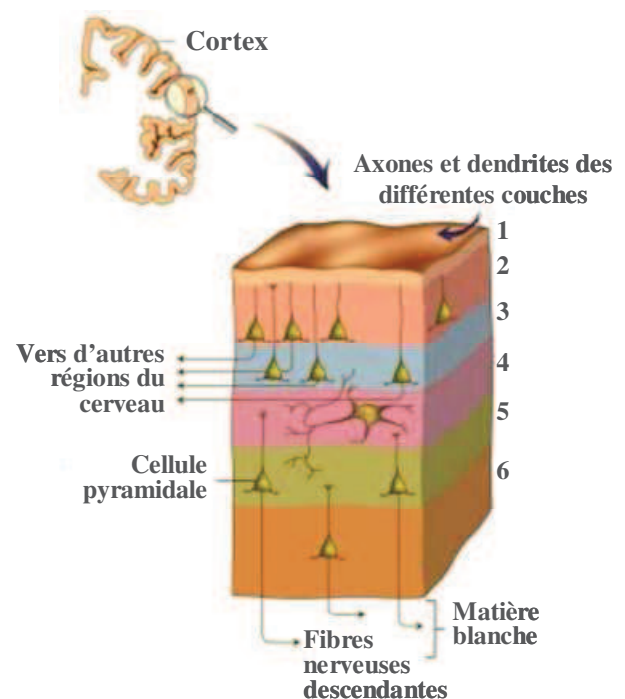


Figure 48 : Schémas des différentes parties de l'encéphale et des différentes couches du cortex

A, l'encéphale. Les quatre lobes cérébraux, le cervelet et la base de la moelle épinière. **B**, les six couches du cortex humain et leurs organisations. (<http://www.blogg.org/blog-50595-date-2006-11.html>).

Dans ces différentes couches, les neurones sont organisés en unités fonctionnelles organisées perpendiculairement à la surface du cortex. Chacune de ces entités fonctionnelles assure une fonction précise. Histologiquement, ces couches ne peuvent être distinguées.

Le cortex cérébral est divisé en quatre aires primaires que sont le lobe frontal, pariétal, occipital et temporal. Chacun de ces lobes assure des fonctions bien définies détaillées ci-dessous. Par ailleurs, le cortex cérébral est aussi constitué d'aires d'associations

situées autour des aires primaires. Ces aires secondaires permettent de coordonner les fonctions de base des aires primaires. La perception et la compréhension d'images et de sons font par exemple intervenir les aires d'association. Enfin, il existe également des aires tertiaires pariéto-temporale et frontale. Ces aires reçoivent des informations de plusieurs aires secondaires et remplissent des fonctions dites « supérieures » comme le langage, les fonctions intellectuelles, le contrôle du comportement et l'organisation dans l'espace.

Le lobe frontal :

Ce lobe est responsable du comportement de l'individu. Il contient également l'aire motrice qui commande tous les mouvements volontaires.

Le lobe pariétal :

Ce lobe est situé en haut de chaque hémisphère cérébral. Cette zone permet à un individu d'avoir des repères spatiaux et de contrôler les gestes et les mouvements. Le lobe pariétal comprend essentiellement l'aire sensitive, la partie du cortex qui assure la réception des informations en provenance de la périphérie du corps.

Le lobe temporal :

Situé latéralement en dessous du lobe pariétal, il participe aux différentes fonctions du cerveau telles que le goût, l'audition, l'olfaction, la mémoire, le langage. Ce lobe intervient enfin dans le fonctionnement végétatif, c'est-à-dire le fonctionnement automatique des viscères.

Le lobe occipital :

Il se situe à l'arrière du cortex et contient l'aire visuelle et reçoit donc les informations en provenance des globes oculaires transmises par la rétine.

Le cervelet

Le cervelet est situé à l'arrière du crâne sous le cerveau neuroanatomique (diencephale et télencéphale), directement sous le lobe occipital (Figure 48). A la manière du cortex cérébral, il est subdivisé en plusieurs lobes : un lobe médian, le vermis cérébelleux, et deux lobes latéraux, les lobes cérébelleux.

Le cervelet joue un rôle majeur dans la coordination des mouvements en général. Il régule les différents réflexes musculaires qui contrôlent les gestes, la station « debout » et les changements de postures.

Le cortex cérébelleux est composé de trois couches distinctes (Figure 49) :

1. La couche externe, ou moléculaire, contient les arborisations dendritiques des cellules de Purkinje provenant de la couche moyenne sous-jacente, les axones des cellules olivaires (fibres grimpantes) et granulaires (fibres parallèles) ainsi que des neurones étoilés et des neurones « en panier ».
2. La couche moyenne contient les corps cellulaires des neurones de Purkinje répartis en une seule assise.
3. La couche interne, ou granulaire, comprend en majorité les corps cellulaires de cellules de petites tailles (cellules granulaires). Les axones de ces neurones remontent dans la couche moléculaire et se divisent en T formant ainsi les fibres parallèles. Cette couche contient aussi les neurones de Golgi. Leurs dendrites ascendantes reçoivent des contacts de fibres parallèles et leurs dendrites descendantes reçoivent des contacts des fibres moussues. Leur axone est très ramifié et forme des contacts complexes avec les axones des cellules granulaires.

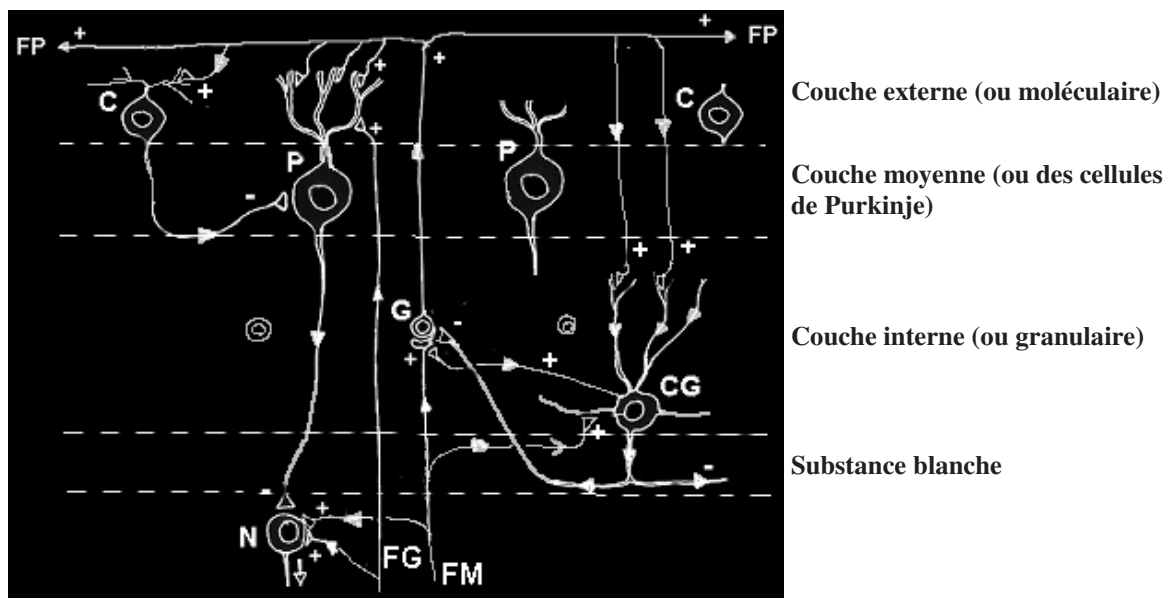


Figure 49 : Schéma fonctionnel des trois couches du cervelet et de ses connexions.

C : neurone en panier, CG : neurone de Golgi, FG : fibre grimpante, FM : fibre moussue, FP : fibre parallèle, G : neurone granulaire, N : neurone de noyau profond, P : neurone de Purkinje, + : stimulation, - : inhibition. (<http://www.inrp.fr/Acces/biotic/gpe/dossiers/mutcer/html/orgcerv.htm>).

En dessous de ces trois couches se trouve la substance blanche cérébelleuse. Dans ce centre médullaire du cervelet passent les fibres afférentes, moussues, grimpantes et efférentes, et s'y trouvent les « noyaux profonds ».

Le cervelet est une structure cérébrale riche en neurones puisqu'elle est constituée d'environ 50 % des neurones du système nerveux central alors qu'elle n'occupe, en volume, que 10 % de la boîte crânienne.

L'hippocampe

L'hippocampe se présente sous la forme d'une structure bilatérale (une partie dans chaque hémisphère cérébral) symétrique faisant partie du système limbique. Le système limbique regroupe les structures du cerveau jouant un rôle important dans le comportement et les émotions (agressivité, peur, plaisir...). Le système limbique est composé de l'amygdale, du Gyrus cingulaire, du fornix, de l'hypothalamus et de l'hippocampe (Figure 50). L'hippocampe est composé de deux structures, le *gyrux dentatus* (dentelé) et l'hippocampe à proprement parler, ou Corne d'Ammon. C'est une zone du cerveau primordial pour l'apprentissage et la mémoire à long terme.

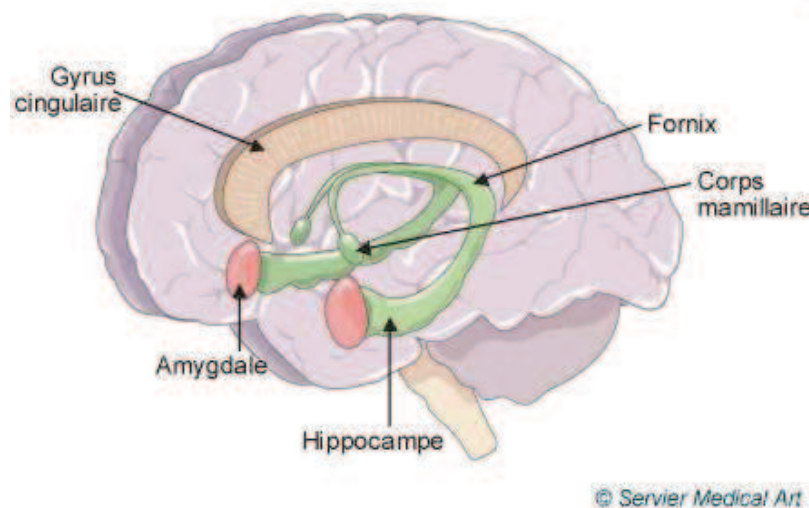


Figure 50 : Représentation schématique du système limbique

Histologiquement, l'hippocampe est constitué d'une seule couche de neurones. Il s'agit d'une structure corticale repliée sur elle-même et située sur la face médiane du lobe temporal. L'hippocampe possède par ailleurs une activité de neurogénèse. La Corne

d'Ammon est divisée en trois modules fonctionnels dont les rôles ne sont pas encore bien établis.

D'un point de vue fonctionnel, l'hippocampe participe à la construction de mémoires événementielles par l'établissement de liens entre des informations reçues des couches corticales éloignées les unes des autres et qui convergent vers la structure hippocampique. Ces « souvenirs » pourraient être restitués pendant les phases de sommeil pour un stockage définitif au niveau du cortex.

Développement embryonnaire du système nerveux central

Le système nerveux central apparaît très tôt (dès le 18^{ème} jour) au cours du développement. Avant le 18^{ème} jour, l'embryon est composé de trois couches cellulaires concentriques (embryon à trois feuillets) : l'ectoderme, le mésoderme et l'endoderme. L'ectoderme donnera notamment la peau et le système nerveux central et périphérique, le mésoderme se différenciera en muscles, os et système vasculaire, et l'endoderme donnera les viscères. Le système nerveux émerge à la suite de signaux envoyés du mésoderme vers l'ectoderme. Ce phénomène est appelé l'induction neurale et se traduit par la transformation d'une partie de l'ectoderme en tissu nerveux primitif : le neurectoderme. La neurulation est la période pendant laquelle se forme le neurectoderme sur la face dorsale de l'embryon.

Le neurectoderme évolue ensuite, sur une période d'environ une semaine, suivant quatre étapes pour donner les ébauches des systèmes nerveux périphériques (Figure 51) :

1. Plaque neurale. Une fois formé le neurectoderme s'épaissit en son centre pour former la plaque neurale qui est marquée d'un sillon visible sur toute la face dorsale de l'embryon.
2. Gouttière et crêtes neurales. La plaque neurale se creuse, s'invagine afin de former la gouttière neurale dont les bords prolifèrent et donnent les crêtes neurales.
3. Le tube neural primaire. Les crêtes neurales ferment la gouttière en se rejoignant. Les crêtes se séparent ensuite du tube nouvellement formé et forment un filet continu au dessus du tube neural.
4. Le tube neural secondaire. Une fois complètement fermé, le tube neural grossit et délimite en son centre une cavité qui donnera les cavités encéphaliques et le canal de l'épendyme. La région centrale du tube donnera la substance grise (ou se trouveront les corps cellulaires des neurones) et la paroi du tube donnera la substance blanche (les fibres nerveuses). Par ailleurs, le filet des crêtes neurales se segmente, migre et donnera par la suite le système nerveux périphérique (ganglions rachidiens et du système neurovégétatif).

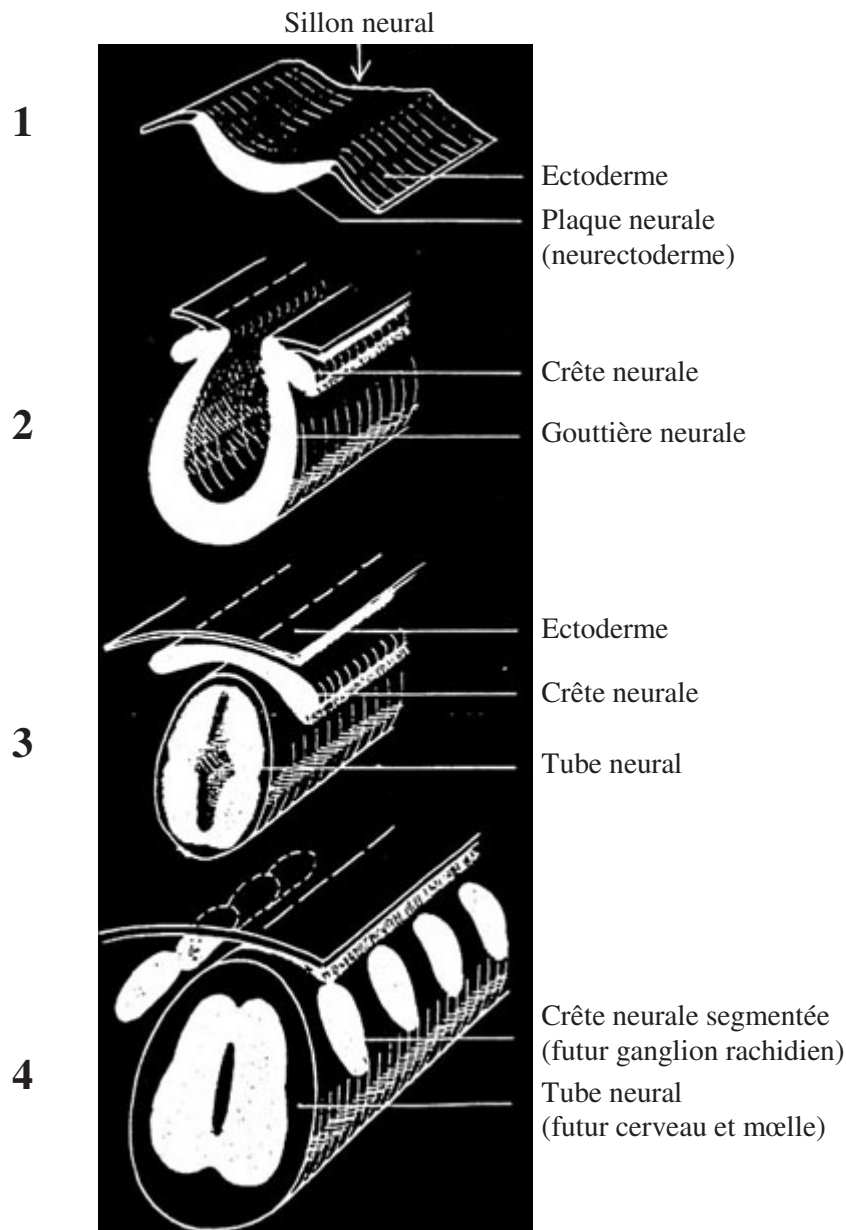


Figure 51 : Développement du neurectoderme et du tube neural

Les 4 étapes de la formation du tube neural. La face dorsale est située sur le dessus des dessins, la face antérieure est située vers l'avant des dessins. (<http://schwann.free.fr/neurogenese.html>).

Le neurectoderme est par ailleurs constitué de deux types de cellules :

1. Les neuroblastes, des neurones embryonnaires qui peuvent se diviser et se différencier en neurones matures incapables de divisions futures (à l'exception de ceux du bulbe olfactif et éventuellement de l'hippocampe).
2. Les précurseurs gliaux, qui donneront les astrocytes, les oligodendrocytes, les cellules microgliales, les cellules choroïdiennes et les cellules épendymaires dans le système nerveux central, et les cellules de Schwann dans le système nerveux périphérique.

La partie postérieure du tube neural va former la moelle épinière. La paroi du tube neural médullaire se subdivise en substance blanche périphérique et en substance grise centrale. La moelle épinière assure donc un rôle d'interface entre le système nerveux central et le système nerveux périphérique.

L'avant du tube neural (tube neural encéphalique) constituera le cerveau. Cette structure est amenée à se développer énormément chez l'Homme. Le tube neural encéphalique se transforme en un système ventriculaire et se divise initialement en trois ventricules : le proencéphale (à l'avant), le mésencéphale (au milieu) et le rhombencéphale (à l'arrière) (Figure 52A). La croissance de ces trois vésicules se fait par multiplication cellulaire classique jusqu'à atteindre une nouvelle étape où le proencéphale se divise en télencéphale et diencéphale, et où le rhombencéphale se segmente en métencéphale et myélocéphale. A ce stade, cinq ventricules sont visibles (Figure 52B) et délimitent le système cavitaire cérébral.

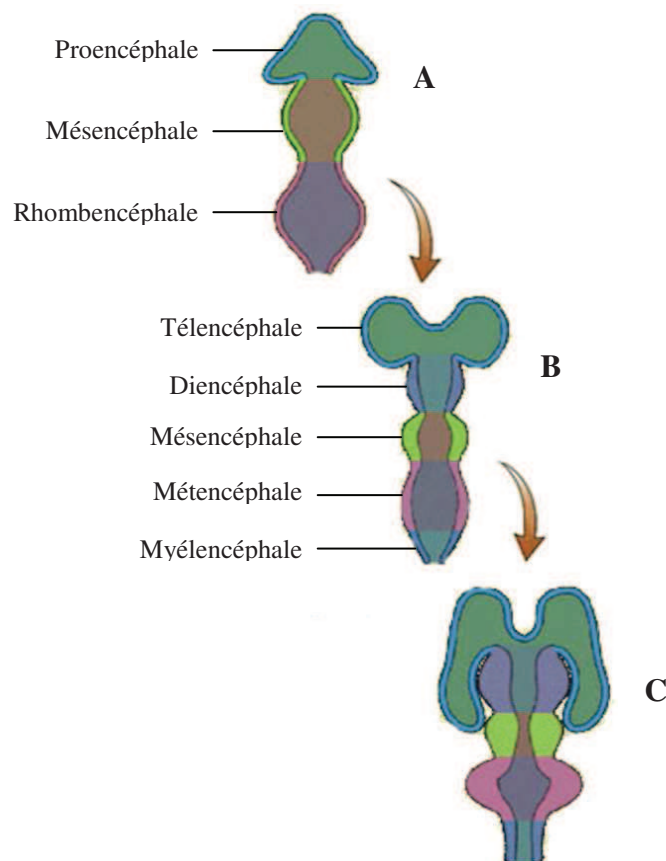


Figure 52 : Schéma du développement du tube neural

La face antérieure est située sur le dessus des dessins. **A**, stade trois ventricules. **B**, stade cinq ventricules, le télencéphale commence à bourgeonner. **C**, les futurs hémisphères cérébraux se développent. (http://lecerveau.mcgill.ca/flash/i/i_09/i_09_cr/i_09_cr_dev/i_09_cr_dev.html).

Le télencéphale se développe de façon importante en se repliant sur l'arrière et finit par recouvrir complètement le diencéphale. La différenciation ultérieure du télencéphale conduit à l'apparition de plusieurs structures que sont le cortex cérébral, le système limbique (dont l'hippocampe) et le striatum.

Le cortex continu à se développer en produisant les circonvolutions qu'on lui connaît de manière à augmenter sa surface sans augmenter son volume dans la future boîte crânienne. Il se divise également en deux hémisphères et quatre lobes cérébraux.

Le diencéphale donnera naissance aux rétines, pour les futurs appendices oculaires, le thalamus, l'hypothalamus, l'épiphyse et la neurohypophyse.

Le mésencéphale donnera les tubercules quadrijumeaux, du côté dorsal, qui sont des relais pour les informations auditives et visuelles, les pédoncules cérébraux, du côté ventral, qui sont des relais pour les informations afférentes ou efférentes du cervelet, et enfin les noyaux des nerfs crâniens III et IV.

Le métencéphale est à l'origine, notamment, du cervelet et des noyaux des nerfs crâniens V à VIII.

Enfin, le myélencéphale se développera pour donner, du côté ventral uniquement, les noyaux des nerfs crâniens IX à XII et le bulbe rachidien.

L'épithélium périphérique des ventricules contient les cellules souches à partir desquelles les neurones sont formés. Les cellules pro-neuronales se divisent dans un premier temps symétriquement avant de se multiplier assymétriquement. Après division, l'une des deux cellules migre à l'extérieur de la zone ventriculaire pour atteindre le cortex où elle se différencie en neurone. L'autre cellule continue de se diviser tout en restant dans la zone des ventricules. Les cellules gliales se multiplient selon un processus similaire, leurs cellules progénitrices sont cependant différentes de ceux des cellules neuronales.

Transcriptional regulation of the human *ST6GAL2* gene in cerebral cortex and neuronal cells

Sylvain Lehoux · Sophie Groux-Degroote · Aurélie Cazet · Claire-Marie Dhaenens · Claude-Alain Maurage · Marie-Laure Caillet-Boudin · Philippe Delannoy · Marie-Ange Krzewinski-Recchi

Received: 29 June 2009 / Revised: 26 August 2009 / Accepted: 1 September 2009
© Springer Science + Business Media, LLC 2009

Abstract The second human β -galactoside α -2,6-sialyltransferase (hST6Gal II) differs from hST6Gal I, the first member of ST6Gal family, in substrate specificity and tissue expression pattern. While *ST6GAL1* gene is expressed in almost all human tissues, *ST6GAL2* shows a restricted tissue-specific pattern of expression, mostly expressed in embryonic and adult brain. In order to understand the mechanisms involved in the transcriptional regulation of *ST6GAL2*, we first characterized the transcription start sites (TSS) in SH-SY5Y neuroblastoma cells. 5' RACE experiments revealed multiple TSS located on three first alternative 5' exons, termed EX, EY and EZ, which are unusually close on the genomic sequence and are

all located more than 42 kbp upstream of the first common coding exon. Using Taqman duplex Q-PCR, we showed that the *ST6GAL2* transcripts initiated by EX or EY are mainly expressed in both brain-related cell lines and human cerebral cortex, testifying for the use of a similar transcriptional regulation *in vivo*. Furthermore, we also showed for the first time hST6Gal II protein expression in the different lobes of the human cortex. Luciferase reporter assays allowed us to define two sequences upstream EX and EY with a high and moderate promoter activity, respectively. Bioinformatics analysis and site-directed mutagenesis showed that NF- κ B and NRSF are likely to act as transcriptional repressors, whereas neuronal-related development factors Sox5, Pur α and Olf1, are likely to act as transcriptional activators of *ST6GAL2*. This suggests that *ST6GAL2* transcription could be potentially activated for specific neuronal functions.

S. Lehoux · S. Groux-Degroote · A. Cazet · P. Delannoy · M.-A. Krzewinski-Recchi
Structural and Functional Glycobiology Unit, UMR CNRS 8576, University of Sciences and Technologies of Lille, 59655 Villeneuve d'Ascq, France

C.-M. Dhaenens · C.-A. Maurage · M.-L. Caillet-Boudin
INSERM, U837, place de Verdun, 59045 Lille, France

C.-M. Dhaenens · C.-A. Maurage · M.-L. Caillet-Boudin
Faculté de Médecine, Institut de Médecine prédictive et de Recherche thérapeutique, Centre de Recherches Jean-Pierre Aubert, University of Lille 2, place de Verdun, 59045 Lille, France

M.-A. Krzewinski-Recchi (✉)
Unité de Glycobiologie Structurale et Fonctionnelle, UMR CNRS n°8576, Bâtiment C9, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq, France
e-mail: marie-ange.recchi@univ-lille1.fr

Keywords β -galactoside α 2,6-sialyltransferase · Cerebral cortex · NF- κ B · *ST6GAL2* · Transcriptional expression

Abbreviations

DSiaT	drosophila sialyltransferase
Gal	β -D-galactopyranosyl
GalNAc	β -D-N-acetylglucosaminyl
GlcNAc	β -D-N-acetylglucosaminyl
HPRT	hypoxanthine phosphoribosyltransferase
LDN	LacdiNAc, GalNAc β 1-4GlcNAc
LN	LacNAc, Gal β 1-4GlcNAc
NF- κ B	nuclear factor-kappa B
NRSF	neuron-restrictive silencer factor
Olf1	olfactory neuron-specific binding protein
Pur α	purin-riche element binding protein alpha
Q-PCR	quantitative polymerase chain reaction

RACE	rapid amplification of cDNA end
Sia	sialic acid
SLDN	siaLacdiNAc, Sia α 2-6GalNAc β 1-4GlcNAc
SLN	sialyl-LacNAc, Sia α 2-6Gal β 1-4GlcNAc
Sox5	SRY box 5
SP1	specificity protein 1
TSS	transcriptional start site
UTR	untranslated region

Introduction

Cell surface carbohydrate structures containing sialic acid play a crucial role in cellular interaction mechanisms such as cell-to-cell, cell-to-pathogen and cell-to-protein recognitions. Sialylated structures are expressed in a tissue-, time- and cell type-specific manner and imply that sialylation should be finely regulated. The transfer of sialic acid (Sia) from the activated donor substrate, CMP-sialic acid, to the terminal position of glycans of glycoproteins and glycolipids is catalyzed by sialyltransferases, a subset of glycosyltransferases gathered in CAZy family 29 [1]. Twenty different sialyltransferases have been identified in mammals. Each of the enzymes exhibits a specificity for acceptor substrate and for the linkage formed: α 2-3 to a β -D-galactopyranosyl (Gal) residue; α 2-6 to a Gal, β -D-N-acetylgalactosaminyl (GalNAc) or β -D-N-acetylglucosaminyl (GlcNAc) residue, or α 2-8 to another sialic acid residue (for review [2–4]).

Among the 20 human sialyltransferases, only two members have been identified in the ST6Gal family. ST6Gal I was identified more than 20 years ago and is certainly one of the most studied glycosyltransferases. It has been characterized in details and its cDNA cloned from different species including rat [5], bovine [6] and Human (hST6Gal I) [7]. In contrast, the cDNA encoding the second human ST6Gal (hST6Gal II) has been cloned in 2002 and its enzymatic properties have been only characterized *in vitro* [8,9]. Both enzymes share 48 % amino acid sequence identity in the catalytic domain and transfer a sialic acid residue in α 2,6 linkage to Gal β 1-4GlcNAc (LacNAc, LN) disaccharide sequence, to form the sialyl-LacNAc (SLN) [8,9]. The corresponding genes (*ST6GAL1* & *ST6GAL2*) have a similar genomic organization of the coding region with 5 coding exons and phylogenetic analyses indicate that they have evolved from a common ancestor [4].

Despite these similarities, *in vitro* assays have underlined slight differences in their enzymatic properties. Using synthetic acceptor substrates, hST6Gal I shows a 4-fold decreased activity towards GalNAc β 1-4GlcNAc (LacdiNAc, LDN) compared to LN, whereas hST6Gal II exhibits a 2-fold increased activity towards this motif, suggesting that hST6Gal II is the enzyme involved in Sia α 2-

6GalNAc β 1-4GlcNAc (SLDN) biosynthesis [10]. Interestingly, a similar observation was also made for DSiaT, the only sialyltransferase identified in *Drosophila melanogaster*, which also prefers LDN as an acceptor substrate [11]. Moreover, sequence alignments also confirmed a close relationship of DSiaT to ST6Gal family, suggesting that mammals ST6Gal and DSiaT proceed from a common ancestral gene [2,11].

While *ST6GAL1* gene is expressed in almost all mammal tissues with a notable lower expression in brain and testis [12], *ST6GAL2* gene exhibits a restricted tissue-specific expression pattern. Expression of *ST6GAL2* has been detected mainly in adult and embryonic brains, and to a lower extent in thyroid gland, small intestine, colon, and testis [8,9]. Moreover, a macro-array analysis of gene expression levels of 140 glycosyltransferases and glycosidases in the cerebral cortex from developing and adult mice has shown an increased expression of *st6gal2* during mouse embryo development, followed by a decrease expression in adult brain. On the contrary, the expression of most of sialyltransferase genes, including *st6gal1*, remains low and unchanged during brain development [13]. Since DSiaT is only expressed during embryonic development of the central nervous system [11], we can hypothesize that *ST6GAL2* compensates the low expression of *ST6GAL1* in brain for a precise unknown function, which could be conserved from invertebrate to human. However, pituitary hormones, glycodelin and urokinase, are the only glycoproteins known to exhibit terminal SLDN motifs [14–16]. Moreover, α 2,8-sialylation is the most, if not the only, type of sialylation depicted for neuronal functions such as cell surface interactions, neurogenesis or neuronal plasticity [17,18] and terminal α 2,6-sialylation has not been yet reported in brain-related glycoconjugates.

Sialyltransferases are known to be expressed in a tissue- and time-specific manner [12]. The expression of these enzymes is mostly regulated at the transcriptional level through the usage of alternative untranslated 5' exons under the control of tissue-specific promoters. In particular, it is well-established that *ST6GAL1* expression is regulated by three major promoters leading to the expression of an ubiquitous transcript [19], a mature B lymphocytes-specific transcript [20] and a liver-specific mRNA [21]. In contrast, nothing is known about *ST6GAL2* transcriptional regulation, but a specific control seems to occur in brain, especially during embryonic development.

In order to clarify the mechanisms involved in the regulation of *ST6GAL2* expression, we first identified and characterized its 5' ends organization in SH-SY5Y neuroblastoma cell line. 5' RACE experiments showed that at least three *ST6GAL2* mRNAs, distinct in their 5' untranslated ends, were expressed in SH-SY5Y cells. The relative expression of these transcripts was analyzed using duplex

Taqman Q-PCR strategy in different brain-related cell lines and in different lobes of the human cortex. We showed that the *ST6GAL2* specific transcripts identified in cultured cells were also expressed in the human cortex. Moreover, the hST6Gal II protein was also detected in the different areas of the human cortex. Luciferase assays allowed us to define two promoter regions that do not contain any canonical TATA or CAAT boxes, but contain general and neuronal-related transcription factor binding sites. The role of these factors was investigated by PCR site-directed mutagenesis.

Materials and methods

Cell culture

Human embryonic carcinoma NT2/D1, neuroblastoma SH-SY5Y, neuroblastoma Kelly, astrocytoma/glioblastoma U118 and glioblastoma T98G cell lines were provided by Dr Luc Buée (INSERM, U837, Lille, France). The NT2/D1, SH-SY5Y, U118 and the T98G cells were cultured in Dulbecco's modified Eagle's medium (DMEM) containing 4500 mg/l of glucose (BioWhittaker, Lonza, Belgium). Human neuroblastoma Kelly cells were cultured in RPMI 1640 medium (BioWhittaker). Media were supplemented with 10 % heat-inactivated fetal bovine serum (FBS) (Eurobio), 2 mM of L-glutamine, 100 U/ml penicillin and 10 µg/ml streptomycin, and incubated at 37°C under 5 % CO₂.

Amplification of 5' cDNA Ends, 5' RACE

Total RNA from SH-SY5Y cells were extracted using the NucleoSpin[®] RNA II kit according to the manufacturer's recommendations (Macherey-Nagel, Hoerd, France). 5' RACE amplifications were performed with the First Choice RLM-RACE kit (Ambion, Huntingdon, UK) according to the manufacturer's instructions. Initial reverse transcription was performed with random decamers oligonucleotides using 10 µg of total RNA from SH-SY5Y cells. After the synthesis of the first strand cDNA, PCR was performed with provided Outer Primer (OP) and gene-specific primers GSP1, GSP1.X, GSP1.Y or GSP1.Z (Table 1) using the following PCR conditions: 94°C for 2 min; 40 cycles of 94°C for 30 sec, 62°C for 30 sec and 72°C for 3 min. Nested-PCR were performed with provided Inner Primer (IP) and gene-specific primers GSP2.1 or GSP2.2, or exon-specific primers GSP2.X, GSP2.Y or GSP2.Z in the following PCR conditions: 94°C for 2 min; 40 cycles of 94°C for 30 sec, 72° for 30 sec and 72°C for 2 min. PCR amplifications were performed with the Taq polymerase from Taq PCR Core Kit (Qiagen GmbH, Germany) according to provider's instructions. Nested-PCR products obtained were size separated by agarose gel electrophoresis, subcloned into

pCR2.1 TOPO TA cloning (Invitrogen, Carlsbad, CA) and sequenced by Genoscreen (Lille, France).

Human cerebral cortex control samples and RNA preparation

Brains were obtained at autopsy from 8 subjects without any brain pathologies. The brain tissues were obtained from the Lille Neurobanque (France) in accordance with the local ethics committee protocol. Post-mortem delay was <36 h for all samples. Tissue samples from cortical areas were cut off frontal, temporal, parietal, occipital regions. Total RNA was isolated using the RNeasy Total RNA Isolation System according to the manufacturer's recommendations (Promega, Madison, USA).

SYBR[®] Green and duplex Taqman Q-PCR

First strand cDNA synthesis were achieved using 2 µg of total RNA from cultured cell lines (SH-SY5Y, NT2/D1, U118, T98G and Kelly) or brain samples with First-Strand cDNA Synthesis Kit (Amersham Biosciences, Freiburg, Germany), according to the manufacturer's instructions, using random hexameric primers (pd(N)₆) and oligodT in a 3:1 ratio. SYBR[®] Green Q-PCR and subsequent analysis were performed using the MX4000 Multiplex Quantitative System (Stratagene, La Jolla, CA, USA) equipped with Version 3.0 software. The 25 µl final volume was composed of 12.5 µl of the Brilliant[®] SYBR[®] Green QPCR Master Mix Kit (Stratagene), 0.3 µM of primers *ST6GAL2* 5', *HPRT* 5' and *HPRT* 3' primers (previously described by Zhang *et al.* [22]) and 0.6 µM of *ST6GAL2* 3' primer and 4 µl of cDNA diluted 1:40. The thermal cycling profile was 95°C for 10 min, then 40 PCR cycles as follows: 95°C for 1 min, 51°C for 1 min and 72°C for 1 min. Q-PCR assays were performed in triplicate in 96-well plates. Standard curves were generated for each pair of primers using serial dilutions of cDNA. Standard curves efficiency used are: 101.3 % for *ST6GAL2* and 99.6 % for *HPRT*. Duplex Taqman Q-PCR and subsequent analysis were performed using the MX4000 Multiplex Quantitative System (Stratagene) equipped with Version 3.0 software. For the *ST6GAL2* exon-specific amplifications, forward primers specific of each exon sequence were paired with a common reverse primer, exon1-3'Taq, hybridizing within the exon 1 sequence. *HPRT* primers were designed to hybridize within the exon 1 and exon 2 sequences. Taqman probes were synthesized by Applied Biosystems. The *ST6GAL2* probes were FAM labelled and the normalizing gene *HPRT* probe was VIC labelled (Table 1). PCR reactions were performed with 12.5 µl of the QuantiTect Multiplex PCR Kit Mix (Qiagen), 0.4 µM of each primer, 0.4 µM of Taqman probe and 4 µl of cDNA diluted 1:40 in a final volume of 25 µl.

Table 1 Nucleotide sequences of 5'-RACE adapter and primers, Q-PCR primers and Taqman probes, and of primers used for PCR site-directed mutagenesis. Oligonucleotides were synthesized by Eurogentec (Seraing, Belgium)

Primer	Sequence
5'-RACE Adapter	5'-GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUGAAA-3'
5' Outer Primer	5'-GCTGATGGCGATGAATGAACACTG-3'
5' Inner Primer	5'-CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG-3'
GSP1	5'-CCTCTTCTTCACCCGCCTC-3'
GSP1.X	5'-GGAGCCGGGCACTGGGAAGTTGGG-3'
GSP1.Y	5'-CTGGAGTCCAAGGCGTGC GTGTCA-3'
GSP1.Z	5'-TACTGAGCCACAGGGGCACA ACTG-3'
GSP2.1	5'-AGCGGGGTTGCTGTCTCGGT-3'
GSP2.2	5'-AGGCGGGGAGGGCTCATG-3'
GSP2.X	5'-AGACTTGGGCTGCGTGGGGCGGTGC-3'
GSP2.Y	5'-TCCTCCTGCCACCGTGGGGTCTCG-3'
GSP2.Z	5'-CCCCAAAGCGCAAGCCGTCTCCC-3'
<i>SYBR® Green primer</i>	<i>Sequence</i>
ST6GAL2 5'	5'-ACGCTGCTGATTGACTCTTCT-3'
ST6GAL2 3'	5'-CACATACTGGCACTCATCTAA-3'
HPRT5' [22]	5'-GACCAGTCAACAGGGGACAT-3'
HPRT3' [22]	5'-AACACTTCGTGGGGTCTTTTC-3'
<i>Taqman Primer</i>	<i>Sequence</i>
HPRT5'Taq	5'-GCCCTGGCGTCTGTGATT-3'
HPRT3'Taq	5'-CTCAGCATAATGATTAGGTATGCAAAA-3'
X5'Taq	5'-GCAGCCCAACTTCCCAGTG-3'
Y5'Taq	5'-GGTGTCGTGCGCTTCGA-3'
Z5'Taq	5'-ACGGGCGGATACAGTTGTG-3'
Exon1-3'Taq	5'-TCAGCACCTTGTGTCTTAATGCA-3'
<i>Taqman probe</i>	<i>Sequence</i>
HPRT	5'-VIC-TGATGATGAACCAGGTTAT-TAMRA-3'
X	5'-FAM-CCAGCTAGGTTTATTCT-MGB-3'
Y	5'-FAM-AAGAAGGCTGGTTCATT-MGB-3'
Z	5'-FAM-CAGTAAAGGTTTATTCTGCCA-MGB-3'
<i>Primer</i>	<i>Sequence</i>
NFκB-5'	5'-CAGGCGCCTTTACTCAATCCTGG-3'
NFκB-3'	5'-CCAGGATTGAGTAAAGGCGCCTG-3'
NRSF-5'	5'-CCGAGTGGGCGATTGGTGTCCA-3'
NRSF-3'	5'-TGGACACCAATCGCCCACTCGG-3'
Olf1-3'	5'-GCTGCCAAGCTTTCCCGTCAAATGCCGCTG-3'
Purα-5'	5'-GGATGGAGGCTTACACCGGGG-3'
Purα-3'	5'-CCCCGGTGTAAGCCTCCATCC-3'
Sox5-5'	5'-GCCCAGAGATATGAGGCGTGC-3'
Sox5-3'	5'-GCACGCCTCATACTCTCGGGC-3'
SP1.1-5'	5'-CTGCCGAAGAAAGGGACCCG-3'
SP1.1-3'	5'-CGGGTCCCTTTCTTCGGCAG-3'
SP1.2-5'	5'-CCCGCACTTTTCCACGCAG-3'
SP1.2-3'	5'-CTGCGTGGAAAAGTGGGG-3'
SP1.3-5'	5'-GAGCTCGGGTGTGGCGC-3'
SP1.3-3'	5'-GCGCCAAACACCCGAGCTC-3'
SP1.4-5'	5'-GCCCGCTAAATGCGCTTG-3'
SP1.4-3'	5'-CAAGCGCATTTAGCGGGGC-3'

After 95°C for 15 min, 40 PCR cycles were performed as follows: 95°C for 1 min, 60°C for 1 min (hybridization and elongation). Q-PCR assays were performed in triplicate in 96-well plates using MX4000 apparatus (Stratagene). Standard curves were generated for each trio of primers and probe of a *ST6GAL2* specific transcript in duplex with primers and probe of the normalizing gene *HPRT*, using 2-fold serial dilutions of cDNA. Standard curves efficiencies used are: 99.5 % and 97.8 % for EX specific transcript and *HPRT*, 100.4 % and 100.5 % for EY specific transcript and *HPRT* and 98.0 % for EZ specific transcript and *HPRT*.

SDS-PAGE and Western blotting analysis

Brain tissues were desintegrated on ice with a glass homogenizer, in RIPA Buffer (Pierce, Rockford, USA), containing CHAPS (0.5 % w/v) and a cocktail of protease and phosphatase inhibitors (Roche, Meylan, France). Subconfluent SH-SY5Y cell cultures, were harvested by adding a solution of 4 mM EDTA in PBS, pelleted and resuspended in lysis buffer (150 mM NaCl, 50 mM Tris-HCl, pH 7.5, 1% Triton X-100), containing a cocktail of protease and phosphatase inhibitors (Roche). The protein concentration of extracted proteins (from brain tissues or cells) was determined with the Micro BCATM Protein Assay Reagent kit (Pierce). 40 µg of total proteins of each sample were boiled for 10 min in reducing Laemmli Sample buffer and resolved by SDS/PAGE on 8 % mini-gels (Bio-Rad, Richmond, USA). After transfer onto a nitrocellulose membrane (200 mA, 2 h), blocking was performed using TBS (Tris Buffer Saline) containing 0,05 % Tween 20 and 5 % (w/v) non-fat dried milk for 1 h at room temperature (RT). Primary antibody incubations with anti-hST6Gal II (Sigma, St Louis, USA) or anti-actin (Santa Cruz Biotechnology inc., Europe) antibodies were performed overnight at 4°C in TBS, 0,05 % Tween 20 and 5 % (w/v) non-fat dried milk, at 1:1000 dilution for the anti-hST6Gal II antibody or 1:10000 dilution for the anti-actin antibody. After three washing steps of 10 min in TBS, 0,05 % Tween 20, incubation of 1 h at RT with horseradish peroxidase-conjugated goat anti-rabbit IgG was performed at 1:10000 dilution in TBS, 0,05 % Tween 20 and 5 % (w/v) non-fat dried milk. Membranes were finally washed three times for 10 min in TBS, 0,05 % Tween 20 and detection was achieved using enhanced chemiluminescence (ECL⁺® advance Western blotting detection reagents, Amersham Biosciences, Little Chalfont, Buks, U.K.).

Bioinformatics analysis

In silico analysis were performed with Matinspector 2.2 (www.genomatix.de) using TRANSAC matrices 4.0 [23]

with “core similarity: 0.95” and “matrices similarity: Optimized”.

Construction of reporter plasmids for luciferase assays

Genomic DNA of SH-SY5Y cells was extracted with the Nucleospin[®] Blood kit (Macherey-Nagel) following manufacturer’s instructions. The -1851 / +263 genomic region upstream EX was amplified from genomic DNA of SH-SY5Y using primers 5'-CGCTTGCCACGGTACCCGGT CATT TTTGT-3' and 5'-AGACTTAAGCTTCGTGGGGCG GTGCGGG-3'. Underlined nucleotides represent the restriction sites that were incorporated, *KpnI* and *HindIII* respectively. PCR fragments were subcloned into pGL3-Basic vector (Promega) upstream of the *Firefly* luciferase gene at *KpnI/HindIII* sites, generating the pGL3_-1851/+263 vector. The pGL3_-1851/+263 was submitted to 5' deletions using natural restriction enzyme sites. Digested ends were blunted with DNA Terminator End Repair Kit (Euro-medex, Mundolsheim, France) following manufacturer’s instructions. Then, blunted ends were ligated with the Quick T4 DNA ligase (Biolabs, Ozyme, Saint Quentin Yvelines, France). *HindIII* and *BglII* were used to obtain pGL3_-752/+263, *HindIII* and *XhoI* were used to obtain pGL3_-498/+263, *HindIII* and *NheI* were used to obtain pGL3_-121/+263 and *HindIII* and *SmaI* were used to obtain pGL3_-92/+263. Following the same strategy, the -121 / +1452 genomic region was amplified from genomic DNA of SH-SY5Y using primers 5'-TCACCGGACTTGGCGAGCTA-3' and 5'-GCCGCCTG AGATCTACATAAG-3'. Underlined nucleotides represent the *BglII* restriction site that was incorporated. Using the *NheI* natural site of the amplified sequence and the *BglII* incorporated site, the PCR fragments were subcloned into the pGL3-Basic vector at the *NheI/BglII* sites, generating the pGL3_-121/+1452 vector. Plasmids pGL3_+517/+1452, pGL3_+773/+1452, pGL3_+465/+772, pGL3_-121/+772, pGL3_-121/+464 and pGL3_-121/+91 were generated by enzymatic digestions, ends blunting and ligation. The plasmid pGL3_+1046/+1452 was generated using primers 5'-GGAGGGTACCGGGTAAGGGTT-3' and 5'-CCGTACC TTAATAAGCTTTCTACATA-3'. Underlined nucleotides represent the restriction sites that were incorporated, *KpnI* and *HindIII* respectively. The amplified fragment was cloned into the pGL3-Basic upstream of the *Firefly* luciferase gene at the *KpnI/HindIII* sites, generating the pGL3_-1046/+1452 vector. The different constructions were sequenced to ensure the absence of mutation (Genoscreen).

PCR site-directed mutagenesis

The base substitution mutations for the predicted binding sites NF-κB, NRSF, Olf1, Purα, Sox5, and the four SP1

sites, named SP1.1, SP1.2, SP1.3 and SP1.4, were introduced by PCR site-directed mutagenesis with the HotStar HiFidelity polymerase kit (Qiagen) following manufacturer's instructions. Using the pGL3_{-121/+772} plasmid as a template, two first rounds of PCR using a common forward primer, 5'-CGATAGGTACCGAGCTCTACGCG-3' (containing a *KpnI* site), and a reverse mutated site-specific primer (See Table 1), and a common reverse primer, 5'-TGCCAAGCTTACTTAGATCGCAGATCTC-3 (containing a *HindIII* site), and a forward mutated site-specific primer (See Table 1) were performed for each mutated site in the PCR conditions: 95°C for 5 min; 30 cycles of 95°C for 1 min, 53-60°C for 1 min and 72°C for 1 min. The expected amplified PCR fragments were purified on agarose gel and used as templates with the common forward and reverse primers for the final PCR using identical conditions. Olf1 mutated site was generated with the common forward primer and the Olf1-3' primer (containing a *HindIII* site) (Table 1) using the following PCR conditions: 95°C for 5 min ; 30 cycles of 95°C for 1 min, 53°C for 1 min and 72°C for 1 min. The final PCR fragments were subcloned into the pGL3-Basic vector at the *KpnI/HindIII* sites and sequenced (Genoscreen) to ensure the absence of any other mutations. The pGL3_{-121/+91NF-κB} plasmid was obtained by enzymatic digestion at the *SmaI* site from the pGL3_{-121/+772NF-κB} plasmid.

Transient transfection and luciferase assays

Transfections were performed with the Cell Line Nucleofector® Kit V (AMAXA, Europe) according to manufacturer's instruction for the SH-SY5Y cell line grown on 6 well plates. Briefly, 2.10⁶ SH-SY5Y cells were harvested at ~70 % confluence, centrifuged 5 min at 1000 rpm and resuspended in 100 μl of Nucleofector® Kit V solution (AMAXA). The cell suspensions were mixed with 2 μg of pGL3-Basic plasmid constructions and 40 ng of control *Renilla* plasmid. Cells and plasmids were transferred in an AMAXA cuvette and submitted to nucleofection using G-004 program that has been optimized by the manufacturer for the SH-SY5Y cells. Cells were then plated on 6-well plates with 2 ml of culture media and incubated for 24 h. Cell survival and transfection efficiency with AMAXA protocol were checked by GFP expression plasmid and were 50 % and 70 %, respectively, consistent with the manufacturer's data. Cells were lysed in 250 μl of Passive Lysis Buffer (PLB, Dual Luciferase Reporter Assay System, Promega) and 20 μl were used for luciferase assay. *Firefly* and *Renilla* luciferase activities were measured using Dual Luciferase Reporter Assay System (Promega). Luminescence was measured with the Centro luminometer (Berthold Technologies)

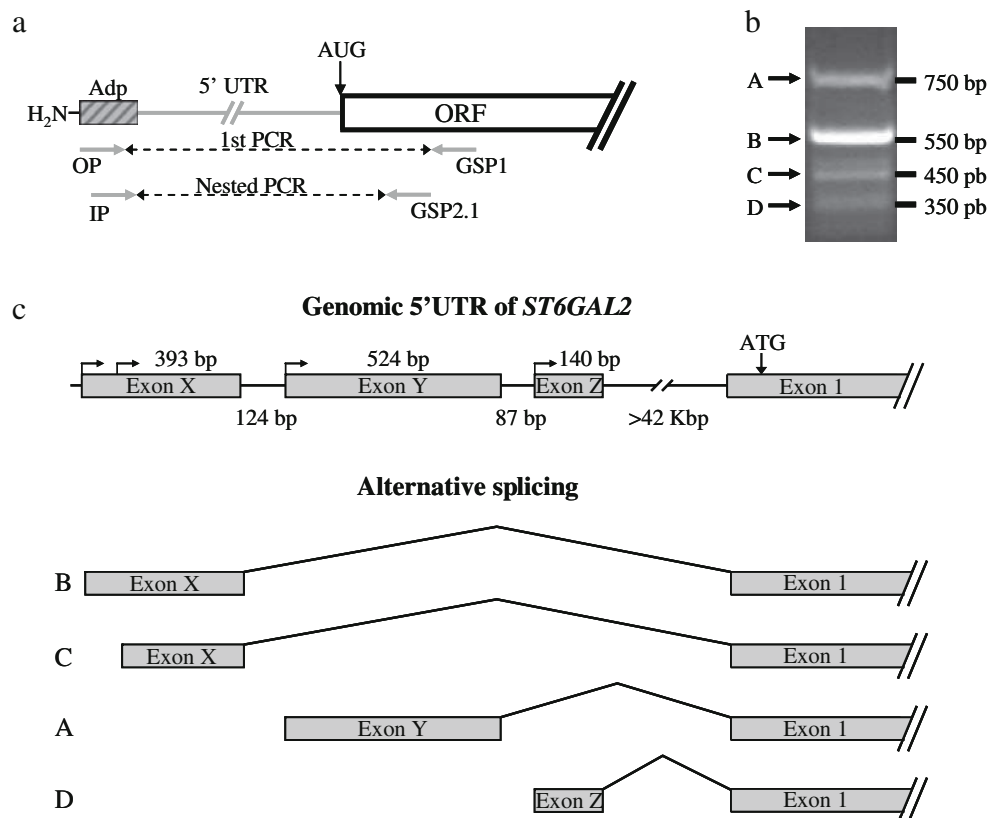
Results

Identification of alternative 5' UTR by 5' RACE

The coding sequence of *ST6GAL2* gene is divided in 5 exons spanned over 85 kbp on chromosome 2. Mature transcripts are more than 8 kb long including a 5 kb 3'-untranslated region (UTR) and incomplete sequences located 42 kbp upstream of the first coding exon were previously reported in 5' UTR [8,9]. To determine the transcription start sites (TSS) of the human *ST6GAL2* gene and potential alternative 5' UTR, capped mRNA 5' RACE was performed using total RNA isolated from SH-SY5Y human neuroblastoma cells. A first PCR was performed with Outer Primer and GSP1, which hybridized within the sequence of the first coding exon E1. A nested PCR was then performed with Inner Primer and GSP2.1 (Fig. 1a). Experimental conditions were optimized for both first-round and nested PCR as described in Materials & Methods section. As shown in Fig. 1b, which is representative of at least three independent amplifications, four bands from 350 to 700 bp were obtained after the nested PCR. The PCR products, named A, B, C and D, were subcloned, sequenced and the sequences were BLAST-aligned (<http://www.ncbi.nlm.nih.gov/BLAST/>) against the human genome. The amplified products were assigned to *ST6GAL2* transcripts, including at the 5'-end three alternative unique sequences termed EX, EY and EZ (GenBank accession # FJ433883, FJ433884, FJ433885), which were located on chromosome 2 more than 42 kbp upstream exon E1 and separated from each other by 124 and 87 bp, respectively (Fig. 1c). Splicing junctions between these exons and exon E1 were in agreement with the GT-AG splicing rule. The sequence of band A contained the sequence of EY, band B corresponded to the sequence of EX and band C corresponded to an 80 bp shorter 5'-truncated sequence of EX. Finally, the sequence of band D contained the sequence of EZ. Moreover, the presence of EX, EY or EZ was mutually exclusive among *ST6GAL2* 5' UTR (Fig. 1c), in agreement with the fact that the 5' ends of EY and EZ do not contain splicing acceptor sites.

Another series of 5' RACE experiments using exon-specific primer pairs were performed to determine precisely the transcription start site (TSS) of each exon. Exon-specific primers hybridizing within the 3'-end sequence of each exon (Table 1) were used in pair with Inner Primer. Results of the nested PCR are shown in Fig. 2a. For EY and EZ specific amplifications, only one band was observed, but as expected, two band sizes were revealed for EX. The amplified fragments were subcloned and sequenced. Twenty nine clones corresponding to different sizes of EX, from 315 bp to 393 bp, were sequenced (Fig. 2b), defining two groups of sequences: a first group with sizes of 315, 318, 320 and

Fig. 1 Identification of *ST6GAL2* transcript 5' ends by 5' RACE. **(a)** Schematic representation of 5' RACE procedure. Adp, oligonucleotide adapter provided by FirstChoice® RLM-RACE kit; ORF, Open Reading Frame; OP, Outer Primer; IP, Inner Primer. **(b)** Agarose gel (2 %) analysis of nested PCR products with primer pair I.P. / GSP2.1. **(c)** Genomic organization of the 5' -end of the *ST6GAL2* gene. Oriented arrows represent the transcription start site (TSS) of each exon. Numbers on the top of exons and numbers at the bottom of the introns indicate their respective lengths in bp



326 bp, and a second group with sizes of 380, 381 and 393 bp. These two groups of sequences corresponded to the two amplified products observed (Fig. 2a, lane EX) and also to bands B and C in Fig. 1b. In parallel, 8 clones from EY were sequenced and only a one bp variation of the TSS was observed, EY length being 523 or 524 bp (Fig. 2b). Finally, among the 6 different EZ clones sequenced, no variation was observed, EZ length being 140 bp (Fig. 2b).

Quantification of transcripts expression by duplex Taqman Q-PCR in cultured cells

The relative expression of *ST6GAL2* exon-specific transcripts was determined by duplex Taqman Q-PCR using *HPRT* as a normalizing gene. A common reverse primer that hybridizes within the exon E1 of *ST6GAL2* cDNA sequence and three exon-specific forward primers hybridizing within EX, EY or EZ, respectively, were designed (Table 1). Taqman probes were designed within exon E1 / exon EX, EY or EZ junction sequence. For *HPRT*, primers and probe were designed to hybridize within exon E1 / exon E2 sequences (Table 1). In order to improve cDNA synthesis and to decrease threshold cycle (Ct) values for both *HPRT* and exon-specific *ST6GAL2* transcripts, a mix of random hexameric primers and oligo-dT primers was used in a ratio 3:1 as primers for retro-transcription. The relative expression of specific transcripts was determined

for several glial and neuronal cell types, including two glioblastoma U-118 and T98G cells, two neuroblastoma SH-SY5Y and Kelly cells and the neuroprogenitor cells NT2/D1. In SH-SY5Y cell line, our results clearly showed a predominant expression of transcripts containing EX or EY, while EZ specific transcript was expressed at low levels (Fig. 3). A similar pattern of expression with a predominant expression of EX and EY specific transcripts was also detected in NT2/D1 cells (Fig. 3), whereas a low level of expression of these transcripts and no expression of EZ specific transcript was observed in the Kelly cell line, and almost no expression of all transcripts in T98G and U-118 cell lines. The level of expression of the specific transcripts was consistent with the total expression of *ST6GAL2* (Data not shown), indicating that no other transcript with alternative 5' UTR seems to be expressed in these cell lines.

Quantification of total and transcript-specific *ST6GAL2* expression in human cerebral cortex

We next investigated the *in vivo* expression of *ST6GAL2* in the temporal, occipital, frontal and parietal lobes of the human cerebral cortex. For each region, up to five different samples were available for quantification. We first quantified the total expression of *ST6GAL2* (Fig. 4a). In all samples, the expression of *ST6GAL2* was in average above the level of expression in the SH-SY5Y cell line. Furthermore, *ST6GAL2*

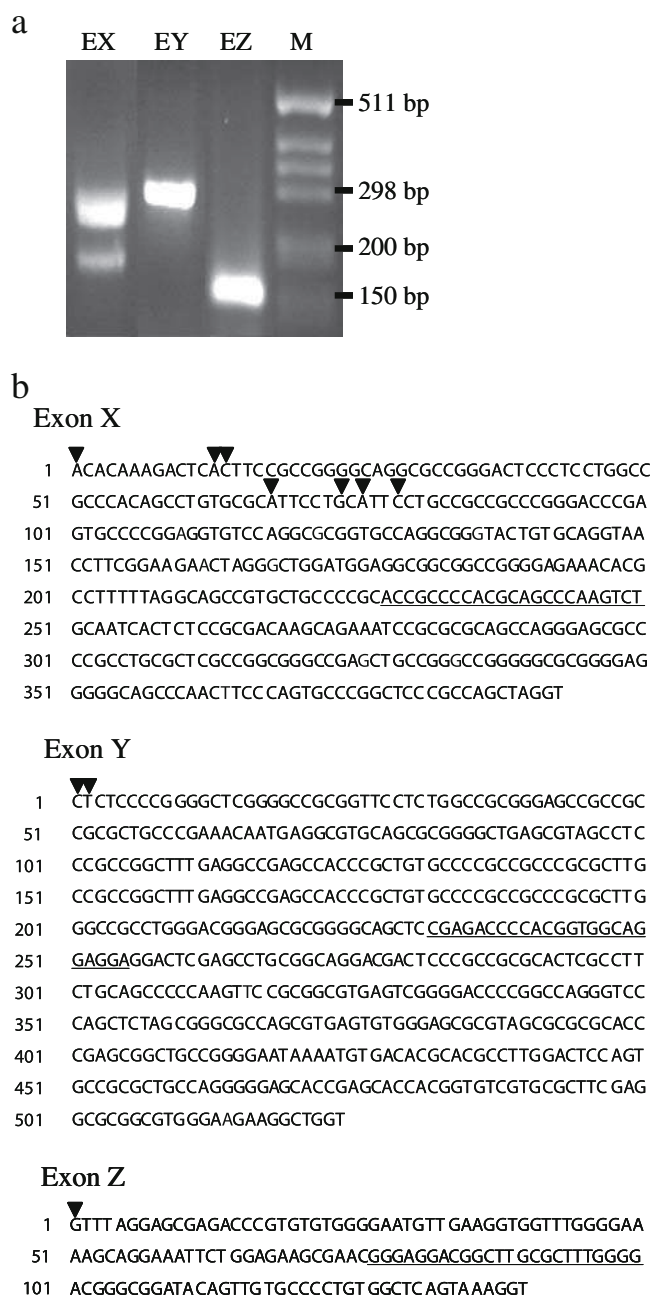


Fig. 2 Exon-specific 5' RACE of *ST6GAL2* transcripts. **(a)** Agarose gel (2 %) analysis of 5' RACE exon-specific nested PCR products. Lanes EX, EY and EZ show amplifications with primer pairs I.P. / GSP2.X, I.P. / GSP2.Y and I.P. / GSP2.Z, respectively. M, molecular weight marker (Invitrogen 1 kb). **(b)** Nucleotide sequence of EX, EY and EZ. Arrowheads indicate the different transcription start sites identified by 5' RACE. The sequences of exon-specific primers used for the three exon-specific 5' RACE are underlined

expression was heterogeneous among the different tissue samples for each subject and among the different subjects for each cortex area. Specific transcript expression identified in cultured cells by 5' RACE was also considered. All brain samples expressed transcripts that contained EX, EY or EZ, confirming the results obtained with cultured cells. Moreover

in each sample, the level of expression of EX, EY and EZ specific transcripts corresponded to the level of *ST6GAL2* total expression suggesting that no other transcript is expressed in these samples (Fig. 4b). In cortical areas, the ratio of EZ/EX containing transcripts was increased compared with SH-SY5Y, NT2/D1 and Kelly cells. Similar variations were observed for the ratio EZ/EY. In spite of these slight differences, our results suggest that the transcriptional regulation of *ST6GAL2* in the human cerebral cortex is similar to that of the cultured cell lines.

hST6Gal II protein expression in the human cerebral cortex areas

Western blotting specificity of the anti-hST6Gal II antibody was checked using protein extracts from COS-7 cells transfected with the expression plasmid p3xFLAG-CMV-10 containing either the full length sequence of hST6Gal I or hST6Gal II, or the empty plasmid. Using the anti-ST6Gal II antibody, a single band was detected only in protein extracts from cells transfected with p3xFLAG-CMV-10-hST6Gal II, whereas the anti-FLAG antibody revealed one band in protein extracts from cells transfected with p3xFLAG-CMV-10-hST6Gal II or p3xFLAG-CMV-10-hST6Gal I. The apparent molecular weights of the proteins detected with the anti-hST6Gal II and anti-FLAG antibodies were similar and consistent with the predicted molecular weight of 3xFLAG-tagged hST6Gal II (62 kDa), confirming that the anti-hST6Gal II antibody is specific for hST6Gal II (data not shown). hST6Gal II protein expression was subsequently analysed by Western blotting in the four cerebral cortex lobes of three subjects and in SH-SY5Y cells. The results presented in Fig. 4c show a single band around 60 kDa in all brain samples, in agreement with the predicted molecular weight of hST6Gal II (58 kDa). The

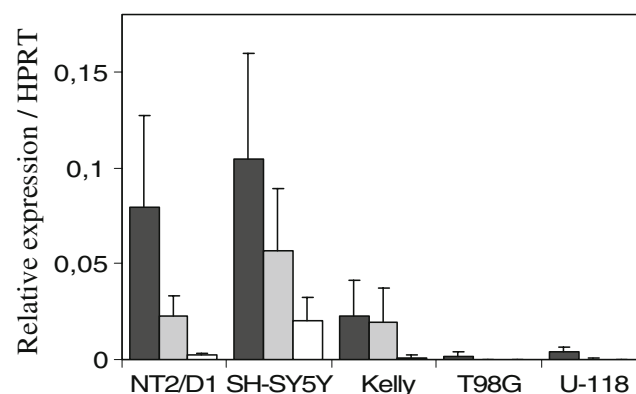
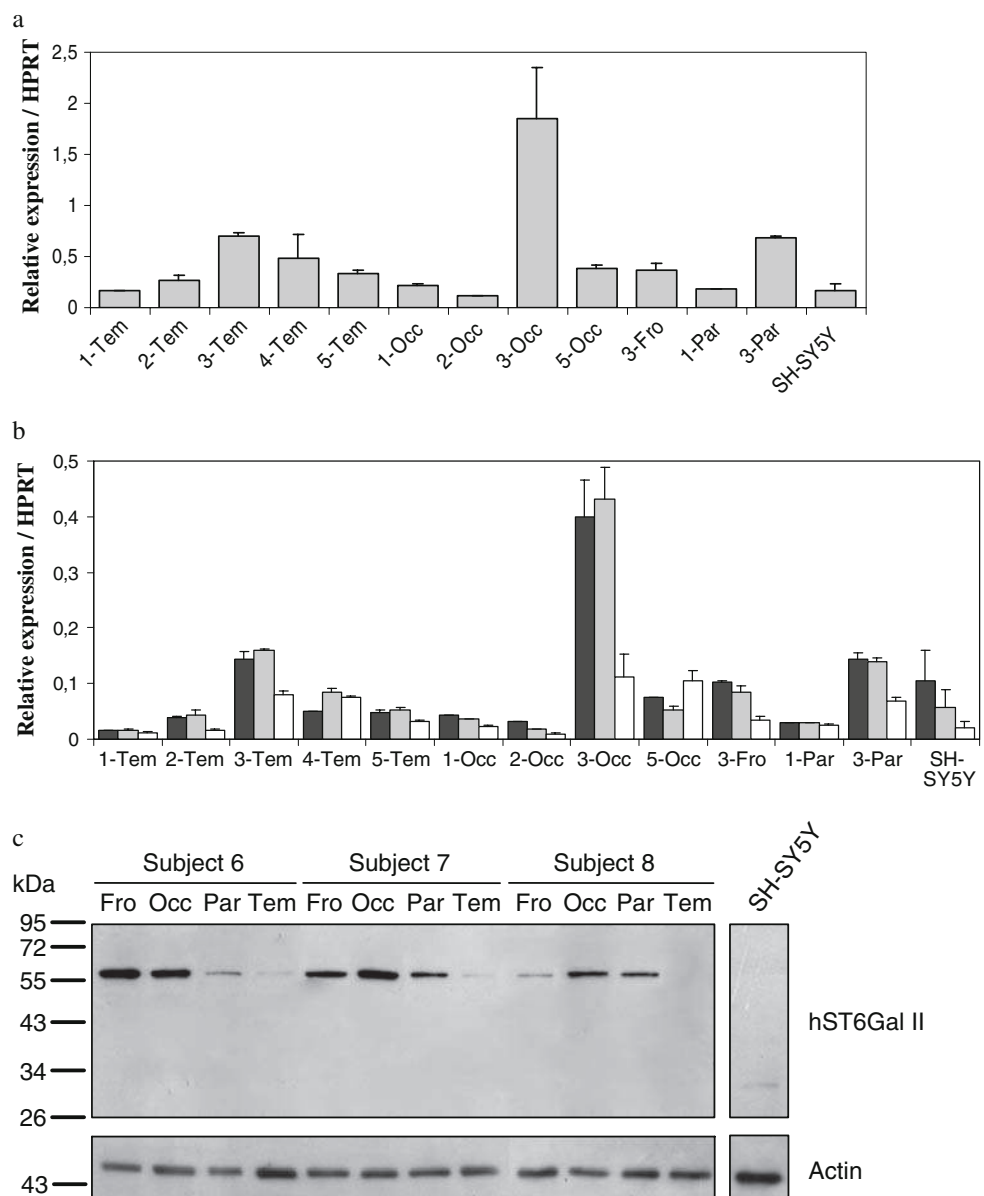


Fig. 3 Duplex Taqman Q-PCR analysis of *ST6GAL2* transcripts relative expression in various human cell lines. Black bars, EX-specific transcript; grey bars, EY-specific transcript; open bars, EZ-specific transcript. Values are normalized to *HPRT* gene expression. Data are means \pm S.D. of $n \geq 3$ experiments

Fig. 4 Protein and mRNA expression of *ST6GAL2* in different lobes of human cerebral cortex. **(a)** Expression of total *ST6GAL2* transcripts by Q-PCR using SYBR® Green. **(b)** Expression of *ST6GAL2* exon-specific transcripts by Duplex Taqman Q-PCR. Black bars, EX-specific transcript; grey bars, EY-specific transcript; open bars, EZ-specific transcript. Values are normalized and expressed relatively to *HPRT* gene. Data are means \pm S.D. of $n=3$ measurements. **(c)** hST6Gal II protein expression visualized by Western blotting. Numbers 1 to 8 correspond to different subjects; Tem., Occ., Fro., Par. refer to Temporal, Occipital, Frontal, Parietal regions of a given sample, respectively. Actin expression visualized by Western blotting was used as a loading control. Numbers 1 to 8 refer to different subjects. Tem, Occ, Fro, Par refer to Temporal, Occipital, Frontal and Parietal regions of a given subject, respectively



expression of the protein was heterogeneous among the different samples tested, with a lower expression in the temporal lobe for the three subjects. In addition, as observed for the gene expression, inter-individual variations were also detected (Fig. 4a). In these experimental conditions, hST6Gal II protein was not detectable in SH-SY5Y.

In silico analysis of the 3.6 kb sequence upstream and including EX, EY and EZ

Preliminary bioinformatics analysis predicted a 600 bp sequence upstream EX as the most probable promoter region for *ST6GAL2* (www.genomatic.de). We submitted to further *in silico* analysis the genomic region from 1.9 kb upstream EX to 200 bp downstream EZ: -1851 to +1452,

nucleotide +1 referred as the first nucleotide of EX longest form identified by 5' RACE (Fig. 5a). The results for the 1.9 kb sequence upstream EX as well as the results for the 1.7 kb sequence overlapping the three exons are shown in Fig. 5a and b. These analyses did not reveal any canonical TATA or CAAT boxes, but several putative binding sites for general transcription factors such as SP1, NF- κ B, GATA1 and MZF1, and also putative binding sites for transcription factors involved in neuronal functions such as NRSF, Pit1, MYT1, Pur α , Olf1 and Sox5 (Fig. 5a and b).

SH-SY5Y transient transfections and luciferase assays

To determine the minimal promoter region upstream EX, the sequence -1851 / +263 was subcloned in the pGL3-

Fig. 5 *In silico* analysis of the nucleotide sequence $-1851 / +1452$. **(a)** Nucleotides are numbered with the nucleotide +1 as the first TSS. Arrowheads indicate the different TSS of EX, open triangles indicate the different TSS of EY and the asterisk indicates the TSS of EZ. Potential transcription factor binding sites are underlined with arrows, left arrow when the binding site is on the $-$ strand and right arrow when the binding site is on the $+$ strand. **(b)** Schematic diagram of the $-1851 / +1452$ region. Upper transcription factors are located on the $+$ strand and bottom transcription factors are located on the $-$ strand

Basic vector upstream the luciferase gene and named pGL3_{-1851/+263}. This plasmid or 5' deleted constructions: pGL3_{-752/+263}, pGL3_{-498/+263}, pGL3_{-121/+263} and pGL3_{+92/+263} were transfected into SH-SY5Y cells for luciferase assays. The results presented in Fig. 6 show a 7-fold increase of luciferase activity for the plasmid pGL3_{-121/+263} compared to pGL3-Basic, whereas the luciferase activity for constructs pGL3_{-1851/+263}, pGL3_{-752/+263} and pGL3_{-498/+263} only show a 2- to 3-fold increased activity. Luciferase activity for the plasmid pGL3_{+92/+263} is even lower than the activity for the pGL3-Basic vector. These data strongly suggest the existence of a promoter region within the sequence $-121 / +92$. We also analyzed the $-121 / +1452$ region encompassing EX, EY and EZ. The plasmid pGL3_{-121/+1452} and its truncated forms: pGL3_{+517/+1452}, pGL3_{-121/+772}, pGL3_{+773/+1452}, pGL3_{+1046/+1452}, pGL3_{-121/+464}, pGL3_{+465/+772} and pGL3_{-121/+91} were generated. The 5' end of the genomic sequence inserted into the plasmid

pGL3_{+517/+1452} starts 8 nucleotides downstream from the first TSS of EY. The genomic sequence inserted into the plasmid pGL3_{-121/+91} corresponds to the $-121 / +91$ sequence. This sequence ends 12 nucleotides downstream of the last identified TSS of EX. These different plasmids were transfected into SH-SY5Y cells and luciferase activity was determined. The plasmids pGL3_{-121/+1452}, pGL3_{+517/+1452}, pGL3_{+773/+1452}, pGL3_{+1046/+1452} and pGL3_{-121/+464} did not show any significant increase in luciferase activity (Fig. 6) whereas pGL3_{-121/+772} and pGL3_{+465/+772} showed a 8- and 13-fold increased activity compared to pGL3-Basic, respectively (Fig. 6), suggesting that the sequence $+465 / +772$ corresponds to a promoter region. Moreover, the plasmid pGL3_{-121/+91}, corresponding to the previous pGL3_{-121/+263} construct deleted of 181 bp at the 3' end, showed a dramatic 35-fold increase of luciferase activity, indicating a high promoter activity within this 212 bp sequence upstream EX. Finally, the sequence $+92 / +464$ is likely to contain negative regulatory elements since luciferase activity decreased for plasmids pGL3_{-121/+772} and pGL3_{-121/+464} compared to pGL3_{+465/+772} and pGL3_{-121/+91}. Altogether, the results presented in Fig. 6 suggest a series of positive and negative regulatory elements within a 900 bp sequence including EX and EY. This region starts with a strong promoter sequence ($-121 / +91$) followed by a region containing negative regulatory elements ($+92 / +464$) and then a second promoter sequence ($+465 / +772$).

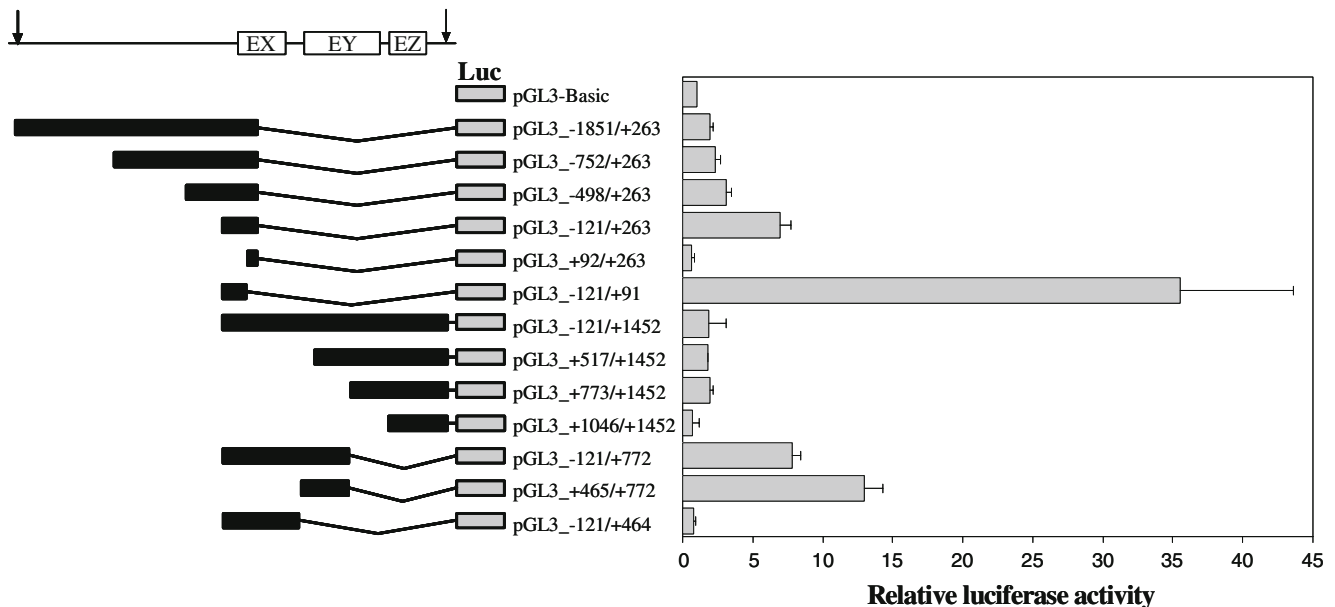


Fig. 6 Luciferase assays after transfections of different plasmid constructs in SH-SY5Y neuroblastoma cells. On the left, a schematic representation of the different constructs inserted in pGL3-Basic reporter vector upstream of the luciferase gene. On the right: the

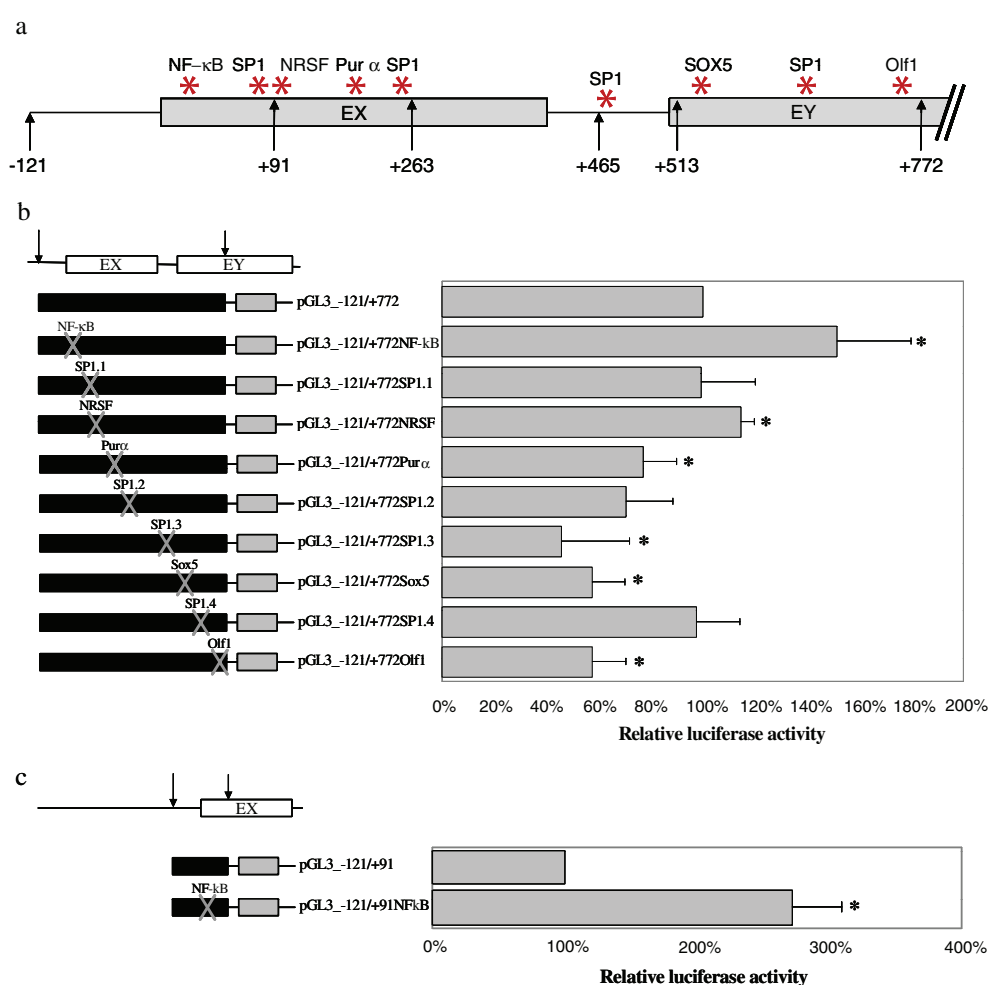
results of luciferase assays. Transfection efficiencies are normalized with the co-transfected plasmid expressing *Renilla* luciferase and luciferase activities are expressed compared to pGL3-Basic vector activity. The data are means \pm S.D. of $n \geq 3$ experiments

Identification of transcription factor binding sites by PCR site-directed mutagenesis

According to the results shown in Fig. 6, the genomic sequence -121 / +772 is likely to include positive and negative regulatory elements involved in *ST6GAL2* transcriptional regulation. Considering both general and brain-related transcription factor binding sites predicted by bioinformatics analysis, we mutated by PCR site-directed mutagenesis the NF- κ B, NRSF, Olf1, Pur α , Sox5, SP1.1, SP1.2, SP1.3 and SP1.4 binding sites (Fig. 7a). Thus, plasmids pGL3_{-121/+772}NF- κ B, pGL3_{-121/+772}NRSF, pGL3_{-121/+772}Olf1, pGL3_{-121/+772}Pur α , pGL3_{-121/+772}Sox5, pGL3_{-121/+772}SP1.1, pGL3_{-121/+772}SP1.2, pGL3_{-121/+772}SP1.3 and pGL3_{-121/+772}SP1.4 were obtained by substitution mutation of the plasmid pGL3_{-121/+772}. Mutated plasmids were transfected into SH-SY5Y cells and luciferase activity was compared to the activity of the pGL3_{-121/+772} transfected cells (Fig. 7b). The pGL3_{-121/+772}SP1.1, pGL3_{-121/+772}SP1.2 and pGL3_{-121/+772}SP1.4 constructs showed no significant

difference of luciferase activity with the plasmid pGL3_{-121/+772} (Fig. 7b), indicating that these SP1 predicted binding sites do not play a role in *ST6GAL2* gene regulation in SH-SY5Y cells. Luciferase activity of the plasmids pGL3_{-121/+772}NRSF and pGL3_{-121/+772}NF- κ B showed a 15 % and 50 % increase, respectively, compared to the pGL3_{-121/+772} plasmid (Fig. 7b), suggesting a repressive role for both factors. The plasmids pGL3_{-121/+772}Pur α , pGL3_{-121/+772}Sox5, pGL3_{-121/+772}Olf1 and pGL3_{-121/+772}SP1.3 showed a decrease in luciferase activity compared to the non-mutated plasmid, suggesting that Pur α , Sox5, Olf1 and the third SP1 binding sites and the corresponding transcription factors play an activator role in this gene regulation. Considering the result obtained with the mutated plasmid pGL3_{-121/+772}NF- κ B and since the plasmid pGL3_{-121/+91} includes the NF- κ B site, we constructed the mutated plasmid pGL3_{-121/+91}NF- κ B and tested it for luciferase assays. The result exposed in Fig. 7c showed a strong increase of luciferase activity of the mutated construct compared to control, confirming that NF- κ B would act as an inhibitory factor.

Fig. 7 Luciferase assays after transfections of the mutated plasmids in SH-SY5Y neuroblastoma cells. **(a)** Schematic representation of the -121 / +772 nucleotide sequence. Asterisks indicate the transcription factor binding sites that were mutated by PCR site-directed mutagenesis. Arrows indicate the position of restriction sites used for plasmid constructions. **(b)** On the left: a schematic representation of the different plasmid constructs. On the right: the results of luciferase assays. **(c)** On the left: a schematic representation of the different plasmid constructs. On the right: the results of luciferase assays. Transfection efficiencies are normalized with the co-transfected plasmid expressing *Renilla* luciferase and luciferase activities are expressed as a percentage of the pGL3_{-121/+772} **(b)** or pGL3_{-121/+91} **(c)** vector activity. The data are mean \pm S.D. of $n \geq 3$ experiments. *, $P < 0.05$ vs. pGL3_{-121/+772} **(b)** or pGL3_{-121/+91} **(c)**



Discussion

5' UTR of *ST6GAL2* mRNA in SH-SY5Y cells

In this study, we identified the 5' UTR of *ST6GAL2* mRNA in the SH-SY5Y neuroblastoma cell line. 5' RACE experiments showed that transcription leads to three different forms of mRNA regarding to their 5' ends. Each transcript contains one different and unique non-coding exon named EX, EY and EZ. The three exons are separated from the first common coding exon E1 by more than 42 kbp, which is, on one hand, a common feature for sialyltransferase genes [24–26]. On the other hand, these three exons are unusually close on the genomic sequence since EX and EY are separated by 124 bp and EY and EZ are separated by only 87 bp (Fig. 1c). We accurately defined the transcription start sites of EX, EY and EZ with a series of exon-specific 5' RACE experiments. While the TSS of EY and EZ are almost not submissive to variation, we showed that the transcription of EX starts in at least seven TSS over 78 bp (Fig. 2b), which is a common attribute for TATA-less promoter [27]. Moreover, the sequences of EX, EY and EZ that we have identified by 5' RACE match with partial sequences in the EST databank (GenBank accession # AB058780, AB059555, BC008680, BE612797 and BX353569). Alternative promoters and alternative 5' non-coding exons are usual for sialyltransferase genes and reflect the need for tissue- and/or time-specific expression [4,12,28], but such a closeness for the first alternative untranslated exons is rather unusual and raises the question of the biological significance of this genomic organisation. A promoter model based on more than 1.7 million human full-length cDNA 5'-end sequences, established that alternative TSS separated by at least 500 bp are theoretically sufficient to constitute mutually alternative promoters and brain was classified as the second-richest origin of putative tissue-specific alternative promoters [29]. The EX cluster of TSS being separated by 427 bp from the EY TSS cluster, itself separated by 608 bp from the EZ TSS, independent promoters could be expected. However, analysis of specific transcripts in cell lines and brain samples is not in favour of a specific regulation by independent promoters but rather indicates that alternative TSS are driven from the same large promoter region, sharing most of the regulatory cis-elements.

Expression of EX, EY and EZ containing transcripts in brain-derived cell lines and in cerebral cortex regions

The relative expression of transcripts containing either EX, EY or EZ was determined using duplex Taqman Q-PCR with the *HPRT* normalizing gene. In SH-SY5Y cells, EX and EY containing transcripts were more expressed than the

EZ containing transcript (Fig. 3). Consistent with the total expression of *ST6GAL2* gene (Data not shown), duplex Q-PCR also revealed that the highest expression of these three transcripts was detected in neuroblastoma SH-SY5Y cells and in proneuronal NT2/D1 cells, whereas their expression was low in the neuroblastoma Kelly cells and was hardly detected in the T98G and U118 glioblastoma cells. We have also shown that the three *ST6GAL2* specific transcripts were expressed in the different lobes of human cerebral cortex (Fig. 4b). In spite of differences in the expression level of *ST6GAL2* in the samples that we examined, the relative expression of *ST6GAL2* exon-specific transcripts was similar to that of neuronal cell lines (with EX>EY>EZ) suggesting a similar transcriptional regulation of *ST6GAL2* in human cerebral cortex and cultured cell lines, which is in favour of a common promoter region controlling the three transcription start sites. These results also confirm that the SH-SY5Y cell line is an adequate and relevant model for the analysis of *ST6GAL2* gene brain-specific transcriptional regulation. Previous studies reported that *ST6GAL2* is mostly expressed in brain with a low expression in the different lobes of the cerebral cortex, but no data are available concerning its expression in the different cell types [8,9]. Prior the identification of *ST6GAL2*, it had been reported that normal human neurons and glial cells did not express *ST6GAL1* [30,31]. Considering the highest expression of *ST6GAL2* in SH-SY5Y and NT2/D1 among the brain-related cell lines tested, our data suggest that *ST6GAL2* would be expressed preferentially in neurons *in vivo* and could compensate the low expression of *ST6GAL1* for the synthesis of specific α 2,6-sialylated glycoconjugates. However, glycoconjugates exhibiting SLDN or SLN terminal structures have not been yet reported in brain tissues. Further studies with cultured primary neurons and glial cells would be helpful to shed light on *ST6GAL2* cell-type specific expression. In addition, *ST6GAL2* transcripts were found to be expressed in the embryonic carcinoma NT2/D1 cells. Human and murine fetal tissues also express *ST6GAL2* [8,9]. The NT2/D1 and SH-SY5Y cell lines are well-known models for studying neuronal differentiation under retinoic acid exposition [32,33]. In such a context, it could be interesting to monitor variations in α 2,6-sialylation and in particular the presence of SLDN motif, as well as the expression of both *ST6GAL1* and *ST6GAL2* genes through neuronal differentiation. An alternative approach to answer this question would be to use murine primary neurons in culture.

Expression of hST6Gal II in brain-derived cell lines and in cerebral cortex regions

In this study, we demonstrated for the first time that hST6Gal II protein is expressed in the different lobes of the

cerebral cortex. Independently of inter-individual variations, we also showed that hST6Gal II is less expressed in the temporal lobe compared to the other cortex areas (Fig. 4c). However, the correlation between *ST6GAL2* mRNA expression and hST6Gal II protein expression remains to be established since Western blotting and Q-PCR analysis could not be performed on the same brain samples. In parallel, the protein was almost not detectable in SH-SY5Y cell line, suggesting that the level of expression of hST6Gal II in cultured cells is lower compared to human brain. Two of our previous studies concerning glycosyltransferase expression have already underlined the inter- and intra-individual variations of mRNA expression observed in this study [34,35]. Human cerebral cortex is divided into areas that are different from one to another in terms of cytoarchitecture and chemoarchitecture, functions and patterns of gene and protein expressions [36,37]. Therefore, variations of expression between the different cerebral cortex samples could be expected. Moreover, genetic variations including male-female differences, epigenetic mechanisms and aging are known to influence mRNA production [38–40] and may also explain inter-individual variations observed in our study.

Promoter regions and transcription factors

Three alternative 5' non-coding exons suggest three alternative promoters. However, taking into account the unexpected closeness of EX, EY and EZ, this would imply that regulatory elements lie within EX and/or EY sequence. Using luciferase reporter assays in SH-SY5Y cells, we showed that two regions (-121 / +91 and +465 / +772) upstream EX and within the 5' part of EY, respectively, possess a promoter activity. The first region showed the highest luciferase activity when transfected into SH-SY5Y cells. This result is consistent with Q-PCR results, which showed that EX was more expressed than EY and EZ. The second sequence is more controversial since it is partially within EY sequence. A negative regulatory region is likely to reside between these two positive regulatory regions, since the luciferase assays with plasmids containing parts or total region +92 to +464 constantly showed a decreased enzyme activity compared to plasmids without parts of this sequence: pGL3_{-121/+91} vs. pGL3_{-121/+263} or pGL3_{-121/+464} and pGL3_{+465/+772} vs. pGL3_{-121/+772} (Fig. 6). We were not able to show evidence for a promoter region leading to exon EZ transcription with luciferase assays. This may partially reflect the very low endogenous expression of EZ in SH-SY5Y cells. Concerning the potential transcription factors involved in *ST6GAL2* regulation, our investigations suggest that NRSF and NF-κB repress the transcription while Olf1, Sox5, Purα and SP1

would act as activators for the transcription of *ST6GAL2*. The SP1, Sox5 and Olf1 transcription factors are thought to act as transcription activators and are EY-related, whereas NF-κB seems to repress *ST6GAL2* transcription in a sequence upstream EX. Between the two identified promoter regions, NRSF and Purα appear to act respectively as repressor and activator of *ST6GAL2* transcription. In the central nervous system, the NF-κB family of transcription factors are involved in processes such as neuronal plasticity, myelin formation, synaptic transmission and inflammation [41]. The potential NF-κB site in the EX promoter sequence is predicted to bind homodimers of p50 proteins, which do not contain a transactivator domain, and so act as a transcription repressor [42–44]. The mutagenesis data strongly suggest that the homodimer p50/p50 is involved in the inhibition of *ST6GAL2* transcription in SH-SY5Y cells. Nevertheless, transcriptional activator elements involved in the high promoter activity of the sequence -121 / +91 (Fig. 6) remain to be identified. The neuron-restrictive silencer factor (NRSF) essentially represses expression of neuron-specific genes in non neuronal cells or tissues [45,46]. According to our results, NRSF would repress the expression of *ST6GAL2*, suggesting that the expression in SH-SY5Y cells could be disinhibited for precise neuronal-related function(s). Olf1, Sox5 and Purα are transcription factors that have been shown to be involved in neurogenesis and general neuronal functions [19,47–50] and would argue for a specific neuronal expression of *ST6GAL2*. SP1 is generally an activator factor that binds GC-rich sequences and four putative SP1 binding sites lie within the two identified *ST6GAL2* promoter sequences. Our data suggest that only the third site (SP1.3) is significantly implicated (Fig. 7b).

In the SH-SY5Y context, the expression of *ST6GAL2* is likely to be repressed by NRSF and NF-κB factors and at the same time enhanced by Sox5, Olf1, Purα and SP1 factors. As for NRSF, the transcriptional repression of NF-κB p50/p50 could potentially be removed to allow hST6Gal II to fulfil precise functions. In the light of transcription factors potentially involved in *ST6GAL2* regulation, one can hypothesize that hST6Gal II is required during brain development and/or for selective neuronal functions in adult brain. Interestingly, we previously showed that *ST6GAL2* expression is dramatically increased after IL-6 and IL-8 cytokines stimulation in the human bronchial mucosa. In this situation, neither *ST6GAL1* nor other sialyltransferase genes studied showed such an increase in their transcriptional expression [35]. This suggests that hST6Gal II would play a pivotal role that was not previously highlighted in the inflammation process in airways. In the present study, we did not find potential response elements for transcription factors involved in IL-6 and/or IL-8 response pathways. Nevertheless, five potential

NF- κ B binding sites have been identified (nucleotides -675; -485; -439; +32 and +1173). Two of these sites are supposed to bind the repressor homodimer p50/p50, and the three other sites are predicted to bind p50, c-rel and p65, the two latter known to be transcriptional activators in the context of inflammation [51].

Conclusion

In this study, we identified and characterized the 5' UTR and promoter regions of *ST6GAL2* in neuroblastoma SH-SY5Y cell line. Furthermore, we showed that *ST6GAL2* and its different transcripts are expressed in human cerebral cortex. The protein hST6Gal II was also detected in the four areas of the cerebral cortex. In the light of transcription factors potentially involved in *ST6GAL2* regulation, one can hypothesize that hST6Gal II could be required during brain development, for cell-type selective functions in adult brain and possibly in pathological conditions. Indeed, it has been previously shown that the genetic locus linked to autosomal dominant acute necrotizing encephalopathy (ADANE) contains four genes including *ST6GAL2* [52]. Although no mutation was found in the coding regions of these genes, the authors do not exclude mutations in the promoter sequences that could alter genes expression. Although variations of α 2,6-sialylation have not been investigated in ADANE, incorrect sialylation caused by deregulation of *ST6GAL2* expression may not be excluded. This work, along with further experiments, will be helpful to understand mechanisms that finely regulate *ST6GAL2* gene expression.

Acknowledgments Sylvain Lehoux is a recipient of a fellowship from the Ministère de la Recherche et de l'Enseignement. We thank le Centre National de la Recherche Scientifique and the University of Lille I.

We also thank the Lille Neurobanque, (Directed by Dr C. Libersa, France), for providing the human brain tissues samples. We are grateful to Béatrice Teylaert and Fala El-Machhour for their technical assistance.

References

- Coutinho, P.M., Deleury, E., Davies, G.J., Henrissat, B.: An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* **328**(2), 307–317 (2003)
- Harduin-Lepers, A., Mollicone, R., Delannoy, P., Oriol, R.: The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology.* **15**(8), 805–817 (2005)
- Harduin-Lepers, A., Recchi, M.A., Delannoy, P.: 1994, the year of sialyltransferases. *Glycobiology.* **5**(8), 741–758 (1995)
- Harduin-Lepers, A., Vallejo-Ruiz, V., Krzewinski-Recchi, M.A., Samyn-Petit, B., Julien, S., Delannoy, P.: The human sialyltransferase family. *Biochimie.* **83**(8), 727–737 (2001)
- Weinstein, J., Lee, E.U., McEntee, K., Lai, P.H., Paulson, J.C.: Primary structure of beta-galactoside alpha 2,6-sialyltransferase. Conversion of membrane-bound enzyme to soluble forms by cleavage of the NH₂-terminal signal anchor. *J. Biol. Chem.* **262**(36), 17735–17743 (1987)
- Paulson, J.C., Rearick, J.I., Hill, R.L.: Enzymatic properties of beta-D-galactoside alpha2 leads to 6 sialyltransferase from bovine colostrum. *J. Biol. Chem.* **252**(7), 2363–2371 (1977)
- Lance, P., Lau, K.M., Lau, J.T.: Isolation and characterization of a partial cDNA for a human sialyltransferase. *Biochem. Biophys. Res. Commun.* **164**(1), 225–232 (1989)
- Krzewinski-Recchi, M.A., Julien, S., Juliant, S., Teintener-Lelievre, M., Samyn-Petit, B., Montiel, M.D., Mir, A.M., Cerutti, M., Harduin-Lepers, A., Delannoy, P.: Identification and functional expression of a second human beta-galactoside alpha2, 6-sialyltransferase, ST6Gal II. *Eur. J. Biochem.* **270**(5), 950–961 (2003)
- Takashima, S., Tsuji, S., Tsujimoto, M.: Characterization of the second type of human beta-galactoside alpha 2, 6-sialyltransferase (ST6Gal II), which sialylates Galbeta 1, 4GlcNAc structures on oligosaccharides preferentially. Genomic analysis of human sialyltransferase genes. *J. Biol. Chem.* **277**(48), 45719–45728 (2002)
- Rohfritsch, P.F., Joosten, J.A., Krzewinski-Recchi, M.A., Harduin-Lepers, A., Laporte, B., Juliant, S., Cerutti, M., Delannoy, P., Vliegthart, J.F., Kamerling, J.P.: Probing the substrate specificity of four different sialyltransferases using synthetic beta-D-Galp-(1->4)-beta-D-GlcpNAc-(1->2)-alpha-D-Manp-(1->O)(CH₂)₇CH₃ analogues general activating effect of replacing N-acetylglucosamine by N-propionylglucosamine. *Biochim. Biophys. Acta.* **1760**(4), 685–692 (2006)
- Koles, K., Irvine, K.D., Panin, V.M.: Functional characterization of Drosophila sialyltransferase. *J Biol Chem.* **279**(6), 4346–4357 (2004)
- Kitagawa, H., Paulson, J.C.: Differential expression of five sialyltransferase genes in human tissues. *J Biol Chem.* **269**(27), 17872–17878 (1994)
- Ishii, A., Ikeda, T., Hitoshi, S., Fujimoto, I., Torii, T., Sakuma, K., Nakakita, S., Hase, S., Ikenaka, K.: Developmental changes in the expression of glycogenes and the content of N-glycans in the mouse cerebral cortex. *Glycobiology.* **17**(3), 261–276 (2007)
- Bergwerff, A.A., Van Oostrum, J., Kamerling, J.P., Vliegthart, J.F.: The major N-linked carbohydrate chains from human urokinase. The occurrence of 4-O-sulfated, (alpha 2-6)-sialylated or (alpha 1-3)-fucosylated N-acetylgalactosamine(beta 1-4)-N-acetylglucosamine elements. *Eur. J. Biochem.* **228**(3), 1009–1019 (1995)
- Dell, A., Morris, H.R., Easton, R.L., Panico, M., Patankar, M., Oehninger, S., Koistinen, R., Koistinen, H., Seppala, M., Clark, G.F.: Structural analysis of the oligosaccharides derived from glycodefin, a human glycoprotein with potent immunosuppressive and contraceptive activities. *J. Biol. Chem.* **270**(41), 24116–24126 (1995)
- Weisshaar, G., Hiyama, J., Renwick, A.G., Nimtz, M.: NMR investigations of the N-linked oligosaccharides at individual glycosylation sites of human lutropin. *Eur. J. Biochem.* **195**(1), 257–268 (1991)
- Bonfanti, L.: PSA-NCAM in mammalian structural plasticity and neurogenesis. *Prog. Neurobiol.* **80**(3), 129–164 (2006)
- Durbec, P., Cremer, H.: Revisiting the function of PSA-NCAM in the nervous system. *Mol. Neurobiol.* **24**(1–3), 53–64 (2001)
- Wang, X., Vertino, A., Eddy, R.L., Byers, M.G., Jani-Sait, S.N., Shows, T.B., Lau, J.T.: Chromosome mapping and organization of the human beta-galactoside alpha 2, 6-sialyltransferase gene. Differential and cell-type specific usage of upstream exon sequences in B-lymphoblastoid cells. *J. Biol. Chem.* **268**(6), 4355–4361 (1993)

20. Lo, N.W., Lau, J.T.: Transcription of the beta-galactoside alpha 2, 6-sialyltransferase gene in B lymphocytes is directed by a separate and distinct promoter. *Glycobiology*. **6**(3), 271–279 (1996)
21. Aas-Eng, D.A., Asheim, H.C., Deggerdal, A., Smeland, E., Funderud, S.: Characterization of a promoter region supporting transcription of a novel human beta-galactoside alpha-2,6-sialyltransferase transcript in HepG2 cells. *Biochim. Biophys. Acta*. **1261**(1), 166–169 (1995)
22. Zhang, X., Ding, L., Sandford, A.J.: Selection of reference genes for gene expression studies in human neutrophils by real-time PCR. *BMC Mol. Biol* **6**(1), 4 (2005)
23. Quandt, K., Frech, K., Karas, H., Wingender, E., Werner, T.: MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* **23**(23), 4878–4884 (1995)
24. Svensson, E.C., Soreghan, B., Paulson, J.C.: Organization of the beta-galactoside alpha 2, 6-sialyltransferase gene. Evidence for the transcriptional regulation of terminal glycosylation. *J. Biol. Chem.* **265**(34), 20863–20868 (1990)
25. Taniguchi, A., Hioki, M., Matsumoto, K.: Transcriptional regulation of human Galbeta1, 3GalNAc/Galbeta1, 4GlcNAc alpha2, 3-sialyltransferase (hST3Gal IV) gene in testis and ovary cell lines. *Biochem. Biophys. Res. Commun.* **301**(3), 764–768 (2003)
26. Yoshida, Y., Kurosawa, N., Kanematsu, T., Kojima, N., Tsuji, S.: Genomic structure and promoter activity of the mouse polysialic acid synthase gene (mST8Sia II). Brain-specific expression from a TATA-less GC-rich sequence. *J. Biol. Chem.* **271**(47), 30167–30173 (1996)
27. Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki, Y., Nakamura, Y., Suyama, A., Sugano, S.: Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites. *EMBO Rep.* **2**(5), 388–393 (2001)
28. Taniguchi, A.: Promoter structure and transcriptional regulation of human beta-galactoside alpha2, 3-sialyltransferase genes. *Curr. Drug Targets.* **9**(4), 310–316 (2008)
29. Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T., Sugano, S.: Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome. Res.* **16**(1), 55–65 (2006)
30. Kaneko, Y., Yamamoto, H., Colley, K.J., Moskal, J.R.: Expression of Gal beta 1, 4GlcNAc alpha 2, 6-sialyltransferase and alpha 2, 6-linked sialoglycoconjugates in normal human and rat tissues. *J. Histochem. Cytochem.* **43**(9), 945–954 (1995)
31. Kaneko, Y., Yamamoto, H., Kersey, D.S., Colley, K.J., Leestma, J.E., Moskal, J.R.: The expression of Gal beta 1, 4GlcNAc alpha 2, 6 sialyltransferase and alpha 2, 6-linked sialoglycoconjugates in human brain tumors. *Acta Neuropathol.* **91**(3), 284–292 (1996)
32. Andrews, P.W.: Retinoic acid induces neuronal differentiation of a cloned human embryonal carcinoma cell line in vitro. *Dev. Biol.* **103**(2), 285–293 (1984)
33. Preis, P.N., Saya, H., Nadasdi, L., Hochhaus, G., Levin, V., Sadee, W.: Neuronal cell differentiation of human neuroblastoma cells by retinoic acid plus herbimycin A. *Cancer Res.* **48**(22), 6530–6534 (1988)
34. Recchi, M.A., Harduin-Lepers, A., Boilly-Marer, Y., Verbert, A., Delannoy, P.: Multiplex RT-PCR method for the analysis of the expression of human sialyltransferases: application to breast cancer cells. *Glycoconj. J.* **15**(1), 19–27 (1998)
35. Groux-Degroote, S., Krzewinski-Recchi, M.A., Cazet, A., Vincent, A., Lehoux, S., Lafitte, J.J., Van Seuning, I., Delannoy, P.: IL-6 and IL-8 increase the expression of glycosyltransferases and sulfotransferases involved in the biosynthesis of sialylated and/or sulfated Lewisx epitopes in the human bronchial mucosa. *Biochem. J.* **410**(1), 213–223 (2008)
36. Sur, M., Rubenstein, J.L.: Patterning and plasticity of the cerebral cortex. *Science* **310**(5749), 805–810 (2005)
37. Rash, B.G., Grove, E.A.: Area and layer patterning in the developing cerebral cortex. *Curr. Opin. Neurobiol.* **16**(1), 25–34 (2006)
38. Yang, X., Schadt, E.E., Wang, S., Wang, H., Arnold, A.P., Ingram-Drake, L., Drake, T.A., Lusis, A.J.: Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome. Res.* **16**(8), 995–1004 (2006)
39. Cowley, M.J., Cotsapas, C.J., Williams, R.B., Chan, E.K., Pulvers, J.N., Liu, M.Y., Luo, O.J., Nott, D.J., Little, P.F.: Intra- and inter-individual genetic differences in gene expression. *Mamm. Genome.* **20**(5), 281–295 (2009)
40. Kerber, R.A., O'Brien, E., Cawthon, R.M.: Gene expression profiles associated with aging and mortality in humans. *Aging Cell.* **8**(3), 239–250 (2009)
41. Meffert, M.K., Baltimore, D.: Physiological functions for brain NF-kappaB. *Trends Neurosci.* **28**(1), 37–43 (2005)
42. Grundstrom, S., Anderson, P., Scheipers, P., Sundstedt, A.: Bcl-3 and NFkappaB p50-p50 homodimers act as transcriptional repressors in tolerant CD4+ T cells. *J. Biol. Chem.* **279**(9), 8460–8468 (2004)
43. Kastenbauer, S., Ziegler-Heitbrock, H.W.: NF-kappaB1 (p50) is upregulated in lipopolysaccharide tolerance and can block tumor necrosis factor gene expression. *Infect. Immun.* **67**(4), 1553–1559 (1999)
44. Kostadinova, R.M., Nawrocki, A.R., Frey, F.J., Frey, B.M.: Tumor necrosis factor alpha and phorbol 12-myristate-13-acetate down-regulate human 11beta-hydroxysteroid dehydrogenase type 2 through p50/p50 NF-kappaB homodimers and Egr-1. *FASEB J.* **19**(6), 650–652 (2005)
45. Ballas, N., Mandel, G.: The many faces of REST oversee epigenetic programming of neuronal genes. *Curr. Opin. Neurobiol.* **15**(5), 500–506 (2005)
46. Lunyak, V.V., Rosenfeld, M.G.: No rest for REST: REST/NRSF regulation of neurogenesis. *Cell.* **121**(4), 499–501 (2005)
47. Darbinian, N., White, M.K., Khalili, K.: Regulation of the Puralpha promoter by E2F-1. *J. Cell Biochem.* **99**(4), 1052–1063 (2006)
48. Khalili, K., Del Valle, L., Muralidharan, V., Gault, W.J., Darbinian, N., Otte, J., Meier, E., Johnson, E.M., Daniel, D.C., Kinoshita, Y., Amini, S., Gordon, J.: Puralpha is essential for postnatal brain development and developmentally coupled cellular proliferation as revealed by genetic inactivation in the mouse. *Mol. Cell Biol.* **23**(19), 6857–6875 (2003)
49. Lai, T., Jabaudon, D., Molyneaux, B.J., Azim, E., Arlotta, P., Menezes, J.R., Macklis, J.D.: SOX5 controls the sequential generation of distinct corticofugal neuron subtypes. *Neuron.* **57**(2), 232–247 (2008)
50. Wang, S.S., Tsai, R.Y., Reed, R.R.: The characterization of the Olf-1/EBF-like HLH transcription factor family: implications in olfactory gene regulation and neuronal development. *J. Neurosci.* **17**(11), 4149–4158 (1997)
51. Li, Q., Verma, I.M.: NF-kappaB regulation in the immune system. *Nat. Rev. Immunol.* **2**(10), 725–734 (2002)
52. Neilson, D.E., Feiler, H.S., Wilhelmsen, K.C., Lynn, A., Eiben, R.M., Kerr, D.S., Warman, M.L.: Autosomal dominant acute necrotizing encephalopathy maps to 2q12.1-2q13. *Ann. Neurol.* **55**(2), 291–294 (2004)

NOUVEAU CHAPITRE DE LA THESE

Session 2009

Sujet de thèse :

« Etude de l'expression et des mécanismes de
régulation transcriptionnelle tissu-spécifique
du gène *ST6GAL2* »

Présentée par

Sylvain LEHOUX

Université des Sciences et Technologies de Lille

Laboratoire CNRS / UMR8576

Directeurs de thèse :

Pr Philippe Delannoy
Dr Marie-Ange Krzewinski-Recchi

Soutenance le vendredi 10 juillet 2009

1. Présentation de la thèse

La glycobiole

La glycobiole est le champ de la biologie s'intéressant aux sucres, à la synthèse des structures glycaniques (enchaînement de sucres) et à leurs fonctions. Si le rôle des protéines en tant que macromolécules structurales et de l'ADN en tant que support de l'hérédité et de l'information nécessaire à la fabrication des protéines ont été appréhendés au milieu du 20^{ème} siècle, la glycobiole et l'importance des structures glycaniques n'ont émergé qu'au début des années 80s. Cependant, cette « nouvelle » discipline est sûrement l'une des plus prometteuses de ces dernières décennies en biologie.

Les protéines sont les éléments structuraux principaux des cellules et des tissus vivants. Néanmoins, pas moins de 50 % de ces protéines, que l'on pourrait comparer grossièrement à un tronc d'arbre, sont plus ou moins ramifiées par des sucres. Ces enchaînements de sucres, de quelques unités à plusieurs centaines, modifient l'interaction des protéines avec leur environnement. Ainsi, la présence de chaînes glycaniques sur des protéines exprimées à la surface des cellules va moduler de nombreuses interactions de ces protéines avec d'autres cellules, avec des protéines circulantes ou avec des agents pathogènes. La nature de ces sucres, leur nombre et l'ordre de leur enchaînement sont autant d'éléments accroissant la diversité des structures glycaniques et modifiant leurs propriétés.

Si cet aspect de la biologie avait été considéré comme superflu il y a 15 ans, son rôle primordial ne fait plus aucun doute maintenant. Des domaines médicaux aussi variés que le cancer, les infections virales, le paludisme ou les protéines recombinantes sont de plus en plus directement concernés par la glycobiole. L'étude des enzymes capables de synthétiser ces chaînes, des interactions sucres / protéines et de la structure même des complexes glycaniques sont désormais des domaines de recherche émergeant dont les retombés scientifiques, économiques et médicales sont plus que considérables.

[Pour informations complémentaires, ce documentaire sur la glycobiole, son importance et son futur diffusé sur Arte en mai 2008 : <http://video.google.fr/videoplay?docid=-7663079290156002221>]

Mon laboratoire

En 2003, l'équipe dans laquelle j'ai effectué ma thèse a mis en évidence et caractérisé chez l'Homme une enzyme capable de réaliser le transfert d'un acide sialique (un sucre) sur des protéines, selon certains critères précis. Etant donné sa récente identification, peu d'informations sont actuellement disponibles concernant le rôle physiologique et l'importance de cette sialyltransférase (enzyme qui transfère un acide sialique) pour l'Homme. Il a cependant été montré que *ST6GAL2*, le gène qui code cette sialyltransférase, est quasi exclusivement exprimé au niveau du système nerveux central embryonnaire et adulte.

Le laboratoire avait décidé peu avant mon arrivée d'orienter leurs recherches concernant *ST6GAL2* vers l'étude de la régulation transcriptionnelle, autrement dit : étudier et comprendre les mécanismes qui permettent l'expression de ce gène, en particulier au niveau du système nerveux central.

Mon sujet de thèse

Le sujet de ma thèse a été directement bâti à la suite des résultats obtenus pendant mon Master 2 Recherche. Au cours de mon M2R, j'avais identifié et isolé les séquences d'ADN par lesquelles débute la transcription du gène *ST6GAL2* (L'ADN est exprimé sous forme d'ARN messager avant d'être traduit en protéine) dans la lignée cellulaire SH-SY5Y (une lignée cellulaire humaine issue d'un cancer de cellules neuronales). J'avais aussi commencé la construction d'outils de biologie moléculaire qui me seraient utiles pendant la thèse. Mon sujet de thèse avait pour objectifs : (i) de caractériser plus en détails les séquences d'ADN à partir desquelles démarre la transcription du gène *ST6GAL2* dans notre lignée cellulaire modèle (SH-SY5Y) mais aussi pour tout un panel d'autres lignées cellulaires, (ii) d'identifier les séquences d'ADN qui influencent l'expression de *ST6GAL2* (régions d'ADN régulatrices) et (iii) d'identifier et de caractériser les protéines susceptibles de moduler cette régulation.

Ma thèse

Après 4 années sur les bancs de l'université à apprendre un savoir, certes essentiel mais trop théorique de mon point de vue, je découvrais le côté pratique. Je voyais enfin l'intérêt et l'application des cours (du moins certains) que j'avais suivis. Ce nouveau mode de travail me convenait bien d'avantage et était beaucoup plus stimulant : un sujet personnel à faire avancer et à défendre. Je voyais et je vois toujours la thèse comme une excellente transition entre le monde universitaire et le monde du travail : je ne quittais pas de façon nette l'université mais je me rapprochais un peu plus de la vie active avec un sujet d'étude dans lequel il faut s'investir, une mise en pratique des cours reçus jusque là et un salaire.

J'avais effectué mon Master 2 Recherche au sein de cette unité (CNRS / UMR8576) dans le laboratoire du professeur Philippe Delannoy sous la tutelle du docteur Marie-Ange Krzewinski-Recchi. Le Master 2 avait été pour moi un grand changement. Malgré mon inexpérience du monde du laboratoire, je n'avais jusqu'alors fait aucun stage dans un laboratoire, j'assimilais et je progressais rapidement. En fin d'année, je pus présenter un sujet, la suite de mon sujet de M2R, pour obtenir une bourse de thèse financée par le Ministère de la Recherche et continuer dans ce laboratoire. Les contacts que nous avons pris au cours de la première année (avec l'U837, basée au CHR et travaillant sur la maladie d'Alzheimer) nous ont permis d'obtenir d'une part différents outils et « matières premières » (lignées cellulaires que nous ne possédions pas, extraits protéiques, ARN...), et d'autre part d'avoir un autre point de vue sur mon sujet avec une approche plus « cérébrale » et moins « glycobiochimie ». Voici une partie des aboutissements de ces 3 ans de thèse.

2. Les enjeux de la thèse

Scientifiques

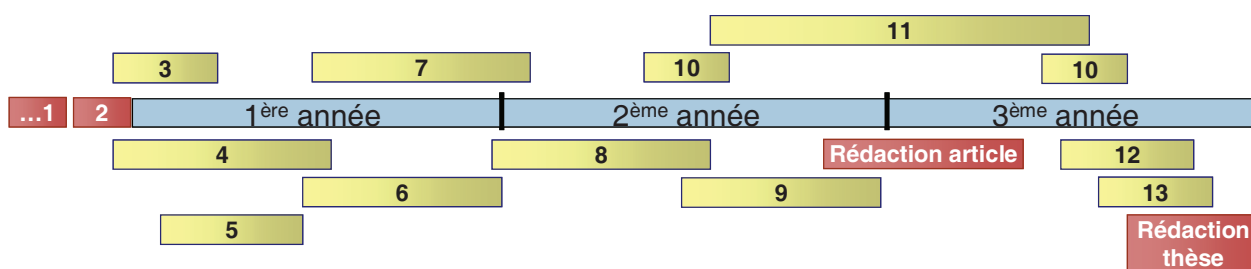
A la suite de mes travaux de M2R, ma thèse avait pour objectif de définir les régions de l'ADN indispensable au contrôle de l'expression du gène *ST6GAL2* ainsi que les protéines intervenant dans cette régulation. Mon modèle d'étude principal était les cellules neuronales humaines appelées SH-SY5Y. La collaboration avec l'U837 me permit d'obtenir plusieurs autres sources de matériel biologique (cellules et tissus). Dans une des lignées cellulaires obtenues (NT2/D1), les études que j'ai menées ont révélé que le gène *ST6GAL2* était très

fortement surexprimé lors de la différenciation en neurone de cette lignée. Cette forte surexpression du gène laissait pour la première fois présager d'un rôle important lors d'un processus biologique, ce qui donnait ici un autre intérêt à l'étude de la régulation de l'expression de *ST6GAL2* et un aspect un peu plus fonctionnel à l'étude de ce gène au cours de ma thèse.

Un autre enjeu scientifique de ma thèse était la publication dans un journal scientifique d'au moins un article exposant mes travaux, un pré requis de l'école doctorale avant toute soutenance de thèse en biologie.

Echéancier de la thèse

Etapes	Dénomination	Durée (mois)
Initialisation	1 Production, assimilation et conclusion des résultats du M2R	12
	2 Planification et organisation de la thèse	1
Réalisation	3 Finalisation des travaux menés en M2R	2
	4 Construction d'outils en biologie moléculaire - 1 ^{ère} partie	6
	5 Etudes et analyses bioinformatiques	3
	6 Tests sur cellules en culture à partir des 1 ^{ers} outils construits	6
	7 Mise au point de la technique de PCR en temps réel avec la technologie Taqman - Application aux cellules en culture	6
	8 Construction d'outils en biologie moléculaire - 2 ^{ème} partie	6
	9 Tests sur cellules en culture à partir des seconds outils construits	6
	10 Application de la technique de PCR en temps réel avec les technologies SYBR Green et Taqman aux tissus cérébraux	3
	11 Induction de la différenciation en neurones des cellules NT2 et étude de ses résultats	10
	12 Mise au point et analyses protéiques de protéines extraites de cellules en culture et de tissus cérébraux	4
	13 Mise au point d'une technique de biologie moléculaire de précipitation de l'ADN (ChIP)	2
Bilan	14 Rédaction de thèse / Article scientifique	7



Il s'agit ici de l'échéancier réel qui diffère quelque peu de celui établi lors de la présentation du projet pour l'obtention des bourses de thèses. Certaines choses n'ont pas été faites, mais d'autres, grâce notamment à la collaboration avec l'U837, ont pu être réalisées et ont permis à ma thèse de prendre une tournure plus intéressante et concrète avec la différenciation en neurones des cellules NT2 et l'étude des ARN et des protéines issus d'échantillons de cerveau notamment.

Budget de la thèse

	Coût (€)
Ressources Humaines	112517
Doctorant	83575
Encadrant	28942
Matériel expérimental	29995
Consommables Biologie Moléculaire	16069
Consommables Biologie Cellulaire	9934
Consommables Protéomique	2392
Amortissement appareils	1600
Matériel de bureau	400
Ordinateurs, logiciels	300
Fourniture, papier, imprimantes, cartouches...	100
Frais de déplacement et représentation	2950
Déplacement	2200
Congrès	750
Frais de formation	400
Autres, communication, impression posters, courrier	100
TOTAL	146362 €

Le chiffre total peut paraître important au regard des retombés scientifiques directes (publications, brevets, avancées scientifiques majeures, etc...). Néanmoins, mes travaux de thèse contribuent indiscutablement à une meilleure compréhension et une Connaissance plus approfondie des processus biologiques chez l'Homme. L'importance de la glycosylation dans le fonctionnement de nombreux mécanismes et interactions cellulaires n'étant plus à démontrer, mon travail pose de solides bases concernant l'étude de la régulation de l'expression d'une enzyme fortement pressentie comme impliquée dans le développement du système nerveux central.

Les 2/3 des dépenses financières sont consommées en ressources humaines. Dans cette optique, ce coût représente aussi le coût de ma formation scientifique durant ces 3 années, aussi bien sur le plan du savoir-faire technique et expérimental que sur le plan du savoir et du savoir être (réflexion scientifique, rédaction, communication orale). Par ailleurs, au cours de ma thèse le coût global des appareils expérimentaux utilisés reste relativement faible compte tenu de leurs utilisations, communes à l'unité, et donc avec un coût d'amortissement faible.

3. Compétences acquises et valorisées

Au cours de cette thèse j'ai appris et développé de nombreuses compétences, qui se divisent en deux catégories : un champ de compétence directement lié à la glycobiologie et aux techniques expérimentales en rapport avec mon sujet de thèse : connaissances approfondies en glycobiologie, techniques de manipulations expérimentales, savoir-faire théorique et pratique associé aux méthodologies développées au cours de ma thèse... et un deuxième champ, moins liée à mon sujet de thèse en biologie, mais plus liée ma formation personnelle : rigueur scientifique, expression et communication, vie d'équipe, développement de projets à court et moyen termes, confiance en soit...

1. COMPETENCES SCIENTIFIQUES

« Glycoscience »

Jusqu'à mon Master 2 Recherche, mon parcours scientifique était plutôt accès génétique et moins, voire pas du tout, biochimie et glycobiologie. Cependant, au cours de cette thèse j'ai découvert tout un aspect de la biologie que je ne connaissais pas vraiment. La

glycobiologie est un domaine encore relativement récent par comparaison d'autres facettes de la biologie comme la génétique ou la protéomique. Ici au laboratoire, j'ai vraiment pu enrichir mes connaissances dans un domaine de recherche attrayant. Même si mes propres recherches sont fondamentales, les connaissances que j'ai acquises me permettent d'avoir **une expertise dans une spécialité pointue**. Mon savoir général en biologie a également progressé bien plus rapidement ici au laboratoire au contact des différentes personnes de l'unité à travers de nombreux échanges.

Techniques expérimentales

Biologie moléculaire :

Manipulation d'acides nucléiques (ADN et ARN). Durant ma thèse, j'ai très largement développé un panel de compétences techniques relatives à l'extraction, la purification et la manipulation des acides nucléiques. Certaines de ces techniques ont nécessité des mises au point, comme la technique de **PCR en temps réel avec la technologie Taqman ou la technique de 5'RACE**. Dans ces deux cas, ceci m'a permis de renforcer mes connaissances de ces techniques en **alliant savoir théorique et savoir-faire**. Par ailleurs, j'ai eu l'occasion de **transmettre en partie ces connaissances** à des étudiants lors de stages au laboratoire.

Biologie cellulaire :

J'ai également beaucoup appris sur les pratiques concernant la biologie cellulaire. En effet, la source de « matériel » nécessaire à mes expérimentations provenait essentiellement de cellules en culture, j'ai donc acquis toutes **les connaissances et compétences nécessaires aux bonnes pratiques de culture cellulaire** : entretien des cellules, entretien de la hôte de culture, règles d'hygiène et sécurité en salle de culture (niveau P2 universitaire) et manipulation de cellules pour différentes expérimentations.

Protéomique (techniques relatives aux protéines) :

Les techniques concernant la protéomique sont parmi les dernières que j'ai apprises, je n'en ai pour le moment qu'une bonne connaissance des bases. Je **maitrise cependant l'extraction de protéines**, provenant de cellules en culture ou de tissus, ainsi que **l'identification de protéines par anticorps par Western Blot** (révélation d'une protéine par l'utilisation d'un anticorps spécifique).

2. COMPETENCES PERSONNELLES

Relationnel et Communication

Communication :

Durant ma thèse, j'ai eu de nombreuses occasions **d'exposer mes travaux** : à l'oral ou sous forme de poster, en français ou en anglais. Si parler en public ne me posait pas de problème, **exposer à l'oral mon travail et le défendre devant une audience ou un jury** m'a permis de **renforcer ma confiance en moi**. Ce type d'exercice m'a également permis **de prendre plus d'assurance**, y compris en dehors du laboratoire ou du domaine « scientifique glycobiochimie ».

Enseignement :

Durant ma première année de thèse j'ai pu m'essayer à l'enseignement. J'ai donné une dizaine d'heures de TD dans le module Agroalimentaire & Nutrition au CNAM de Lille. L'expérience était, quelque peu, particulière puisqu'il s'agissait de donner des cours à des personnes qui, en moyenne, avaient 10 à 15 ans de plus que moi (bien différent du cadre du monitorat où l'on enseigne à des étudiants des premières années de Fac). Cela m'a cependant permis de m'exercer à l'enseignement sur le plan théorique (cours en salle). D'un point de vue pratique, **j'ai eu également l'occasion d'expliquer et de montrer mon savoir-faire** à des étudiants venus en stage au laboratoire. Cela m'a appris à **simplifier et expliquer** l'objet de mes expériences pour que l'étudiant en question puisse comprendre et reproduire la manipulation.

Relations humaines :

Le laboratoire, et l'unité entière même, est composée de personnes de différents horizons. Cette mixité, avec les personnes de mon laboratoire au quotidien principalement, impose ses règles de vie en communauté. La vie au laboratoire a **développé mon esprit d'équipe** et m'a incité à l'échange de compétences. En effet, chacun possède souvent une ou plusieurs spécialités, en plus de l'expérience de l'ancienneté. Dans ce cadre là, j'ai aussi pu réaliser une collaboration, qui m'a non seulement apporté des ressources matérielles mais qui m'a appris à appréhender d'autres points de vue, d'autres personnes et d'autres manières de travailler. Bien que je sois de nature plutôt réservé, ces échanges ne m'ont posé aucun souci particulier et ont été très enrichissants.

Les différents congrès nationaux et internationaux auxquels j'ai pu participer ont aussi été très formateurs dans **l'élargissement de mon horizon scientifique et personnel**. J'ai pu y développer des contacts et améliorer ma relation avec d'autres scientifiques / biologistes.

Organisation et adaptation

Organisation :

Pendant ces 3 années, j'ai du apprendre à **organiser mes idées : développer une stratégie** pour obtenir un résultat précis, à partir d'éléments de départ et ensuite interpréter justement les résultats. La planification de la thèse elle-même ainsi que toute une série d'expérimentations plus ou moins longues (de quelques jours à plusieurs semaines ou mois) ont contribué à **développer mon sens de l'organisation et du timing** (à court et long terme).

Flexibilité, réflexion et ajustement :

En biologie, tout ne se passe pas toujours comme prévu ou pas aussi rapidement que prévu. J'ai appris à **me remotiver** lorsque certaines manipulations ne donnaient pas du tout le résultat attendu et ce même après plusieurs essais. J'ai aussi appris à réfléchir aux résultats obtenus après une expérience, à en **tirer les bonnes conclusions** pour réajuster correctement la prochaine expérience. J'ai acquis cette **rigueur scientifique** indispensable grâce aux nombreux conseils dont j'ai pu bénéficier tout au long de ma thèse par toutes les personnes du laboratoire, aussi bien sur le plan théorique, pratique et personnel.

Capacité de synthèse :

La période où je donnais des cours CNAM m'a aussi été très profitable d'un point de vue de **l'organisation et la synthèse**. Même si les grandes lignes des TD étaient tracées je ne reprenais pas les cours d'une des années précédentes et j'ai donc du préparer mes propres cours à partir d'informations récupérées et/ou trouvées sur Internet par exemple. J'ai ainsi travaillé mes **capacités de synthèse afin de présenter le meilleur condensé** possible.

La rédaction du mémoire de Master 2 Recherche, du projet de thèse, de divers résumés de mon sujet, de l'introduction de ma thèse ou encore les présentations orales dans un temps limité sont autant d'exemples où la capacité de synthèse est de rigueur. C'est donc une qualité que j'ai du développer tout au long de ces années de doctorat.

4. Bilan de ces trois années de thèse

Ces trois années de thèse m'ont beaucoup appris et apporté à plusieurs niveaux.

Bilan scientifique :

Les résultats que j'ai obtenus offrent une meilleure vision de la manière dont est exprimé et régulé le gène *ST6GAL2* dans la cellule SH-SY5Y et dans différents tissus cérébraux. J'ai montré que le gène et la protéine correspondante sont exprimés au niveau des différents lobes du cortex cérébral avec une certaine variation inter-individuelle. Enfin, lors de la différenciation des cellules NT2/D1 en neurones, j'ai pu montrer que *ST6GAL2* et la protéine qu'il code sont fortement surexprimés, laissant présager d'un rôle important lors du développement du système nerveux central.

Au-delà des résultats obtenus, de l'article soumis à un journal et de la thèse en elle-même, j'ai beaucoup appris et énormément progressé du point de vue de la rigueur et de la réflexion scientifique. J'ai un meilleur point de vue et un meilleur recul sur mon sujet mais aussi sur la Recherche scientifique de façon plus générale. Je maîtrise maintenant plusieurs techniques de biologie moléculaire et cellulaire surtout. Dans ces domaines je pense posséder une vraie expertise et être capable de la transmettre.

Bilan humain :

Ces années m'ont permis de m'affirmer d'avantage sur un plan personnel. Je suis convaincu d'avoir développé des compétences et des qualités autres que « scientifico-biologiques ». J'ai acquis une plus grande confiance en moi, notamment grâce aux différentes présentations orales, en français et en anglais, que j'ai pu faire. J'ai également appris à travailler en équipe et à collaborer avec différentes personnes extérieures au laboratoire. Enfin, cette période de la thèse m'aura aidé à définir un peu plus mon projet professionnel, par le biais des Doctoriales et du Nouveau Chapitre de la Thèse plus particulièrement.

5. Objectif professionnel

A courts termes, j'aimerais effectuer un post-doctorat à l'étranger afin d'élargir mes champs de compétences, améliorer mon anglais et profiter d'une expérience dans un autre

environnement scientifique pour y perfectionner et/ou y apprendre de nouvelles techniques. Actuellement, ma meilleur piste pour un post-doctorat se situe aux Etats-Unis au sein d'un laboratoire de glycobiochimie. La concrétisation de cette opportunité me permettrait de rester dans un domaine de pointe en biologie.

A moyen terme, mon souhait est de rejoindre une entreprise privée et d'y intégrer une branche R&D. Je pense que c'est à l'intérieur de ce type de structure que je me sentirai le plus à l'aise. Idéalement, je me vois travailler au sein d'une équipe d'une dizaine de personnes et mener à bien des projets avec des objectifs précis. Par ailleurs, je suis certain que les compétences scientifiques et personnelles développées ou acquises au cours de cette thèse me permettront de poursuivre mon parcours professionnel.

Les premières recherches d'emploi que j'ai pu mener avec cet objectif m'ont permis de trouver plusieurs offres dont le profil correspond à mes compétences et mes attentes. Des sociétés telles que Alcedo, le LFB, Biomerieux, ERYtech pharma ou Imaxio avec des postes d'ingénieur de recherche ou attaché de recherche sont des entreprises susceptibles de m'intéresser.

Sialylation is one of the last step of the biosynthesis of glycan chains carried by glycoproteins and glycolipids. The α 2,6-sialylation of N-acetyllactosaminyl (Gal β 1-4GlcNAc) structures is commonly found at the end of glycan chains and is involved in numerous cell / cell or host / pathogen adhesion and recognition events. In Human, two sialyltransferases synthesise this glycan epitope, namely hST6Gal I and hST6Gal II. They differ from each other in substrate specificity and in tissue-specific pattern of expression. Whereas the gene encoding hST6Gal I, *ST6GAL1*, is expressed in almost all tissues, *ST6GAL2* shows a narrower pattern of tissue expression essentially limited to fetal and adult brain. In addition, hST6Gal II exhibits similarities in terms of substrate specificity and gene expression pattern with the sialyltransferase identified in *D. melanogaster* and therefore, seems to have conserved ancestral properties required for brain function and growing nervous tissue. Several studies have shown that the expression of sialyltransferases is controlled at the transcriptional level by the use of specific promoters that regulate their expression in a tissue-specific fashion. Data about *ST6GAL2* are rather limited; however, it appears the expression of this gene is finely regulated by mechanisms likely conserved through evolution. The aim of this thesis was to identify the 5' non translated regions of the *ST6GAL2* gene and to characterize the associated promoter regions. From a neuroblastoma cultured cell model, we identified by 5'RACE three types of transcripts which are different only in their first non-coding exon. Those exons, named EX, EY and EZ, are located more than 42 kbp upstream of the first common coding exon and are only separated by 124 and 87 bp, respectively. Using Taqman duplex Q-PCR technology we have shown that the transcripts initiated by EX and EY are predominantly expressed compared to EZ both in several cell lines and in human brain tissue samples. We also demonstrated that the hST6Gal II protein is expressed in the different lobes of the human cerebral cortex, the cerebellum and the hippocampus. We isolated different genomic sequences upstream EX and within EX/EY/EZ region and inserted them in a reporter vector for luciferase assays. We could define two promoter sequences upstream EX and ZY. PCR site-directed mutagenesis experiments along with bioinformatics analysis revealed that transcription factors NF- κ B and NRSF are likely to act as transcription inhibitors, whereas the Sox5, SP1, Pura and Olf1 factors would be involved in the transcriptional activation of *ST6GAL2*. The NRSF, Sox5, Pura and Olf1 transcription factors are notably involved in the transcriptional regulation of genes related to neuronal functions and the neuronal development. Eventually, we have shown evidence of a strong increased *ST6GAL2* expression during neuronal differentiation of the NT2/D1 cell line under acid retinoic treatment, suggesting of putative role this enzyme in neuronal differentiation.

Key words: *ST6GAL2*, sialyltransferases, transcriptional regulation, human cerebral cortex, Sia₆LacdiNAc

La sialylation est l'une des dernières étapes de la biosynthèse des chaînes glycaniques des glycoprotéines et des glycolipides. La sialylation en $\alpha 2,6$ des structures N-acétyllactosaminiques (Gal β 1-4GlcNAc) est souvent retrouvée en périphérie des glycannes et est impliquée dans de nombreux mécanismes d'adhésion et de reconnaissance cellule / cellule ou hôte / pathogène. Chez l'Homme, deux sialyltransférases synthétisent ce type d'épitope glycanique : hST6Gal I et hST6Gal II. Elles se distinguent par leur spécificité de substrat accepteur et par leur profil d'expression tissulaire. Alors que le gène *ST6GAL1* codant hST6Gal I est exprimé dans la plupart des tissus, *ST6GAL2* présente une expression tissulaire plus restreinte, se limitant essentiellement au cerveau embryonnaire et adulte. Par ailleurs, hST6Gal II présente des similitudes en termes de spécificité de substrat et d'expression tissulaire avec la sialyltransférase identifiée chez *D. melanogaster* et semble avoir conservé certaines propriétés ancestrales essentielles pour le développement du tissu nerveux. Plusieurs études ont montré que l'expression des sialyltransférases est contrôlée au niveau transcriptionnel par l'utilisation de promoteurs tissulaires régulant l'expression de manière tissu-spécifique. Si les données concernant *ST6GAL2* sont encore limitées, il apparaît cependant que l'expression de ce gène est finement contrôlée par des mécanismes apparemment conservés au cours de l'évolution. Le projet de thèse que nous avons mené a eu pour but d'identifier les régions 5'-non traduites de *ST6GAL2* et de caractériser les régions promotrices associées. A partir d'un modèle cellulaire de neuroblastome en culture, nous avons identifié par 5' RACE trois types de transcrits qui diffèrent par leur premier exon non traduit. Ces exons, appelés EX, EY et EZ, sont situés à plus de 42 kpb du premier exon commun codant et ne sont séparés que de 124 et 87 pb, respectivement. Par Q-PCR en duplex avec le gène normalisateur HPRT, nous avons montré que les transcrits initiés par l'exon EX et EY étaient prépondérants par rapport aux transcrits contenant EZ, à la fois dans plusieurs lignées cellulaires à caractère neuronal et dans des échantillons de tissu cérébral humain. Nous avons également montré que la protéine hST6Gal II est exprimée dans les différents lobes du cortex cérébral, dans le cervelet et dans l'hippocampe. Nous avons isolé différentes régions génomiques situées en amont et à l'intérieur de la région EX/EY/EZ que nous avons sous cloné en amont du gène de la luciférase pour des tests d'activité. Nous avons défini deux régions promotrices, en amont des exons EX et EY. Des expériences de mutagenèse dirigée couplées à des analyses bioinformatiques nous ont révélé que les facteurs de transcription NF- κ B et NRSF sont probablement des répresseurs de la transcription, alors que les facteurs Sox5, SP1, Pur α et Olf1 agirait comme des éléments activateurs de la transcription de *ST6GAL2*. Les facteurs NRSF, Sox5, Pur α et Olf1 régulent notamment la transcription de gènes impliqués dans le fonctionnement et le développement neuronal, suggérant un rôle de *ST6GAL2* dans les fonctions neuronales. Enfin, nous avons mis en évidence une forte augmentation de l'expression *ST6GAL2* au cours de la différenciation en neurones des cellules NT2/D1 sous l'action de l'acide rétinoïque, suggérant un rôle potentiel de cette enzyme au cours de la différenciation neuronale.

Mots clés : *ST6GAL2*, sialyltransférases, régulation transcriptionnelle, cortex cérébral humain, Sia₆LacdiNAc