



**HAL**  
open science

# Analyses biochimiques et fonctionnelles de protéines cibles de POFUT1

Florian Pennarubia

► **To cite this version:**

Florian Pennarubia. Analyses biochimiques et fonctionnelles de protéines cibles de POFUT1. Médecine humaine et pathologie. Université de Limoges, 2018. Français. NNT : 2018LIMO0075 . tel-03065825

**HAL Id: tel-03065825**

**<https://theses.hal.science/tel-03065825v1>**

Submitted on 15 Dec 2020

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

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

## Université de Limoges

ED 615 - Sciences Biologiques et Santé (SBS)

PEIRENE (EA 7500), équipe «Glycosylation et différenciation cellulaire »

Thèse pour obtenir le grade de

### Docteur de l'Université de Limoges

Aspects moléculaires et cellulaires de la biologie

Présentée et soutenue par

**Florian PENNARUBIA**

Le 14 décembre 2018

# Analyses biochimiques et fonctionnelles de protéines cibles de POFUT1

Thèse co-dirigée par le Dr. **Sébastien LEGARDINIER** et le Pr. **Abderrahman MAFTAH**

## JURY

### Président du jury

**Jean-Michel PETIT**, Professeur ; Université de Limoges

### Rapporteurs

**Tony LEFEBVRE**, Professeur ; Université de Lille 1

**Mohamed OUZZINE**, Directeur de Recherche INSERM ; Université de Lorraine

### Examineurs

**Amalia TROUSSON**, Maître de Conférences ; Université Clermont Auvergne

**Sébastien LEGARDINIER**, Maître de Conférences ; Université de Limoges

**Abderrahman MAFTAH**, Professeur ; Université de Limoges





*L'imagination est plus importante que le savoir.*  
**Albert Einstein**



## REMERCIEMENTS

---

Je voudrais tout d'abord remercier Monsieur **Tony LEFEBVRE**, Professeur à l'Université de Lille 1, et Monsieur **Mohamed OUZZINE**, Directeur de recherche INSERM à l'Université de Lorraine de l'honneur qu'ils m'ont fait en acceptant d'être rapporteurs de cette thèse. Je tiens également à remercier Madame **Amalia TROUSSON**, Maître de Conférences à l'Université Clermont Auvergne qui a bien voulu être examinatrice de ce travail, ainsi que Monsieur **Jean-Michel PETIT**, Professeur à l'Université de Limoges, qui a accepté de présider mon jury de soutenance de thèse.

J'adresse mes plus sincères remerciements à mes deux co-directeurs de thèse, Monsieur **Sébastien LEGARDINIER**, Maître de Conférences à l'Université de Limoges et Monsieur **Abderrahman MAFTAH**, Professeur à l'Université de Limoges pour leur encadrement, leur aide et leur disponibilité. Je tiens aussi à souligner leurs qualités humaines et scientifiques qui m'ont permis de réaliser cette thèse dans les meilleures conditions.

J'exprime ma gratitude à Madame **Agnès GERMOT**, Maître de Conférences à l'Université de Limoges pour sa contribution à mes recherches.

Je souhaite aussi remercier Madame **Emilie PINAULT**, Ingénieur d'études à l'Université de Limoges pour son apport important à mes travaux.

Je remercie Monsieur **Bilal AL JAAM**, ancien doctorant de l'Université de Limoges, avec qui j'ai collaboré lors de mes premiers travaux de thèse.

Je tiens à remercier Monsieur **Daniel PETIT**, Professeur émérite à l'Université de Limoges, pour toutes nos discussions et ses conseils.

Je remercie tous les doctorants et post-doctorants avec qui j'ai partagé ces années de thèse pour leur aide. J'aimerais remercier plus particulièrement Monsieur **Julien CHABANAIS** pour ses suggestions et nos échanges scientifiques.

Enfin, je tiens à remercier tous les membres du personnel technique qui ont contribué d'une façon ou d'une autre à cette thèse.

Mes derniers remerciements vont à l'ensemble des membres du laboratoire PEIRENE que j'ai côtoyé durant ces trois années.

Merci à tous.



# TABLE DES MATIERES

---

<b>RESUME</b> .....	<b>17</b>
<b>ABSTRACT</b> .....	<b>19</b>
<b>PUBLICATIONS ET COMMUNICATIONS</b> .....	<b>21</b>
<b>AVANT-PROPOS</b> .....	<b>25</b>
<b>EXPOSE BIBLIOGRAPHIQUE</b> .....	<b>29</b>
<b>Chapitre I. La glycosylation</b> .....	<b>31</b>
I.1. La <i>N</i> -glycosylation .....	33
I.2. La <i>O</i> -glycosylation .....	35
I.3. La <i>C</i> -glycosylation .....	39
<b>Chapitre II. La <i>O</i>-fucosylation</b> .....	<b>41</b>
II.1. La <i>O</i> -fucosylation dépendante de Pofut1 .....	41
II.1.1. Structure de l'enzyme Pofut1 .....	41
II.1.2. Les ELDs comme substrats accepteurs.....	43
II.1.3. L'interaction des ELDs avec Pofut1 .....	45
II.1.4. Activité catalytique de Pofut1.....	47
II.2. La <i>O</i> -fucosylation dépendante de Pofut2.....	51
II.3. Autres domaines <i>O</i> -fucosylables .....	53
II.4. Le rallongement des <i>O</i> -fucoses.....	53
<b>Chapitre III. Les protéines <i>O</i>-fucosylées</b> .....	<b>55</b>
III.1. Le récepteur Notch .....	55
III.1.1. La voie de signalisation induite par Notch .....	57
III.1.2. La <i>O</i> -fucosylation de NOTCH1 .....	59
III.2. Les autres protéines <i>O</i> -fucosylées .....	61
III.3. Les protéines potentiellement cibles de Pofut1 .....	61
III.4. La protéine WIF1.....	63
III.4.1. Un inhibiteur de la voie Wnt .....	63
III.4.2. Structure de la protéine WIF1 .....	65
III.4.3. Le rôle des ELDs de WIF1 .....	67
III.4.4. L'importance biologique de WIF1 .....	67
<b>PROJET DE THESE</b> .....	<b>71</b>
<b>RESULTATS</b> .....	<b>75</b>
Publication N°1 .....	77
Publication N°2 .....	117
Publication N°3 .....	229
<b>DISCUSSION GÉNÉRALE ET PERSPECTIVES</b> .....	<b>305</b>
<b>REFERENCES BIBLIOGRAPHIQUES</b> .....	<b>337</b>
<b>ANNEXE</b> .....	<b>355</b>



## LISTE DES ABREVIATIONS

---

- ADAM:** A Disintegrin and Metallo-proteinase
- AMACO:** A domain-containing protein similar to Matrilin and Collagen
- ANK:** Ankyrin
- APC:** Adenomatous Polyposis Coli
- CAZy:** Carbohydrate-Active enZYmes
- cEGF:** C1r-like EGF
- CHO:** Chinese hamster ovary
- CNX:** Calnéxine
- CRD:** Cysteine Rich Domain
- CRT:** Calréticuline
- CSL:** CBF-1/RBPJ- $\kappa$ /Suppressor of Hairless/Lag-1
- CuAAC:** Copper-Catalyzed Azide-Alkyne Cycloaddition
- DLL:** Delta-like
- DOS:** Delta and OSM-11 like
- Dsh:** Dishevelled
- DSL:** Delta/Serrate/Lag-2
- EGF:** Epidermal Growth Factor
- ELD:** EGF-like domain
- EOGT:** EGF domain-specific *O*-linked *N*-acetylGlucosamine Transferase
- ERAD:** Endoplasmic Reticulum-Associated Degradation
- Fuc:** Fucose
- Gal:** Galactose
- GalNac:** *N*-Acétylgalactosamine
- GDP:** Guanosine 5'-Diphosphate
- Glc:** Glucose
- GlcNac:** *N*-Acétylglucosamine
- GSK-3 $\beta$ :** Glycogen synthase kinase 3
- HD:** Heterodomain
- hEGF:** human EGF-like



**Hes:** Hairy/Enhancer of Split  
**Hey:** Hairy/Enhancer of Split-related with YRPW Motif  
**HSPG:** Glypican heparan sulfate proteoglycans  
**intEGF:** Intégrines EGF-like  
**JAG:** Jagged  
**KDEL:** Lysine-Aspartate-Glutamate-Leucine  
**KO:** Knockout  
**lamEGF:** Laminines EGF-like  
**LFNG:** Lunatic fringe  
**LNR:** Lin-12/Notch Repeats  
**LRP5/6:** Low-density lipoprotein receptor-related proteins 5 or 6  
**Man:** Mannose  
**MAM:** Mastermind  
**MDCS:** Myoblastes Dérivés de Cellules Satellites  
**MEC:** Matrice extracellulaire  
**MRM:** Multiple reaction monitoring  
**NECD:** Notch Extracellular Domain  
**NICD:** Notch Intracellular Domain  
**NLS:** Nuclear Localization Signal  
**NRR:** Negative regulatory region  
**NTR:** Netrin domain  
**Pax:** Paired box protein  
**PDI:** Protéine disulfide isomérase  
**PEST:** Proline-Glutamate-Sérine-Thréonine  
**PNGase:** Peptide *N*-glycosidase  
**Poglut:** Protein *O*-glucosyltransferase  
**Pofut:** Protein *O*-fucosyltransferase  
**PPI:** Protéine prolyl isomérase  
**RAM:** RBPj Associated Molecule  
**RE:** Réticulum Endoplasmique  
**sFRP:** Secreted Frizzled-related protein  
**TAD:** Transactivation Domain



**TCF/LEF:** T Cell factor/Lymphoid enhancer factor

**TMD:** Transmembrane domain

**Tn Ag:** Tn antigen

**TSR:** Trombospondin Type 1 Repeat

**UDP:** Uridine Diphosphate

**UGGT1:** UDP-Glucose Glycoprotein Glucosyltransferase 1

**WIF1:** Wnt inhibitory factor 1

**WT:** Wild type



## LISTE DES FIGURES ET DES TABLEAUX

---

Figure 1 : Schéma simplifié du contrôle du repliement des glycoprotéines dans le RE.....	32
Figure 2 : Schéma simplifié de la biosynthèse des différents types de <i>N</i> -glycanes. ....	34
Figure 3 : Biosynthèse des <i>O</i> -glycanes de type mucine les plus communs. ....	36
Figure 4 : Représentation des <i>O</i> -glycanes les plus communs. ....	36
Figure 5 : Biosynthèse et structure de la <i>C</i> -mannosylation. ....	38
Figure 6 : Schéma de la protéine POFUT1 murine. ....	40
Figure 7 : Structure caractéristique des repliements GT-A et GT-B. ....	42
Figure 8 : Structure cristallisée de la protéine POFUT1 murine. ....	42
Figure 9 : Les différents types d'ELDs.....	42
Figure 10 : Les <i>O</i> -glycosylations des ELDs. ....	44
Figure 11 : Interactions entre les résidus de POFUT1 et de l'ELD (+). ....	44
Figure 12 : Mécanisme catalytique de l'enzyme POFUT1 murine (type $S_N1$ ). ....	48
Figure 13 : Représentation de la <i>O</i> -fucosylation de TSR de type 1. ....	50
Figure 14 : Différence structurale entre les enzymes POFUT1 et POFUT2 humaines.....	50
Figure 15 : Mécanisme catalytique de l'enzyme Pofut2 ( <i>C.elegans</i> ) (type $S_N2$ ). ....	50
Figure 16 : L'extension des <i>O</i> -fucoses par d'autres monosaccharides. ....	52
Figure 17 : Structures des différents récepteurs Notch. ....	54
Figure 18 : Structures des différents ligands canoniques des récepteurs Notch.....	56
Figure 19 : La voie de signalisation Notch. ....	56
Figure 20 : La <i>O</i> -fucosylation du récepteur NOTCH1. ....	58
Figure 21 : Positions des <i>O</i> -fucoses de NOTCH1 favorisant l'interaction avec les ligands DSL. ....	58
Figure 22 : Fringe module l'activation de Notch par ses ligands DSL. ....	60
Figure 23 : Liste des protéines humaines potentiellement cibles de POFUT1.....	60
Figure 24 : Signalisation de la voie canonique Wnt .....	62
Figure 25 : Homologie de structure entre les antagonistes et les protéines associées à Wnt .....	64
Figure 26 : Structure de WIF1 murin .....	64
Figure 27 : Régulation du gradient morphogénétique des Wnts par WIF1. ....	66



## RESUME

---

La *O*-fucosylation, catalysée par Pofut1, est une glycosylation rare qui consiste en l'ajout d'un fucose *O*-lié sur la sérine ou la thréonine d'une séquence consensus (C<sup>2</sup>X<sub>4</sub>(**S/T**)C<sup>3</sup>), portée par un domaine EGF-like (ELD) d'une glycoprotéine membranaire ou sécrétée. Notre analyse de la lignée murine Pofut1<sup>cax/cax</sup>, hypomorphe pour le gène *Pofut1*, a révélé une hypertrophie musculaire post-natale associée à une diminution du pool de cellules satellites. Ce phénotype est en partie associé à un défaut d'interaction entre les récepteurs NOTCH hypo-*O*-fucosylés des myoblastes dérivés de cellules satellites (MDCS) et leurs ligands DSL, ce qui aboutit à une plus faible activation de la signalisation Notch. D'autres protéines potentiellement impliquées dans la myogenèse peuvent également être la cible de POFUT1. C'est notamment le cas de la protéine Wnt inhibitory factor 1 (WIF1), qui dispose de cinq ELDs, dont deux sont potentiellement aptes à recevoir un *O*-fucose (ELDs III et V). Par une approche phylogénétique, nous avons montré la conservation de ces deux sites de *O*-fucosylation et de deux sites de *N*-glycosylation chez la plupart des bilatériens. Nos expériences démontrent l'occupation de tous ces sites, excepté le site de *O*-fucosylation de l'ELD V, chez la protéine WIF1 murine. La capacité de l'ELD III, produit de manière isolée, à recevoir un fucose *O*-lié a été démontrée après *O*-fucosylation *in vitro*, par l'association de cycloaddition azide-alcyne assistée au cuivre (CuAAC) et de spectrométrie de masse en mode MRM. Cette nouvelle approche expérimentale a par la suite été standardisée et sa sensibilité évaluée en comparant deux autres ELDs (ELDs 12 et 26 de NOTCH1) connus pour être *O*-fucosylés mais présentant des affinités différentes pour POFUT1. De façon surprenante, l'ELD V de WIF1 ne peut être *O*-fucosylé, probablement en raison d'un clash stérique entre cet ELD et POFUT1, prévenant ainsi leur interaction. L'analyse de la protéine WIF1 entière a confirmé les résultats obtenus sur les ELDs isolés et démontre l'occupation des deux sites de *N*-glycosylation. Enfin, nos résultats montrent également l'importance de ces deux *N*-glycanes, mais également celle du *O*-fucose de l'ELD III, pour une sécrétion optimale de la protéine WIF1 murine.

**Mots clés:** POFUT1, *O*-fucosylation, EGF-like, WIF1, CuAAC



## ABSTRACT

---

The, Pofut1-catalyzed *O*-fucosylation, is a rare glycosylation which consists of the addition of an *O*-linked fucose to the serine or threonine of a consensus sequence (C<sup>2</sup>X<sub>4</sub>(**S/T**)C<sup>3</sup>), carried by an EGF-like domain (ELD) of a membrane or secreted glycoprotein. Our analysis of the murine line Pofut1<sup>ca<sup>x</sup>/ca<sup>x</sup></sup>, hypomorphic for the *Pofut1* gene, revealed post-natal muscle hypertrophy associated with a decrease in the satellite cell pool. This phenotype was partly associated with a lack of interaction between hypo-*O*-fucosylated NOTCH receptors of satellite cell-derived myoblasts (SCDM) and their DSL ligands, which resulted in a lower activation of Notch signaling. Other proteins potentially involved in myogenesis may also be the target of POFUT1. This is indeed the case for the protein Wnt inhibitory factor 1 (WIF1), which has five ELDs, whose only two are potentially able to receive an *O*-fucose (ELDs III and V). Using a phylogenetic approach, we showed in most bilaterians that these two *O*-fucosylation sites and two *N*-glycosylation sites were conserved. Our experiments showed the occupation of all these sites, except for the *O*-fucosylation site of murine WIF1 protein ELD V. The ability of the ELD III, produced as an isolated protein, to receive *O*-linked fucose was demonstrated after an *in vitro* *O*-fucosylation by combination of copper-catalysed azide-alkyne cycloaddition (CuAAC) and MRM-mass spectrometry. This new experimental approach was then standardized and its sensitivity was evaluated by comparing two other ELDs (NOTCH1 ELDs 12 and 26) known to be *O*-fucosylated but with different affinities for POFUT1. Surprisingly, WIF1's ELD V could not be *O*-fucosylated, probably due to a steric clash between this ELD and POFUT1, thus preventing their interaction. The analysis of the full-length WIF1 protein confirmed our results obtained with isolated ELDs and demonstrated the occupation of the two *N*-glycosylation sites. Finally, our results also showed the importance of these two *N*-glycans, but also the importance of ELD III's *O*-fucose, for optimal secretion of the murine WIF1 protein.

**Keywords:** POFUT1, *O*-fucosylation, EGF-like, WIF1, CuAAC



## **PUBLICATIONS ET COMMUNICATIONS**

---

### **PUBLICATIONS**

**Mouse WIF1 is only *O*-fucosylated on EGF-III despite two evolutionary conserved consensus sites.**

**Pennarubia Florian**, Al Jaam Bilal, Pinault Emilie, Brun Caroline E., Maftah Abderrahman, Germot Agnès and Legardinier Sébastien. **Sci. Rep.** « en révision » (Impact factor en 2016: 4,259).

***In vitro* acellular method to reveal *O*-fucosylation on EGF-like domains**

**Pennarubia Florian**, Pinault Emilie, Maftah Abderrahman and Legardinier Sébastien. **Glycobiology**. 2018 Nov 29. (Impact factor en 2017: 3.664).

**Reduced Notch signaling leads to postnatal skeletal muscle hypertrophy in *Pofut1*<sup>cax/cax</sup> mice.**

Bilal Al Jaam<sup>1</sup>, Katy Heu<sup>1</sup>, **Florian Pennarubia**<sup>2</sup>, Alexandre Segelle, Laetitia Magnol, Agnès Germot, Sébastien Legardinier, Véronique Blanquet and Abderrahman Maftah. **Open Biol** 2016 Sep;6(9). (Impact factor en 2015: 5,36).

### **COMMUNICATIONS**

#### **Communications Orales:**

**Evidence that mouse WIF1 is *O*-fucosylated on one of its EGF-like repeats.**

**Pennarubia Florian**, Pinault Emilie, Germot Agnès, Legardinier Sébastien and Maftah Abderrahman. 11th International Symposium on Glycosyltransferases (GlycoT), Qingdao, Chine, June 19-23, 2018. (Oral poster) - *Best Poster Award GlycoT 2018*

**Click chemistry to reveal *O*-fucosylation of EGF-like domains.**

**Pennarubia Florian**, Pinault Emilie, Maftah Abderrahman and Legardinier Sébastien. 27èmes journées du Groupe Français des Glycosciences (GFG), Nouan le Fuzelier, 21-25 mai 2018.



## **Identification of POFUT1 potential target proteins.**

**Pennarubia Florian**, Maftah Abderrahman, and Legardinier Sébastien. Séminaire thématique biosanté, Pouligny Notre-Dame, 9-10 mars 2017.

## **Communications écrites:**

### **Evidence that mouse WIF1 is *O*-fucosylated on one of its EGF-like repeats.**

**Pennarubia Florian**, Pinault Emilie, Germot Agnès, Legardinier Sébastien and Maftah Abderrahman. 11th International Symposium on Glycosyltransferases (GlycoT), Qingdao, Chine, June 19-23, 2018. *Best Poster Award GlycoT 2018*

### **Click chemistry to reveal *O*-fucosylation of EGF-like domains.**

**Pennarubia Florian**, Pinault Emilie, Maftah Abderrahman and Legardinier Sébastien. 27èmes journées du Groupe Français des Glycosciences (GFG), Nouan le Fuzelier, 21-25 mai 2018.

### **Are molecular mechanisms related to variable phenotypes in Pofut<sup>cax/cax</sup> mice?**

**Pennarubia Florian**, Al Jaam Bilal, Maftah Abderrahman and Legardinier Sébastien. Séminaire thématique biosanté, Pouligny Notre-Dame, 7-9 mars 2016.



## AVANT-PROPOS

---

Les glycosylations sont des modifications post-traductionnelles catalysées par une multitude d'enzymes spécifiques appelées glycosyltransférases. Ces modifications sont retrouvées entre autres sur des lipides et des protéines (glycoconjugués) et consistent à transférer un monosaccharide, pouvant être ou non rallongé par d'autres sucres suite à l'action combinée de glycosyltransférases et de glycosidases. Trois grands types de glycosylation sont recensés pour les glycoprotéines, classés en fonction de la nature de la liaison entre le sucre et l'acide aminé, les *N*-, *O*- et *C*-glycosylations. Les glycanes assurent de nombreuses fonctions importantes notamment pour le repliement, la maturation et l'activité des protéines. Des altérations de ces glycosylations sont souvent associées à diverses pathologies dont les cancers (Varki, 2017).

L'équipe « Glycosylation et différenciation cellulaire » du laboratoire pluridisciplinaire PEIRENE (EA 7500), au sein de laquelle j'ai effectué mes travaux de thèse, a toujours eu pour objectif d'étudier les glycoconjugués et leurs effets biologiques. Ainsi, plusieurs études ont été menées ces dernières années sur différentes glycosyltransférases (fucosyl-, glucosyl-, sialyltransférase, etc...), pour mettre en évidence le rôle de ces enzymes, et par extension des glycanes qu'elles transfèrent, dans la régulation du développement musculaire et plus récemment dans le processus tumoral. Mes travaux de thèse dans cette équipe de recherche ont consisté à étudier la *O*-fucosylation médiée par POFUT1 sous différents aspects tels que sa mise en évidence expérimentale, ses cibles protéiques potentielles, ses effets biologiques et le mécanisme d'action de l'enzyme qui lui est associée.

La *O*-fucosylation des domaines EGF-like (ELD) est une glycosylation rare catalysée par la protéine *O*-fucosyltransférase 1 (Pofut1). Le transfert de *O*-fucoses ne peut être catalysé que sur des ELDs correctement repliés possédant une séquence consensus spécifique de type C<sup>2</sup>-X-X-X-X-(S/T)-C<sup>3</sup> (Müller et al., 2014). Les fucoses *O*-liés sont essentiels à la fonction de certaines glycoprotéines, comme le récepteur NOTCH1 dont la *O*-fucosylation module son interaction avec ses ligands de type DSL (Delta/Serrate/Lag-2) (Kakuda and Haltiwanger, 2017). Ainsi, l'absence de *O*-fucosylation s'avère dramatique pour l'organisme, comme le montre le KO du gène *Pofut1* qui est létal (Shi and Stanley, 2003) mais également les défauts de somitogenèse observés suite à la sous-expression de *Pofut1* chez la lignée murine *Pofut1*<sup>cax/cax</sup> (Schuster-Gossler et al., 2009). Le récepteur NOTCH1, est la glycoprotéine



*O*-fucosylée la plus connue et étudiée et par conséquent la plus analysée dans le contexte d'une altération globale de la *O*-fucosylation. Il existe seulement moins d'une centaine de protéines portant *a minima* une séquence consensus de *O*-fucosylation (Al Jaam, 2016; Schneider et al., 2017). Néanmoins, la présence effective de fucose *O*-lié n'a été démontrée que pour une quinzaine d'entre elles (Li et al., 2017; Schneider et al., 2017). Ainsi plusieurs protéines sont susceptibles d'être impactées par une altération de l'expression de *Pofut1*.

Les objectifs de mes travaux de thèse ont donc été d'étudier l'effet de la sous-expression de *Pofut1* sur la voie Notch dans le contexte de la croissance musculaire post-natale. Ainsi que d'identifier et de caractériser de nouvelles glycoprotéines *O*-fucosylées, potentiellement impliquées dans les processus myogénique et cancéreux.

Dans un premier temps, mes travaux se sont inscrits dans le cadre d'un projet, visant à caractériser l'effet de la sous-expression de *Pofut1* sur la croissance musculaire post-natale chez les souris *Pofut1*<sup>cax/cax</sup> initié lors de la thèse de Bilal Al Jaam (soutenue en 2016). Ma contribution à ce projet, a été de définir l'impact de cette sous-expression, sur la capacité des récepteurs NOTCH présents à la surface des myoblastes dérivés de cellules satellites (MDCS) des souris *Pofut1*<sup>cax/cax</sup> à interagir avec leurs ligands DSL (JAG1 et DLL1).

La majeure partie de mes travaux de thèse a été de caractériser une glycoprotéine potentiellement *O*-fucosylée par POFUT d'après les listes établies chez l'Homme et la souris (Al Jaam, 2016; Schneider et al., 2017) et pouvant être étudiée dans le contexte du processus myogénique et/ou cancéreux. Le choix s'est porté sur la protéine Wnt inhibitory factor 1 (WIF1), protéine pour laquelle une étude phylogénique a été entreprise sur la conservation des sites de *O*-fucosylation et de *N*-glycosylation. Par la suite, plusieurs analyses biochimiques visant à démontrer la présence ou non de fucose *O*-lié *in vitro* et *in cellulo* ont été effectuées. Enfin, le rôle de la *N*-glycosylation et de la *O*-fucosylation dans la sécrétion de WIF1 a été examiné.

Les derniers travaux réalisés au cours de ma thèse se sont focalisés sur la valorisation d'une nouvelle approche expérimentale consistant à mettre en évidence la *O*-fucosylation, par cycloaddition azide-alcyne assistée au cuivre (CuAAC ou chimie click), après transfert de *O*-fucoses *in vitro* par POFUT1.



# EXPOSE BIBLIOGRAPHIQUE



## Chapitre I. La glycosylation

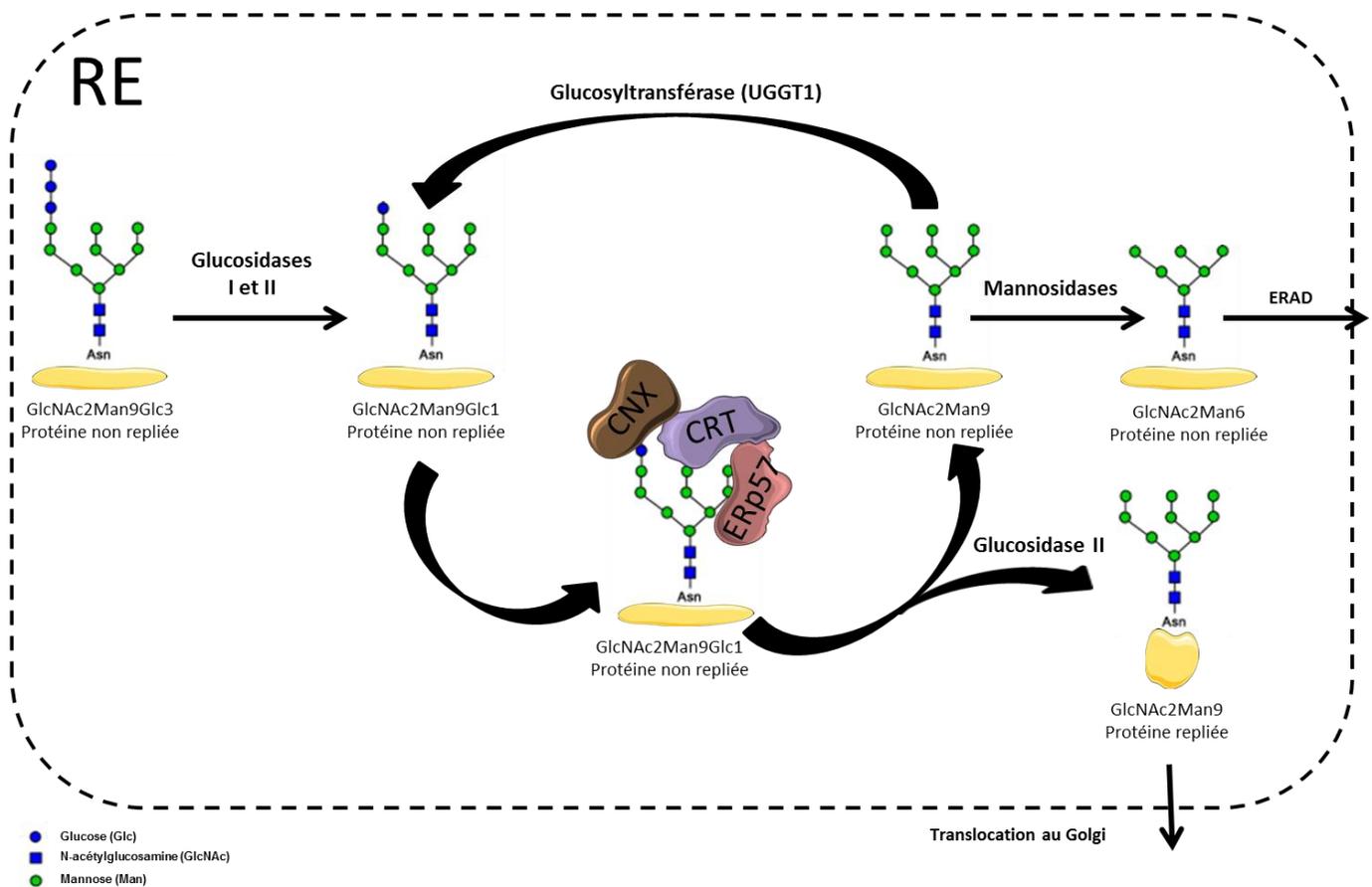
---

La glycosylation est une réaction enzymatique qui consiste en la liaison d'un ou plusieurs monosaccharides sur une chaîne peptidique ou lipidique. Le processus de glycosylation a été fortement préservé au cours de l'évolution puisqu'il est présent chez les archées, les bactéries et les eucaryotes. Malgré cette conservation, le processus de glycosylation chez les eucaryotes est relativement spécifique à ce domaine du vivant et se distingue par d'importantes différences notamment en ce qui concerne les mécanismes enzymatiques et les glycanes qui en résultent (Dell et al., 2010).

Il existe une grande variété de glycosylations chez les mammifères, toutes construites à partir d'un unique monosaccharide (fucose, mannose, galactose, *N*-acétylglucosamine, *N*-acétylgalactosamine,...) qui peut par la suite être rallongé par d'autres sucres pour donner différentes structures glycaniques. Ces réactions sont catalysées dans le réticulum endoplasmique (RE) et/ou l'appareil de Golgi par une grande variété de glycosyltransférases qui utilisent des sucres nucléotidiques donneurs spécifiques. Les glycosyltransférases ne sont pas les seules enzymes impliquées dans le processus de glycosylation. Les glycosidases, responsables de la suppression de certains monosaccharides constituant les précurseurs glycaniques, sont essentielles à la maturation des glycanes au cours de leur transit dans le RE et l'appareil de Golgi. Ainsi, les glycosylations dépendent de l'action complémentaire de ces deux enzymes.

Parmi les glycoconjugués les plus étudiés figurent les glycoprotéines qui sont majoritairement des protéines sécrétées ou membranaires. Les glycanes possèdent de nombreux rôles biologiques. Ils sont notamment impliqués dans la maturation des protéines en participant au repliement, leur oligomérisation éventuelle, leur solubilité (Varki, 2017), leur protection contre la protéolyse (Russell et al., 2009) ou contre les dommages des radicaux libres (Martínek et al., 2010) et dans leurs fonctions (adhésion cellulaire (Zhang and Ten Hagen, 2011), immunité (Rabinovich et al., 2012) interaction ligand/récepteur (Varki, 2017) ou encore trafic intracellulaire (Moremen et al., 2012).

Les glycoprotéines peuvent faire l'objet de différentes glycosylations réparties en trois grandes familles (la *N*-, la *O*- et la *C*-glycosylation) en fonction de la nature de la liaison du premier sucre avec l'acide aminé cible de la protéine.



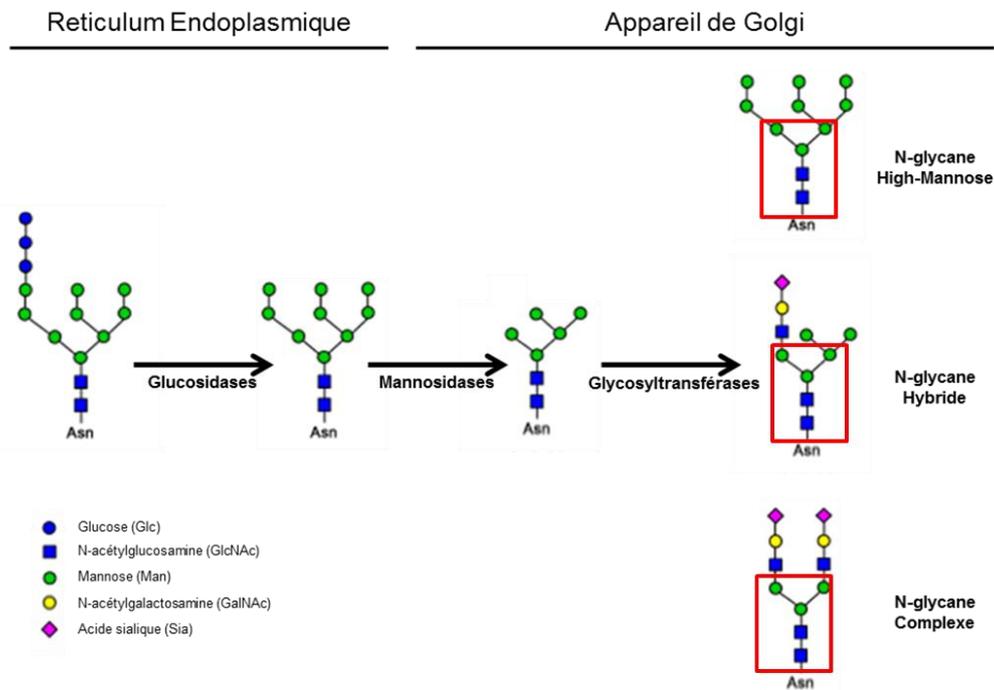
### Figure 1 : Schéma simplifié du contrôle du repliement des glycoprotéines dans le RE.

Dans le RE, le *N*-glycane est attaché à l'asparagine (Asn) d'une glycoprotéine sous la forme d'un précurseur (GlcNAc2Man9Glc3) qui sera immédiatement réduit sous l'action des glucosidases 1 et 2 pour donner un oligosaccharide monoglucosylé (GlcNAc2Man9Glc1) reconnu par les protéines chaperonnes calnexine (CNX) et calréticuline (CRT). Ces protéines chaperonnes associées à l'oxydoréductase ERp57 vont permettre le repliement de la protéine. L'action de la glucosidase II supprime le dernier glucose du précurseur *N*-glycosidique (GlcNAc2Man9), ce qui met un terme à son interaction avec le couple CNX/CRT. Si la protéine est correctement repliée, elle sera translocuée au Golgi. Dans le cas contraire, elle subira l'action d'une glucosyltransférase (UGGT1) qui permet, via l'ajout d'un glucose (GlcNAc2Man9Glc1), de restaurer l'interaction avec le couple CNX/CRT afin de ré-initier le processus de repliement. Si la glycoprotéine ne parvient pas à être repliée, elle sera reconnue par diverses mannosidases réduisant le précurseur de *N*-glycannes (GlcNAc2Man6). Sous cette forme, l'oligosaccharide sera reconnu par des transporteurs, ce qui va permettre le marquage de la glycoprotéine et sa dégradation par le système ERAD. (D'après Anelli and Sitia, 2008 et Goettig, 2016)

## I.1. La *N*-glycosylation

La *N*-glycosylation est caractérisée par la fixation de manière covalente d'un oligosaccharide sur l'azote du groupement amide d'un résidu asparagine. Elle ne peut avoir lieu qu'en présence d'une séquence spécifique, dans la majorité des cas de type « N-X-(S/T) » (avec X représentant n'importe quel acide aminé excepté la proline). Il existe également quelques séquences rares de *N*-glycosylation telles que N-X-C, N-G et N-X-V (Zielinska et al., 2010).

Les *N*-glycanes sont ajoutés aux protéines de manière co-traductionnelle. Ainsi, pendant la translocation de la future glycoprotéine à travers la membrane du RE, une oligosaccharyltransférase va catalyser dans la lumière de cet organite le transfert d'un *N*-glycane sous une forme de précurseur oligosaccharidique de type GlcNAc2Man9Glc3. Ce précurseur sera immédiatement réduit suite à l'action des glucosidases I et II pour aboutir à un oligosaccharide monoglucosylé (GlcNAc2Man9Glc1) capable d'interagir avec deux protéines chaperonnes grâce à leurs domaines lectine ; la calnexine (CNX) et la calréticuline (CRT). Cette interaction engage la protéine dans un cycle qui favorise le repliement de la glycoprotéine et prévient son agrégation *via* l'action du couple CNX/CRT associées à l'oxydoréductase ERp57. Ce repliement est également facilité par l'environnement particulier du RE (environnement Redox et concentration de Ca<sup>2+</sup> spécifique) (Aebi, 2013) ainsi que par la présence d'enzymes impliquées dans la formation des ponts disulfure telles que les protéines disulfide isomérase (PDI) et prolyl isomérase (PPI) (Oka and Bulleid, 2013). L'action de la glucosidase II supprime le dernier glucose (GlcNAc2Man9) ce qui aboutit à la perte d'interaction entre la glycoprotéine et le couple CNX/CRT. Dans le cas où la glycoprotéine est correctement repliée, elle sera transloquée au golgi pour y être maturée. Dans le cas contraire, la glycoprotéine sera reconnue par une glucosyltransférase (UGGT1) ce qui va permettre l'ajout d'un glucose (GlcNAc2Man9Glc1) et le retour de la glycoprotéine dans le cycle de repliement du couple CNX/CRT. Si la glycoprotéine ne peut être correctement repliée, différentes mannosidases réduiront le précurseur *N*-glycanique (GlcNAc2Man9) ce qui permet sa reconnaissance par des lectines spécifiques et sa dégradation ultérieure *via* le système ERAD (Endoplasmic Reticulum-Associated Degradation) (Anelli and Sitia, 2008) (Figure 1).



**Figure 2 : Schéma simplifié de la biosynthèse des différents types de N-glycanes.**

Dans le RE, le N-glycane est attaché à l'asparagine (Asn) d'une protéine sous la forme d'un précurseur qui sera immédiatement réduit sous l'action des glucosidases 1 et 2. Si la protéine est correctement repliée, elle sera transloquée au Golgi, où le N-glycane subira l'action de mannosidases avant d'être modifié par différentes glycosyltransférases pour donner un N-glycane polymannosidique, de type hybride ou de type complexe, tous trois construits à partir d'un noyau pentasaccharidique commun (GlcNAc<sub>2</sub>Man<sub>3</sub>) (encadré rouge). (D'après Goettig, 2016)

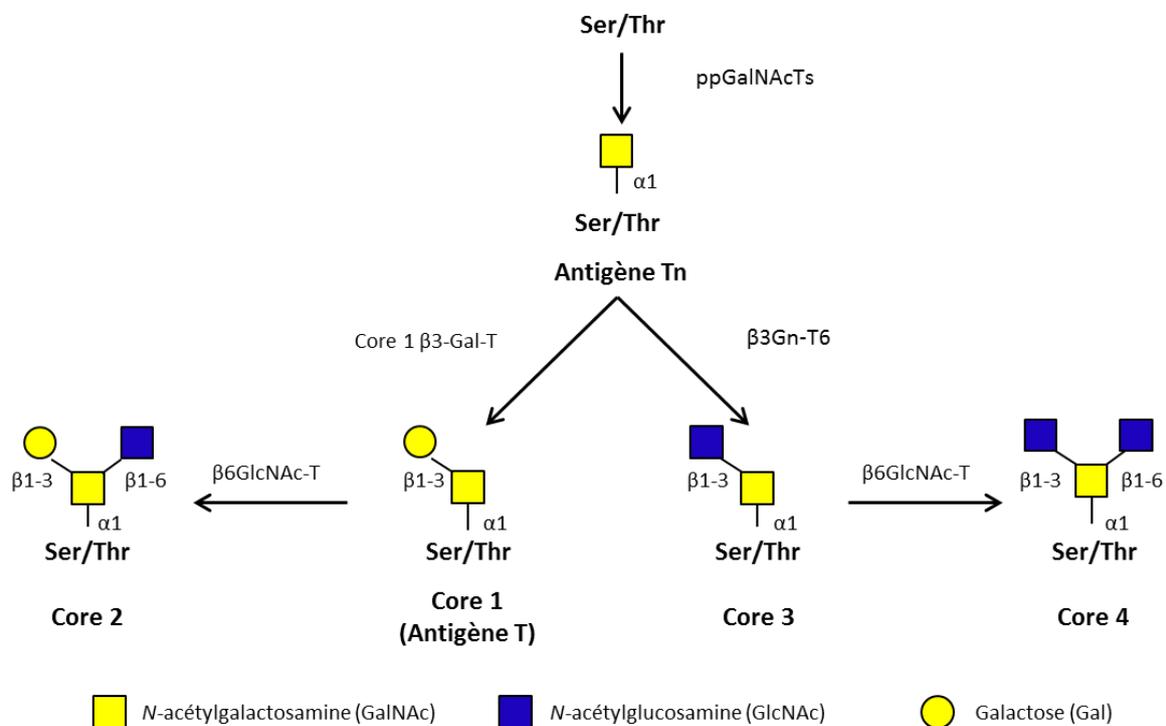
Quand une glycoprotéine transite par le Golgi, le *N*-glycane va subir l'action de mannosidases puis il sera de nouveau modifié par différentes glycosyltransférases pour donner les 3 grandes familles de *N*-glycanes, les polymannosidiques, les types hybrides et les types complexes en fonction de la nature des antennes présentes sur le noyau pentasaccharidique commun GlcNAc2Man3 (Figure 2).

Les *N*-glycanes participent à de nombreux processus biologiques tels que le repliement protéique, la sécrétion, la modulation de la signalisation cellulaire (induite par les récepteurs), la prolifération et la différenciation cellulaire, la rigidité de certaines structures (paroi des cellules végétales et fongiques ainsi que l'exosquelette des arthropodes), le métabolisme (régulation des transporteurs de glucose à la surface cellulaire) et divers processus immunologiques (activation des lymphocytes T, régulation des récepteurs aux cytokines, etc...) (Varki, 2017).

## **I.2. La *O*-glycosylation**

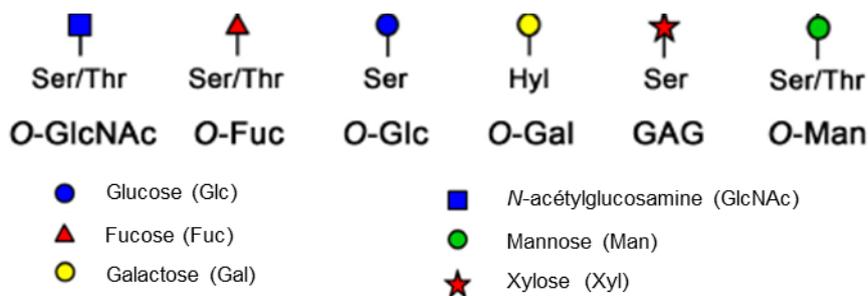
La *O*-glycosylation est l'autre grande famille des glycosylations des protéines et consiste en la liaison d'un monosaccharide sur l'oxygène d'un groupement hydroxyle d'un résidu sérine ou thréonine ou, dans de plus rares cas, sur une tyrosine, une hydroxylysine ou une hydroxyproline (Song and Mechref, 2013).

La *O*-glycosylation de type mucine est la plus fréquemment retrouvée sur les glycoprotéines membranaires et sécrétées, elle est catalysée par des enzymes de la famille des ppGaNTases qui utilisent l'UDP-*N*-acétylgalactosamine (UDP-GalNAc) comme donneur de sucre. Le transfert de GalNAc se fait sur le groupement hydroxyle d'une sérine ou d'une thréonine (GalNAc $\alpha$ 1-S/T) pour former une structure appelée antigène Tn (Tn Ag). Les *O*-GalNAc de type mucine ne sont pas présents sur une séquence consensus clairement établie, mais dans des régions favorables à leur mise en place, caractérisées par une faible structuration et une forte proportion de proline et de glycine. Une fois le motif Tn formé, il peut être rallongé pour former différentes structures glycaniques. La structure de type core 1 (ou antigène T) est majoritaire, elle est définie par l'ajout d'un galactose lié en  $\beta$ 1-3 (Gal $\beta$ 1-3GalNAc $\alpha$ 1-S/T) par une galactosyltransférase. L'ajout d'un *N*-acetylglucosamine (GlcNAc) lié en  $\beta$ 1-3 sur l'antigène Tn (GlcNAc $\beta$ 1-3GalNAc $\alpha$ -S/T) constitue la structure de type core 3 également très abondante. Les structures de type core 1 et 3 peuvent



**Figure 3 : Biosynthèse des *O*-glycanes de type mucine les plus communs.**

La *O*-glycosylation de type mucine se caractérise par le transfert d'un GalNAc sur le groupement hydroxyle d'une serine ou d'une thréonine par des ppGalNAcTases pour former une structure de type antigène Tn. Différentes extensions de la forme antigène Tn, suite à l'action de galactosyltransférase (core  $\beta$ -Gal-T) et/ou *N*-acétylglucosaminyltransférases ( $\beta$ 3Gn-T6), aboutissent à la formation des structures de type Core 1, 2, 3 et 4. (D'après Tian and Ten Hagen, 2009)



**Figure 4 : Représentation des *O*-glycanes les plus communs.**

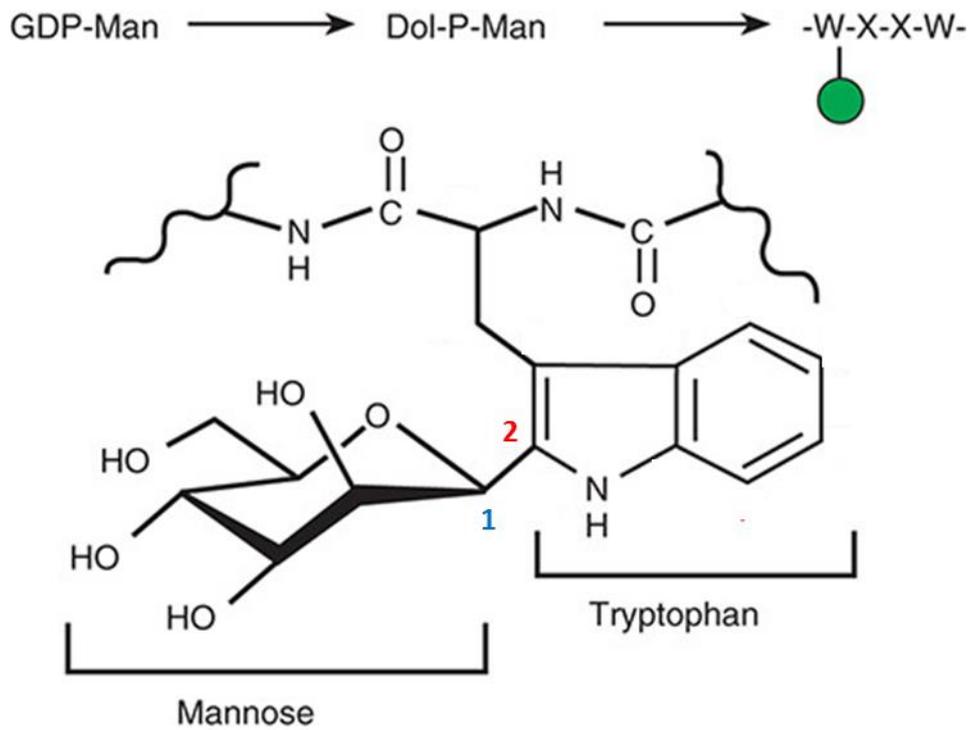
Il existe plusieurs types de sucre pouvant être *O*-liés sur un résidu sérine et/ou thréonine ou hydroxylysine. Les plus courants sont le *N*-acétylglucosamine, le fucose, le glucose, le galactose, le xylose et le mannose. (Goettig, 2016).

respectivement être modifiées en core 2 et 4 suite à l'action d'une  $\beta$ 1-6 *N*-acetylglucosaminyltransférase qui va transférer un GlcNAc sur le GalNAc *O*-lié (Figure 3). Il existe également quatre autres cores construits à partir de l'antigène Tn, le core 5 (GalNAc $\alpha$ 1-3GalNAc $\alpha$ Ser/Thr), 6 (GlcNAc $\beta$ 1-6GalNAc $\alpha$ Ser/Thr), 7 (GalNAc $\alpha$ 1-6GalNAc $\alpha$ Ser/Thr) et 8 (Gal $\alpha$ 1-3GalNAc $\alpha$ Ser/Thr). Ces différentes structures glycaniques peuvent par la suite être étendues par différents monosaccharides pour former des *O*-GalNAc complexes (Brockhausen et al., 2009).

Ces glycanes sont connus pour moduler les phénomènes de reconnaissance, d'interaction et d'adhésion cellulaire, notamment lors de la réponse immunitaire ou de processus développementaux (Tian and Ten Hagen, 2009).

D'autres monosaccharides peuvent être *O*-liés (Figure 4) tels que le fucose que l'on retrouve principalement dans deux types de domaines particuliers, les epidermal growth factor like (ELD) et les thrombospondin type 1 domain (TSR), le *O*-glucose est également retrouvé dans les ELDs tout comme le *O*-GlcNAc. Le xylose, le mannose et le galactose *O*-liés sont quant à eux respectivement et majoritairement présents dans les protéoglycanes, le collagène et l' $\alpha$ -Dystroglycane. Bien que ces six glycanes *O*-liés soient retrouvés sur des glycoprotéines membranaires ou sécrétées, les *O*-GlcNAc se distinguent par leur incorporation fréquente sur des glycoprotéines cytoplasmiques et nucléaires (Hart, 1997). Deux *O*-GlcNAc-transférases, OGT et EOGT, sont respectivement responsables du transfert de GlcNAc *O*-lié sur les glycoprotéines cytoplasmiques ou nucléaires et sur les ELDs de glycoprotéines membranaires ou sécrétées. De la même façon que pour les *O*-GlcNAc des glycosyltransférases spécifiques, utilisant comme substrats différents sucres nucléotidiques donneurs, élaborent les autres *O*-glycosylations et des séquences consensus spécifiques ont été définies pour la plupart d'entre elles.

De nombreux rôles biologiques sont associés à ces différentes *O*-glycanes tels que le repliement, la sécrétion, la protection contre la protéolyse, la modification épigénétique des histones (*O*-GlcNAc) ou encore la modulation de différentes liaisons (HSPGs, MEC) et interactions (ligand/récepteur) (Varki, 2017).

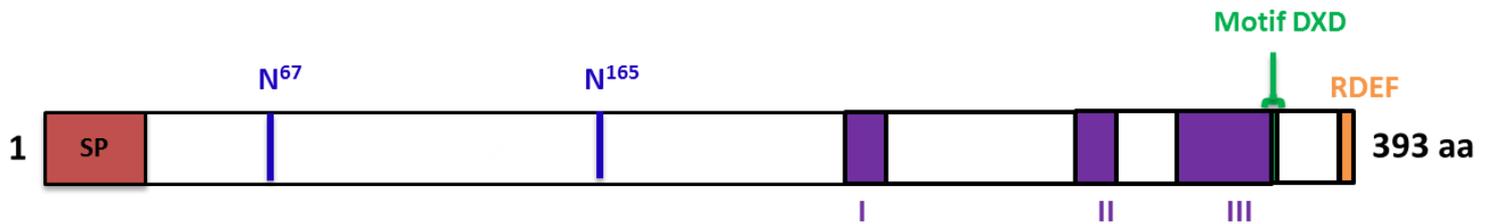


**Figure 5 : Biosynthèse et structure de la C-mannosylation.**

Le Dolichol-P-mannose issu du GDP-mannose est le sucre nucléotidique donneur pour permettre la C-mannosylation. Une C-mannosyltransférase transfère un C-mannose sur le noyau indole d'un tryptophane de la séquence consensus W-X-X-W par l'établissement d'une liaison  $\alpha$  entre le carbone 1 du mannose (en bleu) et le carbone 2 du noyau indole du résidu tryptophane (en rouge). (D'après Haltiwanger et al., 2017)

### I.3. La C-glycosylation

La C-glycosylation ou C-mannosylation est l'une des glycosylations les plus rares et les moins étudiées. Elle consiste en l'addition d'un mannose lié en  $\alpha$  1-2 sur le carbone d'un noyau indole du premier résidu tryptophane d'une séquence consensus W-X-X-W (Hofsteenge et al., 1994; Haltiwanger et al., 2015) (Figure 5) ou sur les trois tryptophanes du motif consensus W-X-X-W-X-X-W-C (Hofsteenge et al., 1999). La C-mannosylation est catalysée par les enzymes de la famille DPY19 (Buettner et al., 2013; Shcherbakova et al., 2017) qui utilise le Dolichol-P-mannose (un dérivé du GDP-mannose) comme nucléotide-sucre donneur. Cette glycosylation est aussi connue pour participer aux processus de repliement et de sécrétion des protéines et module la signalisation cellulaire notamment dans le cadre de la réponse immunitaire (Hofsteenge et al., 1999).



**Figure 6 : Schéma de la protéine POFUT1 murine.**

L'enzyme murine POFUT1 comprend un peptide signal d'adressage au RE (en rouge), trois motifs peptidiques conservés (I, II et III) (en mauve), un motif DXD (en vert) impliqué dans l'activité de liaison du sucre nucléotidique donneur et des cations bivalents ( $Mn^{2+}$ ) et une séquence RDEF (de type KDEL) de rétention au RE (en orange). POFUT1 contient également deux *N*-glycanes (en bleu), impliqués dans la solubilité et l'activité de l'enzyme, localisés sur les asparagines 67 et 165 ( $N^{67}$  et  $N^{165}$ ).

## Chapitre II. La *O*-fucosylation

---

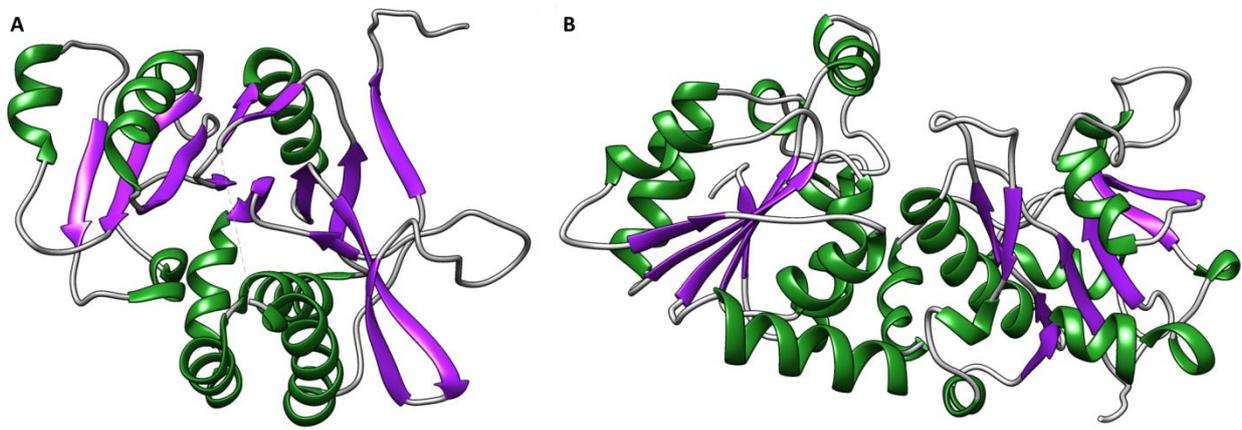
La *O*-fucosylation est une glycosylation rare, qui consiste en la liaison covalente d'un fucose sur le groupe hydroxyle d'un résidu sérine ou thréonine. Cette glycosylation a pour la première fois été mise en évidence en 1975 sur un glycopeptide isolé à partir d'échantillon d'urine humaine (Hallgren et al., 1975). Par la suite, plusieurs études ont montré que la *O*-fucosylation était spécifique de domaines protéiques particuliers, à savoir les ELDs (Kentzer et al., 1990) et les TSRs (Hofsteenge et al., 2001). Il existe deux *O*-fucosyltransférases qui catalysent le transfert de fucose *O*-lié, les protéines *O*-fucosyltransférases 1 et 2 (Pofut1 et Pofut2), qui ciblent respectivement les ELD et les TSR.

### II.1. La *O*-fucosylation dépendante de Pofut1

La protéine *O*-fucosyltransférase 1 (Pofut1) a été partiellement purifiée en 1996 et identifiée comme l'enzyme permettant le transfert de fucose *O*-lié sur certains ELDs à partir du GDP-fucose (Wang et al., 1996). Des études phylogénétiques ont mis en évidence une large distribution de l'enzyme Pofut1 chez les bilatériens, ce qui témoigne d'une forte conservation au cours de l'évolution et donc d'un rôle essentiel (Loriol et al., 2006). A la différence d'autres fucosyltransférases, pour la plupart golgiennes, Pofut1 est localisée dans le RE (Luo and Haltiwanger, 2005). Il a été montré chez *O*-fut1 (l'orthologue de Pofut1 chez *Drosophila melanogaster*) un rôle de protéine chaperonne, qui pourrait être totalement indépendant de son activité *O*-fucosyltransférase (Okajima et al., 2005). Cependant, cela n'a été démontré que chez *Drosophila melanogaster* et cette fonction reste encore controversée chez les mammifères (Stahl et al., 2008).

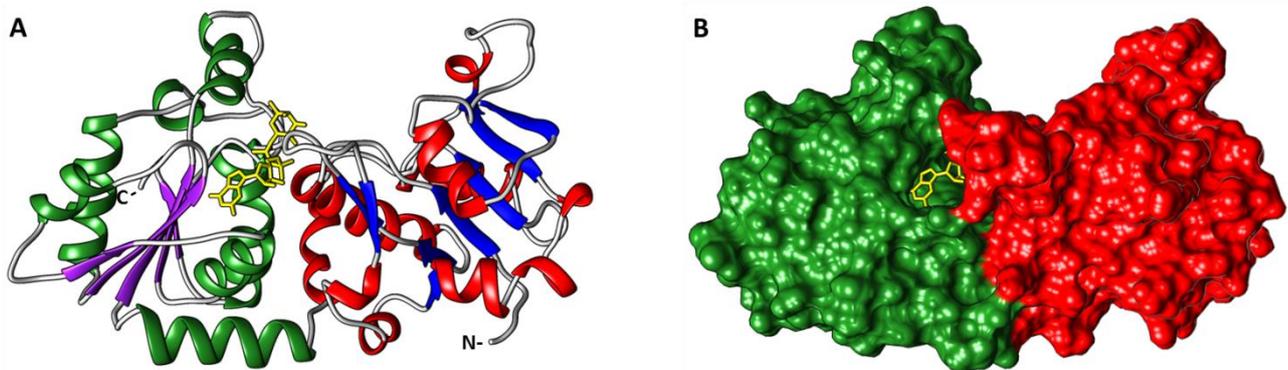
#### II.1.1. Structure de l'enzyme Pofut1

L'enzyme Pofut1 possède plusieurs domaines conservés (Figure 6), un peptide signal permettant l'adressage au RE (Luo and Haltiwanger, 2005), trois domaines peptidiques conservés communs aux  $\alpha$ 1,2- et  $\alpha$ 1,6-fucosyltransférases (motifs I, II et III), un motif DXD impliqué dans la fonction catalytique de l'enzyme (Loriol et al., 2007) et une séquence C-terminale de type KDEL (RDEF) permettant la rétention de Pofut1 dans le RE (Luo and Haltiwanger, 2005).



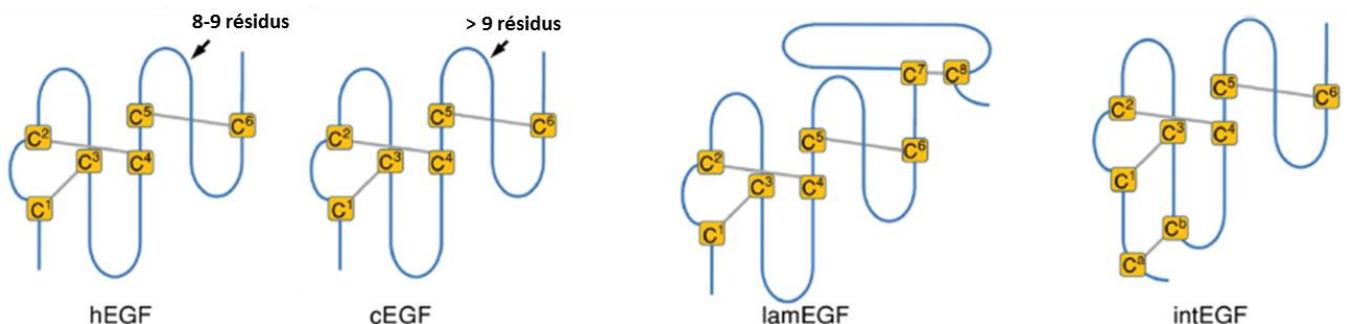
**Figure 7 : Structure caractéristique des repliements GT-A et GT-B.**

(A) Repliement de type GT-A (structure cristallisée de l'alpha-1,4-N-acétylhexosaminyltransférase (EXTL2) murine (PDB 1OMZ)) caractérisée par la présence d'un seul domaine semblable à un domaine de Rossmann (B) Repliement de type GT-B (structure cristallisée de la protéine POFUT1 murine (PDB 5kY4)) qui présente deux domaines de Rossmann séparés par une région de liaison aux différents substrats et un site d'activité catalytique. (D'après Breton et al., 2005).



**Figure 8 : Structure cristallisée de la protéine POFUT1 murine.**

(A) Localisation des feuillets  $\beta$  (bleu et violet) et hélices  $\alpha$  (rouge et vert) des domaines de Rossmann en N et C-terminal de l'enzyme POFUT1 murine séparés par une cavité permettant la liaison du GDP-fucose (jaune). (B) Surface réellement occupée par chacun des 2 domaines de Rossmann en positions N- (rouge) et C-terminales (vert). (PDB 5kY4 ; Li et al., 2017)



**Figure 9 : Les différents types d'ELDs.**

Dans la majorité des cas, les ELDs sont caractérisés par la présence de 6 cystéines connectées par 3 ponts disulfures. Ils sont séparés en 2 familles, les hEGF et les cEGF, en fonction du nombre de résidus compris entre les cystéines 5 et 6. Il existe également des formes plus rares à 8 cystéines, formant 4 points disulfures les lamEGF et les intEGF. (Li et al., 2017)

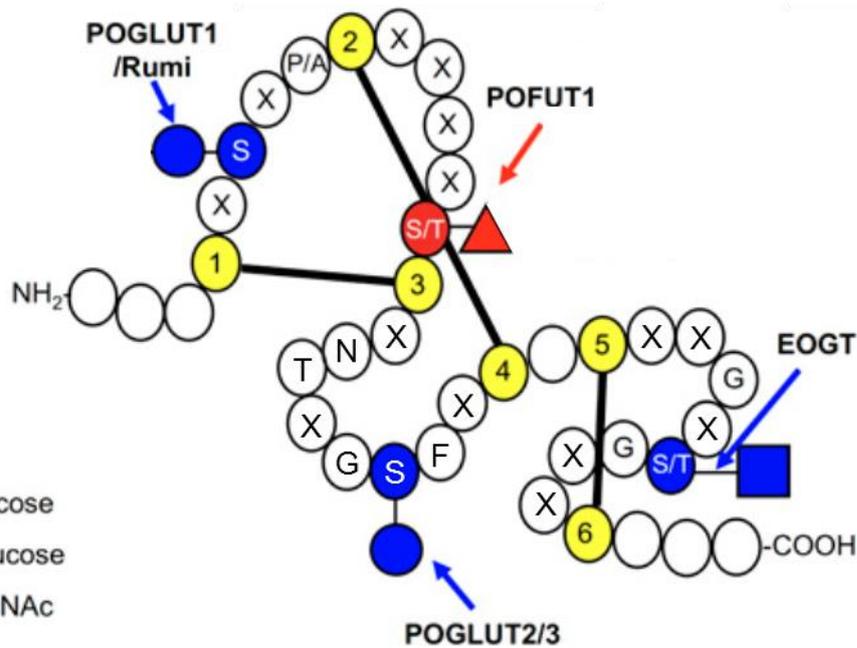
Pofut1 est une glycoprotéine connue pour être *N*-glycosylée (Wang and Spellman, 1998). Il a été montré que la présence de ces deux *N*-glycanes est nécessaire à l'activité de l'enzyme bovine (Loriol et al., 2007). La mutation du site de *N*-glycosylation en position N<sup>65</sup> de l'enzyme Pofut1 bovine (N<sup>67</sup> chez la souris) conduit à une agrégation de l'enzyme ce qui suggère un rôle de ce *N*-glycane dans le repliement de l'enzyme. De la même façon, il a été montré que le *N*-glycane en position N<sup>163</sup> (N<sup>165</sup> chez la souris) est directement impliqué dans l'activité *O*-fucosyltransférase de Pofut1 (Loriol et al., 2007).

L'enzyme Pofut1 possède un repliement de type GT-B, représentatif des glycosyltransférases impliquées dans le transfert du monosaccharide d'ancrage de la structure glycanique. Ce repliement consiste en la présence de deux domaines de Rossman, contrairement au repliement de type GT-A caractérisé par un seul domaine similaire à un domaine de Rossman (Figure 7). Le repliement GT-A est souvent celui des glycosyltransférases impliquées dans l'élongation des glycanes. Chacun de ces domaines de Rossman est caractérisé par la présence de feuillets  $\beta$  centraux entourés d'hélices  $\alpha$ . Ces deux domaines sont séparés par une cavité contenant le site catalytique et les sites de liaison au substrat donneur (le GDP-fucose) ainsi qu'au substrat accepteur (l'ELD) (Breton et al., 2006, 2012; Lira-Navarrete et al., 2011) (Figure 8).

### **II.1.2. Les ELDs comme substrats accepteurs**

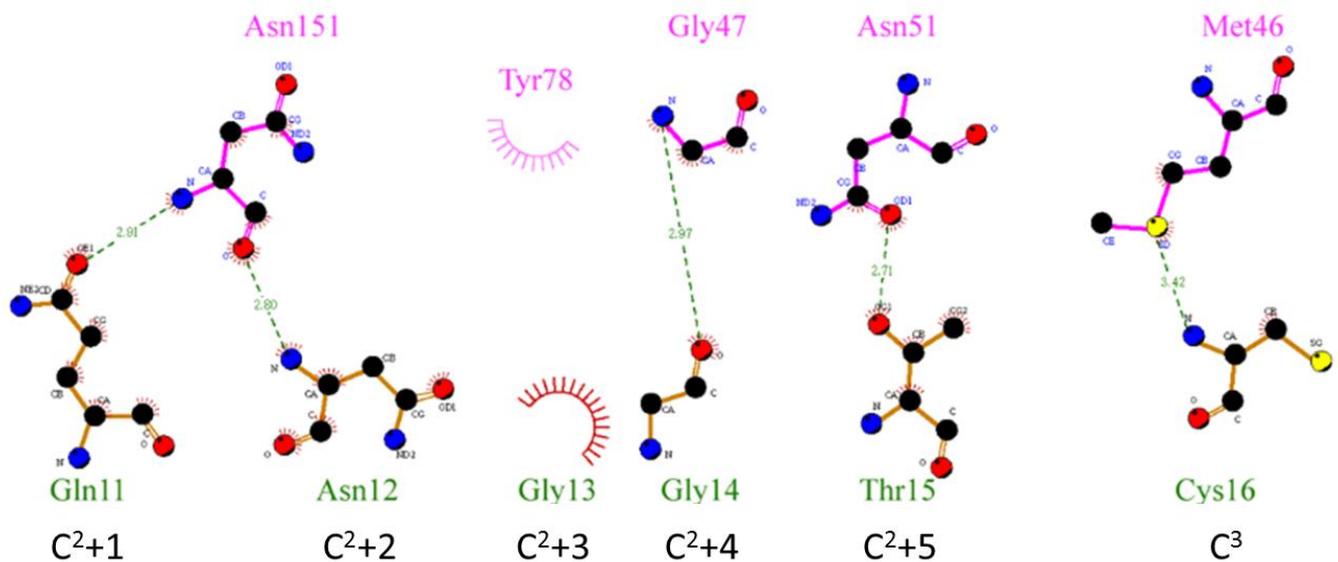
Les ELDs sont de petits domaines composés de 30 à 40 acides aminés, qui sont caractérisés par la présence de six cystéines très conservées formant 3 ponts disulfures. Ces ponts disulfures stabilisent l'ELD en reliant les cystéines 1-3, 2-4 et 5-6. Les ELDs sont répartis en deux grandes familles, les « human EGF-like » (hEGF) et les « C1r-like » (cEGF), la distinction se fait principalement par le nombre de résidus compris entre les cystéines 5 et 6 : entre 8 et 9 pour les hEGF et supérieur à 9 pour les cEGF. Il existe également dans de plus rares cas, des ELDs à 8 cystéines, les « laminines EGF-like » (lamEGF) dont les deux cystéines supplémentaires (C<sup>7</sup> et C<sup>8</sup>) sont reliées par un pont disulfure et les « intégrines EGF-like » (intEGF) qui possèdent un résidu cystéine (C<sup>a</sup>) en amont de la C<sup>1</sup> et une autre (C<sup>b</sup>) entre la C<sup>3</sup> et la C<sup>4</sup>, ces deux cystéines étant également reliées par un pont disulfure (Wouters et al., 2005) (Figure 9).

Plusieurs fonctions sont associées aux ELDs que l'on retrouve en unique ou multiple exemplaires dans des glycoprotéines membranaires ou sécrétées. Plusieurs études ont mis en



**Figure 10 : Les *O*-glycosylations des ELDs.**

Un ELD peut être *O*-fucosylé par POFUT1 sur une séquence C<sup>2</sup>-X-X-X-X-(S/T)-C<sup>3</sup>, *O*-GlcNAcylé par EOGT sur une séquence C<sup>5</sup>-X-X-G-X-(S/T)-G-X-X-C<sup>6</sup> et *O*-glucosylé par POGLUT1 sur une séquence C<sup>1</sup>-X-S-X-P-C<sup>2</sup> ainsi que par POGLUT2 et 3 sur une séquence C<sup>3</sup>-X-N-T-X-G-S-F-X-C<sup>4</sup>. (Takeuchi et al., 2018)



**Figure 11 : Interactions entre les résidus de POFUT1 et de l'ELD (+).**

Les résidus du sous-domaine C<sup>2</sup>-C<sup>3</sup> de l'ELD (en vert) peuvent former des liaisons de type hydrogène ou soufre-hydrogène avec les acides aminés de POFUT1 (en rose). Le sous-domaine C<sup>2</sup>-C<sup>3</sup> est celui de l'ELD(+), généré artificiellement en retenant, pour chaque position de la séquence, le résidu le plus fréquemment retrouvé dans les ELDs possédant une séquence consensus de *O*-fucosylation. (Li et al., 2017)

évidence l'importance de ces ELDs dans des processus d'adhésion (Stenflo, 1991) ou encore dans la signalisation cellulaire (Urban, 2002).

L'enzyme POFUT1 est capable de reconnaître et de se lier aux ELDs de type hEGF correctement repliés (Li et al., 2017). Cependant, la réaction de *O*-fucosylation n'est possible qu'en présence d'une séquence consensus particulière  $C^2\text{-X-X-X-X-(S/T)}\text{-C}^3$  (Müller et al., 2014), située entre les cystéines 2 et 3 ( $C^2\text{-C}^3$ ) où X qui équivaut à n'importe quel résidu (Figure 10).

Les ELDs peuvent également être modifiés par deux autres *O*-glycosylations, à savoir la *O*-GlcNAcylation catalysée par EOGT, qui consiste au transfert d'un GlcNAc *O*-lié sur une séquence consensus  $C^5\text{-X-X-G-X-(T/S)}\text{-G-X-X-C}^6$  et la *O*-glucosylation catalysée par POGLUT1, 2 ou 3 (protéine *O*-glucosyltransférase 1, 2 ou 3) qui permet l'ajout d'un *O*-glucose sur un motif consensus  $C^1\text{-X-S-X-P-C}^2$  (POGLUT1) ou  $C^3\text{-X-N-T-X-G-S-F-X-C}^4$  (POGLUT 2 et 3) (Takeuchi et al., 2018) (Figure 10).

### II.1.3. L'interaction des ELDs avec Pofut1

L'une des étapes clés de la *O*-fucosylation est l'interaction de l'ELD avec Pofut1. Récemment, des études ont révélé que certains résidus de POFUT1 étaient impliqués dans l'interaction avec les ELDs (Li et al., 2017).

Le sous-domaine  $C^2\text{-C}^3$  des ELDs, contenant la séquence de *O*-fucosylation, est le plus important pour l'interaction avec POFUT1. En effet, plusieurs résidus de ce sous-domaine sont susceptibles d'interagir par des liaisons hydrogènes avec des acides aminés clés de l'enzyme. Selon leur nature, les résidus en position  $C^2+1$  et  $C^2+2$  peuvent former des liaisons avec l'asparagine 151 ( $N^{151}$ ) de l'enzyme POFUT1 murine. De la même façon, l'acide aminé en position  $C^2+4$  est capable d'interagir avec la glycine 47 ( $G^{47}$ ). De plus la cystéine 3 ( $C^3$ ) de l'ELD peut établir une liaison soufre-hydrogène avec la méthionine 46 ( $M^{46}$ ) de POFUT1 (Figure 11) (Li et al., 2017).

Le résidu en position  $C^2+3$ , le plus souvent une glycine, est connu pour former des liaisons de Van Der Waals avec l'enzyme. La présence d'un aspartate ou d'un glutamate qui possède une chaîne latérale plus longue que la glycine entraîne une diminution d'interaction du fait d'un clash stérique avec la tyrosine 78 ( $Y^{78}$ ) de POFUT1 (Li et al., 2017).



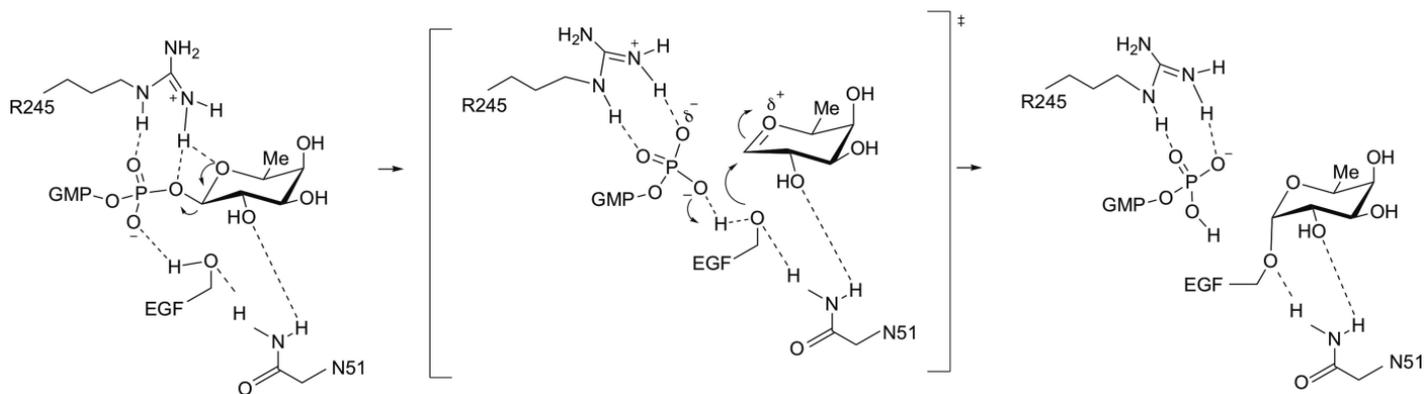
L'acide aminé en position C<sup>2</sup>+5 (position de la sérine ou de la thréonine *O*-fucosylable) est connu pour établir une liaison hydrogène avec l'asparagine 51 (N<sup>51</sup>) de POFUT1. Cependant, la co-cristallisation de l'ELD du facteur de coagulation VII avec POFUT1 a montré une interaction supplémentaire de type apolaire entre le groupement méthyle de la thréonine et les résidus phénylalanine 204 (F<sup>204</sup>) et méthionine 267 (M<sup>267</sup>) de l'enzyme. Cette interaction apolaire est perdue lorsque la thréonine est substituée par une sérine ce qui se traduit par une diminution d'affinité de POFUT1 pour l'ELD. Ainsi, la présence d'une thréonine au lieu d'une sérine en position C<sup>2</sup>+5 favorise légèrement l'interaction avec l'enzyme (Li et al., 2017).

Bien que le sous-domaine C<sup>2</sup>-C<sup>3</sup> soit le plus important, d'autres résidus de l'ELD peuvent être impliqués dans cette interaction. Ainsi la présence de la sérine dans la boucle C<sup>1</sup>-C<sup>2</sup> favorise l'établissement d'une liaison hydrogène avec l'arginine 138 (R<sup>138</sup>), malgré le fait que cette région soit connue pour ne pas interagir fortement avec POFUT1. Selon leur nature, les acides aminés en position C<sup>4</sup>+1 et C<sup>5</sup>+4 sont capables de se lier à l'enzyme par interaction hydrophobe. Par contre, il est clairement établi que la boucle C<sup>3</sup>-C<sup>4</sup> n'intervient pas dans l'interaction avec POFUT1 car celle-ci est exposée au solvant (Li et al., 2017).

#### **II.1.4. Activité catalytique de Pofut1**

L'enzyme Pofut1 est une glycohydrolase et une glycosyltransférase. Son activité hydrolase lui est nécessaire pour cliver le GDP-fucose en absence d'ELD. La présence de ce substrat accepteur induit une transition vers un mécanisme de transfert.

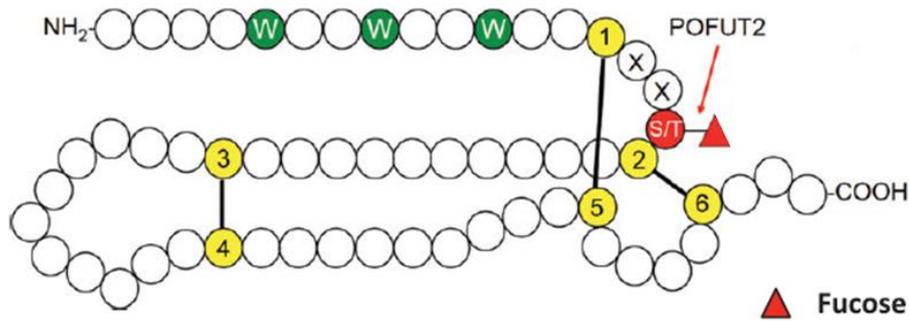
Comme toutes les glycosyltransférases avec un repliement de type GT-B, Pofut1 ne nécessite pas la présence de cations divalents pour être active, à l'inverse des glycosyltransférases au repliement de type GT-A. Cependant, la présence d'un domaine DXD lui permet d'utiliser divers cations divalents pour atteindre son activité optimale (Breton et al., 2006). Pofut1 utilise préférentiellement du manganèse (Mn<sup>2+</sup>), mais d'autres cations métalliques peuvent le substituer avec des effets moins prononcés tels que le cobalt (Co<sup>2+</sup>), le cadmium (Cd<sup>2+</sup>), le magnésium (Mg<sup>2+</sup>) et le nickel (Ni<sup>2+</sup>). A l'inverse, la présence de certains ions dont le fer (Fe<sup>3+</sup>), le cuivre (Cu<sup>2+</sup>) et le zinc (Zn<sup>2+</sup>) inhibent totalement l'activité de l'enzyme (Wang and Spellman, 1998). Il a néanmoins été montré que l'enzyme Pofut1 de *Caenorhabditis elegans* n'était pas dépendante de la présence de Mn<sup>2+</sup> mais qu'au contraire cela avait un effet inhibiteur (Lira-Navarrete et al., 2011).



### Figure 12 : Mécanisme catalytique de l'enzyme POFUT1 murine (type S<sub>N</sub>1).

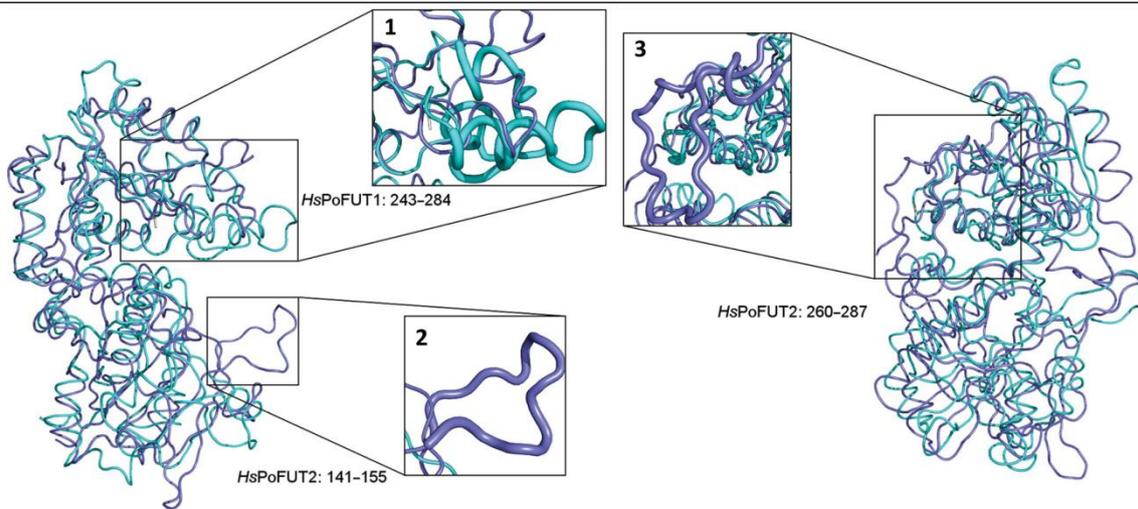
Lors de la *O*-fucosylation médiée par POFUT1, l'arginine 245 (R<sup>245</sup>) et l'asparagine 51 (N<sup>51</sup>) de POFUT1 forment des liaisons hydrogènes avec le groupe β phosphate du GDP et l'ELD (EGF) ce qui permet de stabiliser le complexe et facilite l'hydrolyse du GDP. La présence d'une molécule d'eau à proximité du groupement hydroxyle favorise la catalyse de type S<sub>N</sub>1 en induisant le transfert du proton (H) vers l'oxygène (O) du GDP. Il s'ensuit l'activation du groupement hydroxyle du résidu *O*-fucosylable (sérine ou thréonine) et donc l'attaque nucléophile sur le carbone 1 du fucose. (Lira-Navarrete and Hurtado Guerrero, 2018)

Pofut1 est classée dans la famille 65 de la base de données CAZy (Carbohydrate-Active enZymes) et est caractérisée par un mécanisme d'invertase (Coutinho et al., 2003). En tant que glycosyltransférase de type invertase, Pofut1 devrait utiliser le mécanisme  $S_N2$  (substitution nucléophile bimoléculaire) caractéristique de cette famille d'enzymes. Ce mécanisme fonctionne *via* une base catalytique capable d'activer le groupement hydroxyle du résidu *O*-fucosylable de l'ELD pour l'attaque nucléophile du carbone 1 du fucose. L'absence de base catalytique chez Pofut1 suggère l'utilisation d'un mécanisme similaire au type  $S_N1$  (substitution nucléophile monomoléculaire) mais sans formation d'intermédiaire (cation oxocarbenium) comme c'est le cas pour Pofut1 de *Caenorhabditis elegans* (Lira-Navarrete et al., 2011). Cela semble être confirmé sur POFUT1 murin par la présence d'une liaison hydrogène entre le groupement hydroxyle du résidu *O*-fucosylable et l'asparagine 51 ( $N^{51}$ ) de POFUT1 inadapté à un mécanisme de type  $S_N2$ . Une liaison similaire entre l'arginine 245 ( $R^{245}$ ) et le groupement  $\beta$ -phosphate du GDP stabilise le complexe POFUT1/GDP-fucose et permet le clivage du GDP. De plus, il a été observé, à proximité du groupement hydroxyle du résidu sérine ou thréonine *O*-fucosylable de l'ELD la présence d'une molécule d'eau qui interagit *via* une liaison hydrogène avec l'oxygène  $\beta$ -phosphate du GDP. Ceci permet de fournir un relai protonique pour faire passer le proton issu du groupement hydroxyle de l'accepteur (sérine ou thréonine) sur l'oxygène du groupement  $\beta$ -phosphate du GDP servant ainsi de base catalytique pour un mécanisme de type  $S_N1$ . Cette étape induit l'activation du groupement hydroxyle du résidu *O*-fucosylable permettant la liaison de l'oxygène de ce résidu avec le carbone du fucose (Figure 12). POFUT1 possède également une autre particularité, l'absence de boucle structurée au niveau de la liaison avec le GDP-fucose, qui favoriserait la liaison de l'ELD tout en excluant les molécules d'eau. Ceci aboutit à un complexe POFUT1/GDP-fucose anormalement exposé aux molécules d'eau. Il semblerait donc que l'absence de résidus catalytiques inhérents au mécanisme de type  $S_N2$  empêcherait l'activation du groupement hydroxyle des molécules d'eau lors de l'hydrolyse du GDP-fucose. La fixation de l'ELD dans la cavité de POFUT1 induirait un switch vers un mécanisme proche du type  $S_N1$  pour favoriser la *O*-fucosylation au détriment de l'hydrolyse du GDP-fucose (Li et al., 2017). En effet, la fixation de l'ELD permettrait d'exclure l'eau au niveau du site catalytique de l'enzyme via ses sous-domaines  $C^2$ - $C^3$  et  $C^5$ - $C^6$  ce qui suggère la présence d'un ELD correctement replié. Au-delà de ce rôle d'étanchéité, l'ELD bloquerait le site de liaison de POFUT1 pour le GDP-fucose en enfouissant ce sucre nucléotidique dans la cavité de l'enzyme. Tout cela tend à suggérer que POFUT1 utilise un mécanisme



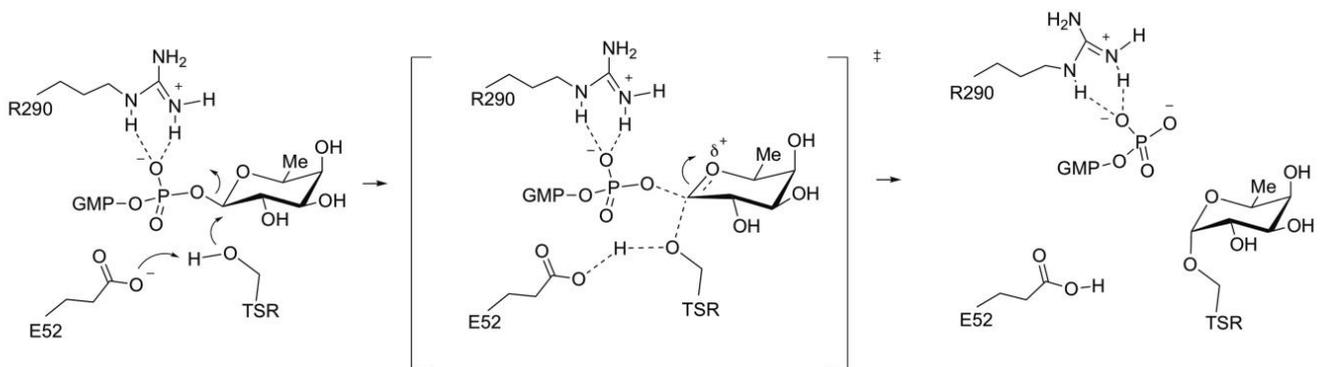
**Figure 13 : Représentation de la *O*-fucosylation de TSR de type 1.**

Les TSRs peuvent être *O*-fucosylés par POFUT2 sur la sérine ou thréonine de la séquence consensus C<sup>1</sup>-X-X-(S/T)-C<sup>2</sup>-G. Il existe également des TSRs de type 2 pouvant être *O*-fucosylés sur une séquence consensus de type ou C<sup>2</sup>-X-X-(S/T)-C<sup>3</sup>-G. (Schneider et al., 2017)



**Figure 14 : Différence structurale entre les enzymes POFUT1 et POFUT2 humaines.**

Les résidus 141 à 155 (encadré 2) et 260 à 287 (encadré 3) forment deux boucles chez POFUT2 (en violet), qui sont absentes chez POFUT1 (en cyan). Aucun rôle catalytique n'a été associé à la boucle composée des résidus 260 à 287. Cependant, la boucle 141-155 est impliquée dans la liaison sélective des TSRs et associée à l'exclusion des ELDs. A contrario, la boucle composée des résidus 243 à 284 (encadré 1) présente chez POFUT1 uniquement, a l'effet inverse et favorise donc la fixation des ELDs au détriment des TSRs. (Lira-Navarrete and Hurtado-Guerrero, 2018)



**Figure 15 : Mécanisme catalytique de l'enzyme Pofut2 (*C.elegans*) (type S<sub>N</sub>2).**

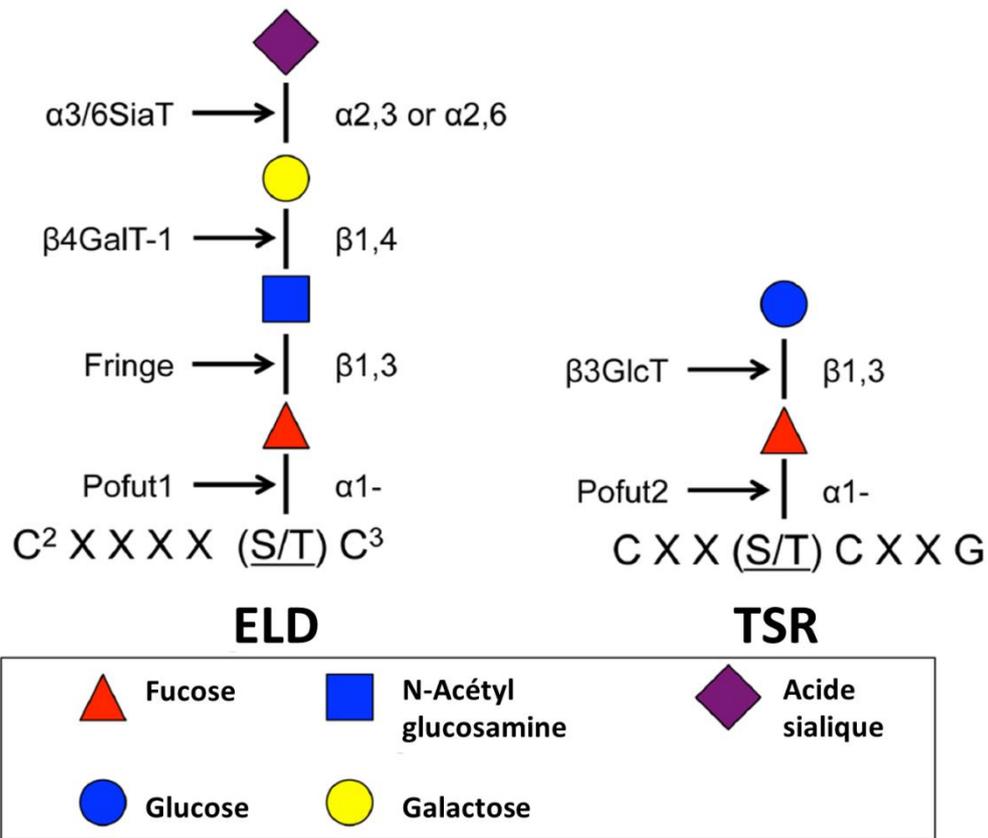
Lors de la *O*-fucosylation médié par POFUT2, l'arginine 290 (R<sup>290</sup>) de l'enzyme interagit avec le groupe  $\beta$  phosphate du GDP. Le glutamate 52 (E52) est la base catalytique caractéristique du système de type S<sub>N</sub>2, qui permet la déprotonation du groupement hydroxyle du résidu *O*-fucosylable du TSR. Cette étape permet l'attaque nucléophile du carbone 1 du fucose nécessaire à la *O*-fucosylation du TSR. (Lira-Navarrete and Hurtado-Guerrero, 2018)

cinétique bi-séquentiel ordonné (fixation du substrat donneur avant le substrat accepteur) (Li et al., 2017).

## II.2. La *O*-fucosylation dépendante de Pofut2

L'autre enzyme responsable de la *O*-fucosylation est la protéine *O*-fucosyltransférase 2 (Pofut2), mais contrairement à Pofut1, elle cible les TSRs (Hofsteenge et al., 2001) qui sont de petits domaines protéiques d'une soixantaine de résidus environ, caractérisés par 6 cystéines conservées formant 3 ponts disulfures. Il existe deux types de TSRs qui se caractérisent par leurs ponts disulfures, les TSRs de type 1 établissent des ponts entre les cystéines C<sup>1</sup>-C<sup>5</sup>, C<sup>2</sup>-C<sup>6</sup> et C<sup>3</sup>-C<sup>4</sup> alors que ces liaisons se font entre les cystéines C<sup>1</sup>-C<sup>4</sup>, C<sup>2</sup>-C<sup>5</sup>, C<sup>3</sup>-C<sup>6</sup> pour les TSRs de type 2 (Leonhard-Melief and Haltiwanger, 2010). Comme pour les ELDs, la *O*-fucosylation des TSRs ne peut se faire qu'en présence d'une séquence consensus de *O*-fucosylation spécifique, C<sup>1</sup>-X-X-(S/T)-C<sup>2</sup>-G ou C<sup>2</sup>-X-X-(S/T)-C<sup>3</sup>-G en fonction du type de TSR (type 1 ou 2) (Figure 13) (Schneider et al., 2017).

Pofut2 présente des similitudes avec Pofut1, notamment pour son activité *O*-fucosyltransférase. On retrouve donc chez ces deux enzymes un repliement caractéristique de type GT-B ainsi qu'une forte conservation des résidus impliqués dans la liaison du GDP-fucose. Cependant, la différence de substrats accepteurs suggère des divergences entre ces deux enzymes. Les enzymes humaines POFUT1 et POFUT2 n'ont qu'une identité relativement faible (28% environ). Ce qui est à corrélé avec la présence de deux régions supplémentaires chez POFUT2 (les boucles 141-155 et 260-287) et une chez POFUT1 (la boucle 243-284) (Figure 14). La boucle comprenant les résidus 141-155 de POFUT2 permet une liaison sélective des TSRs et une exclusion des ELDs. A l'inverse, la boucle supplémentaire de POFUT1 permet d'interagir spécifiquement avec les ELDs et de prévenir la fixation des TSRs. Tout ceci est en adéquation avec des différences de stratégie de liaison du substrat accepteur mais également avec des mécanismes catalytiques différents, un mécanisme de type S<sub>N</sub>1 pour POFUT1 et un mécanisme de type S<sub>N</sub>2 pour POFUT2 (Figure 15) (Lira-Navarrete and Hurtado-Guerrero, 2018).



**Figure 16 : L'extension des *O*-fucoses par d'autres monosaccharides.**

Le *O*-fucose transféré sur un ELD peut être successivement étendu par un *N*-acétylglucosamine, un galactose et un acide sialique suite à l'action d'une  $\beta$ 3-*N*-acétylglucosaminyltransférase de la famille Fringe, d'une  $\beta$ 1,4-galactosyltransférase ( $\beta$ 4GalT-1) et d'une  $\alpha$ 2,3- ou une  $\alpha$ 2,6-sialyltransférase ( $\alpha$ 3/6SiaT). Par contre, le fucose *O*-lié sur les TSR peut être rallongé par un glucose par une  $\beta$ 1-3-glucosyltransferase ( $\beta$ 3GlcT). (Vasudevan and Haltiwanger, 2015)

### II.3. Autres domaines *O*-fucosylables

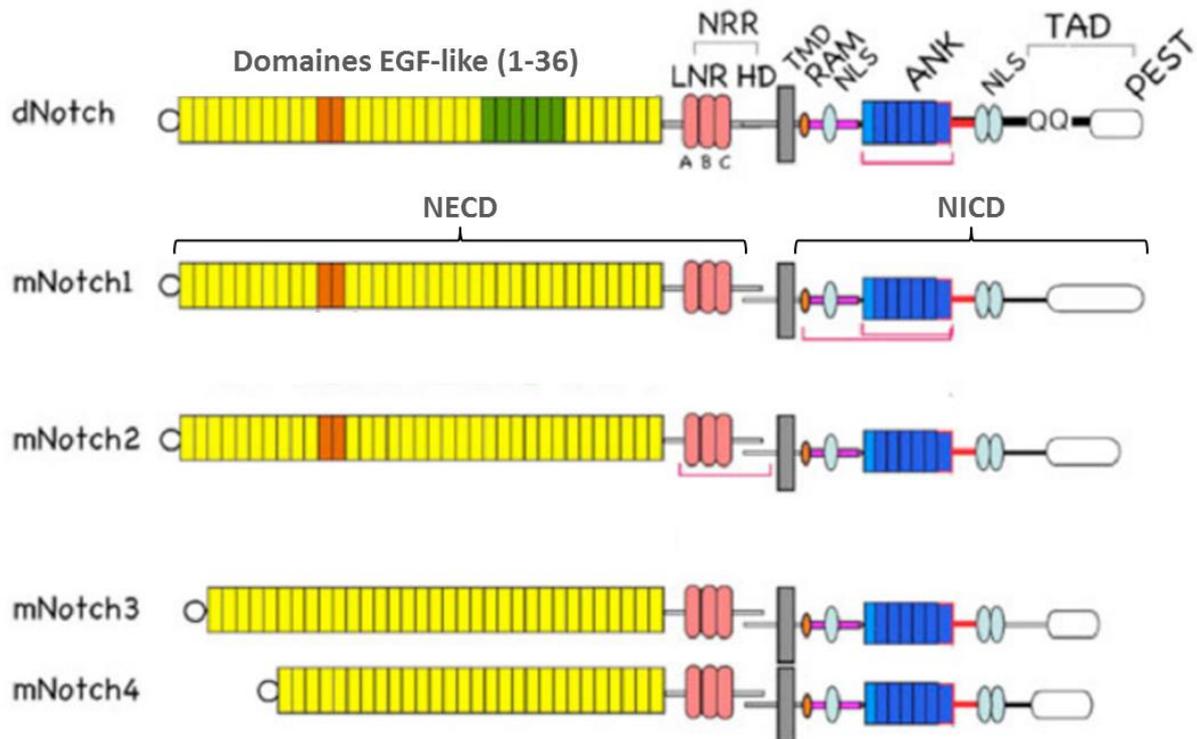
Il semble que les ELDs et les TSR ne soient pas les seuls à pouvoir être *O*-fucosylés. En effet, une étude a montré la présence d'un *O*-fucose sur le second résidu sérine d'une séquence peptidique A-S-Q-G-I-S-R, issue de la chaîne légère d'une IgG1 recombinante (Valliere-Douglass et al., 2008). Ce qui suggère donc que Pofut1 ou Pofut2 seraient capables d'*O*-fucosyler d'autres domaines ou régions protéiques que ceux qui leurs sont attribués (ELD ou TSR) ou qu'il existe encore une ou plusieurs autres protéines *O*-fucosyltransférases encore non identifiées. Cependant, aucune étude récente n'a confirmé la *O*-fucosylation de cette protéine.

### II.4. Le rallongement des *O*-fucoses

Une fois le *O*-fucose transféré sur la sérine ou la thréonine de la séquence consensus, il peut être rallongé par d'autres monosaccharides suite à l'action de différentes glycosyltransférases. Dans le cas d'un *O*-fucose greffé sur un ELD, la première enzyme à intervenir est une  $\beta$ 1-3-*N*-acétylglucosaminyltransférase golgienne de la famille Fringe (comprenant Lunatic, Manic et Radical) qui permet l'ajout en  $\beta$ 1,3 d'un GlcNAc sur le fucose *O*-lié pour former un disaccharide GlcNAc $\beta$ 1-3Fucose (Figure 16) (Brückner et al., 2000; Moloney et al., 2000a; Vasudevan and Haltiwanger, 2014). Cependant, tous les *O*-fucoses ne sont pas nécessairement rallongés par un GlcNAc. De plus, des études récentes ont montré que tous les *O*-fucoses ne sont pas modifiés de la même manière par les enzymes de la famille Fringe. Ainsi, si certains d'entre eux peuvent être modifiés par Lunatic, Manic et Radical Fringe d'autres sont insensibles à l'action d'une ou deux enzymes de la famille Fringe (Kakuda and Haltiwanger, 2017).

Suite à l'action de Fringe, le disaccharide de type GlcNAc $\beta$ 1-3Fucose peut être étendu avec un galactose par une  $\beta$ 1,4-galactosyltransférase (Chen et al., 2001) puis un acide sialique par une  $\alpha$ 2,3- ou une  $\alpha$ 2,6-sialyltransférase pour former un tétrasaccharide sialylé de type Sia $\alpha$ 2-3/6Gal $\beta$ 1-4GlcNAc $\beta$ 1-3Fuc (Moloney et al., 2000a).

Il est intéressant de noter que les *O*-fucoses présents sur les TSRs peuvent également être rallongés par un glucose suite à l'action d'une  $\beta$ 1-3-glucosyltransférase afin de former un dissaccharide Glc $\beta$ 1-3Fucose (Vasudevan and Haltiwanger, 2014).



**Figure 17 : Structures des différents récepteurs Notch.**

Les récepteurs Notch 1 à 4 des mammifères (mNotch1 à 4) et celui de la drosophile (dNotch) se composent d'un domaine extracellulaire (NECD), d'un domaine transmembranaire (TMD) et d'un domaine intracellulaire (NICD). Le NECD comprend entre 29 et 36 ELDs et une région régulatrice négative (NRR) comprenant trois répétitions riches en cystéine de type Lin12-Notch (LNR) et un domaine d'hétérodimérisation (HD). Le NICD se compose d'un module d'association au RBPj $\kappa$  (RAM), de deux séquences d'adressage nucléaire (NLS), de sept répétitions ankyrines (ANK), d'un domaine de transactivation (TAD) et d'une queue C-terminale (PEST). (Kopan and Ilagan, 2009)

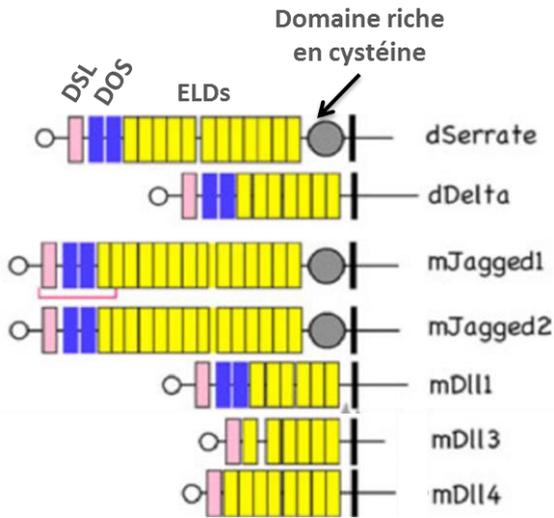
## Chapitre III. Les protéines *O*-fucosylées

---

La *O*-fucosylation d'une protéine est relativement complexe à mettre en évidence du fait de l'absence d'anticorps ou de lectines reconnaissant spécifiquement les *O*-fucoses. De plus, aucune méthode chimique ou enzymatique ne permet de cliver spécifiquement les *O*-fucosylglycanes. Ainsi, moins d'une quinzaine de protéines, pour la plupart recombinantes, ont été démontrées comme étant *O*-fucosylées sur un ou plusieurs de leurs ELDs, principalement par des approches de spectrométrie de masse et de radiomarquage métabolique au  $^{14}\text{C}$ -fucose (Bjoern et al., 1991) ou au  $^3\text{H}$ -fucose (Rampal et al., 2005). Une seule étude fait état d'une méthode alternative pour mettre en évidence les *O*-fucoses, le marquage métabolique utilisant du 6-fucose-alcyne associé à de la chimie click (Al-Shareffi et al., 2013). Ce marquage métabolique, comme le marquage au fucose radioactif, dépend de la voie de récupération ou « sauvetage » du fucose. Cette voie qui représente environ 10% du fucose utilisé par les cellules pour leurs différentes glycosylations consiste en la récupération de L-fucose présent dans le milieu extracellulaire. Ainsi le  $^{14}\text{C}$ -fucose,  $^3\text{H}$ -fucose ou le 6-fucose-alcyne peut être récupéré dans le milieu et incorporé dans des structures glycaniques par les cellules à la place du L-fucose non modifié (Sawa et al., 2006). Dans le cas de l'utilisation de 6-fucose-alcyne, la fixation d'une biotine azido sur le fucose marqué par cycloaddition azide-alcyne assistée au cuivre ou chimie click permet de visualiser les glycoprotéines qui ont incorporé ce sucre modifié. Néanmoins, l'utilisation de lignées spécifiques et/ou de traitements enzymatique ou chimique sont nécessaires pour atténuer le signal de l'incorporation de fucose marqué à la périphérie de *N*- et *O*-glycanes et donc faciliter l'analyse des fucoses *O*-liés.

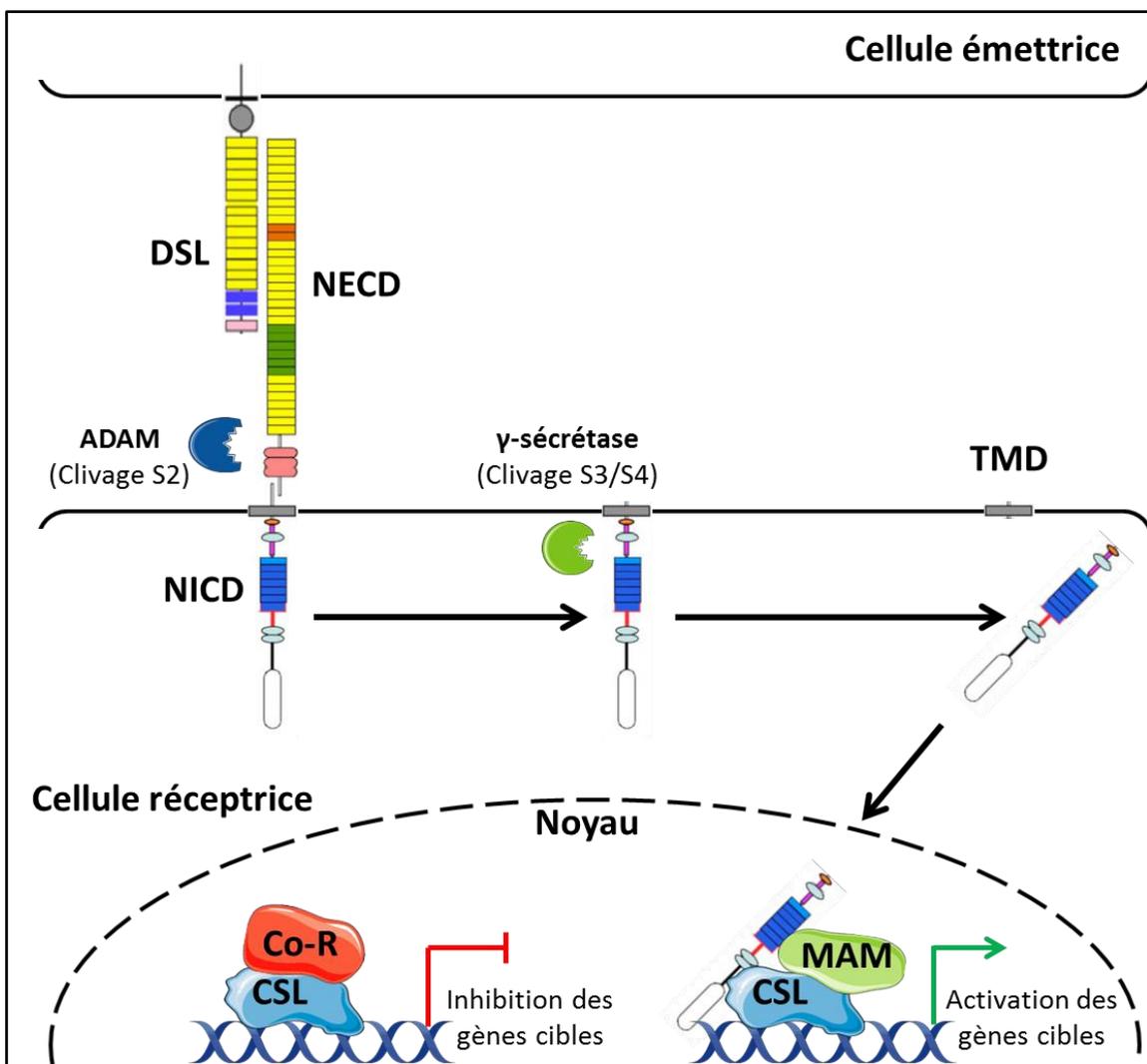
### III.1. Le récepteur Notch

Le gène codant le récepteur Notch a été découvert chez la drosophile en 1919 suite à l'analyse de la mutation d'une région chromosomique induisant l'apparition d'entailles à l'extrémité des ailes chez les femelles hétérozygotes (Mohr, 1919). Présent en un seul exemplaire chez la drosophile, deux et quatre paralogues sont respectivement retrouvés chez *Caenorhabditis elegans* et chez les mammifères (Notch 1 à 4) (Kopan and Ilagan, 2009) (Figure 17). Depuis la découverte de fucoses *O*-liés sur Notch1, ce récepteur demeure la protéine *O*-fucosylée la plus connue et la plus étudiée (Moloney et al., 2000b).



**Figure 18 : Structures des différents ligands canoniques des récepteurs Notch.**

Les ligands Delta et Serrate de drosophile et leurs paralogues mammaliens Delta-like (DII1, 3 et 4) et Jagged (1 et 2) se composent d'un domaine extracellulaire qui contient un domaine d'interaction de type DSL (Delta/Serrate/LAG-2) et plusieurs domaines EGF-like (ELDs). Les deux ELDs N-terminaux forment le domaine DOS (Delta and OSM-11-like proteins) également impliqués dans l'interaction avec Notch. Les ligands de la famille Jagged possèdent en plus une région juxtamembranaire, riche en cystéine permettant de les distinguer des DLLs. (Kopan and Ilagan, 2009)



**Figure 19 : La voie de signalisation Notch.**

L'interaction entre le récepteur Notch d'une cellule réceptrice et un ligand (DSL) d'une cellule émettrice induit le clivage (S2) du NECD par une protéine contenant un domaine désintégrine/métalloprotéinase (ADAM). Le NICD sera séparé du domaine transmembranaire (TMD) par un second clivage (S3/S4) résultant de l'action d'une  $\gamma$ -sécrétase. Ce NICD va être ensuite transloqué au noyau pour interagir avec des protéines de liaison à l'ADN de la famille des CSL et son co-activateur Mastermind (MAM), ce qui va induire la transcription de différents gènes cibles. En absence d'interaction du récepteur Notch avec ses ligands, les protéines de la famille CSL sont complexées avec un co-represseur (Co-R), inhibant ainsi l'activation des gènes cibles de la voie Notch. (D'après Kopan and Ilagan, 2009)

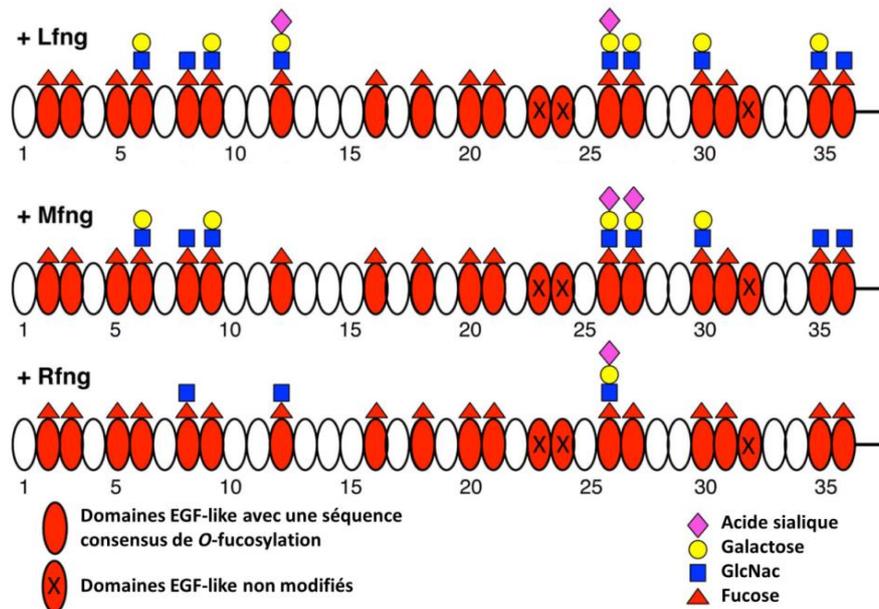
Chez les mammifères, Notch est un récepteur transmembranaire qui contient un domaine extracellulaire (NECD) avec 29 à 36 ELDs (Notch1-4), suivis d'une région de régulation négative (NRR), comportant trois répétitions riches en cystéine de type Lin12-Notch (LNR) et un domaine d'hétérodimérisation (HD), permettant d'inhiber le récepteur en l'absence d'interaction avec un ligand (Figure 17) (Kopan and Ilagan, 2009).

Le domaine transmembranaire (TMD) est prolongé par la partie intracellulaire (NICD) du récepteur, composée d'un module d'association au RBPjk (RAM), de deux séquences d'adressage nucléaire (NLS), de sept répétitions ankyrines (ANK), d'un domaine de transactivation (TAD) et d'une queue C-terminale (PEST) stabilisant le domaine intracellulaire de Notch (NICD) (Figure 17) (Kopan and Ilagan, 2009).

### **III.1.1. La voie de signalisation induite par Notch**

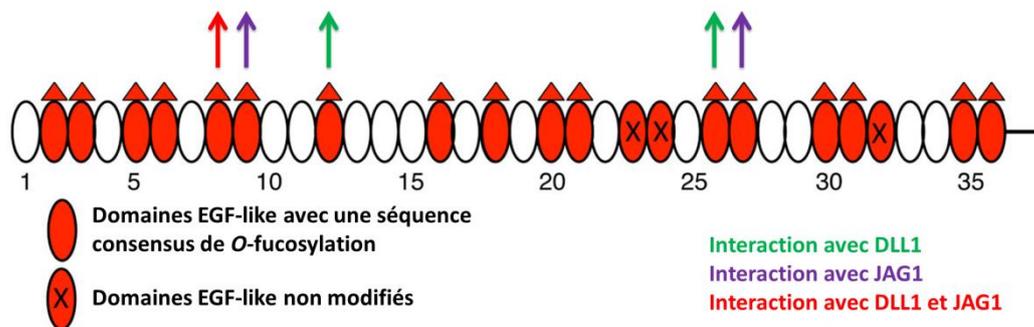
Il existe plusieurs voies de signalisation induites par le récepteur Notch, la principale est la voie canonique dépendant de l'activation du récepteur par les ligands transmembranaires de type DSL (Delta/Serrate/LAG-2) comprenant deux familles, Delta-like (DLL1, 3 et 4) et Jagged (JAG1 et 2), chez les mammifères (Figure 18). La partie extracellulaire de ces ligands de type DSL se compose d'un domaine DSL N-terminal et de plusieurs ELDs (entre 6 et 8 pour les DLLs et 16 pour les JAGs chez les mammifères). Les ligands de type JAG se distinguent des DSLs de type DLL par la présence d'une région riche en cystéine à proximité de leur domaine transmembranaire. L'interaction de ces ligands avec le récepteur Notch se fait principalement par leur domaine DSL et leur domaine DOS (Delta and OSM-11-like proteins) correspondant au duo d'ELDs N-terminaux (Figure 18) (Kopan and Ilagan, 2009).

Lorsque le récepteur Notch d'une cellule réceptrice interagit avec un ligand de type DSL d'une cellule émettrice, un processus d'endocytose du complexe NECD/ligand va provoquer une tension, ce qui va induire un changement conformationnel du récepteur (Meloty-Kapella et al., 2012) permettant le clivage (S2) du NECD par une peptidase de la famille ADAM, protéine comportant un domaine désintégrine/métalloprotéinase. Une fois la partie extracellulaire de Notch clivée, le complexe NECD/ligand DSL sera recyclé suite à son endocytose par la cellule émettrice. Au niveau de la cellule réceptrice, un second clivage (S3/S4) est opéré par la  $\gamma$ -sécrétase, libérant ainsi le NICD du récepteur dans le cytosol ; il sera par la suite transloqué au noyau. Une fois dans le noyau, le NICD clivé va



**Figure 20 : La O-fucosylation du récepteur NOTCH1.**

Sur les 20 ELDs O-fucosylables de NOTCH1, 17 sont réellement porteurs d'un O-fucose et 9 peuvent être prolongés par un N-acétylglucosamine (GlcNAc) suite à l'action de Lunatic Fringe (Lfng) et/ou de Mannic Fringe (Mfng) et/ou de Radical Fringe (Rfng). (Kakuda and Haltiwanger, 2017)



**Figure 21 : Positions des O-fucoses de NOTCH1 favorisant l'interaction avec les ligands DSL.**

Les O-fucoses des ELDs 12 et 26 augmentent l'activation de NOTCH1 par DLL1 (flèches vertes) tandis que ceux présents sur les ELDs 9 et 27 augmentent l'activation induite par JAG1 (flèches mauves). Le O-fucose porté par l'ELD 8 favorise l'activation de NOTCH1 par DLL1 et JAG1 (flèche rouge). (D'après Kakuda and Haltiwanger, 2017)

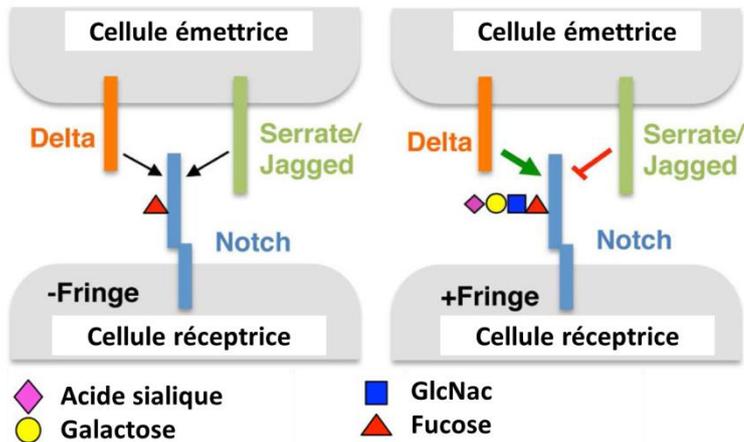
s'associer à des protéines de la famille des CSL (CBF-1/RBPJ- $\kappa$ /Suppressor of Hairless/Lag-1) et à son co-activateur Mastermind (MAM), induisant ainsi la transcription de différents gènes cibles de la famille *Hes* (hairy and enhancer of split-1) et *Hey* (Hairy-related transcription factor) (Figure 19) (Bray, 2006; Hori et al., 2013; Kopan and Ilagan, 2009; Wang et al., 2015).

La voie de signalisation Notch est connue pour son rôle fondamental dans de nombreux processus développementaux tels que la somitogenèse, la myogenèse, la neurogenèse et l'angiogenèse ainsi que dans le maintien de l'homéostasie des tissus adultes (Artavanis-Tsakonas, 1999; Bigas et al., 2010; Conboy and Rando, 2002; Conlon et al., 1995). La voie Notch est particulièrement importante pour le maintien du pool de cellules souches dans un état indifférencié en régulant la balance entre prolifération et différenciation cellulaire (Chen et al., 2014; Conboy and Rando, 2002; Hori et al., 2013). Ainsi une altération de la voie Notch est souvent corrélée à de nombreuses maladies et cancers (Hori et al., 2013).

### **III.1.2. La *O*-fucosylation de NOTCH1**

Parmi les 36 ELDs du récepteur NOTCH1 murin, 20 possèdent une séquence consensus de *O*-fucosylation, ce sont les ELDs 2, 3, 5, 6, 8, 9, 12, 16, 18, 20, 21, 23, 24, 26, 27, 30, 31, 32, 35 et 36. Une étude récente a mis en évidence la présence de *O*-fucoses sur 17 d'entre eux (seuls les ELDs 23, 24 et 32 ne sont pas *O*-fucosylés). De plus, la majorité de ces *O*-fucoses peuvent être allongés par un GlcNAc suite à l'action d'une enzyme de la famille Fringe (Lunatic, Manic, Radical), excepté les ELDs 2, 3, 5, 16, 18, 20, 21 et 31 qui ne semblent être la cible d'aucune glycosyltransférase de cette famille (Figure 20) (Kakuda and Haltiwanger, 2017).

Plusieurs *O*-fucoses du récepteur NOTCH1 sont indispensables à la transduction d'un signal suite à son interaction avec ses ligands DSL. Ainsi la perte du *O*-fucose sur les ELDs 12 et 26 réduit fortement l'activation du récepteur par son ligand DLL1 mais n'a aucun effet sur sa capacité à être activé par JAG1. De manière opposée, l'absence de fucose *O*-lié sur les ELDs 9 et 27 altère l'activation de NOTCH1 induite par JAG1 mais n'a que peu d'effet sur son activation par DLL1. La mutation du site de *O*-fucosylation de l'ELD 8 diminue fortement l'activation du récepteur induite par ses deux ligands (Figure 21). Néanmoins, certains ELDs *O*-fucosylés ne semblent avoir que peu d'effet sur la voie Notch. Par exemple, l'absence de *O*-fucosylation sur les ELDs 6, 30, 35 et 36 ne participe que peu à la capacité des ligands DSL à activer cette voie (Kakuda and Haltiwanger, 2017).



**Figure 22 : Fringe module l'activation de Notch par ses ligands DSL.**

En présence d'un *O*-fucose non rallongé l'activation de la voie Notch se fait sans distinction par Delta et Jagged. Lorsque le *O*-fucose est étendu suite à l'action de Fringe, la capacité de Delta à activer le récepteur augmente tandis que celle de Jagged diminue. (Kakuda and Haltiwanger, 2017)

Gene	Protein	Gene	Protein
AGRIN	Agrin	LTBP2	Latent-transforming growth factor beta-binding protein 2
ATRAID	All-trans retinoic acid-induced differentiation	MEGF6	Multiple epidermal growth factor-like domains protein 6
CELSR1	Cadherin EGF LAG seven-pass G-type receptor 1	MEGF8	Multiple epidermal growth factor-like domains protein 8
CELSR2	Cadherin EGF LAG seven-pass G-type receptor 2	MEGF10	Multiple epidermal growth factor-like domains protein 10
CELSR3	Cadherin EGF LAG seven-pass G-type receptor 3	MEGF11	Multiple epidermal growth factor-like domains protein 11
CD93	Complement component C1q receptor	MMRN1	Multimerin-1
CD97	CD97 antigen	NCAN	Neurocan core protein
CFC1	Cryptic protein	NELL1	Protein kinase C-binding protein NELL1
CFC1B	Cryptic family protein 1B	NID2	Nidogen-2
CNTNAP5	Contactin-associated protein-like 5	NOTCH1	Neurogenic locus notch homolog protein 1
CRB1	Protein crumbs homolog 1	NOTCH2	Neurogenic locus notch homolog protein 2
CRB2	Protein crumbs homolog 2	NOTCH2NL	Notch homolog 2 N-terminal-like protein
CSPG2	Versican core protein	NOTCH3	Neurogenic locus notch homolog protein 3
CUBN	Cubilin	NOTCH4	Neurogenic locus notch homolog protein 4
DLK1	Protein delta homolog 1	PEAR1	Platelet endothelial aggregation receptor 1
DLK2	Protein delta homolog 2	PGBM	Basement membrane-specific heparan sulfate proteoglycan core protein
DLL1	Delta-like protein 1	PGCB	Brevican core protein
DLL3	Delta-like protein 3	PROC	Vitamin K-dependent protein C
DLL4	Delta-like protein 4	PROZ	Vitamin K-dependent protein Z
DNER	Delta and Notch-like epidermal growth factor-related receptor	RAMP	Regeneration-associated muscle protease homolog
EDIL3	EGF-like repeat and discoidin I-like domain-containing protein 3	RELN	Reelin
EGF	Pro-epidermal growth factor	SLIT1	Slit homolog 1 protein
EGFL7	Epidermal growth factor-like protein 7	SLIT2	Slit homolog 2 protein
EGFLAM	Pikachurin	SLIT3	Slit homolog 3 protein
EMR1	Adhesion G protein-coupled receptor E1	SNED1	Sushi, nidogen and EGF-like domain-containing protein 1
EMR2	Adhesion G protein-coupled receptor E2	SREC2	Scavenger receptor class F member 2
EYS	Protein eyes shut homolog	STAB1	Stabilin-1
FA7	Coagulation factor VII	STAB2	Stabilin-2
FA9	Coagulation factor IX	SUSD1	Sushi domain-containing protein 1
FA12	Coagulation factor XII	SVEP1	Polydom protein
FAT1	Protocadherin Fat 1	TEN1	Teneurin-1
FAT2	Protocadherin Fat 2	TEN2	Teneurin-2
FAT3	Protocadherin Fat 3	TEN4	Teneurin-4
FAT4	Protocadherin Fat 4	TIE1	Tyrosine-protein kinase receptor Tie-1
FBLN1	Fibulin-1	TPA	Tissue-type plasminogen activator
FBLN7	Fibulin-7	TSP3	Thrombospondin-3
FBN2	Fibrillin-2	UMOD	Uromodulin
FBN3	Fibrillin-3	UMODL1	Uromodulin-like 1
HABP2	Hyaluronan-binding protein 2	UROK	Urokinase-type plasminogen activator
HGFAC	Hepatocyte growth factor activator	VASN	Vasorin
JAG1	Protein jagged-1	VWDE	von Willebrand factor D and EGF domain-containing protein
JAG2	Protein jagged-2	VWA2	AMACO
LRP1	Pro-low-density lipoprotein receptor-related protein 1	WIF1	Wnt inhibitory factor 1
LRP1B	Low-density lipoprotein receptor-related protein 1B		

**Figure 23 : Liste des protéines humaines potentiellement cibles de POFUT1.**

Il existe 87 protéines humaines comportant 1 à 20 ELDs contenant chacun une séquence consensus de *O*-fucosylation de type C<sup>2</sup>-X-X-X-X-(S/T)-C<sup>3</sup> et donc potentiellement cibles de POFUT1. En vert, les protéines dont la présence d'un *O*-fucose a été démontrée et en rouge celles dont le rôle de ce *O*-fucose est connu. (D'après Schneider et al., 2017)

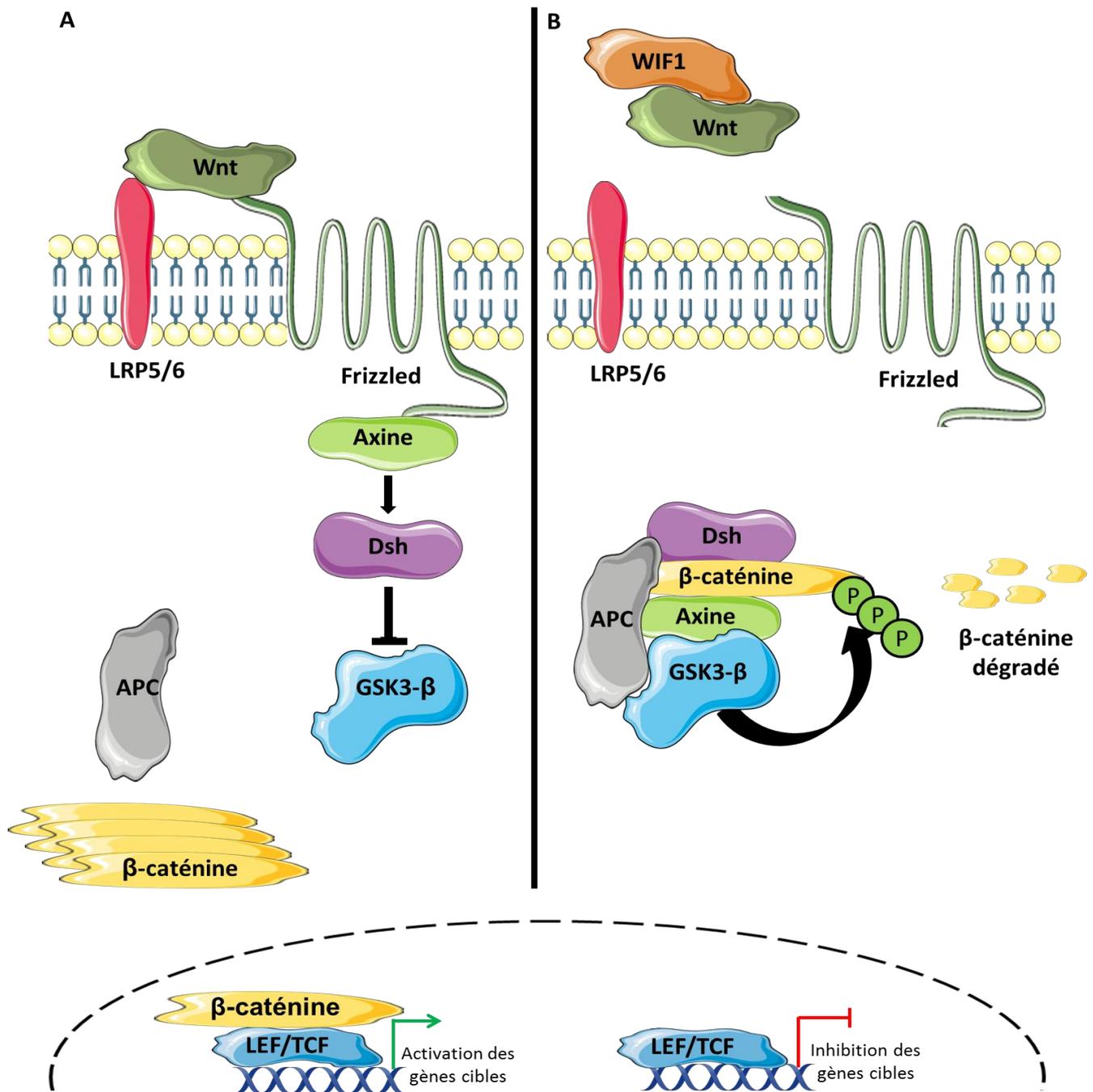
L'extension des *O*-fucoses par Fringe module l'activité du récepteur Notch. Ainsi, il a été montré que la présence d'un *O*-fucose rallongé sur les ELDs 8 et 12 favorise l'activation de la voie Notch par DLL1. De plus, l'action de Fringe sur les fucoses *O*-liés des ELDs 6 et 36 inhibe la capacité du récepteur à être activé par JAG1. D'une façon générale, il semblerait que la présence d'un *O*-fucose allongé favorise l'activation du récepteur par DLL1, directement en augmentant l'affinité du récepteur pour ce ligand ou indirectement en limitant la capacité de JAG1 à activer la voie Notch (Figure 22) (Kakuda and Haltiwanger, 2017).

### **III.2. Les autres protéines *O*-fucosylées**

Outre le récepteur NOTCH1 et ses ligands, seulement dix protéines ont été démontrées comme étant *O*-fucosylées. Quatre d'entre elles sont des protéines naturelles présentes en quantités suffisantes, ce qui a permis de les isoler et de déterminer la présence de fucose *O*-lié par spectrométrie de masse. C'est le cas des facteurs de coagulation VII (Bjoern et al., 1991), IX (Nishimura et al., 1992) et XII (Harris et al., 1992) ainsi que l'UROKINASE (Kentzer et al., 1990). Cependant, le rôle de la *O*-fucosylation pour ces protéines reste indéterminé. Pour les autres protéines (y compris le récepteur NOTCH1), la mise en évidence de la *O*-fucosylation a été réalisée sur des protéines recombinantes entières ou tronquées telles que les ligands de Notch, JAG1 et DLL1 (Panin et al., 2002), NOTCH3 (Arboleda-Velasquez et al., 2005), AMACO (Gebauer et al., 2008), CRIPTO (Schiffer et al., 2001) et l'AGRINE (Kim et al., 2008). Pour toutes ces protéines le rôle du ou des *O*-fucoses demeure inconnu excepté pour l'AGRINE et CRIPTO pour lesquelles la *O*-fucosylation permet respectivement de limiter l'agrégation des récepteurs nicotiques à l'acétylcholine (Kim et al., 2008) et de favoriser la signalisation de Nodal (Schiffer et al., 2001), bien que cet effet soit remis en question (Shi et al., 2007).

### **III.3. Les protéines potentiellement cibles de Pofut1**

Récemment, une liste de protéines cibles potentielles de POFUT1, basée sur la présence d'ELDs contenant une séquence consensus de *O*-fucosylation de type C<sup>2</sup>-X<sub>4</sub>-(**S/T**)-C<sup>3</sup>, fait état de 87 protéines potentiellement *O*-fucosylables chez l'Homme (Schneider et al., 2017) (Figure 23). Ces protéines contiennent entre 1 et 20 ELDs susceptibles de recevoir un *O*-fucose. . En dehors des 11 protéines porteuses de fucose *O*-liés, des récepteurs NOTCH2 et 4 et des ligands DLL3, 4 et JAG2 qui sont susceptibles d'être *O*-fucosylés due à leur forte identité respective avec NOTCH1, DLL1 et JAG1 (Kopan and Ilagan, 2009) aucune donnée ne nous permet de statuer sur la *O*-fucosylation des autres protéines.



**Figure 24 : Signalisation de la voie canonique Wnt**

(A) Lorsque la protéine Wnt se fixe sur son récepteur Frizzled, ceci entraîne une trimérisation du complexe Frizzled/Wnt avec la protéine LRP5/6. Ce trimère permet le recrutement de l'axine et de la protéine Dishevelled (Dsh). Dsh inhibe la phosphorylation de la  $\beta$ -caténine par inhibition de la GSK3- $\beta$ . En absence de phosphorylation, la  $\beta$ -caténine sera stabilisée et adressée au noyau pour interagir avec les différents facteurs de transcription de la famille TCF/LEF. (B) En présence de WIF1, Wnt ne peut se fixer sur Frizzled induisant ainsi une absence de trimérisation avec LRP 5/6 et donc de recrutement de l'axine et de Dsh, qui se retrouveront dans un complexe incluant l'APC, la GSK3- $\beta$  et la  $\beta$ -caténine. La GSK3- $\beta$  n'est pas inhibée et peut phosphoryler librement la  $\beta$ -caténine, qui sera alors polyubiquitinylée et adressée au protéasome pour y être dégradée.

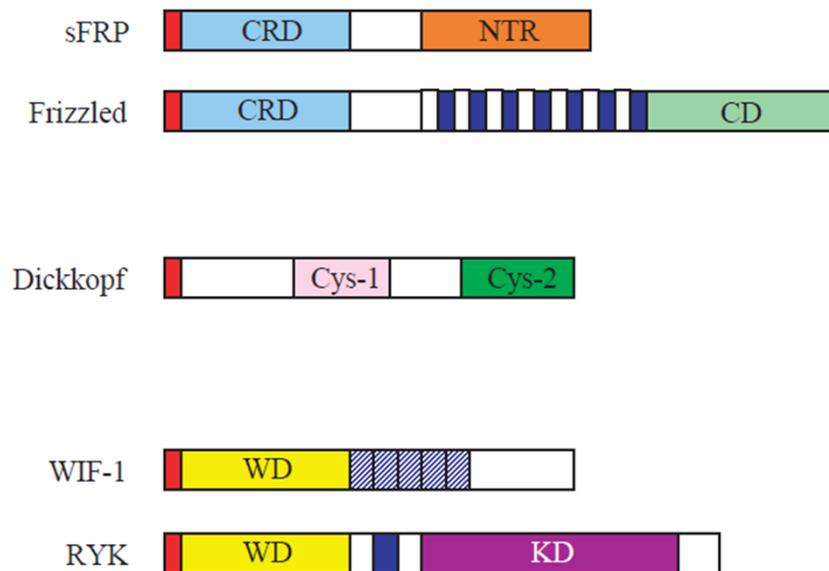
### **III.4. La protéine WIF1**

La protéine WIF1 (Wnt inhibitory factor 1) est une protéine conservée entre les bilatériens, qui a été identifiée dans la rétine humaine puis isolée à partir d'échantillons de souris, de xénope et de poisson zèbre (Hsieh et al., 1999). Il s'agit d'une protéine sécrétée dans le milieu extracellulaire, qui exerce une activité inhibitrice sur la voie Wnt.

#### **III.4.1. Un inhibiteur de la voie Wnt**

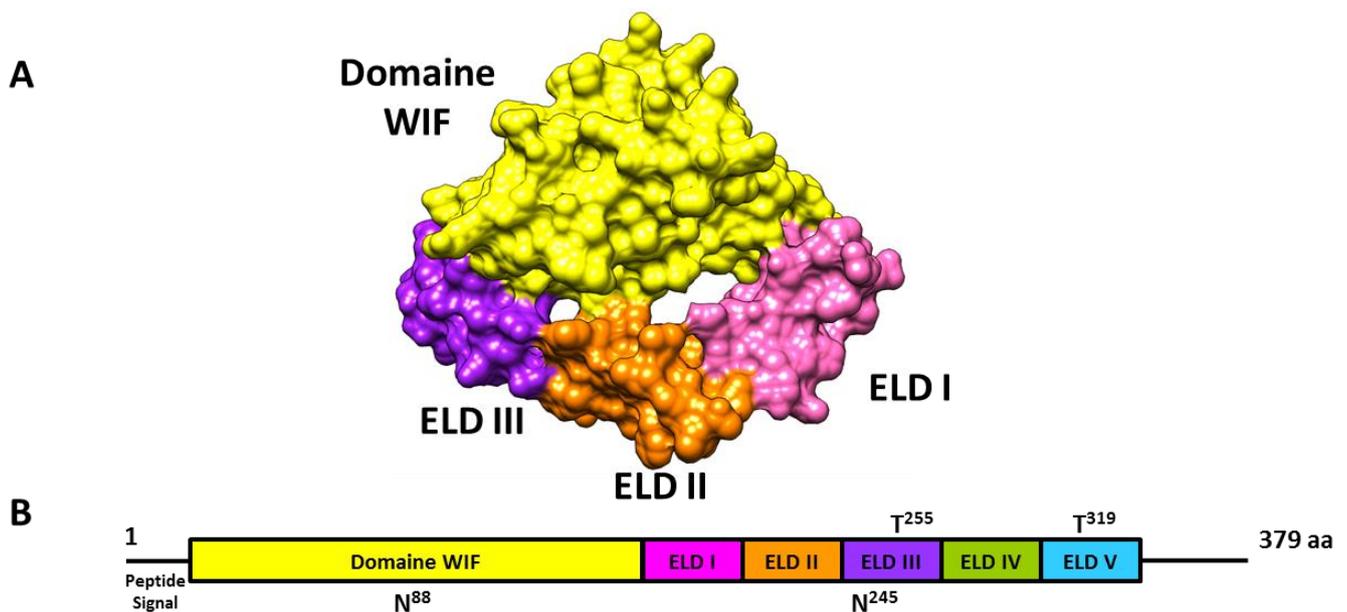
La protéine WIF1, tout comme Cerberus (Piccolo et al., 1999) et d'autres membres des sFRP (secreted Frizzled-related protein) (Rattner et al., 1997) est un inhibiteur des voies canoniques et non canoniques de Wnt (Kawano and Kypta, 2003). L'action de ces inhibiteurs se fait par liaison directe aux protéines de la famille Wnt (interaction préférentielle de WIF1 avec Wnt 3a, 4, 5a, 7a, 9a, 11) (Surmann-Schmitt et al., 2009) pour les empêcher de se fixer à leur récepteur Frizzled et co-récepteur LRP5/6 (low-density lipoprotein receptor-related proteins 5 or 6) (Malinauskas and Jones, 2014). Il existe une autre famille d'inhibiteurs, la famille Dickkopf, qui agit par interaction avec le récepteur LRP5/6 bloquant ainsi la voie canonique Wnt sans aucun effet sur la voie non canonique, uniquement dépendante de la fixation de Wnt à Frizzled. En l'absence de WIF1, les protéines Wnt vont pouvoir former un complexe trimérique avec leur récepteur Frizzled et LRP5/6, permettant la transduction d'un signal canonique et non canonique de la voie Wnt (Kawano and Kypta, 2003).

En ce qui concerne la voie canonique, le trimère Wnt/Frizzled/LRP5/6 permet le recrutement de l'axine et de la protéine Dishevelled (Dsh), dont le rôle est d'inhiber l'activité phosphatase de GSK3- $\beta$ . L'absence de phosphorylation induite par la GSK3- $\beta$  stabilise la  $\beta$ -caténine puis permet sa translocation au noyau pour activer différents facteurs de transcription de la famille TCF/LEF (T-Cell Factor/Lymphoid Enhancer Factor). Lorsque WIF1 interagit avec Wnt, le complexe trimérique Wnt/Frizzled/LRP5/6 ne peut se former et donc recruter l'axine et Dsh, ce qui aboutit à la phosphorylation de la  $\beta$ -caténine par GSK3- $\beta$ , tous les quatre formants avec l'APC un complexe protéique. Une fois phosphorylée, la  $\beta$ -caténine sera polyubiquitinylée et adressée au protéasome pour y être dégradée (Kawano and Kypta, 2003) (Figure 24).



**Figure 25 : Homologie de structure entre les antagonistes et les protéines associées à Wnt**

Les sFRPs partagent avec le récepteur Frizzled un domaine riche en cystéine (CRD), ce qui permettrait à cet antagoniste de rentrer en compétition avec Frizzled pour la liaison des Wnts. Les sFRPs possèdent également un domaine nétrine (NTR) impliquée dans leur liaison aux HSPGs. Les domaines riches en cystéine de Dickkopf (Cys-1 et 2) sont nécessaires à son interaction avec le co-récepteur LRP5/6. Les inhibiteurs WIF1 interagissent avec les Wnts via leur domaine WIF (WD) qu'ils partagent avec les récepteurs à tyrosine kinase de type RYK, potentiellement impliqués dans la voie Wnt. Ce qui suggère une compétition possible entre WIF1 et RYK pour la liaison des Wnts. Le domaine cytoplasmique de Frizzled et le domaine tyrosine kinase du récepteur RYK sont respectivement annotés CD et KD. Les peptides signaux, les domaines transmembranaires et les ELDs sont respectivement visibles en rouge, bleu et bleu hachuré. (Kawano and Kypta, 2003)



**Figure 26 : Structure de WIF1 murin**

(A) La structure cristallisée de WIF1 montre une interaction hydrophobe entre le domaine WIF et l'ELD III. (PBD 2YGQ ; Malinauskas et al., 2011) (B) La protéine WIF1 se compose d'un peptide signal N-terminal, d'un domaine d'interaction protéique (domaine WIF), de cinq ELDs (ELD I à V) et d'une queue C-terminale hydrophile. WIF1 comprend deux sites de *N*-glycosylation respectivement sur le domaine WIF (N<sup>88</sup>) et sur l'ELD III (N<sup>245</sup>) et deux sites de *O*-fucosylation respectivement sur les ELD III (T<sup>255</sup>) et V (T<sup>319</sup>).

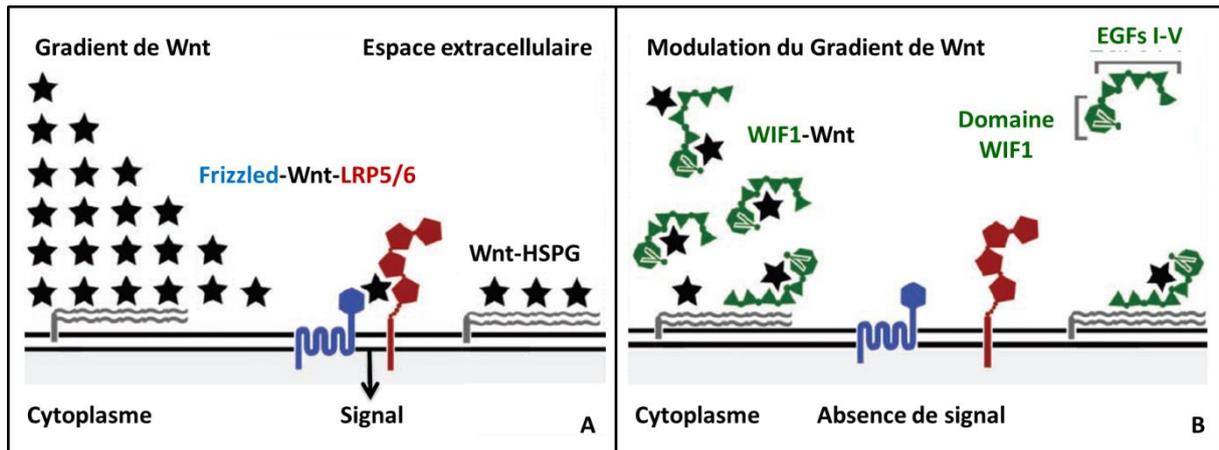
### III.4.2. Structure de la protéine WIF1

La protéine WIF1 est structurellement très différente des autres inhibiteurs de la voie Wnt puisqu'elle ne possède ni domaine riche en cystéine (retrouvé chez d'autres inhibiteurs des Wnts), ni de domaine nétrine (NTR) (présent chez les sFRP). Les sFRP et le récepteur Frizzled possèdent tous deux un domaine conservé riche en cystéine (CRD) qui permet aux inhibiteurs de la famille sFRP de rentrer en compétition avec Frizzled pour interagir avec les Wnts. L'activité inhibitrice des sFRPs est accentuée par leur capacité à interagir avec les protéines Wnts *via* leur domaine NTR. Les protéines de la famille Dickkopf sont également pourvues de domaines riches en cystéine impliqués dans leur liaison avec le co-récepteur LRP5/6 afin de prévenir sa trimérisation avec le complexe Frizzled/Wnt (Kawano and Kypta, 2003; Malinauskas and Jones, 2014) (Figure 25).

Le précurseur de WIF1 se compose d'un peptide signal N-terminal, d'un domaine d'interaction protéique en feuillet  $\beta$  appelé domaine WIF (Liepinsh et al., 2006), de cinq ELDs de type hEGF (ELD I à V) comprenant entre 31 et 33 résidus et d'une queue C-terminale hydrophile (Figure 26). Deux sites distincts du domaine WIF1 permettent son interaction avec les protéines de la famille Wnt, un site de liaison alkyle (interagissant fortement avec les Wnts due à leur palmitoylation) (Malinauskas, 2008) et un second site situé à l'opposé, proche de la partie C-terminale de la protéine (Kerekes et al., 2015).

De façon intéressante, le domaine WIF est également porté par la partie extracellulaire des récepteurs à tyrosine kinase de la famille RYK (Patthy, 2000) potentiellement impliquée dans la voie canonique (Inoue et al., 2004) et/ou non canonique des Wnts (Cadigan and Liu, 2006; Kim et al., 2008). Ainsi, WIF1 pourrait entrer en compétition avec ces récepteurs pour la fixation des Wnts, de la même façon que les sFRPs avec Frizzled.

Comme beaucoup de protéines sécrétées, WIF1 possède un certain nombre de sites potentiels de *N*- et *O*-glycosylation, particulièrement dans ses ELDs. Ainsi, deux sites potentiels de *N*-glycosylation ont été identifiés : un dans le domaine WIF (en position N<sup>88</sup>) et l'autre dans l'ELD III (en position N<sup>245</sup>). Deux sites potentiels de *O*-fucosylation ont également été trouvés respectivement dans les ELD III et V (en position T<sup>255</sup> et T<sup>319</sup>) (Figure 25)



**Figure 27 : Régulation du gradient morphogénétique des Wnts par WIF1.**

(A) En absence de WIF1, les protéines Wnt se fixent sur leurs récepteurs Frizzled et LRP5/6 pour induire la transduction du signal. La fixation des protéines Wnts aux HSPGs permet de réguler leur gradient morphogénétique. (B) En présence de WIF1, celui-ci est capable de se fixer directement aux protéines Wnt dans le milieu extracellulaire. WIF1 est aussi apte à interagir avec les HSPGs pour se positionner au niveau du gradient morphogénétique de Wnt afin de le réguler. (Malinauskas et al., 2011)

### III.4.3. Le rôle des ELDs de WIF1

Malgré le fait que le domaine WIF1 soit suffisant pour donner à la protéine WIF1 son activité inhibitrice (Hsieh et al., 1999), les domaines ELD I à V renforcent l'interaction WIF1/Wnt en se liant directement aux protéines Wnts et sont donc nécessaires pour conférer à la protéine une activité optimale (Malinauskas et al., 2011).

De plus, les domaines ELD I à IV présentent un certain nombre de résidus chargés positivement (lysines et arginines) pouvant établir des liaisons ioniques avec les HSPGs (glypican heparan sulfate proteoglycans), qui sont chargés négativement et présents à la surface des cellules (Malinauskas et al., 2011). Cette liaison aux HSPGs *via* les ELDs de WIF1 permet de contrôler le gradient morphogénétique des Wnts et donc de réguler sa voie de signalisation (Malinauskas et al., 2011) (Figure 26).

Un contrôle du gradient morphogénétique est également décrit pour les inhibiteurs de la famille sFRP, qui sont capables d'interagir avec les HSPGs *via* leur domaine NTR ce qui témoigne de l'importance de ce mécanisme de contrôle (Üren et al., 2000).

### III.4.4. L'importance biologique de WIF1

En tant qu'inhibiteur de la voie Wnt et comme les autres antagonistes de cette voie, WIF1 joue un rôle fondamental dans de nombreux processus développementaux (Hunter et al., 2004; Ng et al., 2014; Surmann-Schmitt et al., 2009) et participe également au maintien de l'homéostasie des tissus adultes (Lu et al., 2013; Vaes et al., 2005). Ainsi, dès sa découverte en 1999, sa sur-expression a révélée des défauts de somitogenèse suite à son action sur la voie Wnt (Hsieh et al., 1999). Les sFRPs, dont l'effet est similaire à celui de WIF1, induisent un défaut de prolifération et de différenciation myogénique *in cellulo* suite à l'inhibition de la voie Wnt (Descamps et al., 2008). Bien que le rôle de WIF1 dans la myogenèse n'ait jamais été clairement établi, son effet sur la voie Wnt et son fonctionnement proche de celui des sFRPs suggère une implication probable dans ce processus physiologique.

Plusieurs cancers ont été associés à une dérégulation de WIF1, le plus souvent une sous-expression induite par l'hyperméthylation de son promoteur. Ce phénomène a été observé dans de nombreux cancers (cancer de la vessie, du poumon, du cerveau, colon-rectum) (He et al., 2005; Mazieres et al., 2004; Urakami, 2006; Wu et al., 2012). La sous-expression de WIF1 induirait une sur-activation de la voie Wnt et donc des altérations de



la prolifération et de la différenciation cellulaires. Des études visant à mettre en place de nouvelles approches anti-tumorales ont d'ailleurs montré que l'apport exogène de WIF1 inhibe l'angiogenèse et la croissance tumorale *via* son effet antagoniste sur la voie Wnt (Hu et al., 2009). De la même manière, il a été montré dans le cancer colorectal que la déméthylation du promoteur de WIF1 (par l'apport de genisteine) permettait de restaurer son expression ce qui a pour effet de prévenir l'invasion et la migration tumorale (Zhu et al., 2018).



## PROJET DE THESE

---

La *O*-fucosylation des domaines EGF-like (ELD) retrouvés dans certaines glycoprotéines membranaires ou sécrétées, est une glycosylation rare catalysée par la protéine *O*-fucosyltransférase 1 (Pofut1) réticulaire. Le transfert de *O*-fucose par Pofut1 ne peut se faire qu'en présence d'un ELD correctement replié et portant une séquence consensus particulière de type C<sup>2</sup>-X-X-X-X-(**S/T**)-C<sup>3</sup>. La présence de fucoses *O*-liés sur un ou plusieurs ELDs est essentielle à la fonction de certaines protéines : c'est notamment le cas du récepteur NOTCH1, dont les *O*-fucoses modulent la capacité à induire un signal après avoir interagi avec ses ligands canoniques DSL (DLL et JAG).

Mon projet de thèse s'est défini autour de la compréhension des mécanismes biochimiques et des rôles biologiques en lien avec la *O*-fucosylation induite par l'enzyme POFUT1. Trois axes de recherche ont émergé de ce projet, à savoir l'étude du rôle de la *O*-fucosylation induite par POFUT1 dans le développement musculaire post-natal chez la souris, la mise en évidence de nouvelles protéines cibles de POFUT1 et la valorisation d'une nouvelle approche expérimentale d'analyse de la présence de fucoses *O*-liés.

Le premier axe s'est inscrit dans la continuité de travaux réalisés avant mon arrivée au sein de mon équipe d'accueil. Ils avaient pour objectif d'étudier l'effet de la sous-expression de *Pofut1* sur l'activation de la voie Notch dans la lignée myoblastique murine C2C12 (Der Vartanian et al., 2015). Mes premiers travaux de thèse ont consisté à participer à l'étude de la lignée murine *Pofut1*<sup>cax/cax</sup>, connue pour sous-exprimer de façon hypomorphe le gène *Pofut1* (Schuster-Gossler et al., 2009). Cette lignée murine a été choisie comme modèle d'étude pour une thèse initiée précédemment au laboratoire, dont l'objectif était de définir l'effet de la sous-expression de *Pofut1* sur le développement musculaire post-natal (Al Jaam, 2016). Ma contribution s'est particulièrement portée sur la caractérisation de myoblastes dérivés de cellules satellites (MDCS), issus de muscles de souris *Pofut1*<sup>cax/cax</sup> et sauvages C3H. Plus précisément, mon intérêt s'est porté sur la capacité des récepteurs NOTCH, exprimés à la surface des MDCS *Pofut1*<sup>cax/cax</sup> ou sauvages, à interagir avec leurs ligands canoniques JAG1 et DLL1, apportés de manière exogène (Al Jaam et al., 2016). L'objectif était de voir si la diminution d'activation de la voie Notch pouvait être reliée à une moindre interaction de ces ligands avec les MDCS *Pofut1*<sup>cax/cax</sup>, exprimant des récepteurs Notch hypo-*O*-fucosylés.



Le second axe de mon projet de thèse a été de mettre en évidence de nouvelles protéines *O*-fucosylées, en se basant sur des listes de protéines potentiellement cibles de POFUT1 (Al Jaam, 2016; Schneider et al., 2017). Afin de rester cohérent avec les orientations thématiques de l'équipe, à savoir le développement musculaire et le cancer colorectal, le choix s'est porté sur la protéine sécrétée Wnt inhibitory factor 1 (WIF1), porteuse de deux séquences consensus de *O*-fucosylation au sein de ses ELDs III et V. Avant de tester biochimiquement les glycosylations de WIF1, une approche phylogénétique a été menée pour définir la conservation des sites de *N*-glycosylation et de *O*-fucosylation chez les bilatériens. Puis, des tests biochimiques visant à démontrer la présence de *O*-fucose et/ou de *N*-glycane sur la protéine WIF1 murine ont été effectués. Pour cela, des ELDs III et V de WIF1 ont été produits de manière isolée dans un système procaryote afin de tester leur capacité à recevoir *in vitro* un *O*-fucose. La protéine WIF1 entière, sous sa forme sauvage et différentes formes mutées, a été produite dans un système d'expression mammalien stable (CHO) de façon recombinante pour en étudier la *N*-glycosylation et la *O*-fucosylation *in cellulo*. La difficulté technique liée à la mise en évidence de la *O*-fucosylation a été contournée par la combinaison de la cycloaddition azide-alcyne assistée par le cuivre (CuAAC ou chimie click) et la spectrométrie masse en mode MRM, après transfert de *O*-fucose *in vitro* par POFUT1 murin. Enfin, une approche fonctionnelle a consisté à analyser les effets de la *N*-glycosylation et de la *O*-fucosylation sur la sécrétion de WIF1. L'objectif de cette étude était de mettre en évidence une nouvelle protéine *O*-fucosylée, WIF1. Puis, d'effectuer une caractérisation plus globale de sa glycosylation et de l'impact de la présence de *N*-glycane et de *O*-fucose sur la sécrétion de cette protéine.

Le dernier axe de mon projet de thèse s'est effectué en parallèle de l'étude de la *O*-fucosylation de WIF1, et a consisté à standardiser une nouvelle approche expérimentale de mise en évidence de la présence des *O*-fucoses au sein d'ELD cibles. Pour cela, les ELDs 12 et 26 de NOTCH1 connus pour être *O*-fucosylés par POFUT1 mais présentant des affinités différentes pour cette enzyme (Li et al., 2017) ont été sélectionnés et produits dans un système d'expression procaryote. Le but de ces travaux étant de démontrer que notre approche biochimique, basée sur l'utilisation de la chimie click, pouvait être une alternative rapide et efficace aux techniques classiques de mise en évidence de la *O*-fucosylation (spectrométrie de masse, radiomarquage métabolique, etc...).



# RESULTATS



**Reduced Notch signalling leads to postnatal skeletal muscle hypertrophy in *Pofut1*<sup>cax/cax</sup> mice.**

Al Jaam Bilal<sup>1</sup>, Heu Katy<sup>1</sup>, **Pennarubia Florian**<sup>2</sup>, Segelle Alexandre, Magnol Laetitia, Germot Agnès, Legardinier Sébastien, Blanquet Véronique and Maftah Abderrahman.

**Open Biology** 2016 Sep;6(9). <http://dx.doi.org/10.1098/rsob.160211>



De précédentes études réalisées au sein de mon équipe d'accueil ont mis en évidence le rôle essentiel de la *O*-fucosylation, médiée par la protéine *O*-fucosyltransférase 1 (POFUT1), dans la différenciation des cellules musculaires. L'utilisation de la lignée myoblastique murine C2C12, reproduisant les étapes clés de la myogenèse (activation des cellules progénitrices, prolifération et différenciation cellulaire), a permis d'étudier les conséquences d'une sous-expression de *Pofut1*, induite de manière constitutive. Il a été montré que cette sous-expression entraînait une diminution du pool de progéniteurs myogéniques (Pax7+, MyoD-) et une différenciation précoce, conséquences directes attribuées à la dérégulation de la voie Notch. En effet, la quantité de NOTCH clivé intracellulaire (NICD) est significativement diminuée, ce qui signifie une plus faible activation de la signalisation Notch (Der Vartanian et al., 2015).

Une partie de mon travail de thèse s'inscrit dans l'exploration du rôle de POFUT1 dans le développement musculaire post-natal d'une lignée murine. Etant donné que le KO du gène *Pofut1* entraîne une létalité au cours du développement embryonnaire (Shi and Stanley, 2003), nous avons choisi d'étudier des souris sous-exprimant de manière significative l'enzyme POFUT1, la lignée murine *Pofut1*<sup>cax/cax</sup> (lignée généreusement fournie par l'équipe du Pr. Achim Gossler, Allemagne). Ces souris *Pofut1*<sup>cax/cax</sup> présentent un allèle hypomorphe pour le gène *Pofut1* résultant de l'insertion spontanée d'un élément rétro-transposable (IAP) dans l'intron 4 de ce gène. La sous-expression de *Pofut1* perturbe essentiellement la voie de signalisation Notch et aboutit chez ces souris à des défauts de somitogenèse, se traduisant par des déformations squelettiques d'intensité variable (Schuster-Gossler et al., 2009). C'est à partir de ce modèle murin que la différenciation myogénique et la croissance musculaire post-natale ont pu être étudiées par des approches *in vivo* et *ex vivo*.

L'approche *in vivo* a consisté à prélever chez des souris *Pofut1*<sup>cax/cax</sup> différents types de muscles squelettiques; à fibres majoritairement mixtes (*Quadriceps*, *Gastrocnemius*), rapides (*Tibialis*) ou lentes (*Soleus*). La détermination du poids relatif de ces muscles, de la taille des fibres qui les composent, et de leur contenu en cellules satellites (cellules souches musculaires) nous ont permis de montrer une altération du développement musculaire chez cette lignée murine. Ces souris présentent en effet une hypertrophie musculaire postnatale associée à une réduction du pool de cellules satellites, mais aucune hyperplasie n'a été observée par comparaison à des souris sauvages de souche C3H.



L'approche *ex vivo* a consisté à isoler des myoblastes dérivés de cellules satellites (MDCS) provenant des muscles des pattes postérieures des souris *Pofut1<sup>cax/cax</sup>*. L'analyse de ces cellules en cultures primaires a révélé que la sous-expression de *Pofut1* entraînait une dérégulation du programme myogénique des MDCS, se traduisant par une déplétion en progéniteurs Pax7+/MyoD- et une fusion avancée de 48h, sans aucune incidence sur la prolifération. Ma contribution s'est particulièrement portée sur la mise en évidence de la dérégulation de la voie Notch induite par la sous-expression de *Pofut1* comme principale cause de ces événements. Pour cela, des tests d'interaction *in vitro* entre les récepteurs NOTCH exprimés à la surface des MDCS et ses ligands canoniques DLL1 et JAG1 (apportés de manière exogène) ont été mis en place. Ces tests nous ont montré une forte diminution d'interaction de DLL1 avec les MDCS *Pofut1<sup>cax/cax</sup>* comparées aux MDCS sauvages (aucune différence significative n'a été observée pour JAG1). Nous avons également montré que cette diminution d'interaction avec le ligand DLL1 était corrélée à une diminution d'activation de la signalisation Notch (moins de clivage du NICD et une moindre activation des gènes cibles, *Hes/Hey*).

Nous sommes donc arrivés à la conclusion que la sous-expression de *Pofut1* chez la lignée murine *Pofut1<sup>cax/cax</sup>* induit bien une diminution d'interaction entre les récepteurs NOTCH et le ligand DLL1, du fait d'une moindre *O*-fucosylation du récepteur NOTCH plutôt qu'à une diminution de sa quantité en surface. Ceci conduit à une activation plus faible de la signalisation Notch et par conséquent à une différenciation précoce associée à une diminution du pool de progéniteurs. Cela se traduit chez les souris *Pofut1<sup>cax/cax</sup>* par une hypertrophie musculaire postnatale précoce et une diminution du pool de cellules satellites.





**Cite this article:** Al Jaam B, Heu K, Pennarubia F, Segelle A, Magnol L, Germot A, Legardinier S, Blanquet V, Maftah A. 2016 Reduced Notch signalling leads to postnatal skeletal muscle hypertrophy in *Pofut1*<sup>cax/cax</sup> mice. *Open Biol.* **6**: 160211. <http://dx.doi.org/10.1098/rsob.160211>

Received: 16 July 2016

Accepted: 24 August 2016

#### Subject Area:

cellular biology/biochemistry/molecular biology/biotechnology

#### Keywords:

POFUT1, Notch, satellite cells, skeletal muscles, hypertrophy

#### Author for correspondence:

Sébastien Legardinier  
e-mail: [sebastien.legardinier@unilim.fr](mailto:sebastien.legardinier@unilim.fr)

<sup>†</sup>B.A.J. and K.H. are considered co-first authors and contributed equally to this work.

<sup>‡</sup>Present address: Laboratoire de parasitologie et mycologie, Centre Hospitalier de Cayenne, Guyane.

<sup>§</sup>S.L., V.B. and A.M. are considered co-last authors and contributed equally to this work.

Electronic supplementary material is available online at [doi:10.6084/m9.figshare.c.3461754](https://doi.org/10.6084/m9.figshare.c.3461754).

# Reduced Notch signalling leads to postnatal skeletal muscle hypertrophy in *Pofut1*<sup>cax/cax</sup> mice

Bilal Al Jaam<sup>†</sup>, Katy Heu<sup>†,‡</sup>, Florian Pennarubia, Alexandre Segelle, Laetitia Magnol, Agnès Germot, Sébastien Legardinier<sup>§</sup>, Véronique Blanquet<sup>§</sup> and Abderrahman Maftah<sup>§</sup>

Univ. Limoges, INRA, UMR 1061, UGMA, 87060 Limoges, France

SL, 0000-0001-5923-5689

Postnatal skeletal muscle growth results from the activation of satellite cells and/or an increase in protein synthesis. The Notch signalling pathway maintains satellite cells in a quiescent state, and once activated, sustains their proliferation and commitment towards differentiation. In mammals, POFUT1-mediated O-fucosylation regulates the interactions between NOTCH receptors and ligands of the DELTA/JAGGED family, thus initiating the activation of canonical Notch signalling. Here, we analysed the consequences of downregulated expression of the *Pofut1* gene on postnatal muscle growth in mutant *Pofut1*<sup>cax/cax</sup> (*cax*, compact axial skeleton) mice and differentiation of their satellite cell-derived myoblasts (SCDMs). *Pofut1*<sup>cax/cax</sup> mice exhibited muscle hypertrophy, no hyperplasia and a decrease in satellite cell numbers compared with wild-type C3H mice. In agreement with these observations, *Pofut1*<sup>cax/cax</sup> SCDMs differentiated earlier concomitant with reduced *Pax7* expression and decrease in PAX7<sup>+</sup>/MYOD<sup>-</sup> progenitor cells. *In vitro* binding assays showed a reduced interaction of DELTA-LIKE 1 ligand (DLL1) with NOTCH receptors expressed at the cell surface of SCDMs, leading to a decreased Notch signalling as seen by the quantification of cleaved NICD and Notch target genes. These results demonstrated that POFUT1-mediated O-fucosylation of NOTCH receptors regulates myogenic cell differentiation and affects postnatal muscle growth in mice.

## 1. Introduction

Skeletal muscles are composed of post-mitotic multinucleated myofibres and satellite cells, which are mononucleated cells located at the periphery of myofibres between the plasma membrane and basal lamina [1]. Satellite cells are precursor stem cells, which are involved in after-birth muscle growth and muscle regeneration after injuries, either by exercise or by disease [2]. These cells are maintained in a quiescent state in normal adult muscles, and upon activation their asymmetric division leads to formation of two cell types: differentiated myogenic cells involved in myofibre formation or repair and self-renewing cells able to replenish the stem pool [3–5]. Postnatal muscle growth can be achieved by either an increase in the number of myofibres (hyperplasia), an increase in their size (hypertrophy), or a combination of both as in rats [6]. In mice, postnatal muscle growth mainly occurs by myofibre hypertrophy and not hyperplasia, with a steady increase in the number of myonuclei per myofibre from birth to three weeks after birth (P21) [7,8]. This myonuclear accretion in myofibres has been shown to be the result of cell fusion between satellite cell-derived myoblasts (SCDMs) and pre-existing myofibres, thus increasing their width and length [9,10]. From P21 to adulthood and thereafter, muscle growth by hypertrophy does not occur by a satellite cell-dependent myogenic fusion [11], but mainly by increased protein synthesis,



which provokes an enlargement of myofibres without formation of new myonuclei [8]. Satellite cells are thus the main source of new myonuclei in growing myofibres [12], before their quiescence around P21 [8]. They remain in this state thereafter in normal mature muscle [13], except during regeneration in which satellite cells are re-activated [2].

The IGF/Akt/mTOR and myostatin/Smad pathways are the two major signalling pathways, which control muscle mass increase by hypertrophy during postnatal life after P21, promoting protein synthesis and limiting protein degradation [11]. The genes of the Notch signalling pathway, which are involved in the control of the stem cell's fate and behaviour, are essential for the myogenic progress [14,15]. However, little is known about the role of Notch signalling in perinatal/juvenile muscle growth because *Notch1*<sup>-/-</sup> mice mutants [16,17] or embryos lacking components required for Notch signalling such as its main transcriptional repressor RBP-Jk [18] die at midgestation or around birth with severe defects in somitogenesis, vasculogenesis, cardiogenesis and/or neurogenesis [19]. Because global inactivation of Notch signalling leads to drastic and advanced effects on somitogenesis, only conditional mutagenesis of some Notch regulators provides information on the involvement of Notch signalling in myogenesis after midgestation. Indeed, conditional mutations of *Rbp-Jk* [20] and the gene encoding the NOTCH-ligand *DLL1* [21] lead to mutant mice exhibiting severe muscle hypotrophy during embryonic development, owing to uncontrolled differentiation of progenitor cells generating a rapid and significant depletion of the progenitor cell pool.

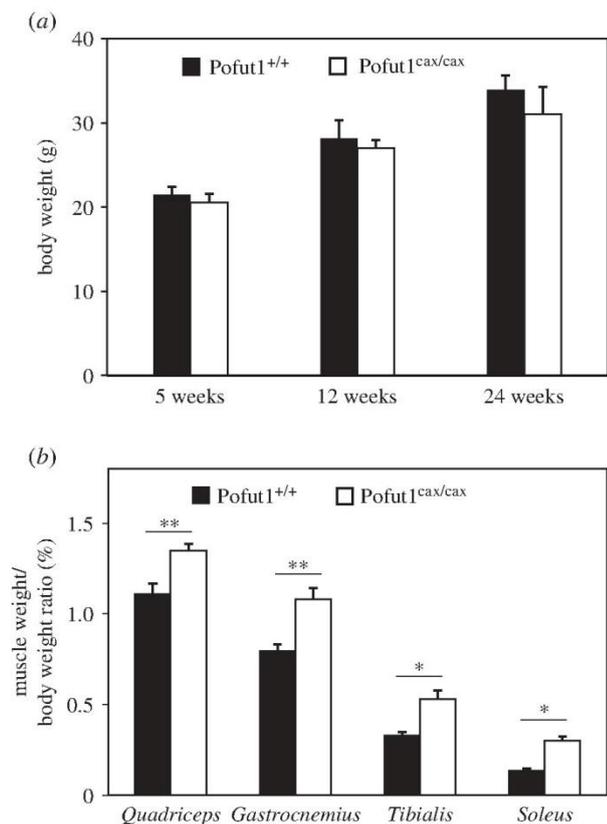
Canonical Notch signalling is initiated by interaction of the extracellular domain of ligands (*DLL1*, *-3*, *-4* and *JAGGED1* and *-2*) with their counterparts on one of the four receptors (*NOTCH1*–*4*), leading to sequential proteolytic cleavages by ADAM proteases and the  $\gamma$ -SECRETASE complex of the NOTCH receptor. Once cleaved, the latter releases its NOTCH intracellular domain (NICD), which translocates to the nucleus where it interacts with RBP-Jk by displacing co-repressors [22]. This allows the recruitment of coactivators such as MASTERMIND-LIKE-1 (MAML1) [23] to induce transcriptional activation of specific target genes, including *Hes* and *Hey* family genes [24,25]. By activating the expression of target genes such as *Hes1*, Notch inhibits differentiation by repressing the expression of *MyoD* [26], which belongs to the family of myogenic regulating factors (MRFs) including MYF5, MYOGENIN (or MYOG) and MRF4 (or MYF6) [27]. During postnatal muscle growth and muscle regeneration, activated satellite cells coexpress *Pax7* and *MyoD* [28]. While most of them proliferate, myoblasts from activated satellite cells downregulate *Pax7* leading to their differentiation in myocytes, whose fusion gives rise to myogenin-expressing multinucleated myotubes [29]. Some of those proliferating myoblasts (*PAX7*<sup>+</sup>/*MYOD*<sup>+</sup>) revert to a quiescent state by repressing *MyoD* expression [30]. Thus, the expression of *Pax7* maintains proliferation and prevents a precocious differentiation, without promoting quiescence [28]. Overexpressed NICD upregulates *Pax7* through a RBP-Jk-dependent binding to its promoter, resulting in enhanced self-renewal of satellite cells, whereas inhibition of Notch signalling leads to a downregulation of *Pax7*, resulting in satellite cell depletion and improved terminal differentiation [30]. Indeed, the loss of *Pax7* expression leads to a complete absence of satellite cells in postnatal skeletal muscles [31].

NOTCH receptors and ligands are glycoproteins, whose extracellular domains are subjected to several glycosylations such as *O*-fucosylation mediated by protein *O*-fucosyltransferase 1 (POFUT1). POFUT1 is a resident *N*-glycosylated glycosyltransferase of the endoplasmic reticulum [32,33], which adds fucose on S or T within the consensus site C<sup>2</sup>X<sub>4</sub>(S/T)C<sup>3</sup> (where C<sup>2</sup> and C<sup>3</sup> are the second and third cysteines, respectively) of epidermal growth factor (EGF)-like repeats present in numerous proteins [34], including NOTCH receptors and ligands. The POFUT1-mediated *O*-fucosylation of extracellular domains of NOTCH receptors controls the receptor–ligand interactions which are critical for activation of Notch signalling [35]. Among the 36 EGF-like repeats of murine NOTCH1 receptor, EGF repeat 12, for example, was necessary and sufficient for ligand binding [36] and mutation of its *O*-fucosylation site affected Notch–ligand interactions [37]. In contrast to NOTCH receptors, *O*-fucosylation of DELTA/JAGGED ligands is poorly documented and recent studies report that *O*-fucosylation of murine *DLL1* is not required for ligand–receptor interactions [38]. POFUT1-mediated *O*-fucosylation of receptors is therefore essential for canonical Notch signalling by *DLL1* or *JAGGED1* [39]. Notch signalling is also modulated by elongation of *O*-linked fucoses on some EGF-like repeats with *N*-acetylglucosamine (GlcNAc) and then with galactose residues, owing to the action of a  $\beta$ 1,3-*N*-acetylglucosaminyltransferase of the Fringe family [40] and a  $\beta$ 1,4-galactosyltransferase [41,42], respectively. The resulting trisaccharide *O*-fucosylglycan (Gal $\beta$ 1-4GlcNAc $\beta$ 1-3fucose) can be terminated with  $\alpha$ 2–3 or  $\alpha$ 2–6 linked sialic acids, whose involvement in Notch signalling has not been yet reported [43].

In a recent *in vitro* study, we showed that *Pofut1* knock-down reduces Notch signalling and affects differentiation of the mouse myoblast cell line C2C12. The expression patterns of *PAX7* and *MYOD* are modified under these conditions and induce earlier cell differentiation [44]. *In vivo*, knock-out of *Pofut1* is lethal: mice embryos die at E9.5 with a phenotype similar to that of mice in which NOTCH receptor signalling is inactivated [19]. In 2009, a spontaneous mutation in *Pofut1* gene called *Pofut1*<sup>cax</sup> was described in a mouse strain [45]. *Pofut1*<sup>cax/cax</sup> mice have an insertion of an intracisternal A particle (IAP) in the fourth intron of the *Pofut1* gene, leading to a hypomorphic allele and a decrease in gene expression without any change in protein structure and activity. Homozygous *Pofut1*<sup>cax/cax</sup> mice display defects in the axial skeleton consistent with the known patterning functions of Notch in somitogenesis. Nevertheless, no detailed phenotyping was performed on skeletal muscles of *Pofut1*<sup>cax/cax</sup> mice.

In this study, we report the consequences of the *Pofut1* hypomorphic mutation on postnatal growth of skeletal muscles in *Pofut1*<sup>cax/cax</sup> mice. Immunostaining studies on isolated *Pofut1*<sup>cax/cax</sup> skeletal muscles showed a slight but significant muscular hypertrophy with myonuclear accretion compared with wild-type controls. In addition, the number of *PAX7*<sup>+</sup> satellite cells was significantly reduced in *Pofut1*<sup>cax/cax</sup> mice. Analyses of *ex vivo* *Pofut1*<sup>cax/cax</sup> SCDMs revealed a depletion of *PAX7*<sup>+</sup>/*MYOD*<sup>-</sup> progenitor cells, a decrease in *Pax7* expression and disruption of the myogenic programme, leading to earlier *Pofut1*<sup>cax/cax</sup> SCDM differentiation. These observations could explain the accrued muscle mass occurring in the first weeks of postnatal life in *Pofut1*<sup>cax/cax</sup> mice, as a result of increased fusion of SCDMs with pre-existing myofibres.





**Figure 1.** Body weight and muscle/body weight ratios in mice. (a) Body weights (g) in Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> mice ( $n = 6$ ) at three different ages (5, 12, 24 weeks). (b) Muscle weight/body weight ratios as a percentage for four different muscles of five week old mice ( $n = 6$ ). Means  $\pm$  s.e.m. are shown (two-tailed  $t$ -test, with a significance level of  $*p < 0.05$ ,  $**p < 0.01$ ,  $***p < 0.001$ ).

## 2. Results

### 2.1. Pofut1<sup>cax</sup> mutation induces postnatal muscle hypertrophy and decrease in the satellite cell pool

As previously described [45], Pofut1<sup>cax/cax</sup> mice showed either a normal phenotype or shortened bodies with kinky or absent tails. About 40% of Pofut1<sup>cax/cax</sup> mice had shortened kinky tails ( $n = 19$ ) with a length of  $6.16 \text{ cm} \pm 0.68$  versus  $8.50 \text{ cm} \pm 0.20$  in Pofut1<sup>+/+</sup> mice but showed unchanged body size compared with their wild-type littermates (data not shown). Additional morphometric analyses did not reveal a statistically significant difference ( $n = 6$  per genotype and per age) in body weight regardless of the age (5, 12, 24 weeks) of Pofut1<sup>cax/cax</sup> mice compared with Pofut1<sup>+/+</sup> mice (figure 1a).

To determine whether the hypomorphic mutation of Pofut1<sup>cax/cax</sup> mice affected postnatal muscle growth, skeletal muscles with fast-twitch (*Tibialis*) or slow-twitch (*Soleus*) or mixed (*Gastrocnemius* and *Quadriceps*) myofibres were weighed and analysed by immunohistochemistry in Pofut1<sup>cax/cax</sup> mice and compared with those from wild-type mice after weaning at five weeks (figures 1b and 2) and long after sexual maturity at 12 and 24 weeks (electronic supplementary material, table S1). The analysis of muscle weight/body weight (M/B) ratios for these four skeletal muscles from five week old mice revealed higher values in Pofut1<sup>cax/cax</sup> mice than in Pofut1<sup>+/+</sup> mice (figure 1b). Similar results were found with 12 and 24 week old mice (electronic supplementary material, table S1). These results showed that higher M/B ratios in Pofut1<sup>cax/cax</sup> mice seem to be unrelated to the type of muscle metabolism and could be due

to myofibre hypertrophy, hyperplasia or both. To test these hypotheses, immunohistochemical analyses were carried out to evaluate myofibre mean area and number using cross sections. Immunostaining using anti-laminin antibody (figure 2a) showed significant increased myofibre mean area (figure 2b) in Pofut1<sup>cax/cax</sup> mice compared with Pofut1<sup>+/+</sup> mice but no hyperplasia (figure 2c), because the myofibre number per field, although with a lower mean, was not significantly different regardless of muscle type and mouse age (electronic supplementary material, table S1). To determine whether myofibre hypertrophy was only due to increased protein synthesis and/or myonuclear accretion, the number of DAPI-labelled nuclei (figure 2a) was evaluated. More nuclei per cross-sectional myofibre were found in each of the four muscles from Pofut1<sup>cax/cax</sup> mice (figure 2d), suggesting greater myonuclear accretion in muscles from Pofut1<sup>cax/cax</sup> mice.

The pool of satellite cells at the periphery of myofibres was evaluated using anti-Pax7 immunostaining (figure 2a) and the number of satellite cells per cross-sectional myofibre for the four selected muscles was determined (figure 2e and electronic supplementary material, table S1). Regardless of muscle type and mouse age, the number of satellite cells per cross-sectional myofibre was significantly lower in Pofut1<sup>cax/cax</sup> mice than in wild-type controls. The decrease in satellite cells ranged from 35% to 66% if considering the four muscles from five week old Pofut1<sup>cax/cax</sup> mice compared with those from wild-type mice (figure 2e and electronic supplementary material, table S1), whereas it was more pronounced (from 33% to 87%) in muscles of 12 and 24 week old mice (electronic supplementary material, table S1).

### 2.2. Pofut1<sup>cax/cax</sup> satellite cell-derived myoblasts display reduced expression of Pofut1 and Pax7

As the number of Pax7<sup>+</sup> satellite cells was significantly reduced in Pofut1<sup>cax/cax</sup> mice compared with the wild-type controls, primary cultures of SCDMs were analysed from skeletal muscles to highlight molecular mechanisms involved in satellite cell depletion.

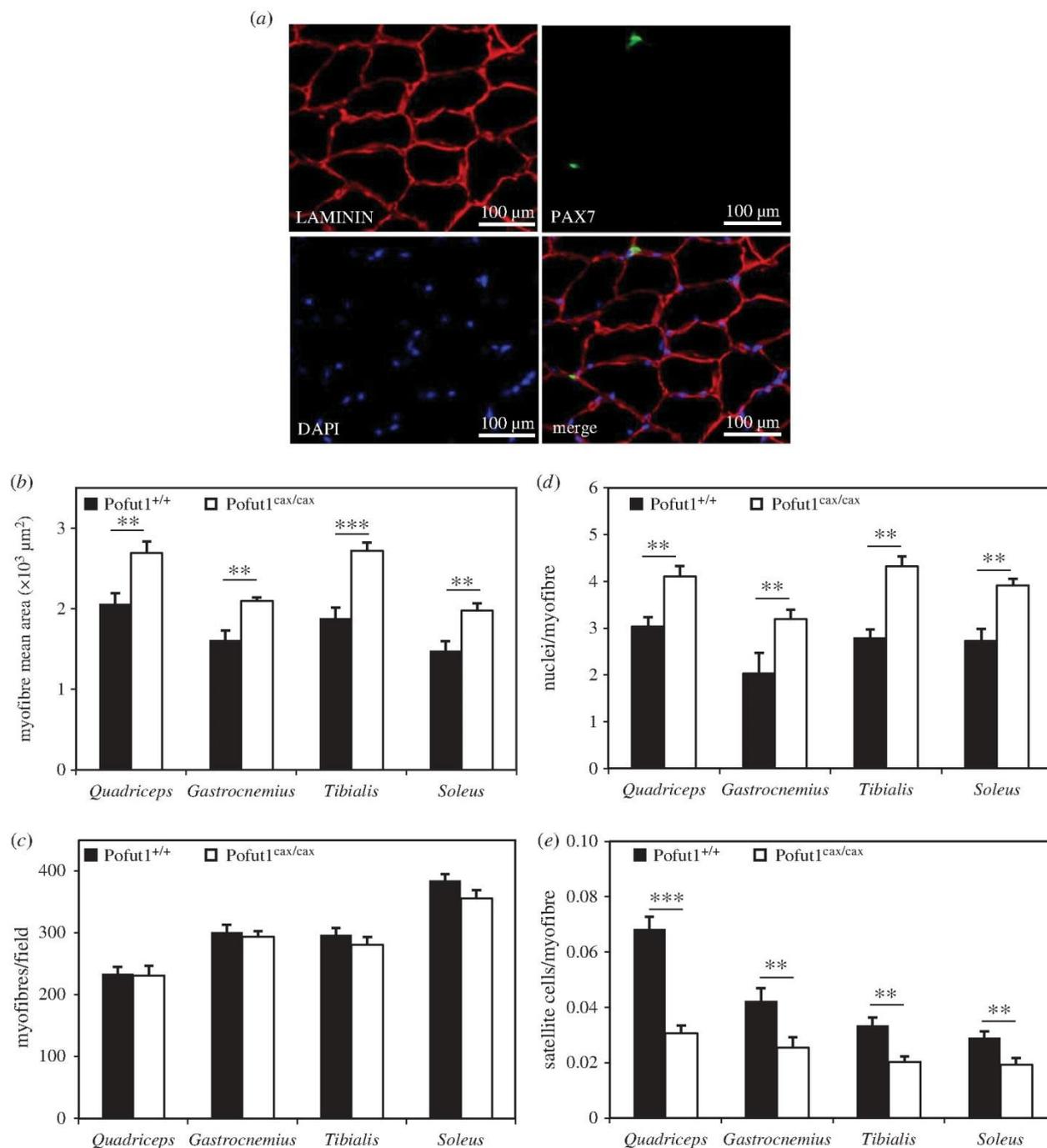
Primary cultures of SCDMs from Pofut1<sup>cax/cax</sup> and Pofut1<sup>+/+</sup> mice were prepared from a pool of four skeletal muscles of hind legs (*Quadriceps*, *Gastrocnemius*, *Tibialis*, *Soleus*), mixed on the basis of a similar reduction by 80–90% of Pofut1 mRNA quantities in each of the four selected muscles, as shown in figure 3a. As expected, western blot analyses showed a reduction of POFUT1 quantity, thus, validating the *ex vivo* model. This decrease in protein expression was approximately 40% in Pofut1<sup>cax/cax</sup> SCDMs, when compared with Pofut1<sup>+/+</sup> SCDMs (figure 3b).

Then, expression of PAX7, which plays a key role in activation and quiescence of satellite cells *in vivo*, was evaluated by western blot. A significant 45% reduction in protein expression was obtained for PAX7 in Pofut1<sup>cax/cax</sup> SCDMs, similarly to POFUT1 (figure 3b). These results might suggest that PAX7<sup>+</sup> satellite cells, less numerous in muscles of Pofut1<sup>cax/cax</sup> mice, express reduced PAX7 protein levels; this could be consequent to decreased expression of POFUT1.

### 2.3. Pofut1<sup>cax/cax</sup> satellite cell-derived myoblasts differentiate earlier than wild-type cells

Before analysing *in vitro* differentiation, we determined purity of the myoblast cell population in SCDM preparations by





**Figure 2.** Characterization of immunostained cross sections of skeletal muscles from five week old mice. (a) Representative field of *Quadriceps* cryosection from five week old *Pofut1*<sup>+/+</sup> mice. LAMININ (red), DAPI (blue) and PAX7 (green) immunostaining showed basal lamina of myofibres, nuclei and satellite cells respectively. (b) Myofibre mean areas, (c) number of myofibres per field, (d) nuclei per myofibre and (e) satellite cells per myofibre were determined for skeletal muscles with mixed (*Quadriceps*, *Gastrocnemius*), fast-twitch (*Tibialis*) or slow-twitch (*Soleus*) myofibres. Means  $\pm$  s.e.m. ( $n = 6$ ) are shown (two-tailed  $t$ -test, with a significance level of \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

calculating percentage of DESMIN-expressing cells relative to total cells (DAPI<sup>+</sup>). Results showed that primary cultures of *Pofut1*<sup>+/+</sup> and *Pofut1*<sup>cax/cax</sup> SCDMs contained more than 95% of myoblasts expressing DESMIN (figure 4a), which is considered as a muscle-specific protein.

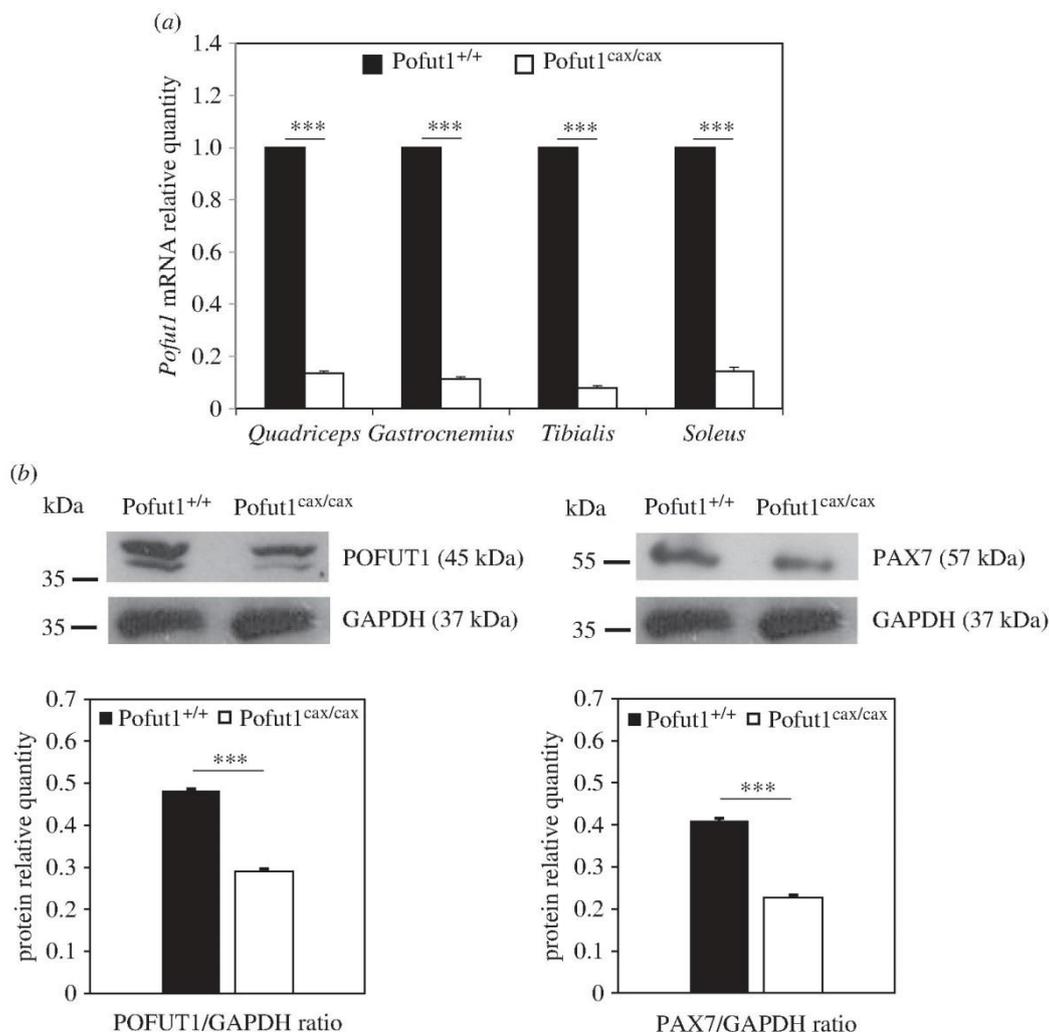
As attested by MTS assay (figure 4b), the decreased expression of *Pofut1* did not significantly modify proliferation of *Pofut1*<sup>cax/cax</sup> SCDMs from T0 (where cells were seeded at a low density in growth medium (GM)) to 96 h. Indeed, the production of formazan (absorbance at 490 nm) by viable cells is nearly identical for both SCDMs at each time point.

Prior to inducing differentiation by switching the culture medium (differentiation medium (DM) instead of GM),

*Pofut1*<sup>+/+</sup> and *Pofut1*<sup>cax/cax</sup> SCDMs exhibiting comparable purity and proliferation rates were thus seeded at the same cell density. To characterize the differentiation process of both SCDMs, fusion indexes and expression patterns of *Cdkn1a* (cyclin-dependent kinase inhibitor 1a, also called *p21*), *Pofut1*, *Pax7* and three myogenic marker genes (*Myf5*, *MyoD*, *Myog*) were determined during 240 h of differentiation. Finally, the impact of decreased *Pofut1* was evaluated by counting the three different cell populations as follows: PAX7<sup>+</sup>/MYOD<sup>-</sup> (self-renewing cells), PAX7<sup>+</sup>/MYOD<sup>+</sup> (proliferating cells) and PAX7<sup>-</sup>/MYOD<sup>+</sup> (differentiating cells).

Following immunostaining of SCDM nuclei and myosin heavy chain (MyHC), the fusion index was calculated by





**Figure 3.** Quantification of *Pofut1* and *Pax7* transcripts and proteins in muscles and SCDMs. (a) *Pofut1* mRNA relative quantities in four skeletal muscles from five week old *Pofut1<sup>cax/cax</sup>* mice compared with those of *Pofut1<sup>+/+</sup>* controls set to 1. Quantities of *Pofut1* mRNA were normalized to *18S RNA* and *Gapdh* mRNA quantities. (b) Detection of POFUT1, PAX7 and GAPDH by western blotting on total extracted proteins of proliferating *Pofut1<sup>+/+</sup>* and *Pofut1<sup>cax/cax</sup>* SCDMs. Histograms represent POFUT1/GAPDH and PAX7/GAPDH ratios. Means  $\pm$  s.e.m. ( $n = 3$ ) are shown (two-tailed *t*-test, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

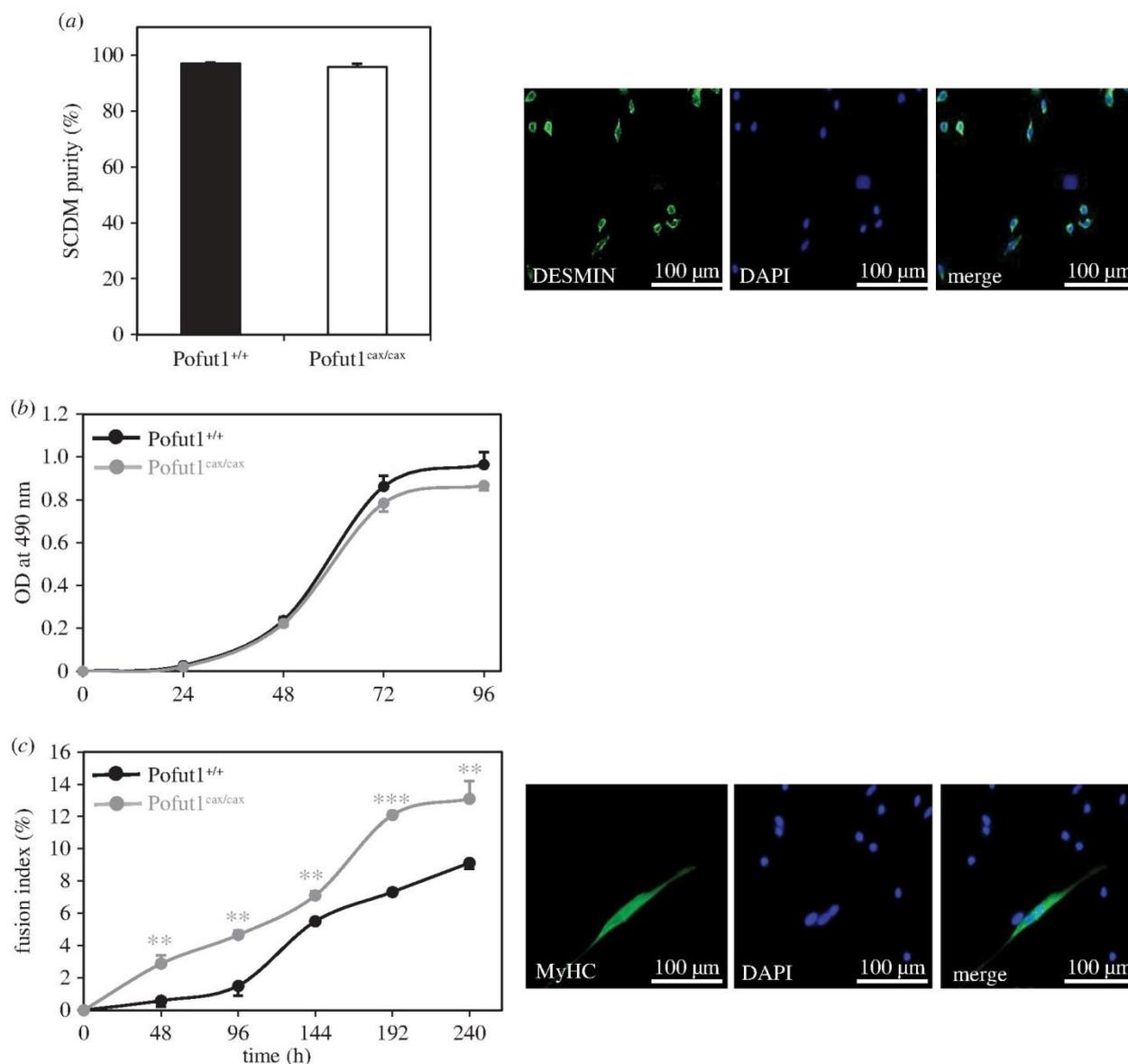
determining the percentage of nuclei in myotubes compared with the total number of nuclei during 240 h of differentiation (figure 4c). When seeded at about 50% confluence and placed in DM (T0), *Pofut1<sup>cax/cax</sup>* SCDM differentiated significantly faster than *Pofut1<sup>+/+</sup>* cells with an advance of about 48 h. At 240 h of differentiation, 13% of cell fusion was reached in *Pofut1<sup>cax/cax</sup>* SCDMs versus only 9% in wild-type cells. This indicates that the earliest appearance of myotubes in *Pofut1<sup>cax/cax</sup>* SCDMs was not due to a difference in proliferation (as demonstrated in figure 4b) but more likely to a disruption in the myogenic programme.

To understand the origin of this premature differentiation, expressions of *Cdkn1a*, *Pofut1* and *Pax7* as well as of MRFs was followed by qPCR during differentiation for both SCDMs (figure 5). First, differentiation was followed by expression of *Cdkn1a*, a marker for cell cycle arrest. A non-significant variation in the expression of this marker was observed during the first 48 h of differentiation in both SCDMs relative to T0, suggesting that most cells continued to proliferate in spite of induced differentiation. However, its expression was significantly higher at 96 h compared with T0, which might signal a cell cycle exit. This phase (0–96 h), represented in grey in all graphs, was considered as a step where most wild-type cells continued to proliferate.

The second phase (96–240 h) where high expression levels of *Cdkn1a*, especially in *Pofut1<sup>cax/cax</sup>* SCDMs, were monitored was considered as the main differentiation stage. Globally, expression of *Pofut1*, *Pax7* and MRFs (*Myf5*, *MyoD*, *Myog*) was reduced in *Pofut1<sup>cax/cax</sup>* SCDMs, compared with the wild-type cells. Expression profiles evolved in the same manner during the differentiation process for cells of the two genotypes, with increases and decreases occurring at the same points of the time course.

In *Pofut1<sup>+/+</sup>* SCDM and compared with T0, the expression of *Pofut1*, *Pax7* and *MyoD* was significantly increased at 48 h, whereas that of *Cdkn1a* did not change, in agreement with maintenance of a proliferative state in most cells. The expression of *Pofut1* globally decreased up to the end of differentiation. Interestingly, *Pax7* expression was significantly higher at 96 h compared with T0, whereas that of *MyoD* strongly decreased over time and became not statistically different from T0. From 144 to 192 h, *Cdkn1a* expression increased, as a result of prior induction by *MyoD*, which was optimal at 144 h [46]. Variation of these markers (*Pax7*, *MyoD*, *Cdkn1a*) was in favour of a cell cycle exit and commitment of a majority of these cells into differentiation at 96 h. However, as shown by the low fusion index (only 1.5% fusion at 96 h), differentiation was slowed down by the





**Figure 4.** Characterization of purity, proliferation and differentiation of SCDMs. (a) Co-immunostaining for DESMIN (green) and DAPI (blue) distinguishes myoblasts from non-muscular cells in primary cultures of Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> SCDMs. Histogram represents percentage of Desmin<sup>+</sup> cells related to total DAPI<sup>+</sup> cells for both SCDM preparations. (b) SCDM cell proliferation was determined using an MTS assay, which relies on absorbance of formazan at 490 nm, a product of bioreduced MTS. Test was performed on both SCDMs every 24 h until 96 h after seeding cells at a low density. (c) Commitment to differentiation was first evaluated by calculating fusion indexes of Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> SCDMs. Fusion indexes were determined by immunostaining of nuclei with DAPI (blue) and myotubes with MyHC (green) and were calculated from the percentage of nuclei within MyHC<sup>+</sup> myotubes compared with total nuclei. Statistical analyses were performed by comparing both SCDMs at each differentiation time (48–240 h). (a–c) Data were obtained from three independent experiments in triplicate and error bars indicate standard errors of the mean (two-tailed *t*-test, with a significance level of \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001).

strong expression of *Myf5*. After 96 h, *Myf5* and *Pax7* expression decreased, whereas that of *MyoD* increased up to a maximum at 144 h. Most cells then expressed *Myog* between 144 and 240 h, consistent with fusion of most cells.

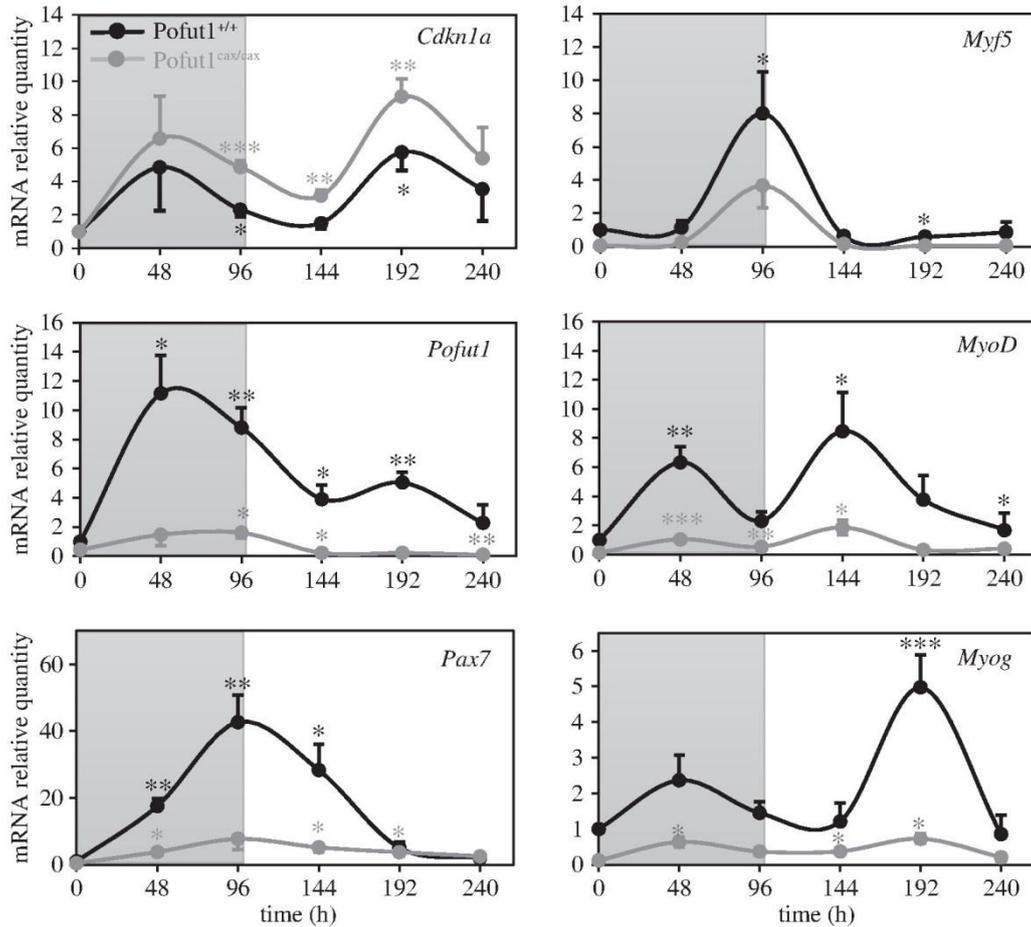
In Pofut1<sup>cax/cax</sup> SCDMs, most cells expressed a significantly higher level of *MyoD* at 48 h as well as at 96 h compared with T0, whereas *Pax7* was weakly expressed at those two time points. While *Pax7* and *Myf5* expression was not significantly high at 96 h, *MyoD* expression was maintained up to 144 h resulting in earlier induction of *Myog*. This is in agreement with a premature commitment to differentiation, as shown by the fusion index.

The transition from proliferative to differentiated state is thus under the control of a *Pax7* and *MyoD* expression equilibrium. In order to further study this transition, SCDMs were seeded at confluence and differentiation was induced 4 h

later. We then monitored by immunostaining the expression of PAX7 and MYOD in SCDMs during 96 h of differentiation and counted the three cell populations as follows: PAX7<sup>+</sup>/MYOD<sup>-</sup>, PAX7<sup>+</sup>/MYOD<sup>+</sup> and PAX7<sup>-</sup>/MYOD<sup>+</sup> cells (figure 6a). During differentiation, the amount of PAX7<sup>+</sup>/MYOD<sup>-</sup> progenitor cells decreased for the two genotypes in favour of PAX7<sup>-</sup>/MYOD<sup>+</sup> differentiating cells (figure 6b), reflecting commitment of SCDMs to differentiation. Remarkably, the amount of PAX7<sup>+</sup>/MYOD<sup>-</sup> progenitor cells was drastically decreased by 53%, 56% and 83% in Pofut1<sup>cax/cax</sup> compared with Pofut1<sup>+/+</sup> SCDMs at 0 h, 48 h and 96 h, respectively. In parallel, the proportion of PAX7<sup>-</sup>/MYOD<sup>+</sup> differentiating cells increased.

To conclude, all these results showed an early commitment to differentiation by Pofut1<sup>cax/cax</sup> SCDMs. Decreased expression of *Pofut1* resulted in a lower *Pax7* expression





**Figure 5.** Gene expression for *Cdkn1a*, *Pofut1*, *Pax7* and MRFs during SCDM differentiation. mRNA relative quantities of *Cdkn1a*, *Pofut1*, *Pax7* and myogenic MRF genes (*Myf5*, *MyoD* and *Myog*) in *Pofut1*<sup>+/+</sup> and *Pofut1*<sup>cax/cax</sup> SCDMs. Area in grey represents the pre-differentiation state. For each cell type, statistical analyses were performed by comparison of each differentiation time point relative to T0. Quantities of target mRNA were normalized to *18S RNA* and *Gapdh* mRNA and then calibrated with T0 using the  $\Delta\Delta Ct$  method. Data were obtained from three independent experiments in triplicate and error bars indicate standard errors of the mean (two-tailed *t*-test, with a significance level of \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001).

associated with depletion of PAX7<sup>+</sup>/MYOD<sup>-</sup> progenitor cells and a lower expression of MRFs in spite of similar expression profiles compared with *Pofut1*<sup>+/+</sup> SCDMs.

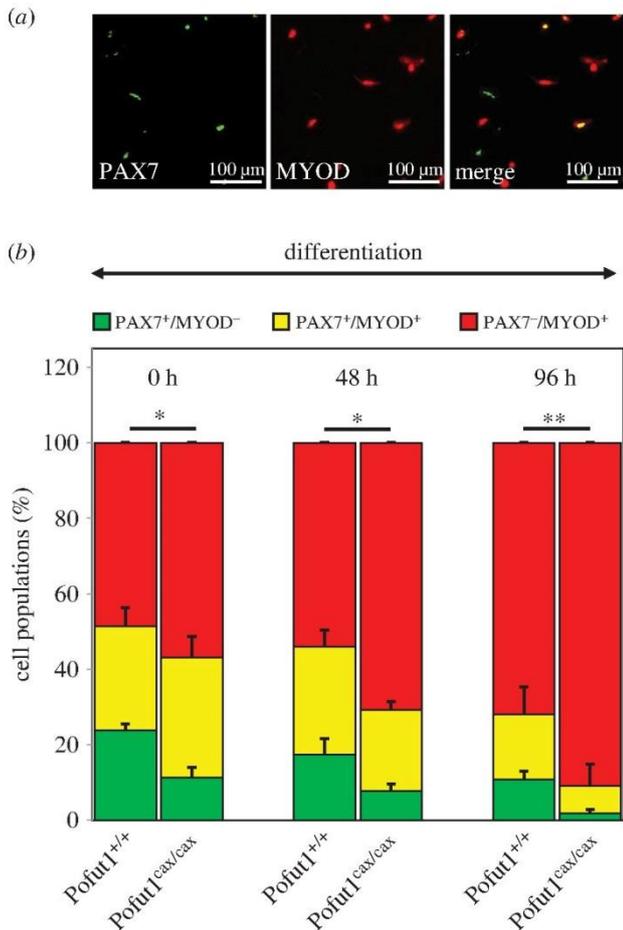
#### 2.4. *Pofut1*<sup>cax/cax</sup> satellite cell-derived myoblasts exhibit reduced Notch–ligand interactions leading to lowered Notch signalling

The Notch signalling pathway members, which control the behaviour and fate of stem cells, are essential for the proper progress of myogenesis [14,15]. POFUT1-mediated *O*-fucosylation of extracellular domains of NOTCH receptors controls receptor–ligand interactions, which are critical for the activation of Notch signalling [35]. In addition, the FRINGE-mediated elongation or not of *O*-linked fucose with GlcNAc residues defines the type of ligands from the DELTA/JAGGED family that can interact with NOTCH receptors [47]. To determine the consequences of decreased *Pofut1* expression in *Pofut1*<sup>cax/cax</sup> SCDMs on Notch binding to its ligands, recombinant fusion proteins for two Notch ligands were produced, as previously done for *Drosophila* DELTA [48]. The extracellular domains of murine DLL1 and JAGGED1 including eight and 16 EGF-like repeats respectively were fused to human placental alkaline phosphatase (AP; figure 7a). Using an anti-AP antibody, recombinant Ctrl-AP, Dll1-AP and Jag1-AP from

concentrated supernatants of transfected COS-7 cells were detected at about 60, 120 and 175 kDa, respectively (figure 7b). These apparent molecular weights were higher than expected on the basis of amino acid sequences (54.6, 109.9 and 166.8 kDa, respectively). These results suggested the presence of post-translational modifications such as *N*- and *O*-glycosylations, as predicted by *in silico* analyses and known for the ligands [43].

Proliferating *Pofut1*<sup>+/+</sup> and *Pofut1*<sup>cax/cax</sup> SCDMs were incubated simultaneously for 1 h 30 with the same quantities of recombinant proteins (Ctrl-AP, Dll1-AP or Jag1-AP) at three different doses, corresponding to an absorbance at 405 nm of 40, 80 or 160. Proteins were previously concentrated by ultrafiltration from supernatants of transfected COS7 cells and quantified on the basis of AP activity. The specific binding of these proteins to NOTCH receptors on SCDMs was determined using an AP assay by subtracting Ctrl-AP activity (background). For both SCDMs, Dll1-AP binds specifically to cells and its binding increased in a dose-dependent manner, while binding of Jag1-AP was very low ( $A^{405nm} < 0.08$ ), regardless of the dose of stimulating ligand (figure 8). Moreover, the specific binding of Dll1-AP was significantly decreased in *Pofut1*<sup>cax/cax</sup> SCDMs, varying between –32% and –37% according to the dose. Because binding of Jag1-AP on 200 000 cells was very low and did not significantly increase according to the dose of ligand, it could not be considered as specific. For this reason, only binding of Dll1-AP was followed





**Figure 6.** Immunostaining for PAX7 and MYOD during SCDM differentiation. (a) Co-immunostaining for PAX7 (green) and MYOD (red) in Pofut1<sup>+/+</sup> SCDMs. (b) Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> SCDMs were seeded at confluence 4 h before induction of differentiation and then immunostained to distinguish three cell populations: self-renewing cells (PAX7<sup>+</sup>/MYOD<sup>-</sup>), proliferating cells (PAX7<sup>+</sup>/MYOD<sup>+</sup>) and differentiating cells (PAX7<sup>-</sup>/MYOD<sup>+</sup>). Percentages of each population were calculated at three times during differentiation (0, 48, 96 h) and mean  $\pm$  s.e.m. ( $n = 3$ ) are shown. The significance of the  $t$ -test (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ), indicated above histograms for each time, was the same for each analysed cell population.

during 144 h of SCDM differentiation (figure 9). To avoid centrifugation during multiple washing steps that could prevent optimal binding of recombinant proteins, confluent adherent SCDMs were seeded in multi-well plates and incubated with the same dose of recombinant proteins (Dll1-AP, Ctrl-AP) diluted in HBAH at the same concentrations on the basis of their AP activity. A significant reduction in specific Dll1-AP binding was clearly observed in both SCDMs during the time course; this result is likely to be related to the reduction in Pofut1 expression during differentiation (figure 5).

Because decreased ligand–receptor interactions in Pofut1<sup>cax/cax</sup> proliferating SCDMs (figure 8) could be due to lower amounts of NOTCH receptors at the cell surface and/or to reduced POFUT1-mediated O-fucosylation of NOTCH receptors, the quantity of NOTCH1 receptors present at the cell surface of SCDM was determined. The amount of biotinylated extracellular domain of NOTCH1 receptor (N1ECD) retained by avidin agarose was reported relative to the amount of PAN CADHERIN (figure 10). First, we showed that the antibody raised against PAN-CADHERIN (recognition of all cadherin members) detected

specific bands at about 140 kDa in the pool of proteins eluted from avidin agarose (figure 10a, lanes 3) only when proteins (input) were previously biotinylated (figure 10a, lanes 1). The cytosolic GAPDH protein was mainly detected for the input and not retained protein samples (figure 10a, lanes 1 and 2). Only a small amount of GAPDH was found in the eluted sample (figure 10a, lanes 3), showing that biotinylated proteins, retained by avidin–agarose, were mostly membrane proteins expressed at the cell surface.

After SCDM cell surface biotinylation and capture of biotinylated proteins by avidin agarose, N1ECD and PAN-CADHERIN were specifically detected by western blot (figure 10b, lanes 1–3) in all the samples (input, unretained proteins and proteins eluted from avidin agarose). N1ECD/PAN-CADHERIN ratios were similar in Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> SCDMs, demonstrating that the decrease in ligand–receptor interactions in Pofut1<sup>cax/cax</sup> SCDM is in fact the consequence of lower O-fucosylation of NOTCH receptors.

To study the effects of decreased Pofut1 expression on Notch signalling, relatively cleaved NICD quantities in total protein extracts from Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> SCDMs were compared by western blot (figure 11a) because they reflect NOTCH cleavage after ligand binding to NOTCH receptors. Results showed a significantly lower cleaved NICD/GAPDH ratio of approximately 50% in Pofut1<sup>cax/cax</sup> SCDMs compared with wild-type cells, consistent with reduced interaction of NOTCH receptors with Dll1-AP.

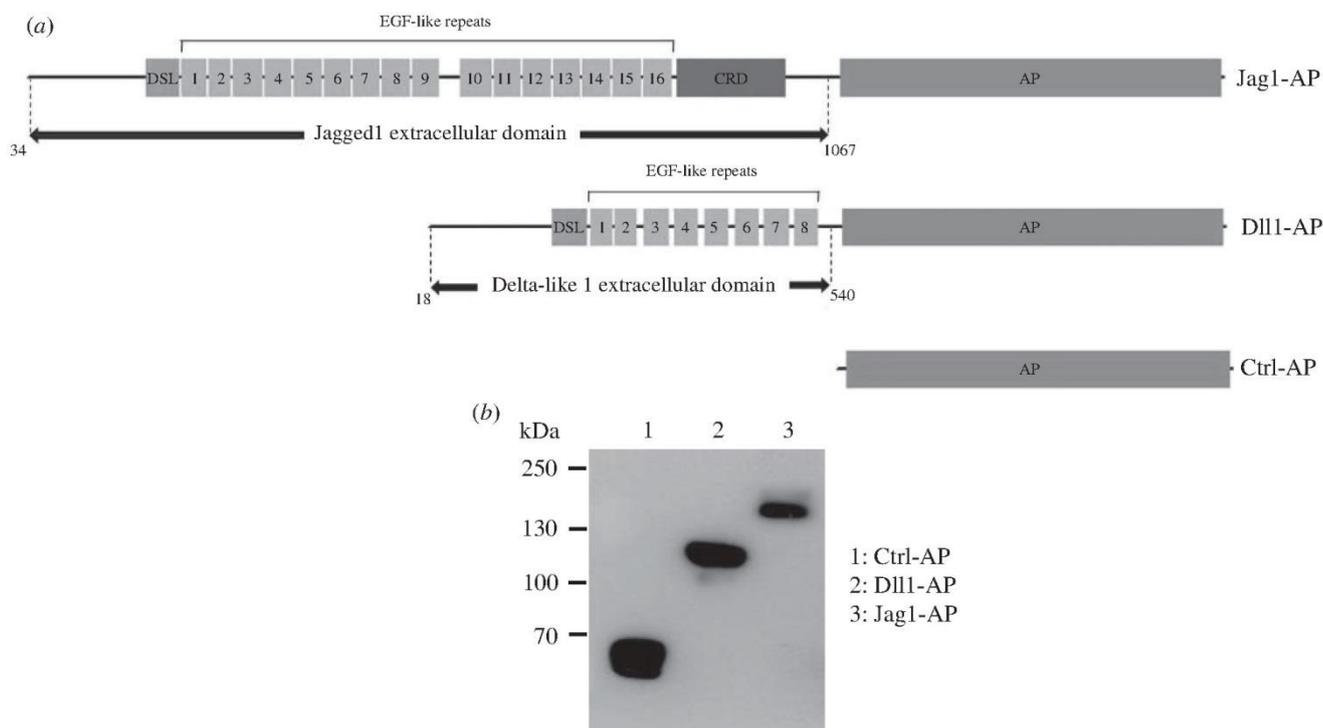
Furthermore, we analysed the expression of three members of the Hes/Hey family (*Hes1*, *Hey1*, *Heyl*), which are recognized as target genes of Notch and regulated by NICD [49] and *Hes6*, which does not appear to be regulated by the Notch pathway [50]. Expression levels of *Hes1*, *Hey1* and *Heyl* were significantly reduced (50–75%) in Pofut1<sup>cax/cax</sup> SCDMs, whereas *Hes6* expression was slightly less (figure 11b). Together, these results reflect the reduction in Notch signalling, induced by decreased Pofut1 expression in Pofut1<sup>cax/cax</sup> SCDMs.

### 3. Discussion

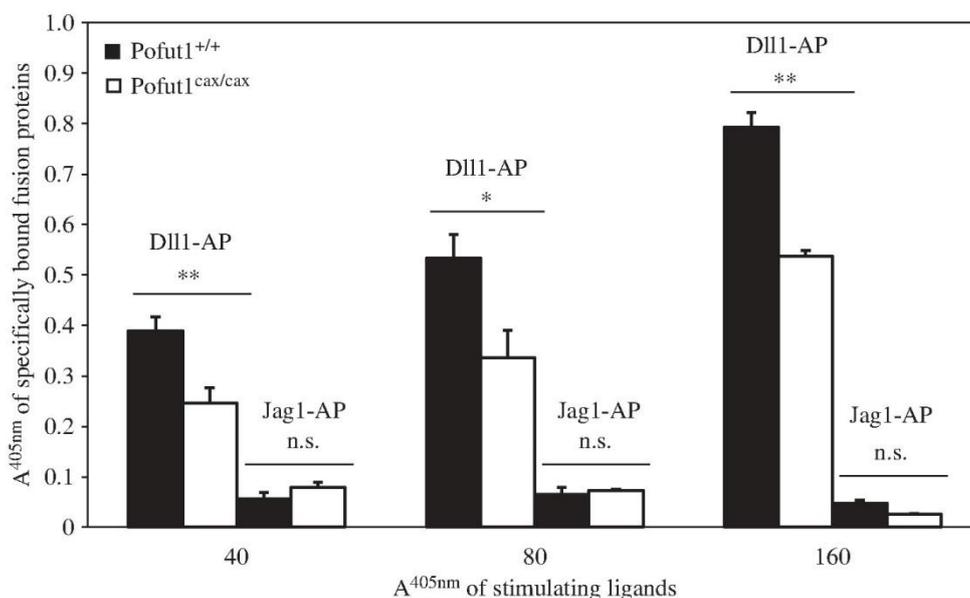
Pofut1-null mouse embryos exhibit severe retarded growth and die at midgestation ( $\approx$ E10) with major defects in many developmental processes such as somitogenesis and neurogenesis, owing to a global inactivation of Notch signalling [19]. Pofut1<sup>cax/cax</sup> mice are characterized by decreased expression of Pofut1 mRNA and protein, resulting in reduced Notch signalling [45]. These mice with hypomorphic Pofut1 alleles are viable and fertile but exhibit skeletal defects of variable intensity, affecting the length and shape of the body and tail, subsequent to defects in somitogenesis. Our present data show for the first time that the downregulated expression of Pofut1 also affects postnatal muscle growth in mice leading to a moderate muscular hypertrophy of skeletal muscles with myonuclear accretion and a significant decrease in the satellite cell pool. We did not observe any hyperplasia in muscles from Pofut1<sup>cax/cax</sup> mice, regardless of mouse age. This result is in agreement with previous studies showing that postnatal muscle growth in mice occurs by hypertrophy and not hyperplasia [7], because the myofibre number is already established around birth in many muscles [51].

All satellite cells of skeletal muscles which express Pax7 derived from PAX3<sup>+</sup>/PAX7<sup>+</sup> muscle progenitor cells of central





**Figure 7.** Production of recombinant proteins. (a) Extracellular domain of murine JAGGED1 (Jag1) and DELTA-LIKE1 (DII1) were fused to alkaline phosphatase (AP) to obtain recombinant fusion proteins (Jag1-AP and DII1-AP), which were secreted in culture medium of transfected COS-7 cells, as was the recombinant control protein (Ctrl-AP). EGF-like repeats, Delta-Serrate-Lag domain (DSL) and cysteine-rich domain (CRD) are shown. (b) Western blot analysis of recombinant concentrated Ctrl-AP and fusion proteins produced in COS-7 cells, detected with anti-AP antibody.

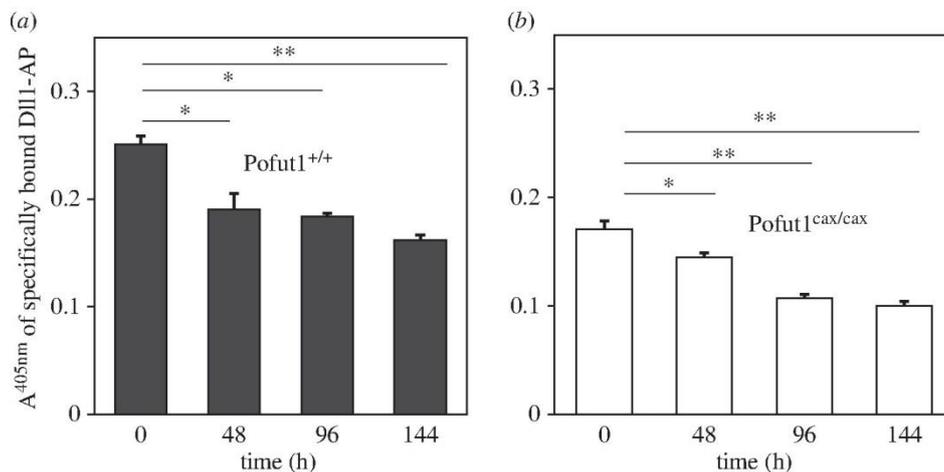


**Figure 8.** Binding assays of recombinant fusion proteins at the cell surface of proliferating SCDMs. Specific binding of DII1-AP and Jag1-AP at cell surfaces of proliferating Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> suspension cells of SCDMs. AP assays based on absorbance at 405 nm ( $A^{405nm}$ ) of pNPP substrate were performed on cell lysates after incubation of SCDM suspensions with three increasing doses of concentrated recombinant proteins (DII1-AP, Jag1-AP or Ctrl-AP), tested as stimulating ligands and quantified on the basis of their AP activity. Absorbance values of Ctrl-AP were subtracted. Mean absorbance values  $\pm$  s.e.m. ( $n = 3$ ) were compared between both types of SCDMs for each fusion protein (two-tailed  $t$ -test, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , n.s., not significant).

dermomyotome [52]. They adopt their satellite position just before birth (E16.5–18.5) [53,54]. Once this first pool of satellite cells is constituted, their number per myofibre steadily decreases from P6 to reach the adult level at P21, whereas the number of myonuclei steadily increases during the same period in murine *extensor digitorum longus* muscles [8], which contains slow-twitch myofibres such as the *Soleus* muscle used in our study. After P21 in mouse, the number of satellite

cells and myonuclei is relatively stable in the physiological state and muscle growth occurs essentially by protein anabolism resulting in a myofibre volume increase without satellite cell activation and addition of new myonuclei. After weaning, in five week old mice (P35), we showed that all skeletal muscles from Pofut1<sup>cax/cax</sup> mice displayed myofibre hypertrophy with myonuclear accretion but less satellite cells compared with wild-type mice. Consequently, we hypothesize that the





**Figure 9.** Quantification of specific binding of Dll1-AP at cell surface of SCDMs during differentiation. The specific binding of Dll1-AP at cell surfaces of Pofut1<sup>+/+</sup> (a) and Pofut1<sup>cax/cax</sup> (b) was determined every 48 h until 144 h of differentiation, using adherent SCDMs seeded at confluence. AP assays based on absorbance at 405 nm (A<sup>405nm</sup>) of pNPP substrate were performed on cell lysates after incubation of SCDMs with unconcentrated proteins (Dll1-AP or Ctrl-AP), quantified on the basis of their AP activity. Absorbance values of Ctrl-AP were subtracted to determine binding specificity of Dll1-AP. Mean absorbance values  $\pm$  s.e.m. ( $n = 3$ ) were compared with those at T0 for each cell type (two-tailed  $t$ -test, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

number of satellite cells decreased faster in muscles of Pofut1<sup>cax/cax</sup> mice between P6 and P21, at the same time as a more rapid increase in the number of post-mitotic myonuclei as a result of more myoblast fusion related to hypertrophy. However, we cannot rule out that a reduced number of PAX7<sup>+</sup> satellite cells is already present at birth in Pofut1<sup>cax/cax</sup> mice compared with Pofut1<sup>+/+</sup> controls.

In Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> mice, myofibre mean areas of most skeletal muscles continued to significantly increase between 5 and 12 weeks of age in the same proportions. It then became stable between 12 and 24 weeks except for *Quadriceps*. These results suggest that IGF/Akt/mTOR and myostatin/Smad pathways that govern the postnatal muscle growth by increased protein synthesis from P21 to adulthood were not significantly affected by reduced Notch signalling in Pofut1<sup>cax/cax</sup> mice. However, Notch signalling may affect the IGF/Akt/mTOR pathway, by regulating the transcription of the gene encoding the phosphatase PTEN, which prevents activation of Akt (also named protein kinase B) by dephosphorylating PIP3 [55]. Indeed, Notch signalling leads to an activation of target genes *Hes1* and *c-Myc*, which act as transcriptional repressors of *Pten* expression [56]. Conversely, RBP-Jk-dependent Notch signalling induces *Pten* upregulation [55,57,58]. It is therefore difficult to predict how downregulation of *Pofut1* which leads to reduced Notch signalling might impact IGF/Akt/mTOR pathway through regulation of the *Pten* expression. However, it was previously shown that inhibition of Notch signalling by DAPT treatment induces myotube hypertrophy, through myostatin decrease that promotes activation of the Akt pathway [59].

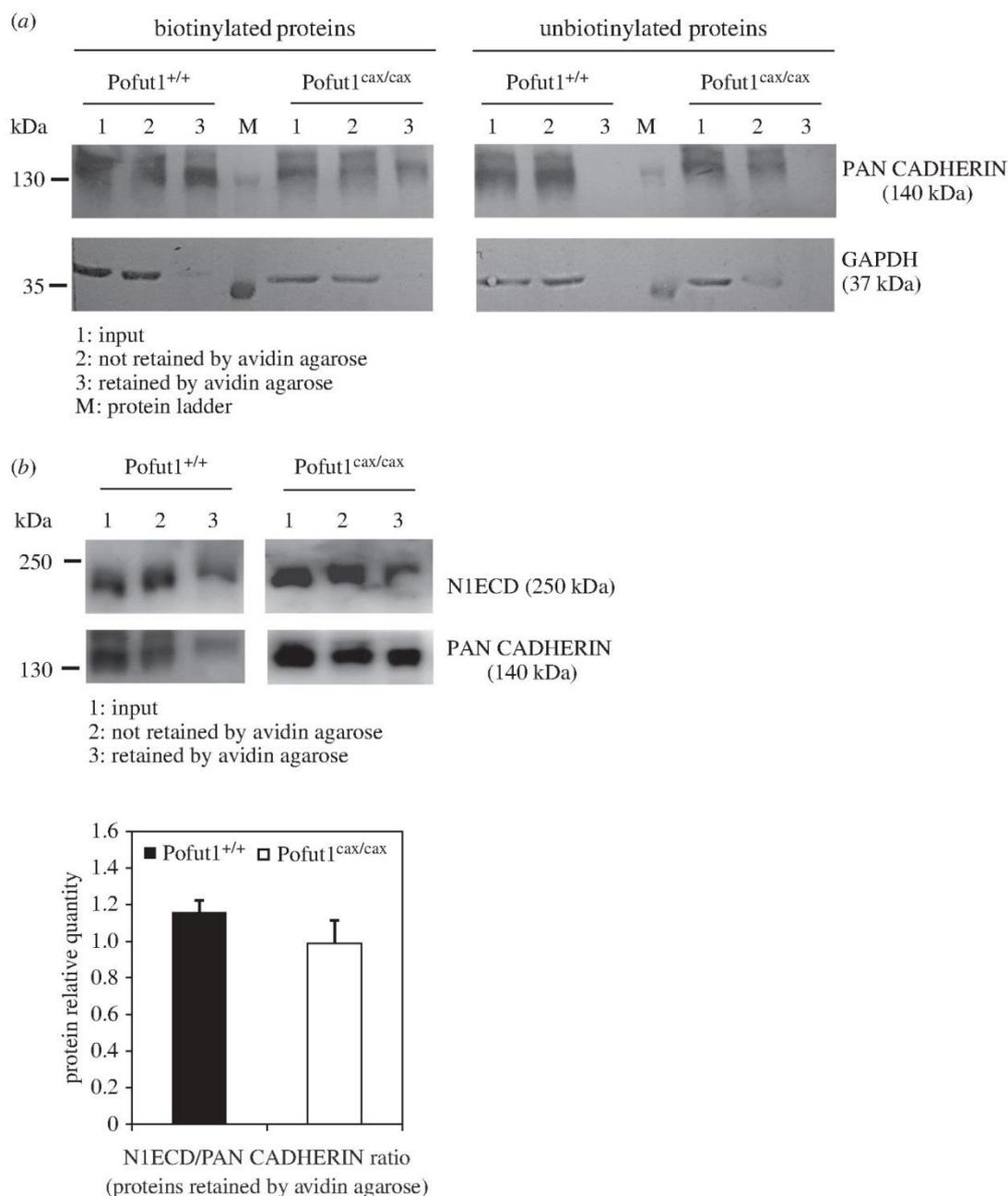
Notch signalling regulates myogenic differentiation and satellite cell fate [15]. Primary cultures of SCDMs were settled from skeletal muscles of Pofut1<sup>cax/cax</sup> mice to better understand the origin of postnatal muscular hypertrophy. First, we showed comparable purities and proliferation rates of SCDM preparations and validated the decreased expression of *Pofut1* in Pofut1<sup>cax/cax</sup> SCDMs, compared with wild-type SCDMs. We also showed an earlier commitment to differentiation of Pofut1<sup>cax/cax</sup> SCDMs, subsequent to a global decrease in Notch signalling provoked by reduced amounts of POEUT1. Indeed, reduced POEUT1 expression

was directly correlated with reduced binding of the Dll1-AP fusion protein to NOTCH receptors at the cell surface of Pofut1<sup>cax/cax</sup> SCDMs, leading to a lesser amount of cleaved NICD and reduced expression of several Notch target genes from the *Hes/Hey* family and *Pax7* [30]. Our results strongly suggest that reduction of these ligand–receptor interactions was only due to decreased POEUT1-mediated *O*-fucosylation of NOTCH receptors, because the relative quantity (RQ) of NOTCH receptors expressed at the cell surface was unaffected in Pofut1<sup>cax/cax</sup> SCDMs, as also previously observed in mouse C2C12 myogenic cell knockdown for *Pofut1* [44]. This might strengthen the hypothesis that POEUT1 does not have a chaperone role for mammalian NOTCH receptors, in contrast to what had been previously described in *Drosophila* [60]. However, only NOTCH1 receptors were quantified at the cell surface out of the four mammalian ones.

Among 20 predictive *O*-fucosylation sites within the mouse NOTCH1 extracellular domain composed of 36 EGF-like repeats [35], more than 50% of sites are known to be really modified by *O*-linked fucoses [61]. Recently, the majority of predicted *O*-fucosylation sites were shown to be occupied on *Drosophila* Notch by fucose residues, whose degree of elongation by Fringe was variable [62]. Thus, even though NOTCH receptors would have less *O*-fucosylated EGF-like repeats in Pofut1<sup>cax/cax</sup> SCDMs than in wild-type cells, we cannot rule out that some EGF-like, such as the highly conserved EGF-like repeat 12 [36], are *O*-fucosylated in priority and even elongated by Fringe owing to their involvement in modulation of Notch receptor–ligand interactions.

We showed that both proliferating SCDMs specifically bound Dll1-AP to their cell surfaces in a dose-dependent manner, whereas the very low binding of Jag1-AP was not considered as relevant. This might be related to a high activity of glycosyltransferases of the Fringe family, which allow GlcNAc extension of *O*-linked fucose residues on NOTCH EGF repeats and thus promote binding to Delta and alter binding to Jagged, as previously shown [63]. In addition, we showed a gradual decrease in Dll1-AP binding during differentiation in both SCDMs, which could be explained by a gradual reduced expression of *Pofut1* throughout this process.





**Figure 10.** Quantification of NOTCH1 receptors at the cell surface of SCDMs. (a) Western blot using an antibody directed against members of the cadherin family (PAN-CADHERIN), which are membrane proteins and not considered as target proteins of POFUT1. PAN-CADHERIN and GAPDH were monitored during purification of cell surface membrane proteins on avidin agarose after biotinylation or not of scraped SCDMs. Purified fractions containing initial extracted proteins (input, lanes 1), unretained proteins (lanes 2) or proteins eluted from avidin agarose (lanes 3) were analysed for their content of PAN-CADHERIN and GAPDH. (b) After cell surface biotinylation of proliferating Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> SCDMs and purification of biotinylated surface proteins using avidin agarose, all fractions described above were analysed by western blot using antibodies against NOTCH1 extracellular domain (N1ECD) and PAN-CADHERIN. Histograms represents the N1ECD/PAN-CADHERIN ratio, determined after quantification of these two proteins in lanes 3. Means  $\pm$  s.e.m. ( $n = 3$ ) are shown (two-tailed  $t$ -test, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

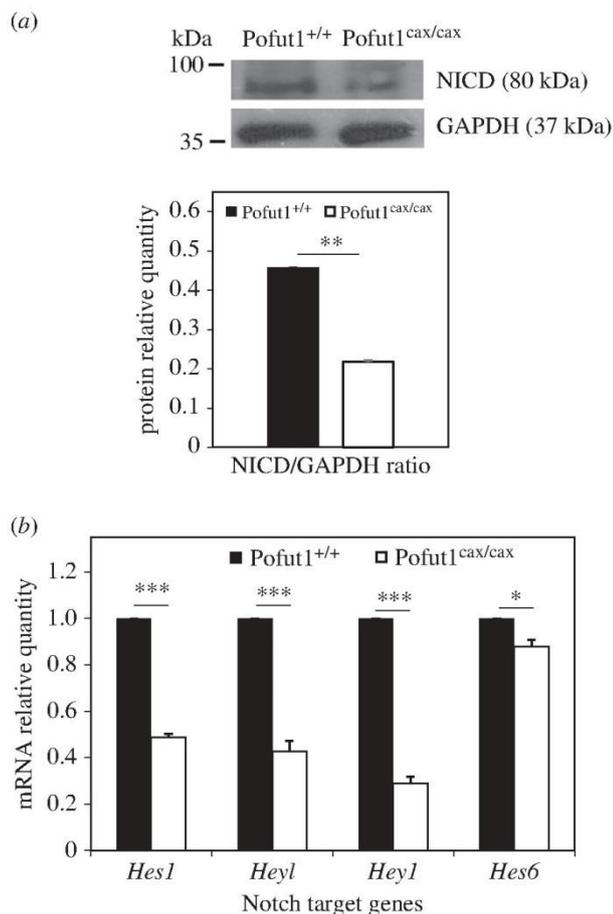
As in Pofut1 knockdown C2C12 cell line (called Po<sup>-</sup>) [44], we showed earlier differentiation of Pofut1<sup>cax/cax</sup> SCDMs owing to reduced Notch signalling and a significant lowering in PAX7<sup>+</sup>/MYOD<sup>-</sup> progenitor cells in favour of more PAX7<sup>-</sup>/MYOD<sup>+</sup> cells committed to differentiation. Thus, *in vitro* and *ex vivo* cell models represented by the C2C12 Po<sup>-</sup> cell line and Pofut1<sup>cax/cax</sup> SCDMs, respectively, gave similar results, highlighting the essential role of POFUT1 in myogenic differentiation through Notch signalling regulation.

While NOTCH receptors and their canonical ligands are already known targets of POFUT1, other proteins could be directly or indirectly affected by the decreased expression of Pofut1 gene in Pofut1<sup>cax/cax</sup> mice. Indeed, approximately

100 proteins are predicted to be *O*-fucosylated [34], some of which might be involved in one of the main signalling pathways responsible for postnatal muscle growth. Except by mass spectrometry, it is difficult to demonstrate that a natural protein is actually *O*-fucosylated as no antibody or specific lectin recognizing *O*-linked fucose residues is available and no enzyme is known to specifically cleave *O*-fucose. Thus, it is experimentally difficult to know which sites are precisely *O*-fucosylated on NOTCH receptors expressed at the cell surface of SCDMs and how decreased expression of Pofut1 could affect *O*-fucosylation of these sites.

Finally, it would be of interest to test the impact of a more drastic downregulation of Pofut1 on postnatal muscle growth





**Figure 11.** Notch signalling pathway activation. (a) Detection of cleaved NOTCH intracellular domain (NICD) and GAPDH by western blotting on total extracted proteins of proliferating Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> SCDMs. Histograms represent Notch signalling activation (calculated by cleaved NICD/GAPDH ratio). (b) Gene expression levels of Notch target genes (*Hes1*, *Hey1*, *Hey1*) and *Hes6* during proliferation. Target mRNAs were normalized to *18S RNA* and *Gapdh* mRNA. Means  $\pm$  s.e.m. ( $n = 3$ ) are shown (two-tailed *t*-test, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

of heteroallelic Pofut1<sup>null/cax</sup> mice and to analyse the functional consequences of muscular hypertrophy on muscle strength and locomotor activity in these murine models.

## 4. Material and methods

### 4.1. Animals

Pofut1<sup>cax</sup> is a spontaneous mutation on a C3H background, which results from the insertion of an IAP retrotransposon into the fourth intron of the *pofut1* gene [45]. Mice were maintained under this background by intercrossing Pofut1<sup>+/cax</sup> animals. Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> mice were genotyped from genomic DNA by PCR as previously described [45]. All mice were bred and housed in the animal facility of Limoges University under controlled specific pathogen free conditions (21°C, 12 h light/12 h dark cycle, environmental enrichment) with free access to standard mouse chow and tap water.

### 4.2. Isolation of satellite cell-derived myoblasts

Primary cultures of myoblasts were obtained from mice of five weeks of age. SCDMs were isolated from skeletal

muscles of hind legs after enzymatic digestion by pronase, as previously described [64]. Cells were plated at a density of 15 000 cells cm<sup>-2</sup> on Matrigel<sup>®</sup>-coated Petri dishes (BD Biosciences, Franklin Lakes, NJ) in Ham's F10 (Invitrogen, Thermo Fisher Scientific, Waltham, MA) supplemented with 20% horse serum, 100 units ml<sup>-1</sup> penicillin and 100  $\mu$ g ml<sup>-1</sup> streptomycin. Cells were maintained for 48 h at 37°C and 5% CO<sub>2</sub> then washed with Ham's F10 before being placed in the GM, i.e. Ham's F10 supplemented with 5 ng ml<sup>-1</sup> basic fibroblast growth factor (bFGF, Invitrogen, Thermo Fisher Scientific), 20% heat-inactivated horse serum, 100 units ml<sup>-1</sup> penicillin and 100  $\mu$ g ml<sup>-1</sup> streptomycin. The myoblast population was enriched by eliminating fast-adherent fibroblasts using serial 30 min preplating procedures. To induce differentiation, SCDMs were incubated at 40–50% confluence in a DM consisting of Ham's F10 containing 10% heat-inactivated horse serum, 100 units ml<sup>-1</sup> penicillin and 100  $\mu$ g ml<sup>-1</sup> streptomycin. For estimation of MYOD and/or PAX7 expressing cell populations and for binding assays during the time course of differentiation, SCDMs were seeded at confluence and differentiation was induced four hours later. SCDM purity was determined by the proportion of cells expressing DESMIN (an intermediate filament protein located in proliferating skeletal myoblasts) compared with total cell number (DAPI+ cells).

### 4.3. Proliferation

SCDMs were seeded at 5000 cells per well in GM into 96-well plates and 20  $\mu$ l of MTS (Cell Titre 96 Aqueous Non-Radioactive cell proliferation assay; Promega Corp., Madison, WI) in a 200  $\mu$ l final volume were added at 0, 24, 48, 72 and 96 h. The plates were then incubated for 1 h at 37°C and absorbance of formazan, a product from the bioreduced MTS, was measured at 490 nm using an ELISA plate reader (FLUOstar Omega; BMGLabtech, Ortenberg, Germany). Six replicates were analysed at each time point and absorbance values of background (GM) were subtracted.

### 4.4. Cell culture and transient transfections of COS-7 cells

COS-7 cells (ATCC-CRL 1651) were maintained at 37°C in DMEM supplemented with 10% heat-inactivated fetal bovine serum, 100 units ml<sup>-1</sup> penicillin and 100  $\mu$ g ml<sup>-1</sup> streptomycin. COS-7 cells were seeded in 100 mm tissue culture dishes and transfected at 75% confluence using 18  $\mu$ l of XtremeGene9 transfection reagent (Roche Applied Science, Mannheim, Germany) with 6  $\mu$ g of each construct in serum-free DMEM supplemented with 1% penicillin/streptomycin. Supernatants were recovered by centrifugation at 96 h after transfection.

### 4.5. Ligand constructs

Dll1-AP and Jag1-AP were constructed by cloning the sequences encoding amino acids 1–540 of mouse DELTA-LIKE PROTEIN 1 (NP\_031891.2) and amino acids 1–1067 of mouse JAGGED1 (NP\_038850.1), respectively, in frame with human placental AP from pAPtag-2 (GenHunter Corp., Nashville, TN). Total cDNA from mouse tissues was used as a template for PCR to generate modified cDNA ends with



*Hind*III site in 5' and *Bgl*II or *Bam*HI site in 3'. After subcloning into pGEM-Teasy (Promega Corp.) and enzymatic digestion with appropriate restriction enzymes, purified fragments were directly cloned into pAPtag-2 using *Hind*III and *Bgl*II cloning sites to obtain the constructs referred to as pAPtag2-Dll1 and pAPtag2-Jag1. According to a previously described cloning technique [65], the pAPtag-2 was modified with prehybridized oligonucleotides encoding a signal peptide promoting secretion of recombinant Ctrl-AP. Eight prehybridized oligonucleotides encoding the same Igk-chain leader sequence as found in commercial pSecTag/FRT/V5-His-TOPO (Invitrogen, Thermo Fisher Scientific) were cloned between *Hind*III and *Bgl*II cloning sites downstream from the cytomegalovirus promoter in pAPtag-2. The plasmidic constructs were sequenced and were used to transiently transfect COS-7 cells.

#### 4.6. RNA extraction, reverse transcription and gene expression analysis

Total RNA was extracted from SCDMs and hind limb skeletal muscles of five week old *Pofut1*<sup>+/+</sup> and *Pofut1*<sup>cax/cax</sup> mice using RNeasy mini kit (Qiagen, Inc., Hilden, Germany). Quality and quantity of total RNA were measured using an Agilent 2100 bioanalyzer (Santa Clara, CA) and a Nanodrop 1000 spectrophotometer (Wilmington, DE), respectively. The high capacity cDNA reverse transcription kit (Invitrogen, Thermo Fisher Scientific) was used to convert 2 µg of total RNA into single-stranded cDNA. Semi-quantitative PCR was performed from 2 ng total cDNA in an ABI Prism 7900 Sequence Detector System (Applied Biosystems, Thermo Fisher Scientific) using 40 cycles at 95°C for 15 s followed by 60°C for 1 min. Taqman<sup>®</sup> primers and probe sets used in this study were as follows: 18S (Hs99999901\_s1), *Gapdh* (Mm99999915\_g1), *Cdkn1a* (Mm00432448\_m1), *Pofut1* (Mm00475567\_m1), *Pax7* (Mm03053796\_s1), *MyoD1* (Mm00440387\_m1), *Myog* (Mm00446194\_m1), *Myf5* (Mm00435125\_m1), *Hes1* (Mm00468601\_m1), *Hey1* (Mm00516555\_m1), *Hey1* (Mm00468865\_m1) and *Hes6* (Mm00517097\_g1). As described previously [44], the  $\Delta\Delta C_t$  method was used to quantify the relative abundance of each mRNA. RQ values, calculated only for a threshold cycle (*Ct*) lower than 37, reflected expression changes in the sample of interest compared with the calibrator sample, after normalization with 18S and *Gapdh* reference genes. Statistical analyses were performed by comparison of each differentiation time (from 48 to 240 h) relative to T0 for wild-type cells, which was set as 1. Expression of *Pofut1* in different skeletal muscles (*Quadriceps*, *Gastrocnemius*, *Tibialis* and *Soleus*) from *Pofut1*<sup>cax/cax</sup> mice was also studied, using wild-type muscles as a calibrator.

#### 4.7. Protein extraction and western blot

Proteins were isolated from SCDMs with RIPA extraction buffer (50 mM Tris-HCl, 150 mM NaCl, 0.5% sodium deoxycholate, 1% NP-40, 0.1% SDS, pH 8) containing a protease inhibitor cocktail (Roche Applied Science). Protein lysates were then centrifuged at 12 000g for 10 min at 4°C and soluble proteins from supernatants were quantified using a bicinchoninic acid protein assay (Sigma-Aldrich Corp., St Louis, MO) with bovine serum albumin (BSA) as a standard. Equal amounts of extracted proteins (30–100 µg) were separated under denaturing and reducing conditions on SDS-

polyacrylamide gels (6–12%) and then transferred to a Hybond C-extra nitrocellulose membrane (GE Healthcare, Buckinghamshire, UK). Then, membranes were blocked using 5% non-fat dried milk (w/v) in TBST (50 mM Tris-HCl, 150 mM NaCl, 0.1% Tween-20, pH 7.4) for 1 h at room temperature, followed by incubation overnight at 4°C with specific primary antibodies diluted in 2.5% non-fat dried milk (w/v) in TBST. The following primary antibodies were used for immunoblotting: 1:500 dilution of anti-Pofut1 purified antibody used in a previous study [44], 1:50 dilution of anti-Pax7 antibody (Developmental Studies Hybridoma Bank, University of Iowa, IA), 1:2000 dilution of anti-GAPDH (R&D Systems Inc., Minneapolis, MN), 1:300 dilution of anti-cleaved NOTCH1 (Val1744; Cell Signaling Technology, Danvers, MA). After three washes in TBST, membranes were incubated for 1 h at room temperature with 1:1000 dilution of secondary HRP conjugate antibodies (Dako, Glostrup, Denmark) in 2.5% non-fat dried milk (w/v) in TBST. After three more washes in TBST, immunoblots were developed using BM Chemiluminescence western blotting substrate (peroxidase, POD; Roche Applied Science) and exposed (Hyperfilm ECL, GE Healthcare). For relative quantification, analysis of band intensities was carried out using IMAGEJ software (Wayne Rasband, National Institutes of Health, USA).

#### 4.8. Production, quantification and characterization of recombinant proteins

Recombinant proteins (Ctrl-AP, Dll1-AP and Jag1-AP) from harvested supernatants of transfected COS-7 cells were directly used for binding assays or concentrated 25- to 30-fold by ultrafiltration in binding buffer HBAH (Hanks' balanced salt solution, 0.5 mg ml<sup>-1</sup> BSA, 0.1% sodium azide, 20 mM Hepes, pH 7). To estimate protein quantities, we quantified the enzymatic activity of AP by measuring the absorbance of para-nitrophenylphosphate, whose dephosphorylation led to a coloured product, according to recommendations of GenHunter (GenHunter Corp., Nashville, TN). The integrity of concentrated fusion protein and Ctrl-AP was checked by western blot using a 1:1000 dilution of anti-AP antibody (GenHunter Corp.). For AP assay, 120 µl pure or diluted protein samples were added to 120 µl of 2X AP reagent (2 M diethanolamine, 1 mM MgCl<sub>2</sub>, 1 mg ml<sup>-1</sup> BSA, 24 mM *p*-nitrophenylphosphate, pH 9.8). After mixing and incubation at 37°C for 10 min, the reaction was stopped by addition of 240 µl 0.5 N NaOH. The reaction was done in duplicate, and the absorbance was read at 405 nm.

#### 4.9. Biotinylation of cell surface proteins and Notch detection

*Pofut1*<sup>+/+</sup> and *Pofut1*<sup>cax/cax</sup> SCDMs were plated in 100 mm tissue culture dishes, collected 2 days later with a cell scraper and washed three times with PBS (pH 8). For each of the cell lines, 7 × 10<sup>6</sup> cells were incubated in 600 µl PBS (as a control) or PBS containing 1.66 mM EZ-link sulfo-*N*-hydroxysuccinimide-biotin (sulfo-NHS-biotin; Pierce Chemical Company, Rockford, IL) at 4°C for 45 min. The biotinylation reagent was then removed by centrifugation, and cells were washed three times with 1 ml of 100 mM glycine-PBS (pH 8) for 5 min to quench the biotinylation reaction. After cell lysis for 45 min at 4°C in 50 µl RIPA buffer, supernatants were



recovered by centrifugation (12 000g for 10 min) at 4°C, and 400 µg soluble total proteins (input) were incubated overnight at 4°C with 50 µl avidin agarose (Pierce Chemical Company) previously equilibrated with RIPA buffer. Unretained proteins were recovered by centrifugation at 200g for 5 min, and avidin agarose was washed three times with diluted RIPA buffer for 20 min at 4°C. Retained biotinylated proteins were eluted from the Pierce avidin agarose by boiling for 5 min in the presence of 40 µl SDS-PAGE sample Laemmli buffer under reducing conditions. Amounts of loaded proteins were 25 µg for samples corresponding to input and unretained proteins. For eluted proteins, 15 µl supernatant from boiled avidin agarose was loaded. All samples were subjected to SDS-PAGE using 6% or 12% running gels and then transferred to Immobilon-P<sup>®</sup> PVDF membranes (Bio-Rad, CA). After membrane blocking as previously described, a 1:500 dilution of an anti-NOTCH1 antibody (H-131) raised against amino acids 20–150 within the extracellular domain of human NOTCH1 (Santa Cruz Biotechnology, Santa Cruz, CA) was used to detect NOTCH1 at the SCDM cell surface. A 1:500 dilution of anti-pan-cadherin (Sigma-Aldrich Corp.) antibody was used to control the correct separation of membrane proteins. Anti-GAPDH (1:2000) was used to control that cytosolic proteins such as GAPDH were not biotinylated unlike membrane proteins at the cell surface.

#### 4.10. Binding assays

Binding assays were performed either with adherent cells seeded at confluence in 12-well plates or with suspension cells placed in microcentrifuge tubes, as previously described [66]. About 500 000 adherent cells per well from Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> SCDMs were washed once with 2.5 ml HBAH and incubated at room temperature for 90 min with 500 µl crude supernatants from transfected COS-7 cells containing either Dll1-AP, Jag1-AP or Ctrl-AP at the same concentrations. The quantities of Dll1-AP, Jag1-AP and Ctrl-AP were normalized based on their AP activity. After gentle shaking every 30 min, cells were washed for 5 min six times with 2.5 ml HBAH. Cells were then lysed in two steps with a total volume of 200 µl TT buffer (10 mM Tris-HCl, 1% Triton X-100, pH 8), and lysates were centrifuged at 16 000g for 5 min. After heat inactivation of cell lysates in a 65°C water bath for 10 min and then incubation on ice for 15 min, we performed AP assay as described above to evaluate specific binding of the fusion proteins, considering Ctrl-AP as a negative control.

Owing to low amount of protein for Jag1-AP in medium from transfected COS7 cells compared with that for Dll1-AP and undetectable binding of Jag1-AP to adherent cells with crude supernatants, we performed binding assays with concentrated proteins from supernatants. Suspension cells were thus used instead of adherent cells to avoid important cell detachment that we had observed when using concentrated supernatants. The same protocol as above was performed with 200 000 suspension cells, previously washed with HBAH and incubated in the presence of 100 µl of concentrated supernatants for Ctrl-AP, Dll1-AP and Jag1-AP assayed at the same concentrations ( $40 < A^{405nm} < 160$ ), based on the AP assay. Washes were performed six times for 5 min with 500 µl HBAH at a low speed spin for 1 min at 2000g in a microcentrifuge. Cells were then lysed in 75 µl

TT buffer and centrifuged at 16 000g for 5 min prior to heat inactivation and AP assay.

#### 4.11. Immunofluorescent staining of tissue sections and cells

Dissected skeletal muscles (*Quadriceps*, *Gastrocnemius*, *Tibialis* and *Soleus*) from six Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> mice at three different ages (5, 12 and 24 weeks) were frozen in liquid nitrogen-cooled isopentane and stored at -80°C before being sectioned. Cryosections (10 µm) were thawed at room temperature and air-dried. Cryosections and SCDMs in primary cultures from Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> mice were fixed for at least 15 min with 4% paraformaldehyde (PFA) in PBS and washed three times in PBS. Only cells were permeabilized for 30 min at 4°C with a buffer pH 7.4 containing 20 mM HEPES, 300 mM sucrose, 50 mM NaCl, 3 mM MgCl<sub>2</sub> and 0.5% Triton X-100. Both cryosections and cells were blocked for 1 h at room temperature in two blocking buffers (BBs): BB1 (10% goat serum, 1% BSA and 0.1% Triton X-100 in PBS) and BB2 (5% fetal calf serum, 2% BSA and 0.2% Triton X-100 in PBS). Then, they were washed with 0.2% BSA in PBS and incubated with primary antibodies diluted in PBS with 1% BSA for 1 h at 37°C. After three washes with PBS, 0.2% BSA, 0.1% Tween-20, slides were incubated for 15 min at 37°C with fluorescent conjugated secondary antibodies diluted 1:1000 in PBS 1% BSA as follows: Alexa Fluor<sup>®</sup> 488 goat anti-mouse IgG (H+L) (Invitrogen, Thermo Fisher Scientific) for anti-MyHC (1:2000; Sigma-Aldrich Corp.) and anti-PAX7 (1:500; Developmental Studies Hybridoma Bank, University of Iowa, IA), Alexa Fluor<sup>®</sup> 488 rabbit anti-goat IgG (H+L) (Invitrogen, Thermo Fisher Scientific) for anti-DESMIN (1:100; Santa Cruz Biotechnology) or Alexa Fluor<sup>®</sup> 546 F(ab')<sub>2</sub> fragment of goat anti-rabbit IgG (H+L) (Invitrogen, Thermo Fisher Scientific) for anti-LAMININ (1:500; Sigma-Aldrich Corp.) and anti-MyoD (1:1000; Santa Cruz Biotechnology). Staining was completed with three PBS washes and incubation for 5 min at room temperature in 1 µg ml<sup>-1</sup> DAPI solution to label cell nuclei. Cells and cryosections were rinsed five times with PBS, mounted with Mowiol 4–88 mounting medium and sealed with glass coverslips.

For cryosections, data were analysed from 12 randomly chosen fields in each of 12 cryosections per muscle for each animal (six mice per genotype and per age) by using an epifluorescence microscope (Leica DMI4000B MM AF Imaging System) powered by MetaMorph (Universal Imaging Corp., Downingtown, PA) equipped with a 40× objective. Triple immunostaining (LAMININ/DAPI/PAX7) revealed the myofibre surface area (LAMININ<sup>+</sup>), all myonuclei (DAPI<sup>+</sup>) and specifically satellite cells (PAX7<sup>+</sup>) located peripherally. Mean fluorescence intensity of captured images was analysed by IMAGEJ. We converted each image into a binary one, then collected the pixel values (one pixel on IMAGEJ = 0.47 µm<sup>2</sup> on MetaMorph) and calculated the total area of each field in µm<sup>2</sup>. By manual counting, we determined the total number of myofibres, myonuclei and satellite cells on cross-sectional field. The mean myofibre area was obtained by dividing the total area of the field (µm<sup>2</sup>) by total number of myofibres. Then, the number of nuclei per myofibre was calculated as well as the number of Pax7<sup>+</sup> satellite cells per myofibre, as recently published cryosection analyses [67].

For fusion indexes of SCDMs, data were analysed from 12 randomly chosen fields per well in duplicate for three



independent experiments ( $n = 3$ ). Images for fusion indexes were obtained by the same method as described above. Fusion index was calculated by dividing the number of myonuclei (DAPI+) contained in MyHC-expressing myotubes by the total number of myonuclei. To count the different cell populations expressing PAX7 and/or MYOD in Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> SCDMs, data were analysed from 12 randomly chosen fields per well in triplicate ( $n = 3$ ).

## 4.12. Statistical analysis

All experiments were performed in biological triplicates and results are reported as the means  $\pm$  s.e.m. Statistical comparisons were performed using two-tailed *t*-tests implemented in PRISM, v. 5.03 (GraphPad Software, Inc., San Diego, CA). A *p*-value of 0.05 or less was considered statistically significant.

**Ethics.** All of the experimental procedures were carried out in accordance with the recommendation in the guidelines of the European

Communities Council (Directive 2010/63/UE). Experiments were approved by the Committee of Ethics on Animal Experiments of the authors' institution, 'Comité Régional d'Éthique de l'Expérimentation Animale' of the Limousin region (no. 13-2014-13).

**Authors' contributions.** S.L., V.B. and A.M. conceived and designed the experiments. B.A.J., K.H., F.P., A.S. and S.L. performed the experiments. B.A.J. and S.L. drafted the manuscript. L.M., A.G., V.B. and A.M. critically revised the manuscript. All the authors have corrected and approved the manuscript for publication.

**Competing interests.** We declare we have no competing interests.

**Funding.** This work was supported by Lebanese Association for Scientific Research (LASER) to Bilal Al Jaam and was co-financed by the European Union (FEDER) and by the Limousin Regional Council.

**Acknowledgements.** We gratefully acknowledge Prof. Dr Achim Gossler for providing us with Pofut1<sup>cax/cax</sup> mice. We also thank Karine Pasquier, from the animal facility (SCA), for her technical help with animals and Dr Jeanne Cook-Moreau for her careful work on the English editing of the manuscript. The Pax7 antibody developed by Atsushi Kawakami was obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD and maintained by The University of Iowa, Department of Biology, Iowa City, IA 52242.

## References

- Mauro A. 1961 Satellite cell of skeletal muscle fibers. *J. Biophys. Biochem. Cytol.* **9**, 493–495. (doi:10.1083/jcb.9.2.493)
- Charge SB, Rudnicki MA. 2004 Cellular and molecular regulation of muscle regeneration. *Physiol. Rev.* **84**, 209–238. (doi:10.1152/physrev.00019.2003)
- Kuang S, Kuroda K, Le Grand F, Rudnicki MA. 2007 Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell* **129**, 999–1010. (doi:10.1016/j.cell.2007.03.044)
- Cheung TH, Rando TA. 2013 Molecular regulation of stem cell quiescence. *Nat. Rev. Mol. Cell Biol.* **14**, 329–340. (doi:10.1038/nrm3591)
- Almeida CF, Fernandes SA, Ribeiro Junior AF, Keith Okamoto O, Vainzof M. 2016 Muscle satellite cells: exploring the basic biology to rule them. *Stem Cells Int.* **2016**, 1078686. (doi:10.1155/2016/1078686)
- Chiakulas JJ, Pauly JE. 1965 A study of postnatal growth of skeletal muscle in the rat. *Anat. Rec.* **152**, 55–61. (doi:10.1002/ar.1091520107)
- Timson BF. 1982 The effect of varying postnatal growth rate on skeletal muscle fiber number in the mouse. *Growth* **46**, 36–45.
- White RB, Bierinx AS, Gnocchi VF, Zammit PS. 2010 Dynamics of muscle fibre growth during postnatal mouse development. *BMC Dev. Biol.* **10**, 21. (doi:10.1186/1471-213X-10-21)
- Yin H, Price F, Rudnicki MA. 2013 Satellite cells and the muscle stem cell niche. *Physiol. Rev.* **93**, 23–67. (doi:10.1152/physrev.00043.2011)
- Ciciliot S, Schiaffino S. 2010 Regeneration of mammalian skeletal muscle. Basic mechanisms and clinical implications. *Curr. Pharm. Des.* **16**, 906–914. (doi:10.2174/138161210790883453)
- Schiaffino S, Dyar KA, Ciciliot S, Blaauw B, Sandri M. 2013 Mechanisms regulating skeletal muscle growth and atrophy. *FEBS J.* **280**, 4294–4314. (doi:10.1111/febs.12253)
- Moss FP, Leblond CP. 1971 Satellite cells as the source of nuclei in muscles of growing rats. *Anat. Rec.* **170**, 421–435. (doi:10.1002/ar.1091700405)
- Schultz E, Gibson MC, Champion T. 1978 Satellite cells are mitotically quiescent in mature mouse muscle: an EM and radioautographic study. *J. Exp. Zool.* **206**, 451–456. (doi:10.1002/jez.1402060314)
- Guruharsha KG, Kankel MW, Artavanis-Tsakonas S. 2012 The Notch signalling system: recent insights into the complexity of a conserved pathway. *Nat. Rev. Genet.* **13**, 654–666. (doi:10.1038/nrg3272)
- Mourikis P, Tajbakhsh S. 2014 Distinct contextual roles for Notch signalling in skeletal muscle stem cells. *BMC Dev. Biol.* **14**, 2. (doi:10.1186/1471-213X-14-2)
- Conlon RA, Reaume AG, Rossant J. 1995 Notch1 is required for the coordinate segmentation of somites. *Development* **121**, 1533–1545.
- Swiatek PJ, Lindsell CE, del Amo FF, Weinmaster G, Gridley T. 1994 Notch1 is essential for postimplantation development in mice. *Genes Dev.* **8**, 707–719. (doi:10.1101/gad.8.6.707)
- Oka C *et al.* 1995 Disruption of the mouse RBP-J kappa gene results in early embryonic death. *Development* **121**, 3291–3301.
- Shi S, Stanley P. 2003 Protein O-fucosyltransferase 1 is an essential component of Notch signaling pathways. *Proc. Natl Acad. Sci. USA* **100**, 5234–5239. (doi:10.1073/pnas.0831126100)
- Vasyutina E, Lenhard DC, Wende H, Erdmann B, Epstein JA, Birchmeier C. 2007 RBP-J (Rbbsuh) is essential to maintain muscle progenitor cells and to generate satellite cells. *Proc. Natl Acad. Sci. USA* **104**, 4443–4448. (doi:10.1073/pnas.0610647104)
- Schuster-Gossler K, Cordes R, Gossler A. 2007 Premature myogenic differentiation and depletion of progenitor cells cause severe muscle hypotrophy in Delta1 mutants. *Proc. Natl Acad. Sci. USA* **104**, 537–542. (doi:10.1073/pnas.0608281104)
- Mumm JS, Kopan R. 2000 Notch signaling: from the outside in. *Dev. Biol.* **228**, 151–165. (doi:10.1006/dbio.2000.9960)
- Wu L, Sun T, Kobayashi K, Gao P, Griffin JD. 2002 Identification of a family of Mastermind-like transcriptional coactivators for mammalian Notch receptors. *Mol. Cell Biol.* **22**, 7688–7700. (doi:10.1128/MCB.22.21.7688-7700.2002)
- Kopan R, Ilagan MX. 2009 The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell* **137**, 216–233. (doi:10.1016/j.cell.2009.03.045)
- Kadesch T. 2004 Notch signaling: the demise of elegant simplicity. *Curr. Opin. Genet. Dev.* **14**, 506–512. (doi:10.1016/j.gde.2004.07.007)
- Kuroda K, Tani S, Tamura K, Minoguchi S, Kurooka H, Honjo T. 1999 Delta-induced Notch signaling mediated by RBP-J inhibits MyoD expression and myogenesis. *J. Biol. Chem.* **274**, 7238–7244. (doi:10.1074/jbc.274.11.7238)
- Aurade F, Pinset C, Chafey P, Gros F, Montarras D. 1994 Myf5, MyoD, myogenin and MRF4 myogenic derivatives of the embryonic mesenchymal cell line C3H10T1/2 exhibit the same adult muscle phenotype. *Differentiation* **55**, 185–192. (doi:10.1046/j.1432-0436.1994.5530185.x)
- Zammit PS, Relaix F, Nagata Y, Ruiz AP, Collins CA, Partridge TA, Beauchamp JR. 2006 Pax7 and myogenic progression in skeletal muscle satellite cells. *J. Cell Sci.* **119**, 1824–1832. (doi:10.1242/jcs.02908)
- Andres V, Walsh K. 1996 Myogenin expression, cell cycle withdrawal, and phenotypic differentiation are temporally separable events that precede cell fusion upon myogenesis. *J. Cell Biol.* **132**, 657–666. (doi:10.1083/jcb.132.4.657)



30. Wen Y, Bi P, Liu W, Asakura A, Keller C, Kuang S. 2012 Constitutive Notch activation upregulates Pax7 and promotes the self-renewal of skeletal muscle satellite cells. *Mol. Cell. Biol.* **32**, 2300–2311. (doi:10.1128/MCB.06753-11)
31. Seale P, Sabourin LA, Girgis-Gabardo A, Mansouri A, Gruss P, Rudnicki MA. 2000 Pax7 is required for the specification of myogenic satellite cells. *Cell* **102**, 777–786. (doi:10.1016/S0092-8674(00)00666-0)
32. Wang Y, Shao L, Shi S, Harris RJ, Spellman MW, Stanley P, Haltiwanger RS. 2001 Modification of epidermal growth factor-like repeats with O-fucose. Molecular cloning and expression of a novel GDP-fucose protein O-fucosyltransferase. *J. Biol. Chem.* **276**, 40 338–40 345. (doi:10.1074/jbc.M107849200)
33. Loril C, Audfray A, Dupuy F, Germot A, Maftah A. 2007 The two N-glycans present on bovine Pofut1 are differently involved in its solubility and activity. *FEBS J.* **274**, 1202–1211. (doi:10.1111/j.1742-4658.2007.05663.x)
34. Rampal R, Luther KB, Haltiwanger RS. 2007 Notch signaling in normal and disease states: possible therapies related to glycosylation. *Curr. Mol. Med.* **7**, 427–445. (doi:10.2174/156652407780831593)
35. Stanley P. 2007 Regulation of Notch signaling by glycosylation. *Curr. Opin. Struct. Biol.* **17**, 530–535. (doi:10.1016/j.sbi.2007.09.007)
36. Rampal R, Arboleda-Velasquez JF, Nita-Lazar A, Kosik KS, Haltiwanger RS. 2005 Highly conserved O-fucose sites have distinct effects on Notch1 function. *J. Biol. Chem.* **280**, 32 133–32 140. (doi:10.1074/jbc.M506104200)
37. Ge C, Stanley P. 2008 The O-fucose glycan in the ligand-binding domain of Notch1 regulates embryogenesis and T cell development. *Proc. Natl Acad. Sci. USA* **105**, 1539–1544. (doi:10.1073/pnas.0702846105)
38. Muller J, Rana NA, Serth K, Kakuda S, Haltiwanger RS, Gossler A. 2014 O-Fucosylation of the Notch ligand mDLL1 by POFUT1 is dispensable for ligand function. *PLoS ONE* **9**, e88571. (doi:10.1371/journal.pone.0088571)
39. Stahl M, Uemura K, Ge C, Shi S, Tashima Y, Stanley P. 2008 Roles of Pofut1 and O-fucose in mammalian Notch signaling. *J. Biol. Chem.* **283**, 13 638–13 651. (doi:10.1074/jbc.M802027200)
40. Haltiwanger RS, Stanley P. 2002 Modulation of receptor signaling by glycosylation: fringe is an O-fucose-beta1,3-N-acetylglucosaminyltransferase. *Biochim. Biophys. Acta* **1573**, 328–335. (doi:10.1016/S0304-4165(02)00400-2)
41. Chen J, Moloney DJ, Stanley P. 2001 Fringe modulation of Jagged1-induced Notch signaling requires the action of beta 4galactosyltransferase-1. *Proc. Natl Acad. Sci. USA* **98**, 13 716–13 721. (doi:10.1073/pnas.241398098)
42. Hou X, Tashima Y, Stanley P. 2012 Galactose differentially modulates Lunatic and Manic Fringe effects on Delta1-induced NOTCH signaling. *J. Biol. Chem.* **287**, 474–483. (doi:10.1074/jbc.M111.317578)
43. Takeuchi H, Haltiwanger RS. 2014 Significance of glycosylation in Notch signaling. *Biochem. Biophys. Res. Commun.* **453**, 235–242. (doi:10.1016/j.bbrc.2014.05.115)
44. Der Vartanian A, Audfray A, Al Jaam B, Janot M, Legardinier S, Maftah A, Germot A. 2015 Protein O-fucosyltransferase 1 expression impacts myogenic C2C12 cell commitment via the Notch signaling pathway. *Mol. Cell Biol.* **35**, 391–405. (doi:10.1128/MCB.00890-14)
45. Schuster-Gossler K, Harris B, Johnson KR, Serth J, Gossler A. 2009 Notch signalling in the paraxial mesoderm is most sensitive to reduced Pofut1 levels during early mouse development. *BMC Dev. Biol.* **9**, 6. (doi:10.1186/1471-213X-9-6)
46. Halevy O, Novitsch BG, Spicer DB, Skapek SX, Rhee J, Hannon GJ, Beach D, Lassar AB. 1995 Correlation of terminal cell cycle arrest of skeletal muscle with induction of p21 by MyoD. *Science* **267**, 1018–1021. (doi:10.1126/science.7863327)
47. Yang LT, Nichols JT, Yao C, Manilay JO, Robey EA, Weinmaster G. 2005 Fringe glycosyltransferases differentially modulate Notch1 proteolysis induced by Delta1 and Jagged1. *Mol. Biol. Cell* **16**, 927–942. (doi:10.1091/mbc.E04-07-0614)
48. Bruckner K, Perez L, Clausen H, Cohen S. 2000 Glycosyltransferase activity of Fringe modulates Notch–Delta interactions. *Nature* **406**, 411–415. (doi:10.1038/35019075)
49. Katoh M. 2007 Integrative genomic analyses on HES/HEY family: Notch-independent HES1, HES3 transcription in undifferentiated ES cells, and Notch-dependent HES1, HES5, HEY1, HEY2, HEYL transcription in fetal tissues, adult tissues, or cancer. *Int. J. Oncol.* **31**, 461–466. (doi:10.3892/ijco.31.2.461)
50. Fior R, Henrique D. 2005 A novel hes5/hes6 circuitry of negative regulation controls Notch activity during neurogenesis. *Dev. Biol.* **281**, 318–333. (doi:10.1016/j.ydbio.2005.03.017)
51. Ontell M, Feng KC, Klueber K, Dunn RF, Taylor F. 1984 Myosatellite cells, growth, and regeneration in murine dystrophic muscle: a quantitative study. *Anat. Rec.* **208**, 159–174. (doi:10.1002/ar.1092080203)
52. Relaix F, Rocancourt D, Mansouri A, Buckingham M. 2005 A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435**, 948–953. (doi:10.1038/nature03594)
53. Lepper C, Fan CM. 2010 Inducible lineage tracing of Pax7-descendant cells reveals embryonic origin of adult satellite cells. *Genesis* **48**, 424–436. (doi:10.1002/dvg.20630)
54. Kassar-Duchossoy L, Giaccone E, Gayraud-Morel B, Jory A, Gomes D, Tajbakhsh S. 2005 Pax3/Pax7 mark a novel population of primitive myogenic cells during development. *Genes Dev.* **19**, 1426–1431. (doi:10.1101/gad.345505)
55. Bermudez Brito M, Goulielmaki E, Papakonstanti EA. 2015 Focus on PTEN regulation. *Front. Oncol.* **5**, 166. (doi:10.3389/fonc.2015.00166)
56. Palomero T *et al.* 2007 Mutational loss of PTEN induces resistance to NOTCH1 inhibition in T-cell leukemia. *Nat. Med.* **13**, 1203–1210. (doi:10.1038/nm1636)
57. Chappell WH, Green TD, Spengeman JD, McCubrey JA, Akula SM, Bertrand FE. 2005 Increased protein expression of the PTEN tumor suppressor in the presence of constitutively active Notch-1. *Cell Cycle* **4**, 1389–1395. (doi:10.4161/cc.4.10.2028)
58. Whelan JT, Forbes SL, Bertrand FE. 2007 CBF-1 (RBP-J kappa) binds to the PTEN promoter and regulates PTEN gene expression. *Cell Cycle* **6**, 80–84. (doi:10.4161/cc.6.1.3648)
59. Rodriguez J, Vernus B, Chelh I, Cassar-Malek I, Gabillard JC, Hadj Sassi A, Seiliez I, Picard B, Bonniou A. 2014 Myostatin and the skeletal muscle atrophy and hypertrophy signaling pathways. *Cell Mol. Life Sci.* **71**, 4361–4371. (doi:10.1007/s00018-014-1689-x)
60. Okajima T, Xu A, Lei L, Irvine KD. 2005 Chaperone activity of protein O-fucosyltransferase 1 promotes Notch receptor folding. *Science* **307**, 1599–1603. (doi:10.1126/science.1108995)
61. Rana NA, Haltiwanger RS. 2011 Fringe benefits: functional and structural impacts of O-glycosylation on the extracellular domain of Notch receptors. *Curr. Opin. Struct. Biol.* **21**, 583–589. (doi:10.1016/j.sbi.2011.08.008)
62. Harvey BM, Rana NA, Moss H, Leonardi J, Jafar-Nejad H, Haltiwanger RS. 2016 Mapping sites of O-glycosylation and fringe elongation on *Drosophila* Notch. *J. Biol. Chem.* **291**, 16348. (doi:10.1074/jbc.M116.732537)
63. Haltiwanger RS, Lowe JB. 2004 Role of glycosylation in development. *Annu. Rev. Biochem.* **73**, 491–537. (doi:10.1146/annurev.biochem.73.011303.074043)
64. Brun C, Perie L, Baraige F, Vernus B, Bonniou A, Blanquet V. 2014 Absence of hyperplasia in Gasp-1 overexpressing mice is dependent on myostatin up-regulation. *Cell Physiol. Biochem.* **34**, 1241–1259. (doi:10.1159/000366335)
65. Legardinier S, Klett D, Poirier JC, Combarrous Y, Cahoreau C. 2005 Mammalian-like nonsialyl complex-type N-glycosylation of equine gonadotropins in Mimic insect cells. *Glycobiology* **15**, 776–790. (doi:10.1093/glycob/cwi060)
66. Flanagan JG, Cheng HJ. 2000 Alkaline phosphatase fusion proteins for molecular characterization and cloning of receptors and their ligands. *Methods Enzymol.* **327**, 198–210. (doi:10.1016/S0076-6879(00)27277-7)
67. Fry CS *et al.* 2015 Inducible depletion of satellite cells in adult, sedentary mice impairs muscle regenerative capacity without affecting sarcopenia. *Nat. Med.* **21**, 76–80. (doi:10.1038/nm.3710)



# SUPPLEMENTARY DATA

		<i>Quadriceps</i>	<i>Gastrocnemius</i>	<i>Tibialis</i>	<i>Soleus</i>	
Muscle weight/Body weight ratios (%)	+/+	1.11 ± 0.06	0.79 ± 0.03	0.33 ± 0.02	0.14 ± 0.01	5 weeks
	cax/cax	1.35 ± 0.03**	1.08 ± 0.06**	0.53 ± 0.04*	0.30 ± 0.02*	
	+/+	1.06 ± 0.04	1.10 ± 0.04	0.29 ± 0.01	0.15 ± 0.01	12 weeks
	cax/cax	1.31 ± 0.06 **	1.48 ± 0.04 **	0.43 ± 0.03 *	0.29 ± 0.01 *	
	+/+	0.98 ± 0.07	0.75 ± 0.03	0.26 ± 0.01	0.15 ± 0.00	24 weeks
	cax/cax	1.30 ± 0.03 **	1.01 ± 0.03 **	0.43 ± 0.03 *	0.27 ± 0.01 *	
Myofibre mean area (x10 <sup>3</sup> µm <sup>2</sup> )	+/+	2.06 ± 0.12	1.61 ± 0.11	1.88 ± 0.13	1.48 ± 0.11	5 weeks
	cax/cax	2.69 ± 0.14**	2.09 ± 0.04**	2.72 ± 0.10 ***	1.98 ± 0.08**	
	+/+	2.56 ± 0.08	2.57 ± 0.09	2.03 ± 0.08	1.77 ± 0.03	12 weeks
	cax/cax	3.34 ± 0.09 **	3.19 ± 0.04 **	2.83 ± 0.02 ***	2.95 ± 0.06 ***	
	+/+	2.86 ± 0.08	2.48 ± 0.06	2.01 ± 0.06	1.81 ± 0.11	24 weeks
	cax/cax	3.80 ± 0.15 **	3.11 ± 0.10 **	2.84 ± 0.17 *	2.55 ± 0.11 **	
Nuclei/Myofibre	+/+	3.03 ± 0.19	2.03 ± 0.43	2.79 ± 0.18	2.73 ± 0.25	5 weeks
	cax/cax	4.10 ± 0.22**	3.20 ± 0.20**	4.32 ± 0.21**	3.91 ± 0.14**	
	+/+	2.72 ± 0.06	2.98 ± 0.10	3.51 ± 0.18	2.69 ± 0.09	12 weeks
	cax/cax	3.90 ± 0.25 **	4.55 ± 0.04 **	4.88 ± 0.12 **	5.26 ± 0.26 ***	
	+/+	2.52 ± 0.06	1.08 ± 0.15	2.30 ± 0.06	3.04 ± 0.20	24 weeks
	cax/cax	3.67 ± 0.10 **	1.96 ± 0.16 **	3.41 ± 0.06 **	4.34 ± 0.11 **	
Myofibres/Field	+/+	234.14 ± 10.97	301.50 ± 11.25	296.56 ± 11.64	384.55 ± 10.16	5 weeks
	cax/cax	231.48 ± 15.27	293.83 ± 8.86	281.25 ± 11.59	356.05 ± 13.48	
	+/+	174.46 ± 7.39	191.86 ± 7.57	205.63 ± 17.45	233.07 ± 24.29	12 weeks
	cax/cax	177.06 ± 7.24	206.89 ± 20.22	212.37 ± 4.90	243.49 ± 21.93	
	+/+	167.64 ± 10.86	211.93 ± 8.46	195.03 ± 2.55	229.75 ± 14.46	24 weeks
	cax/cax	158.23 ± 6.70	218.06 ± 20.71	209.21 ± 9.46	223.29 ± 09.21	
Satellite Cells/Myofibre	+/+	0.068 ± 0.004	0.042 ± 0.004	0.033 ± 0.002	0.029 ± 0.002	5 weeks
	cax/cax	0.031 ± 0.002***	0.025 ± 0.003**	0.020 ± 0.001**	0.019 ± 0.002 **	
	+/+	0.054 ± 0.008	0.068 ± 0.002	0.046 ± 0.002	0.021 ± 0.002	12 weeks
	cax/cax	0.016 ± 0.002 ***	0.009 ± 0.001 ***	0.009 ± 0.005 **	0.005 ± 0.001 **	
	+/+	0.063 ± 0.02	0.034 ± 0.01	0.052 ± 0.002	0.031 ± 0.01	24 weeks
	cax/cax	0.015 ± 0.004 *	0.023 ± 0.009	0.023 ± 0.001 ***	0.004 ± 0.002 *	

**Table S1. Characteristics of skeletal muscles from 5, 12 and 24 week old Pofut1<sup>+/+</sup> and Pofut1<sup>cax/cax</sup> mice.** In addition to muscle weight/body weight ratios, myofibre mean area, number of nuclei per myofibre, number of myofibres per field and number of satellite cells per myofibre were determined for skeletal muscles with fast-twitch (*Tibialis*), slow-twitch (*Soleus*) or mixed (*Quadriceps*, *Gastrocnemius*) myofibres. Mean ± SEM are shown (t-test two-tailed, with a significance level of \*: p<0.05, \*\*: p<0.01, \*\*\*: p<0.001).



**Mouse WIF1 is only *O*-fucosylated on EGF-III despite two evolutionary conserved consensus sites**

**Pennarubia Florian**, Al Jaam Bilal, Pinault Emilie, Brun Caroline E., Maftah Abderrahman, Germot Agnès and Legardinier Sébastien

**Sci. Rep.** « en révision »



L'étude précédente a mis en évidence le rôle de la protéine *O*-fucosyltransférase POFUT1 dans la dérégulation de la voie de signalisation Notch et ses conséquences sur la différenciation myogénique et la croissance musculaire post-natale chez la souris. Cependant, le récepteur NOTCH n'est pas la seule protéine susceptible d'être altérée la diminution de quantité de POFUT1. En effet, il existe un certain nombre d'autres glycoprotéines membranaires et sécrétées connues pour être potentiellement *O*-fucosylées. Néanmoins, aucune d'entre elles ne semble être directement impliquée dans le processus de myogenèse.

L'objectif principal de ma thèse a donc été d'identifier de nouvelles protéines susceptibles d'être *O*-fucosylées par POFUT1, en adéquation avec un des domaines d'études de mon équipe d'accueil (rôle régulateur de la *O*-fucosylation sur la différenciation normale (myogénique) et pathologique (cancer colorectal)). Dans un premier temps, une étude *in silico*, basée sur la recherche de motifs consensus de *O*-fucosylation ( $C^2-X-X-X-X-(S/T)-C^3$ ) au sein de domaines EGF-like (ELDs), a permis d'établir une liste de protéines murines potentiellement modifiées par POFUT1. La publication par une équipe américaine d'une liste similaire, effectuée sur les protéines humaines, a mis en évidence un petit nombre de protéines communes potentiellement *O*-fucosylables chez l'Homme et la souris (Schneider et al., 2017). Au total, il en existe un peu moins d'une centaine chez ces deux espèces, dont la protéine Wnt Inhibitory Factor 1 (WIF1) qui possède deux séquences consensus de *O*-fucosylation (au sein des ELDs III et V). Cet inhibiteur de la voie Wnt est un candidat de choix pour l'étude de la myogenèse et du cancer colorectal (les deux thématiques de recherches de notre équipe) puisqu'il est potentiellement impliqué dans la myogenèse *via* son activité inhibitrice de la voie Wnt et dans différents types de cancers puisque reconnu comme un gène suppresseur de tumeurs.

Nous avons donc mis en place différentes approches dans le but de prouver que WIF1 pouvait être *O*-fucosylé. Dans un premier temps, nous nous sommes intéressés à la distribution et à la conservation de la protéine Wif1 chez différentes espèces. Ce qui nous a permis de dresser un arbre phylogénétique qui montre l'apparition de Wif1 chez les ancêtres bilatériens ainsi qu'une forte pression de sélection pour maintenir la protéine Wif1 au cours de l'évolution suggérant un rôle important. Chez les gnathostomes, faisant partie des deutérostomiens, nous avons mis en évidence une forte conservation des deux sites potentiels de *O*-fucosylation, présents respectivement au sein des ELDs III et V de Wif1 (avec une plus



forte conservation pour l'ELD III). Dans le contexte d'une caractérisation plus globale de la glycosylation, nous avons également mis en évidence la conservation, chez ce même infra-embranchement, de deux sites de *N*-glycosylation (caractérisés par la présence de la séquence consensus (N-X-(S/T)) localisés dans le domaine WIF et l'ELD III. La très forte conservation de l'ensemble de ces sites de glycosylation suggère leur occupation effective par des mono ou oligosaccharides et très probablement un rôle important pour le fonctionnement de la protéine.

Une seconde étape a été de démontrer biochimiquement l'occupation de ces sites. Nous avons produit de manière recombinante les ELDs III et V (sauvages et mutés sur le site de *O*-fucosylation) de WIF1 murins exprimés dans le périplasme bactérien (*E. coli*). De plus, nous avons exprimé la protéine WIF1 murine entière, sous sa forme sauvage et mutée sur un ou plusieurs sites de *O*-fucosylation ou de *N*-glycosylation, de manière sécrétée par une lignée de cellules de mammifères (CHO). L'élimination des *N*-glycanes, par traitement enzymatique à la PNGase F, nous a permis de conclure que les deux sites de *N*-glycosylation conservés étaient réellement occupés dans la protéine WIF1 de souris. La *O*-fucosylation *in vitro* des ELDs III et V de WIF1 avait pour objectif de tester rapidement le caractère *O*-fucosylable ou non de ces ELDs. Ainsi, nous avons démontré à l'aide de deux techniques complémentaires (la spectrométrie de masse en mode MRM et la chimie click) que seul l'ELD III était capable de porter un *O*-fucose. Des résultats complémentaires ont montré que la protéine WIF1 entière était *O*-fucosylable *in vitro* par POFUT1, sur son ELD III uniquement. Ceci prouve que WIF1 est bien une protéine cible de POFUT1 mais aussi que l'expression endogène de cette enzyme ne suffit pas à modifier 100% des protéines. Pour révéler la présence effective du fucose *O*-lié sur la protéine WIF1 sécrétée par les lignées CHO stables, nous avons choisi d'utiliser une autre enzyme produite de manière recombinante, Lunatic Fringe (LFNG). Cette dernière ne rallonge que les *O*-fucoses préexistants en ajoutant un *N*-acétylglucosamine (GlcNAc). Après incubation avec LFNG et révélation par chimie click, nous avons démontré que la protéine WIF1 murine était *O*-fucosylée *in cellulo* sur son ELD III.

Afin de statuer sur le rôle potentiel de la *N*-glycosylation et de la *O*-fucosylation, nous avons d'abord étudié la sécrétion de WIF1 sauvage et muté. Comme attendu, nous avons montré que la présence des deux *N*-glycanes favorisait la sécrétion de WIF1 et de manière plus surprenante que la présence du fucose *O*-lié de l'ELD III avait un effet similaire.



Le dernier aspect de notre étude tente d'apporter une explication moléculaire à l'absence de *O*-fucosylation sur le site très conservé présent dans l'ELD V de WIF1. La co-cristallisation de POFUT1 avec l'ELD 26 de NOTCH1 (connu pour être *O*-fucosylé) (Li et al., 2017) a révélé l'emplacement précis de cet ELD à l'interface des deux domaines de Rossman de POFUT1, ainsi que les résidus clés interagissant principalement par des liaisons hydrogène avec les résidus du sous-domaine C<sup>2</sup>-C<sup>3</sup> de l'ELD 26. L'obtention d'un modèle structural par homologie de l'ELD V de WIF1 et son positionnement dans la cavité de l'enzyme à la place de l'ELD 26 a révélé un clash stérique important entre l'histidine du domaine C<sup>2</sup>-C<sup>3</sup> (en position C<sup>2+3</sup>) de l'ELD V et un résidu de POFUT1, normalement impliqué dans l'interaction avec les ELDs. Ce clash stérique pourrait être suffisamment conséquent pour prévenir l'interaction entre POFUT1 et l'ELD V, rendant impossible la *O*-fucosylation.

Nous avons ainsi pu mettre en évidence une conservation de la protéine WIF1 au cours de l'évolution, ainsi qu'une forte pression de sélection pour maintenir deux sites de *N*-glycosylation et deux sites de *O*-fucosylation. Pourtant, trois d'entre eux seulement sont occupés par des mono- ou oligosaccharides, et sont tous les trois nécessaires pour une sécrétion optimale de la protéine WIF1.



1 **Mouse WIF1 is only *O*-fucosylated on EGF-III despite two evolutionarily conserved consensus sites**

2  
3 **Pennarubia Florian<sup>1</sup>, Al Jaam Bilal<sup>1</sup>, Pinault Emilie<sup>1, 2</sup>, Brun Caroline E.<sup>1, 3</sup>, Maftah**  
4 **Abderrahman<sup>1, #, \*</sup>, Germot Agnès<sup>1, #</sup> and Legardinier Sébastien<sup>1, #</sup>**

5  
6 <sup>1</sup> Univ. Limoges, PEIRENE, EA 7500, Glycosylation and cell differentiation, F-87000 Limoges, France.

7 <sup>2</sup> Univ. Limoges, BISCEM Mass Spectrometry Platform, F-87025 Limoges, France

8 <sup>3</sup> Regenerative Medicine Program, Department of Cellular and Molecular Medicine, Faculty of Medicine,  
9 Sprott Center for Stem Cell Research, Ottawa Hospital Research Institute, University of Ottawa, Canada.

10 <sup>#</sup>A. Maftah, A. Germot and S. Legardinier are considered co-last authors and contributed equally to this  
11 work

12 <sup>\*</sup> corresponding author: Abderrahman Maftah, [abderrahman.maftah@unilim.fr](mailto:abderrahman.maftah@unilim.fr)

13  
14 **Running head:** Wif1 as a target protein of Pofut1

15 **Keywords:** Wif1; phylogeny; Pofut1; *O*-fucosylation; EGF-like domain

16  
17 **ABSTRACT**

18 **The canonical Wnt/ $\beta$ -catenin pathway plays a prominent role in embryogenesis and homeostasis of**  
19 **adult tissues. It is principally controlled by secreted extracellular antagonists such as Wnt**  
20 **Inhibitory Factor 1 (Wif1), which binds to Wnt proteins and inhibits their activity. Structural and**  
21 **functional studies showed that EGF-like domains (EGFs) reinforced the inhibitory activity of Wif1**  
22 **mainly attributed to WIF domain (WD) and suggested their potential interaction with extracellular**  
23 **matrix proteoglycans. In this study, a phylogenetic analysis elucidated the distribution, conservation**  
24 **and evolution of Wif1 proteins in bilaterians. We demonstrated that the two highly conserved**  
25 ***N*-glycosylation sites in deuterostomes are occupied in mouse WIF1 and both are involved in protein**  
26 **secretion. Using click chemistry and mass spectrometry analyses, we proved that only EGF-III is**  
27 ***O*-fucosylated by POFUT1, although EGF-V carries the *O*-fucosylation consensus motif**  
28 **C<sup>2</sup>X4(S/T)C<sup>3</sup>, conserved in gnathostome Wif1. Based on X-ray crystal structures and protein**  
29 **modeling, we showed that EGF-V is probably unable to interact with POFUT1. Similarly to**  
30 ***N*-glycans, the presence of *O*-fucose on EGF-III was likely to promote WIF1 secretion. This study**  
31 **raises the question of the role of *O*-fucosylation in the functional activity of Wif1 and the regulation**  
32 **of the Wnt signaling pathway.**



## 33 INTRODUCTION

34 Wif1 was first identified in human retina and then isolated from mouse, *Xenopus* and zebrafish <sup>1</sup>. Wif1,  
35 like Cerberus and members of secreted Frizzled-related protein (sFRP), is an extracellular antagonist of  
36 both canonical and non-canonical Wnt signaling pathways <sup>2</sup>. Indeed, it can bind to Wnt proteins and  
37 prevents them from interacting with the cysteine-rich domain of the Frizzled receptor <sup>3</sup>. The canonical  
38 Wnt/ $\beta$ -catenin pathway is essential during vertebrate embryonic development <sup>4</sup> and in homeostasis of  
39 almost all adult tissues <sup>5,6</sup>. It plays important roles in cell physiology such as in determination,  
40 proliferation, differentiation and migration <sup>7</sup>. Thus, the deregulation of this signaling pathway is frequently  
41 associated with human diseases such as osteoporosis <sup>7</sup> and cancers <sup>8,9</sup>. If in *Wif1*<sup>-/-</sup> mice no obvious  
42 phenotype has been observed, the loss of WIF1 increases susceptibility to spontaneous and radiation-  
43 induced osteosarcomas <sup>10</sup>. Down-expression of human *WIF1* due to hypermethylation of its promoter is  
44 frequently observed in a broad-spectrum of cancers such bladder <sup>11</sup>, brain <sup>12,13</sup>, colorectal <sup>14</sup> and lung <sup>15</sup>  
45 cancers. This inhibition induces aberrant activation of Wnt signaling, subsequently leading to abrogation  
46 of controlled growth and impairment of cell differentiation <sup>16</sup>.

47 Wif1 contains an N-terminal signal peptide, a  $\beta$ -sandwich WIF domain (WD) <sup>17</sup>, five EGF-like repeats  
48 (EGFs) of the hEGF type <sup>18</sup> comprising 31-33 residues with three conserved disulfide bonds each <sup>19</sup> and a  
49 hydrophilic C-terminus. Although the WIF domain alone is sufficient to confer an inhibitory activity to  
50 Wif1 <sup>1</sup>, the presence of EGFs I-V is required for full activity possibly by strengthen Wif1 binding to Wnt  
51 proteins. They bind to negatively charged cell-bound glypican heparan sulfate proteoglycans (HSPGs)  
52 through electrostatic interactions with lysine and arginine residues of EGF-IV <sup>19</sup>. Similarly to Shifted  
53 (Shf), the *Drosophila* ortholog of Wif1 <sup>20</sup>, which controls the distribution and movements of lipid-  
54 modified Hedgehog (Hh) morphogens by forming Shf-Hh complexes, free or bound to HSPG <sup>21</sup>, Wif1 is  
55 thought to exert such a role in the regulation of Wnt morphogen gradients <sup>19</sup>. However, Shf does not  
56 inhibit signaling by Wingless, the *Drosophila* ortholog of mouse Wnt-1, suggesting contrasting properties  
57 between Shf and Wif1.

58 As a secreted protein, Wif1 is the target of post-translational modifications by *N*- and *O*-glycosylations.  
59 *N*-glycans, linked to asparagine residues within the asparagine-X-serine/threonine consensus sequences  
60 (N-X-S/T, with X never P), are well-known to be involved in protein folding, quality control and  
61 trafficking as well as in protein function, plasma half-life and protection against proteolysis <sup>22,23</sup>. EGF-like  
62 repeats are the target of three rare protein modifications: *O*-fucosylation, *O*-glucosylation and  
63 *O*-GlcNAcylation. In human WIF1, three potential glycosylation sites, N<sup>88</sup>, N<sup>245</sup> and T<sup>255</sup> have been  
64 identified <sup>19</sup>. Among them, T<sup>255</sup> belongs to the consensus motif of *O*-fucosylation C<sup>2</sup>X4(S/T)C<sup>3</sup> (where C<sup>2</sup>  
65 and C<sup>3</sup> are the second and third conserved cysteines, respectively) present in EGF-III. The endoplasmic-  
66 resident protein *O*-fucosyltransferase 1 (Pofut1) is responsible for EGF-like repeat *O*-fucosylation <sup>24</sup> and is



67 involved in numerous processes such as cell proliferation and differentiation <sup>25</sup> during myogenesis <sup>26</sup>. If  
68 about 100 membrane or secreted proteins are potentially *O*-fucosylated in human on the basis of  
69 consensus motif presence <sup>27</sup>, only a few have been described such as Notch receptors and its Delta and  
70 Jagged ligands (listed in <sup>28</sup>), urokinase plasminogen activator <sup>29</sup>, blood coagulation factors (VII, IX, XII)  
71 <sup>30-32</sup>, Cripto <sup>33</sup>, Agrin <sup>34</sup> and Amaco <sup>35</sup>. Pofut1 is widely distributed in animals <sup>36</sup> and its presence is highly  
72 correlated with that of *O*-fucosylable EGFs of the hEGF type, such as those of Wif1 <sup>28</sup>. Three main  
73 regions are involved in the recognition of EGF-like domains, namely the C<sup>1</sup>-C<sup>2</sup> loop, the C<sup>2</sup>-C<sup>3</sup> and C<sup>5</sup>-C<sup>6</sup>  
74 subdomains. Moreover, the unique residue between C<sup>4</sup> and C<sup>5</sup> is also important <sup>28</sup>. The C<sup>2</sup>-C<sup>3</sup> subdomain,  
75 which includes the *O*-fucosylation motif is composed of residues directly interacting with Pofut1. These  
76 latter establish one sulphur-hydrogen and several hydrogen bonds with conserved residues of Pofut1,  
77 located in a deep groove between the two Rossmann-fold domains <sup>28,37</sup>. The C<sup>1</sup>-C<sup>2</sup> loop plays a minor role  
78 in substrate binding in contrast to the C<sup>5</sup>-C<sup>6</sup> subdomain and residue at position C<sup>4</sup>+1, which modulates  
79 substrate-binding affinity through apolar interactions.

80 Except for a role in the regulation of interactions between Notch receptor and its membrane-bound Delta  
81 and Jagged/Serrate ligands <sup>38,39</sup> and in the regulation of agrin functions such as aggregation of  
82 acetylcholine receptors <sup>34</sup>, the function of *O*-linked fucose in EGF-like containing glycoproteins is poorly  
83 documented. Some *O*-fucoses can be extended by a GlcNAc or further elongated to form the *O*-linked  
84 sialylated tetrasaccharide, NeuAc- $\alpha$ 2,3/6-Gal $\beta$ -1,4-GlcNAc- $\beta$ 1,3-Fuc <sup>39</sup>. The Fringe-dependent addition of  
85 this GlcNAc can modulate Notch interaction with Jagged and Delta ligands <sup>40-43</sup>. However, the GlcNAc  
86 transferases of Fringe family (Lunatic, Manic and Radical) do not target all Notch *O*-fucoses <sup>44</sup>.

87 Using a phylogenetic approach, we studied the distribution and evolutionary conservation of Wif1 as of  
88 its potential *O*-fucosylation and *N*-glycosylation sites in metazoans. In gnathostomes, we identified two  
89 potential conserved *O*-fucosylation sites in EGF-III and EGF-V, whereas in protostomes only the first site  
90 was sporadically found. We also showed that the *N*-glycosylation site homologous to *Mus* N<sup>88</sup> was highly  
91 conserved in bilaterians, unlike the one homologous to *Mus* N<sup>245</sup>. However, both were occupied and  
92 promoted secretion of murine WIF1. To detect *O*-fucosylation, we used a strategy based on *in vitro*  
93 POFUT1-mediated *O*-fucosylation of recombinant isolated EGF-like repeats with GDP-azido-fucose or  
94 GDP-fucose. Using copper-assisted azide-alkyne cycloaddition (CuAAC) chemistry and multiple  
95 reactions monitoring-mass spectrometry (MRM-MS), EGF-III was shown to be *O*-fucosylated whereas  
96 EGF-V was not. The same results were obtained with full-length WIF1. Protein modeling was used to  
97 explain the lack of interaction between EGF-V and POFUT1.

98

## 99 RESULTS



100 **Wif1 appeared in bilaterian ancestors.** More than 300 Wif1 orthologs were retrieved from diverse  
101 databases using tblastn and PSI-BLAST programs of the BLAST<sup>®</sup> algorithm. No sponge or diploblastic  
102 species were identified with the characteristic WIF domain and the 5 EGF-like domains. We selected 47  
103 complete sequences, 27 from deuterostomes (exclusively represented by sequences of gnathostomes  
104 because only partial sequences were retrieved from databases for earlier emerging taxa) and 20 from  
105 protostomes, in order to encompass maximal bilaterian taxonomic diversity. Their sizes varied from 373  
106 amino acids (aa) for *Ornithorhynchus anatinus* to 381 for *Astyanax mexicanus* in deuterostomes and from  
107 353 aa for *Onthophagus taurus* to 456 for *Drosophila melanogaster* (due to a longer N-terminal part) in  
108 protostomes. The Maximum likelihood (ML) tree clearly separated protostomes from deuterostomes  
109 (Fig. 1) and supported our current knowledge of animal evolutionary relationships<sup>45</sup>. Monophyly of  
110 amniotes, sarcopterygians, lophotrochozoans and arthropods were supported by bootstrap proportions of  
111 88%, 86%, 70% and 100%, respectively. Bayesian inference of the phylogenetic tree produced the same  
112 topology, except for the teleost paraphyly and their earlier divergence was comparable to chondrichthyes.  
113 This was surprising, considering that the site-heterogeneous category (CAT) mixture model used is able to  
114 overcome LBA (Long Branch Attraction) artefacts<sup>46</sup>. LBA causes systematic errors in phylogeny as  
115 clusters sequences based on their shared dissimilarity (due to mutational saturation of sites) relative to  
116 closely related groups of organisms and consequently does not reveal their true evolutionary relationships.  
117 Percentages of similarity of the selected deuterostome and protostome sequences ranged from 71.1%  
118 (*Astyanax* vs *Ornithorhynchus*) to 96.6% (*Camelus* vs *Orycteropus*) and from 33.8% (*Drosophila* vs  
119 *Mizuhopecten*) to 88.4% (*Megachile* vs *Microplitis*), respectively. It resulted in significantly longer  
120 branches for protostomes compared to deuterostomes, underlying distinct selective pressures and  
121 functional divergences as demonstrated for *Drosophila* Wif1 orthologs Shifted and human WIF1,  
122 targeting Hedgehog or Wingless/Wnt morphogens, respectively<sup>47</sup>. Among partial sequences, Wif1 was  
123 found in the early diverging deuterostome taxon of ambulacrarians, with the echinoderms  
124 *Strongylocentrotus purpuratus* (XP\_003724586.1, XP\_011670169.1, XP\_783155.2) and *Lytechinus*  
125 *variegatus* (JI441084), and the hemichordate *Saccoglossus kowalevskii* (NP\_001161492.1). Surprisingly,  
126 *Strongylocentrotus purpuratus* was the only species for which several different Wif1 sequences were  
127 recovered from databases. No Wif1 ortholog was found in cephalochordates, urochordates, cyclostomes,  
128 platyhelminths or nematodes.

129  
130 **Predicted O-fucosylation and N-glycosylation sites are evolutionarily conserved in deuterostomes.**

131 The consensus sequence C<sup>2</sup>X<sub>4</sub>(S/T)C<sup>3</sup> for O-fucosylation was searched for in the complete and partial  
132 Wif1 sequences (Fig. 2a). It was found in EGF-III and EGF-V for nearly all gnathostomes, except in the  
133 nine-banded armadillo *Dasyurus novemcinctus*, the Eurasian shrew *Sorex araneus* and the squamate *Gekko*



134 *japonicus* where amino acids V for the first and A for the last were present instead of S/T in EGF-V. In  
135 addition, the American pika *Ochonta princeps* and the platypus *Ornithorhynchus anatinus* lacked the  
136 conserved threonine in EGF-III; however, another potential *O*-fucosylation site was found for the platypus  
137 in EGF-I, C<sup>2</sup>RNGGSC<sup>3</sup>. The snakes *Python bivittatus* and *Thamnophis sirtalis* were the only deuterostome  
138 representatives for whom Wif1 was devoid of *O*-fucosylation sites. A single point mutation could account  
139 for the change from T (target of potential *O*-fucosylation) to I. In protostomes, the consensus site was  
140 principally found in EGF-II but never in EGF-V. In lophotrochozoans, it was present in EGF-III and/or  
141 EGF-IV. In endopterygote insects, no consensus site was found and remarkably, EGF-I was absent in the  
142 order of Hymenoptera. Furthermore, in the partial sequences of early emerging deuterostomes, the  
143 ambulacrarians (echinoderms and hemichordates), a consensus site was also found in EGF-II. The most  
144 parsimonious explanation concerning the evolution of *O*-fucosylation sites among the 5 EGF-like domains  
145 of Wif1 is that during bilaterian evolution, EGF-II was the first to contain an *O*-fucosylation motif. In  
146 deuterostomes, after ancestral emergence of ambulacrarians, it was replaced by those of EGF-III and V in  
147 gnathostomes. In protostomes, most had conserved the ancestral situation with the *O*-fucosylation site in  
148 EGF-II, with in some cases a supplementary site in EGF-III or IV for spiralian (mollusks and annelids).  
149 The site in EGF-II would have disappeared and been supplanted by sites in EGF-III and IV for  
150 brachiopods or was never replaced in some hexapods as springtails (*Folsomia*) and endopterygotes  
151 (*Drosophila*, *Tribolium*,...). *Wif1* gene was located on chromosome 10 D2-D3 in *Mus musculus* (Fig. 2b).  
152 It is composed of 10 exons encoding a 379 aa protein. The 5 EGF-like repeats are encoded by a 96-bp  
153 exon each. This organization was probably ancestral in bilaterians since it was also found in gnathostomes  
154 (*Homo sapiens*, *Ornithorhynchus anatinus*, *Taeniopygia guttata*, *Xenopus tropicalis*, *Danio rerio*),  
155 hemichordates (with the partial sequence of *Saccoglossus kowalewskii*) and annelids (*Capitella teleta*).  
156 Genomic sequences of protostomes such as *Acyrtosiphon pisum* and *Drosophila melanogaster* showed a  
157 single exon for EGF-I and II, and for EGF-IV and V. Interestingly, the only splicing sites conserved in  
158 bilaterians were those bordering the exon encoding EGF-III. The potential *O*-fucosylation site in EGF-III  
159 of *Mus musculus* WIF1 was C<sup>2</sup>FNGGTC<sup>3</sup> and corresponded to the consensus site found in gnathostomes.  
160 The second site, present in EGF-V, was C<sup>2</sup>GAHGTC<sup>3</sup> but it was less conserved in gnathostomes. When  
161 EGF-III sequences were compared between gnathostomes and protostomes (Fig. 2b), 12 aa were  
162 significantly different: T<sup>255</sup> in deuterostomes vs K in protostomes but also some amino acids present  
163 between C<sup>1</sup>-C<sup>2</sup>, C<sup>3</sup>-C<sup>4</sup> and C<sup>4</sup>-C<sup>5</sup>. Ten aa were different concerning EGF-V, mostly present between C<sup>1</sup>-C<sup>2</sup>,  
164 C<sup>3</sup>-C<sup>4</sup> and on the last 7 positions of the EGF-like. The *N*-glycosylation site homologous to *Mus* N<sup>88</sup> was  
165 highly conserved in bilaterians. Only *Capitella teleta*, *Saccoglossus kowalewskii* and *Strongylocentrotus*  
166 *purpuratus* XP\_783155.2 Wif1 did not share this one. All gnathostomes possessed the two potential  
167 *N*-glycosylation sites homologous to mouse N<sup>88</sup> and N<sup>245</sup>. In protostomes, one (for *Lingula anatina*) to six



168 (for *Drosophila melanogaster*) sites were found. This higher number in *Drosophila* could be linked to the  
169 protein size (456 aa) and the absence of *O*-fucosylation sites.

170  
171 ***N*-glycans bind to conserved *N*-glycosylation consensus sites of mouse WIF1.** A consensus sequence  
172 for potential *N*-glycosylation is found in WD (N<sup>88</sup>FT) and in EGF-III (N<sup>245</sup>CS) of mouse WIF1, as is the  
173 case for humans<sup>19</sup>. However, using NetNGlyc 1.0 Server, only N<sup>88</sup> was predicted to be *N*-glycosylated. To  
174 reveal the presence of *N*-glycans, wild-type, single (N88Q and N245Q) and double (N88/245Q) mutant  
175 WIF1 proteins with a C-terminal V5 polyhistidine tag were produced in stably transfected Flp-In<sup>TM</sup> CHO  
176 cells and recovered from supernatants of culture media. Using an anti-V5 antibody (Fig. 3), WT WIF1  
177 appeared at a higher apparent molecular weight (53 kDa) than the theoretical calculation based on amino  
178 acid sequence (43,667 Da), suggesting the presence of posttranslational modifications such as *N*- and  
179 *O*-glycosylations. Partial digestion of WT WIF1 by PNGase F led to the appearance of two lower bands at  
180 50 and 47 kDa, consistent with specific removal of one (50 kDa) and two *N*-glycans (47 kDa),  
181 respectively (Fig. 3a). Total digestion of WT and single mutant WIF1 proteins agreed with the presence of  
182 *N*-glycans on both N<sup>88</sup> and N<sup>245</sup>, with a 3-kDa shift per removed *N*-glycan (Fig. 3b). Higher MW than the  
183 theoretical one for double mutant and the totally de-*N*-glycosylated WT and single mutant WIF1 proteins  
184 (47 kDa) could be explained by the presence of potential *O*-fucosylglycans and mostly by the presence of  
185 mucin-type *O*-glycans as predicted by NetOglyc 4.0 server<sup>48</sup> and shown by mass spectrometry and lectin  
186 blot (supplementary Fig.3).

187  
188 **Only WIF1 EGF-III is *O*-fucosylated by POFUT1.** To determine the *O*-fucosylation status of WIF1, we  
189 chose complementary approaches based on click chemistry and MRM-MS after incubation of WT and  
190 mutated EGF-like domains with recombinant secreted POFUT1, purified from stably transfected CHO  
191 cells. First, isolated WT EGF-III and V as well as their mutant counterparts (T255A, T319A) were  
192 produced in *E.coli* and purified. Mouse NOTCH1 EGF26 (N1-EGF26), known to be *O*-fucosylated<sup>28,49</sup>,  
193 was also produced and used as a positive control. The different EGFs were incubated with POFUT1 and  
194 GDP-azido-fucose. Then, the CuAAC reaction was used to detect azido *O*-fucosylated EGFs. The azido-  
195 fucose transfer was only observed on WT N1-EGF26 and WT EGF-III (Fig. 4a). A second slight signal  
196 was detected near 25 kDa that might correspond to the presence of EGF dimers. No signal was detected  
197 for WT EGF-V and all T/A mutant EGFs, although the same protein quantities were loaded, as shown by  
198 Coomassie blue-stained gels (Fig. 4a). Slight differences in apparent MW were observed between EGFs,  
199 which could be attributed to differences in composition of charged amino acids. Indeed, full-scan LC-MS  
200 analyses were performed and measurements of deconvoluted MW of undigested EGF-III and V were  
201 correlated with EGFs lacking signal peptides (Supplementary Fig. 4 and 5). For MRM-MS analyses, the



202 same amounts of the WT, EGF-III and V mutants, were incubated with POFUT1 and non-modified  
203 GDP-fucose. In all cases, peaks corresponding to non-fucosylated peptides were detected (Fig. 4b, upper  
204 panels) whereas a peak representing the fucosylated peptide was only observed for WT EGF-III (Fig. 4b,  
205 lower panels) with a calculated *O*-fucosylated/non fucosylated peptide ratio of  $2.06 \pm 0.09$  (n=3). We  
206 concluded that POFUT1 was able to transfer fucose to WIF1 EGF-III as it did for the control N1-EGF26,  
207 but it was unable to *O*-fucosylate WIF1 EGF-V despite the presence of a conserved *O*-fucosylation  
208 consensus motif.

209  
210 **Full-length WIF1 is *O*-fucosylated in CHO cells.** Full-length WT WIF1 and its counterparts mutated on  
211 *O*-fucosylation sites (T255A, T319A, T255/319A) were produced with a C-terminal V5 polyhistidine tag  
212 and purified from stably transfected CHO cells. Using the same CuAAC technique as above, we first  
213 determined the predisposition of these different WIF1 proteins to be *O*-fucosylated *in vitro*, assuming that  
214 all overexpressed WIF1 proteins were not *O*-fucosylated by endogenous CHO POFUT1. Indeed, specific  
215 signals corresponding to azido-fucose incorporation were only detected for WT WIF1 and T319A mutant  
216 proteins (Fig. 5a). The signal at 53 kDa clearly corresponded to *O*-fucosylated WIF1. A second signal at  
217 49 kDa most likely corresponded to POFUT1, which was also detected. For the same quantities used,  
218 T255A and T255/319A mutant WIF1 proteins were not revealed. Furthermore, we checked that the same  
219 protein quantities, loaded on a silver-stained gel, showed comparable purity for the four purified WIF1  
220 proteins (Fig. 5b). Consistent with previous results obtained from isolated EGFs, full-length WIF1  
221 produced *in cellulo* was only *O*-fucosylable on its EGF-III by recombinant mouse POFUT1. Therefore,  
222 the presence of the *N*-glycan on N<sup>245</sup> did not hinder the action of POFUT1. However, these results  
223 provided no evidence that WIF1 was in fact endogenously *O*-fucosylated by CHO cells. To answer this  
224 question, recombinant mouse lunatic FRINGE (LFNG) was produced and purified from stably transfected  
225 CHO cells. Lunatic Fringe is known to extend *O*-linked fucose attached to an EGF-like domain by  
226 transferring a beta-D-GlcNAc residue from UDP-D-GlcNAc<sup>50</sup>. WT and mutant WIF1 proteins were  
227 incubated with recombinant LFNG and UDP-azido-GlcNAc. Specific signals were detected at expected  
228 MW for WT and T319A WIF1 (Fig. 5c), showing successful LFNG-mediated addition of azido-GlcNAc  
229 to *O*-fucosylated T<sup>255</sup> of WIF1 produced in CHO cells. A second signal at 37 kDa most likely  
230 corresponded to the recombinant LFNG enzyme. In CuAAC experiments, recombinant POFUT1 or LFNG  
231 was also revealed by HRP-streptavidin, probably due to the presence of azido nucleotide sugar still  
232 associated with the enzyme. All these results provided evidence that recombinant mouse WIF1 was  
233 *O*-fucosylated on EGF-III at T<sup>255</sup> site and not at T<sup>319</sup>.

234



235 **Optimal secretion of WIF1 requires both *N*-glycans and *O*-fucose.** The role of *N*-glycans in trafficking  
236 and secretion of glycoproteins is well documented<sup>51-54</sup>. We explored how the loss of *N*-glycosylation sites  
237 could impact WIF1 secretion. In the same way, the consequences of the loss of *O*-fucosylation sites were  
238 examined. Relative quantifications by Western Blot were performed on WT and mutated WIF1 proteins  
239 expressed by CHO cells and recovered from pellets and supernatants of culture media to discriminate  
240 between secretion modifications and differences in protein expression.

241 A significant decrease in quantity (more than 50 %) was observed for the secretion of single mutant  
242 proteins (N88Q, N245Q and T255A) in the supernatants relative to WT WIF1 (Fig. 6a). A more  
243 pronounced decrease, approximately 80%, for the double mutant N88/245Q suggested an additive effect  
244 of the loss of both *N*-glycans. However, secretion of double mutant T255/319A was not statistically  
245 different from that of the single mutant T255A, consistent with occupancy of the *O*-fucosylation site T<sup>255</sup>  
246 only. All these results on culture supernatants were negatively correlated with those obtained for cell  
247 pellets (Fig. 6b), for which WT WIF1 was the lowest retained protein in the intracellular compartment.  
248 We conclude that presence of *N*-glycans and *O*-fucose or *O*-fucosylglycan on WIF1 influences its  
249 intracellular trafficking and secretion.

250  
251 **Steric clash and charge repulsion could explain absence of *O*-fucose on WIF1 EGF-V.** Known X-ray  
252 structures of mouse POFUT1 in complex with EGF26 of NOTCH1<sup>28</sup> were used to give a potential  
253 explanation for unsuccessful *O*-fucosylation of EGF-V, despite the presence of the conserved  
254 *O*-fucosylation motif. Automated homology models were first generated using Swiss-model and X-ray  
255 structure of human WIF1<sup>19</sup> as a reference template. Mouse WIF1 EGF-III and V exhibited 96.88% and  
256 40.63% identity with human WIF1 EGF-III and EGF-I (human EGF-V was not crystallized), respectively.  
257 They were considered as relevant structural models in view of their Global Model Quality Estimation  
258 (GMQE), 0.98 for EGF-III and 0.74 for EGF-V. Other templates proposed for EGF-V with the best scores  
259 of identity (54.84%) and GMQE (0.84) were dismissed because of the presence of an additional residue in  
260 the C<sup>2</sup>-C<sup>3</sup> subdomain. Using MatchMaker of CHIMERA<sup>55</sup>, generated models for EGF-III and V were  
261 superimposed with murine NOTCH1 EGF26 co-crystallized with murine POFUT1<sup>28</sup> (Fig. 7a). The  
262 overall shape of models for EGF-III and V was comparable to that of EGF26, exhibiting correctly  
263 superimposed C<sup>2</sup>-C<sup>3</sup> and C<sup>5</sup>-C<sup>6</sup> subdomains, known to be involved in interactions with POFUT1.  
264 However, more divergence was clearly observed for the flexible hydrophilic loop C<sup>1</sup>-C<sup>2</sup>, known to  
265 “accommodate” numerous different structures. The C<sup>3</sup>-C<sup>4</sup> region, known to be exposed to solvent and not  
266 to be involved in interactions with POFUT1, differed the most because it was 3 residues shorter in  
267 EGF-III or V than murine EGF26. Substitution of EGF26 with WIF1 EGF-III in the deep groove of  
268 POFUT1 between the two Rossmann domains seemed relevant (Fig.7b) since the amino acid sequence in



269 the most important region for interaction with POFUT1, the C<sup>2</sup>-C<sup>3</sup> subdomain, was the same between  
270 EGF-III and EGF26 (C<sup>2</sup>FNGGTC<sup>3</sup>). In addition, other identical or similar amino acids, outside the C<sup>2</sup>-C<sup>3</sup>  
271 subdomain and known to be involved in interactions with POFUT1, were found in the same positions such  
272 as a serine residue of the C<sup>1</sup>-C<sup>2</sup> loop interacting with POFUT1 R<sup>138</sup> and the aliphatic residue in position  
273 C<sup>4</sup>+1 interacting with a group of apolar residues on POFUT1. The interchange between EGF26 and  
274 EGF-V revealed a potential major steric clash between H<sup>317</sup> of C<sup>2</sup>-C<sup>3</sup> subdomain (C<sup>2</sup>GAHGTC<sup>3</sup>) and Y<sup>78</sup>  
275 of POFUT1, with a subsequent binding defect. Tyrosine at the homologous position in other gnathostome  
276 Wif1 could result in the same consequence. Another steric clash and/or potential charge repulsions  
277 between two arginine residues (the R<sup>329</sup> in position C<sup>5</sup>+1 and R<sup>48</sup> of POFUT1) could be also involved. In  
278 addition, residues known to reduce binding affinity<sup>28</sup> such as proline in the C<sup>1</sup>-C<sup>2</sup> loop (strictly conserved  
279 in gnathostomes) and the C<sup>4</sup>+1 glutamine (largely distributed in gnathostomes) were found in EGF-V.  
280 Consequently, T<sup>319</sup> of mouse WIF1 EGF-V could not be *O*-fucosylated, because POFUT1 was unable to  
281 recognize it. Due to some key residue conservation, this could be the same for EGF-V of gnathostome  
282 Wif1 despite the presence of the consensus *O*-fucosylation motif.  
283



284 **DISCUSSION**

285 Murine recombinant WIF1 was shown to be *O*-fucosylated on EGF-III by endogenous POFUT1 when  
286 expressed in CHO cells. The high conservation of the sequence C<sup>2</sup>XNGGIC<sup>3</sup> in gnathostomes suggested  
287 that their EGF-III was certainly prone to be occupied by *O*-fucose, except in rare cases such as in platypus  
288 (*Ornithorhynchus anatinus*) and some snakes for which the *O*-fucosylation site was missing. The C<sup>2</sup>-C<sup>3</sup>  
289 subdomain of EGF-V was more divergent with the following sequence C<sup>2</sup>GX(H/Y)G(S/T)C<sup>3</sup>. H or Y at  
290 C<sup>2</sup>+3 could be responsible for major steric clashes and perturbations of vicinal hydrogen bonds as  
291 previously suggested <sup>28</sup>, leading to a probable inability of POFUT1 to interact with EGF-V regardless of  
292 the considered species of gnathostomes. The arginine in position C<sup>5</sup>+1 was probably involved in another  
293 steric clash and/or charge repulsion in *Mus musculus*. The situation was similar in other gnathostomes, for  
294 which the arginine residue was replaced by lysine, glutamine or histidine (Supplementary Fig. 1).  
295 According to our study, only EGF-III in WIF1 might be a substrate for POFUT1 in the majority of  
296 gnathostomes where the C<sup>2</sup>XNGGIC<sup>3</sup> sequence was found. However, *Ornithorhynchus anatinus* and  
297 ambulacrarians unusually possess an *O*-fucosylation motif in EGF-I and EGF-II respectively, instead of  
298 EGF-III. In platypus and *Saccoglossus kowalevskii*, the sequence of their C<sup>2</sup>-C<sup>3</sup> subdomain  
299 (C<sup>2</sup>XNGG(S/T)C<sup>3</sup>) was very close to the *O*-fucosylable sequence of gnathostome EGF-III, suggesting that  
300 it was a relevant potential target for Pofut1. Automated models suggest that these two EGFs were indeed  
301 prone to be *O*-fucosylated by their own Pofut1 (Supplementary Fig. 6 and 7). If *O*-fucose plays a role in  
302 the biological activity of Wif1, the loss of the *O*-fucosylation site in EGF-III might be offset by the gain of  
303 a new *O*-fucosylation site in EGF-I or EGF-II for these species. Thus, it would be of interest to determine  
304 the impact of a total lack of *O*-fucose, such as in snakes (*Python bivittatus*, *Thamnophis sirtalis*) devoid of  
305 *O*-fucosylation site and in the pika *Ochotona princeps*, where the unique *O*-fucosylation consensus  
306 sequence was found in EGF-V. A large diversity of protostomes only has one potential *O*-fucosylation  
307 motif, located in EGF-II. In lophotrochozoans, an additional site was found in Wif1 EGF-III or IV  
308 (Fig. 2A and supplementary Fig.2), except *Lingula anatina* where *O*-fucosylable EGFs were III and IV.  
309 We focused our attention on three arthropods (*Centruroides sculpturatus*, *Daphnia pulex* and  
310 *Cryptotermes secundus*) and not other protostomes for which Pofut1 was devoid of its ER-retention signal  
311 or for which a *N*-glycan might be present in their EGF-II impairing *O*-fucosylation. These species  
312 exhibited the *O*-fucosylation sequence C<sup>2</sup>MNGG(S/T)C<sup>3</sup>, which was very close to that found in EGF-III  
313 of most gnathostomes. According to structural models, this EGF-II could be *O*-fucosylated by their own  
314 Pofut1 as shown for *Daphnia pulex* (Supplementary Fig. 8), suggesting conservation of Wif1 biological  
315 activities during bilaterian evolution.

316 An EGF-like domain can interact with Pofut1 through its C<sup>2</sup>-C<sup>3</sup> subdomain, by establishing several  
317 hydrogen bonds with specific residues at the bottom of the groove of Pofut1 <sup>28</sup>. If only considering the



318 C<sup>2</sup>-C<sup>3</sup> subdomains, Wif1 EGF-III and V in the mouse and most gnathostomes should *a minima* establish  
319 three important links with mouse POFUT1: a sulphur-hydrogen bond between the EGF-like C<sup>3</sup> residue  
320 and M<sup>46</sup>, a strong hydrogen bond between the C<sup>2</sup>+5 threonine and N<sup>51</sup> and another hydrogen bond between  
321 the C<sup>2</sup>+4 glycine and G<sup>47</sup> <sup>28</sup>. According to their nature, the residue C<sup>2</sup>+1 and/or C<sup>2</sup>+2 can also form  
322 hydrogen bonds with N<sup>151</sup> of POFUT1. In EGF-III, only the asparagine (C<sup>2</sup>+2) can interact with N<sup>151</sup> of  
323 POFUT1; the lack of connection with the residue C<sup>2</sup>+1 has been found in proteins without any  
324 consequences. In EGF-V, the glycine and alanine in positions C<sup>2</sup>+1 and C<sup>2</sup>+2 respectively could be unable  
325 to bind to N<sup>151</sup> of POFUT1. It is likely that hydrogen bonds never occur in *O*-fucosylable EGFs at the  
326 C<sup>2</sup>+3 position, where a glycine is frequently seen and involved in Van der Waals interactions. In  
327 *O*-fucosylable EGF-like domains such as in NOTCH1 EGF12, it was recently shown that an aspartate at  
328 this position (C<sup>2</sup>QNDATC<sup>3</sup>) was responsible for a steric clash leading to a weak interaction with  
329 POFUT1<sup>28</sup>. For EGF-V, we reported here that a more voluminous residue at this position such as histidine  
330 was incriminated in a strong steric clash leading to lack of interaction with mouse POFUT1 rather than a  
331 weakened interaction. Thus, it is tempting to speculate that an amino acid with a bulky side chain at this  
332 position could systematically decrease and even prevent interaction with POFUT1. Outside the C<sup>2</sup>-C<sup>3</sup>  
333 subdomain, other residues can be involved at a lesser extent in binding to POFUT1 and particularly in the  
334 C<sup>5</sup>-C<sup>6</sup> subdomain. Contrary to EGF-III, a positively charged amino acid (R<sup>329</sup>) was found in position C<sup>5</sup>+1  
335 instead of proline. It would participate in the lack of interaction between EGF-V and POFUT1, due to  
336 potential steric clash or charge repulsion with R<sup>48</sup> of POFUT1. In addition, EGF-V contains several amino  
337 acids known to decrease affinity for POFUT1 including proline in the C<sup>1</sup>-C<sup>2</sup> loop and glutamine at  
338 position C<sup>4</sup>+1 <sup>28</sup>. Therefore, the presence of a consensus sequence such as C<sup>2</sup>X4(S/T)C<sup>3</sup> is necessary but  
339 not sufficient for *O*-fucosylation of EGF-like domains. Indeed, POFUT1 is not able to distinguish between  
340 hEGFs having a fucosylable S/T residue and those without one <sup>28</sup>. Thus, the nature of residues in an  
341 EGF-like domain must be considered in the deeply buried interacting C<sup>2</sup>-C<sup>3</sup> subdomain and to a lesser  
342 extent, at C<sup>4</sup>+1 and in C<sup>5</sup>-C<sup>6</sup> subdomain.

343 In this study, we showed that *N*-glycans were involved in secretion of mouse WIF1, as we already  
344 demonstrated for bovine Pofut1 <sup>52</sup>. Similarly, the presence of *O*-fucose on EGF-III added by POFUT1 also  
345 promoted WIF1 secretion. These results are correlated with what is described for *Drosophila* OFut1.  
346 Indeed, OFut1 exerts a Notch chaperone role regardless of its glycosyltransferase activity and was shown  
347 to facilitate Notch folding and trafficking <sup>56</sup>. Thus, it might facilitate secretion of other OFut1 protein  
348 targets. This role has also been previously shown for *O*-fucose added by Pofut2, a paralog of Pofut1,  
349 which acts on thrombospondin type 1 repeats (TSR) in glycoproteins such as ADAMTS proteases <sup>57,58</sup>. In  
350 addition to their role in secretion, the *N*-glycans and *O*-fucose might, together or independently, modulate  
351 the functions of Wif1. They could regulate its interactions with Wnt proteins and/or with HSPGs <sup>19</sup>.



352 Indeed, *O*-linked fucoses widely contribute to protein-protein interactions such as those between Notch  
353 receptors and DSL ligands<sup>40-43</sup>.

354 Our results gave a strong evidence that mouse WIF1 EGF-III is *O*-fucosylated on T<sup>255</sup> unlike EGF-V  
355 which is not in spite of its evolutionarily conserved consensus sequence. Interestingly, we also reported  
356 that this *O*-fucose could be extended with a GlcNAc by LFNG. We wonder if mouse WIF1 can be a  
357 natural target for one or all GlcNAc transferases of the Fringe family and if the *O*-linked GlcNAc-fucose  
358 disaccharide might be extended to form a sialylated *O*-fucosylglycan. The ability of Fringe to reduce the  
359 affinity of Notch for Jagged/Serrate while enhancing that for Delta was previously shown<sup>50,59</sup>. Therefore,  
360 the potential role of extendable *O*-fucoses by Fringe in the modulation of biological activities of WIF1  
361 should be raised, in addition to the role of *O*-fucose itself.

362

## 363 METHODS

364

365 *Phylogenetic reconstruction and sequence conservation analyses.* Wif1 orthologs were retrieved from  
366 GenBank (<https://www.ncbi.nlm.nih.gov>) database using on-site tblastn, blastp and psi-blast facilities,  
367 with mouse WIF1 sequence (NP\_036045.1) as query. The collected homologous sequences (n > 300) were  
368 aligned with MUSCLE<sup>60</sup> implemented in SeaView v.4<sup>61</sup>. Alignment is available upon request. Only  
369 complete sequences covering the maximum taxonomic diversity were selected (n=47) and 301  
370 homologous sites were retained using Gblocks<sup>62</sup>. Phylogenetic analyses were performed with maximum  
371 likelihood (ML) method using PhyML v.3.0<sup>63</sup>, the LG empirical amino acid substitution matrix<sup>64</sup> and  
372 gamma-distribution ( $\Gamma$ ) of among-site rate variation (4 discrete categories)<sup>65</sup> and estimated proportion of  
373 invariant sites, and with Bayesian phylogenetic inference using PhyloBayes v4.1c<sup>66</sup>, LG +  $\Gamma$  evolution  
374 model associated with CAT mixture model<sup>67</sup>, which accounts for across-site heterogeneities in the amino-  
375 acid replacement process. Two independent runs were conducted with a total length of 20,000 cycles.  
376 They were compared to check the convergence of continuous parameters of the models and assess the  
377 convergence in tree space. They satisfactorily converged (maxdiff less than 0.032). The 1,000 initial trees  
378 were discarded as burn-in, and the majority-rule posterior consensus tree was computed from the  
379 remaining sub-sampled trees to collect posterior probabilities. For the ML tree, non-parametric bootstrap  
380 proportions were calculated after 500 replicates. The determination of exon-intron organization of mouse  
381 WIF1 was obtained using ncbi online utility Splign<sup>68</sup> with *Wif1* mRNA (NM\_011915.2) and chromosome  
382 10 DNA (NC\_000076.6). Multiple alignment portions corresponding to potential *N*-glycosylation sites on  
383 mouse EGF-III and V were used to obtain logos for the entire dataset. Subfamily logos<sup>69</sup> were generated  
384 for gnathostome (n=25) and protostome sequences (n=22) by the TEXshade package<sup>70</sup>, as for the Wif1  
385 logos and gnathostome consensus sequence of EGF-III and V.



386 Edition of phylogenetic trees and selection of color portions of the alignment were obtained by MEGA6<sup>71</sup>  
387 and Bioedit v.7.2.5<sup>72</sup>, respectively. Percentages of similarity were calculated online at  
388 [http://www.bioinformatics.org/sms2/ident\\_sim.html](http://www.bioinformatics.org/sms2/ident_sim.html).

389  
390 **Plasmid constructs.** We cloned mouse WIF1 (NP\_036045.1, residues 38-377) cDNA into the  
391 pSecTag/FRT/V5-His-TOPO<sup>R</sup> (pSec vector)(Thermo Fisher Scientific, Waltham, MA, USA), in order to  
392 obtain a secreted protein with a C-terminal V5 polyhistidine tag (V5-His6). WIF1 counterparts mutated  
393 for *N*-glycosylation (N88Q, N245Q, N88/245Q) and *O*-fucosylation (T255A, T319A, T255/319A) sites  
394 were generated using the GENEART Site-Directed Mutagenesis System (Thermo Fisher Scientific).  
395 Mouse POFUT1 (NP\_536711.3, residues 31-389) and beta-1,3-N-acetylglucosaminyltransferase lunatic  
396 Fringe (LFNG) (NP\_032520.1, residues 86-378) cDNAs were both cloned between *KpnI* and *BamHI*  
397 unique restriction sites, into a modified pSec vector, in which a polyhistidine-V5 sequence was directly  
398 fused to the original secretory signal peptide of the vector to produce secreted proteins with an N-terminal  
399 polyhistidine-V5 tag. For isolated WT or mutant (T255A, T319A) EGFs, cDNAs encoding mouse WIF1  
400 amino acids 243-275 (EGF-III) and 307-339 (EGF-V) were cloned between *BamHI* and *XhoI* into the  
401 pET-25b(+) vector (Novagen, Millipore, MA, USA). Similarly, amino-acids 983-1019 (EGF26) of mouse  
402 NOTCH1 (N1) (NP\_032740.3) and its mutant counterpart on its *O*-fucosylation site T997A were cloned  
403 into the pET-25b (+) vector. All the sequences of pET- and pSec-derived constructs were verified, used to  
404 transform BL21 bacteria and to stably transfect mammalian cells, respectively.

405  
406 **Protein expression and purification.** Full-length mouse WIF1, POFUT1 and LFNG were expressed by  
407 stable Flp-In<sup>TM</sup> CHO cells (Thermo Fisher Scientific). After production during 72–96 h in serum-free  
408 Opti-MEM I medium (Thermo Fisher Scientific), proteins were recovered by centrifugation from cell  
409 culture supernatants, concentrated in binding buffer (25 mM Tris-HCl, 500 mM NaCl, 5 mM CaCl<sub>2</sub>, 20  
410 mM imidazole, pH 7.5) and purified on Ni-NTA column by imidazole gradient using AKTA prime system  
411 (GE Healthcare, Piscataway, NJ, USA). Isolated EGFs for mouse WIF1 and NOTCH1 were produced in  
412 BL21 bacteria after 4 h incubation at 37°C in LB broth supplemented with 100 µg.mL<sup>-1</sup> ampicillin and  
413 1 mM IPTG. After lysis with 0.5 mg.mL<sup>-1</sup> lysozyme and sonication, soluble proteins were recovered in  
414 supernatants after centrifugation at 10,000 g for 20 min at 4°C and diluted in binding buffer before being  
415 purified in the same way.

416  
417 **Glycosyltransferase reaction.** For mass spectrometry, 2.5 µg POFUT1 were incubated with 5 µg isolated  
418 EGF and 0.1 mM GDP-fucose in 20 µl of reaction buffer (25 mM Tris, 5 mM CaCl<sub>2</sub>, 10 mM MnCl<sub>2</sub>, pH  
419 7.5) and incubated overnight at 37°C. For click chemistry, glycosyltransferase reactions were carried out



420 in 25  $\mu$ l reaction buffer and incubated at 37°C for 1 h. For *O*-fucosylation reaction, 1  $\mu$ g POFUT1 was  
421 mixed with 2 nmoles GDP-azido-Fucose (R&D Systems Inc., Minneapolis, MN, USA) and 2  $\mu$ g isolated  
422 EGF or 1  $\mu$ g full length WIF1. For *O*-fucose extension, 1  $\mu$ g LFNG was mixed with 2 nmoles UDP-azido-  
423 GlcNAc (R&D Systems Inc.) and 1  $\mu$ g full-length WIF1.

424  
425 ***Click chemistry reactions.*** Copper-assisted azide–alkyne cycloaddition (CuAAC) was performed using  
426 1.25 mM CuCl<sub>2</sub>, 2.5 mM ascorbic acid and 0.125 mM alkynyl biotin (R&D Systems Inc.), directly added  
427 to the glycosyltransferase reaction. The mixture was incubated in the dark for 1 h at room temperature.

428  
429 ***SDS-PAGE and blotting techniques.*** Crude recombinant proteins from culture media were recovered by  
430 centrifugation. Protein from cell pellets were first extracted with RIPA extraction buffer (50 mM  
431 Tris-HCl, 150 mM NaCl, 0.5% sodium deoxycholate, 1% NP-40, 0.1% SDS, pH 8) containing a protease  
432 inhibitor cocktail (Roche Applied Science, Mannheim, Germany). Proteins were quantified using a  
433 bicinchoninic acid (BCA) protein assay (Sigma-Aldrich Corp., St. Louis, MO, USA). When treated with  
434 PNGase F as previously described<sup>73</sup>, crude or purified proteins were separated on 10, 12 or 15%  
435 acrylamide gels (SDS-PAGE) and blotted to nitrocellulose membranes. For Western blot, membranes  
436 were blocked with 5% nonfat milk in TBST (50mM Tris–HCl, 150 mM NaCl, 0.1% Tween-20, pH 7.4)  
437 for 1 h at room temperature and incubated overnight at 4°C with anti-V5-HRP (Thermo Fisher Scientific)  
438 or anti-GAPDH (R&D Systems Inc.) diluted to 1:2000 in 2.5% nonfat milk-TBST. For anti-GAPDH,  
439 membranes were incubated after three washes in TBST for 1 h at room temperature with 1: 2,000 dilution  
440 of secondary HRP conjugate antibodies (Dako, Glostrup, Denmark) in 2.5% nonfat milk-TBST. After  
441 click chemistry, gel separation and transfer, membranes were blocked with 10% nonfat milk-TBST for 10  
442 min, before being incubated with streptavidin-HRP in TBST at 25 ng/ml for 30 min. The membranes were  
443 washed three-times before and after streptavidin-HRP incubation with TBST (15 min per wash). In both  
444 cases, membranes were revealed using enhanced chemiluminescence peroxidase substrate. Signals were  
445 visualized and quantified using Amersham Imager 600 (GE Healthcare).

446  
447 ***Automatic modeling and superimposition with X-ray structures.*** Automatic homology models were  
448 generated for mouse WIF1 EGF-III and EGF-V on the Swiss-model server  
449 (<https://swissmodel.expasy.org>), using the X-ray structure of human WIF1 (PDB 2YGQ)<sup>19</sup> as a reference  
450 template, since both EGF-III and EGF-V shared more than 40% identity with the human protein. Obtained  
451 models were considered as relevant on the basis of the selected template, their Global Model Quality  
452 Estimation (GMQE) score and the presence of the same number of amino-acids in the C<sup>2</sup>-C<sup>3</sup> subdomain  
453 containing the *O*-fucosylation consensus motif. Using MatchMaker of UCSF CHIMERA<sup>55</sup>, mouse



454 N1-EGF26, co-crystallized with murine POFUT1 (PDB 5KY4)<sup>28</sup>, was superimposed with obtained  
455 models for EGF-III and EGF-V. Finally, it was replaced with either EGF-III or EGF-V at the same  
456 location in murine POFUT1 to identify potential steric clashes and charge repulsions.

457  
458 **Mass spectrometry.** Glycosyltransferase reactions were reduced, alkylated and in-solution digested by  
459 trypsin. Resulting peptides were analysed by microLC-MS/MS using a nanoLC 425 in micro-flow mode  
460 (Eksigent, Dublin, CA) system coupled with a TripleTOF 5600+ (Sciex, Framingham, MA). 5µl of each  
461 sample were trapped on C18 Pepmap100 cartridge (300 µm Id x 5 mm, 5µm; Thermo Scientific), and  
462 desalting was carried out at 10µL/min with Loading mobile phase for 5 minutes. Chromatographic  
463 separation was performed on a ChromXP C18 column (150 x 0.3 mm i.d., 120Å, 3 µm; Sciex) at a flow  
464 rate of 3µL/min. The mobile phase was a gradient of water/acetonitrile/formic acid 100/0/0.1% (A) and  
465 5/95/0.1% (B) programed as follows: initial, 5% B, increased to 25% over 20 min, then increased to 95%  
466 B over 2 min, maintained at 95% for 4 min, and finally, decreased to 5% B for reequilibration. The TTOF  
467 5600+ was operated in information-dependent acquisition (IDA) mode with Analyst 1.7TF software  
468 (Sciex). MS and MS/MS data were continuously recorded with up to 20 precursors selected for  
469 fragmentation from each MS survey scan. Precursor selection was based upon ion intensity and whether or  
470 not the precursor had been previously selected for fragmentation (dynamic exclusion). Collision energies  
471 were automatically adjusted to the charge state and  $m/z$  value of the precursor ions. The recombinant  
472 protein sequence database was searched with ProteinPilot 5.0 (SCIEX) (Supplementary Figures 9 and 10),  
473 and the MRM transition list was established using Skyline 3.5.0 (MacCoss Lab, University of  
474 Washington, Seattle, WA, USA) for the WT and mutant non-*O*-fucosylated peptides. *O*-fucosylation  
475 (coded by [dHx] amino acid modification) was added *in silico* at the expected position with PeakView  
476 software (SCIEX) and  $m/z$  of precursor and fragments were calculated. Data were acquired in high-  
477 Resolution MRM mode: product ion scans were collected for the  $m/z$  corresponding to *O*-fucosylated or  
478 non-*O*-fucosylated WT and mutant peptides during 30 min using the same parameters as previously  
479 described in the IDA method. Data were processed with MultiQuant Software 3.0.1 (SCIEX), considering  
480 the six most abundant fragments for each peptide with a resolution of 10,000. The same fragments were  
481 used for non *O*-fucosylated and *O*-fucosylated peptides (Supplementary Tables 1 and 2). Areas were  
482 collected for the same most major fragment of *O*-fucosylated and non-*O*-fucosylated peptides and a ratio  
483 of fucosylation was calculated.

484  
485 **Statistical analysis.** All experiments were performed in biological triplicates and results were reported as  
486 the means ± SEMs. Statistical comparisons were performed using two-tailed *t* tests implemented in Prism,



487 version 5.03 (GraphPad Software, Inc., San Diego, CA, USA). A *p* value of 0.05 or less was considered  
488 statistically significant.

489

## 490 **REFERENCES**

- 491 1 Hsieh, J. C. *et al.* A new secreted protein that binds to Wnt proteins and inhibits their activities.  
492 *Nature* **398**, 431-436,(1999).
- 493 2 Kawano, Y. & Kypta, R. Secreted antagonists of the Wnt signalling pathway. *J Cell Sci* **116**, 2627-  
494 2634,(2003).
- 495 3 Malinauskas, T. & Jones, E. Y. Extracellular modulators of Wnt signalling. *Curr Opin Struct Biol* **29**,  
496 77-84,(2014).
- 497 4 Dequeant, M. L. & Pourquie, O. Segmental patterning of the vertebrate embryonic axis. *Nat Rev*  
498 *Genet* **9**, 370-382,(2008).
- 499 5 Logan, C. Y. & Nusse, R. The Wnt signaling pathway in development and disease. *Annu Rev Cell*  
500 *Dev Biol* **20**, 781-810,(2004).
- 501 6 MacDonald, B. T., Tamai, K. & He, X. Wnt/beta-catenin signaling: components, mechanisms, and  
502 diseases. *Dev Cell* **17**, 9-26,(2009).
- 503 7 Aguilera, O., Munoz, A., Esteller, M. & Fraga, M. F. Epigenetic alterations of the Wnt/beta-  
504 catenin pathway in human disease. *Endocr Metab Immune Disord Drug Targets* **7**, 13-21,(2007).
- 505 8 Ilyas, M. Wnt signalling and the mechanistic basis of tumour development. *J Pathol* **205**, 130-  
506 144,(2005).
- 507 9 Clevers, H. Wnt/beta-catenin signaling in development and disease. *Cell* **127**, 469-480,(2006).
- 508 10 Kansara, M. *et al.* Wnt inhibitory factor 1 is epigenetically silenced in human osteosarcoma, and  
509 targeted disruption accelerates osteosarcomagenesis in mice. *J Clin Invest* **119**, 837-851,(2009).
- 510 11 Urakami, S. *et al.* Epigenetic inactivation of Wnt inhibitory factor-1 plays an important role in  
511 bladder cancer through aberrant canonical Wnt/beta-catenin signaling pathway. *Clin Cancer Res*  
512 **12**, 383-391,(2006).
- 513 12 Wu, J. *et al.* Wnt inhibitory factor-1 regulates glioblastoma cell cycle and proliferation. *J Clin*  
514 *Neurosci* **19**, 1428-1432,(2012).
- 515 13 Yang, Z. *et al.* Downregulation of WIF-1 by hypermethylation in astrocytomas. *Acta Biochim*  
516 *Biophys Sin (Shanghai)* **42**, 418-425,(2010).
- 517 14 He, B. *et al.* Blockade of Wnt-1 signaling induces apoptosis in human colorectal cancer cells  
518 containing downstream mutations. *Oncogene* **24**, 3054-3058,(2005).
- 519 15 Mazieres, J. *et al.* Wnt inhibitory factor-1 is silenced by promoter hypermethylation in human  
520 lung cancer. *Cancer Res* **64**, 4717-4720,(2004).
- 521 16 Moura, R. S., Carvalho-Correia, E., daMota, P. & Correia-Pinto, J. Canonical Wnt signaling activity  
522 in early stages of chick lung development. *PLoS One* **9**, e112388,(2014).
- 523 17 Liepinsh, E., Banyai, L., Patthy, L. & Otting, G. NMR structure of the WIF domain of the human  
524 Wnt-inhibitory factor-1. *J Mol Biol* **357**, 942-950,(2006).
- 525 18 Wouters, M. A. *et al.* Evolution of distinct EGF domains with specific functions. *Protein Sci* **14**,  
526 1091-1103,(2005).
- 527 19 Malinauskas, T., Aricescu, A. R., Lu, W., Siebold, C. & Jones, E. Y. Modular mechanism of Wnt  
528 signaling inhibition by Wnt inhibitory factor 1. *Nat Struct Mol Biol* **18**, 886-893,(2011).
- 529 20 Glise, B. *et al.* Shifted, the *Drosophila* ortholog of Wnt inhibitory factor-1, controls the  
530 distribution and movement of Hedgehog. *Dev Cell* **8**, 255-266,(2005).
- 531 21 Han, C. & Lin, X. Shifted from Wnt to Hedgehog signaling pathways. *Mol Cell* **17**, 321-322,(2005).



532 22 Stanley, P., Taniguchi, N. & Aebi, M. in *Essentials of Glycobiology* (eds rd et al.) 99-111 (2015).

533 23 Lederkremer, G. Z. Glycoprotein folding, quality control and ER-associated degradation. *Curr Opin Struct Biol* **19**, 515-523,(2009).

534 24 Wang, Y. et al. Modification of epidermal growth factor-like repeats with O-fucose. Molecular cloning and expression of a novel GDP-fucose protein O-fucosyltransferase. *J Biol Chem* **276**, 40338-40345,(2001).

535 25 Der Vartanian, A. et al. Protein O-fucosyltransferase 1 expression impacts myogenic C2C12 cell commitment via the Notch signaling pathway. *Mol Cell Biol* **35**, 391-405,(2015).

536 26 Al Jaam, B. et al. Reduced Notch signalling leads to postnatal skeletal muscle hypertrophy in Pofut1cax/cax mice. *Open Biol* **6**,(2016).

541 27 Schneider, M., Al-Shareffi, E. & Haltiwanger, R. S. Biological functions of fucose in mammals. *Glycobiology* **27**, 601-618,(2017).

542 28 Li, Z. et al. Recognition of EGF-like domains by the Notch-modifying O-fucosyltransferase POFUT1. *Nat Chem Biol* **13**, 757-763,(2017).

543 29 Kentzer, E. J., Buko, A., Menon, G. & Sarin, V. K. Carbohydrate composition and presence of a fucose-protein linkage in recombinant human pro-urokinase. *Biochem Biophys Res Commun* **171**, 401-406,(1990).

544 30 Harris, R. J., Ling, V. T. & Spellman, M. W. O-linked fucose is present in the first epidermal growth factor domain of factor XII but not protein C. *J Biol Chem* **267**, 5102-5107,(1992).

545 31 Nishimura, H., Takao, T., Hase, S., Shimonishi, Y. & Iwanaga, S. Human factor IX has a tetrasaccharide O-glycosidically linked to serine 61 through the fucose residue. *J Biol Chem* **267**, 17520-17525,(1992).

546 32 Bjoern, S. et al. Human plasma and recombinant factor VII. Characterization of O-glycosylations at serine residues 52 and 60 and effects of site-directed mutagenesis of serine 52 to alanine. *J Biol Chem* **266**, 11051-11057,(1991).

547 33 Schiffer, S. G. et al. Fucosylation of Cripto is required for its ability to facilitate nodal signaling. *J Biol Chem* **276**, 37769-37778,(2001).

548 34 Kim, M. L. et al. O-fucosylation of muscle agrin determines its ability to cluster acetylcholine receptors. *Mol Cell Neurosci* **39**, 452-464,(2008).

549 35 Gebauer, J. M., Muller, S., Hanisch, F. G., Paulsson, M. & Wagener, R. O-glucosylation and O-fucosylation occur together in close proximity on the first epidermal growth factor repeat of AMACO (VWA2 protein). *J Biol Chem* **283**, 17846-17854,(2008).

550 36 Loriol, C. et al. Molecular evolution of protein O-fucosyltransferase genes and splice variants. *Glycobiology* **16**, 736-747,(2006).

551 37 Lira-Navarrete, E. et al. Structural insights into the mechanism of protein O-fucosylation. *PLoS One* **6**, e25365,(2011).

552 38 Rampal, R., Arboleda-Velasquez, J. F., Nita-Lazar, A., Kosik, K. S. & Haltiwanger, R. S. Highly conserved O-fucose sites have distinct effects on Notch1 function. *J Biol Chem* **280**, 32133-32140,(2005).

553 39 Moloney, D. J. et al. Mammalian Notch1 is modified with two unusual forms of O-linked glycosylation found on epidermal growth factor-like modules. *J Biol Chem* **275**, 9604-9611,(2000).

554 40 Panin, V. M., Papayannopoulos, V., Wilson, R. & Irvine, K. D. Fringe modulates Notch-ligand interactions. *Nature* **387**, 908-912,(1997).

555 41 Rana, N. A. & Haltiwanger, R. S. Fringe benefits: functional and structural impacts of O-glycosylation on the extracellular domain of Notch receptors. *Curr Opin Struct Biol* **21**, 583-589,(2011).



579 42 Haltiwanger, R. S. & Stanley, P. Modulation of receptor signaling by glycosylation: fringe is an O-  
580 fucose-beta1,3-N-acetylglucosaminyltransferase. *Biochim Biophys Acta* **1573**, 328-335,(2002).

581 43 Taylor, P. *et al.* Fringe-mediated extension of O-linked fucose in the ligand-binding region of  
582 Notch1 increases binding to mammalian Notch ligands. *Proc Natl Acad Sci U S A* **111**, 7290-  
583 7295,(2014).

584 44 Kakuda, S. & Haltiwanger, R. S. Deciphering the Fringe-Mediated Notch Code: Identification of  
585 Activating and Inhibiting Sites Allowing Discrimination between Ligands. *Dev Cell* **40**, 193-  
586 201,(2017).

587 45 Telford, M. J., Budd, G. E. & Philippe, H. Phylogenomic Insights into Animal Evolution. *Curr Biol*  
588 **25**, R876-887,(2015).

589 46 Lartillot, N., Brinkmann, H. & Philippe, H. Suppression of long-branch attraction artefacts in the  
590 animal phylogeny using a site-heterogeneous model. *BMC Evol Biol* **7 Suppl 1**, S4,(2007).

591 47 Sanchez-Hernandez, D., Sierra, J., Ortigao-Farias, J. R. & Guerrero, I. The WIF domain of the  
592 human and Drosophila Wif-1 secreted factors confers specificity for Wnt or Hedgehog.  
593 *Development* **139**, 3849-3858,(2012).

594 48 Steentoft, C. *et al.* Precision mapping of the human O-GalNAc glycoproteome through SimpleCell  
595 technology. *EMBO J* **32**, 1478-1488,(2013).

596 49 Shao, L., Moloney, D. J. & Haltiwanger, R. Fringe modifies O-fucose on mouse Notch1 at  
597 epidermal growth factor-like repeats within the ligand-binding site and the Abruptex region. *J*  
598 *Biol Chem* **278**, 7775-7782,(2003).

599 50 Moloney, D. J. *et al.* Fringe is a glycosyltransferase that modifies Notch. *Nature* **406**, 369-  
600 375,(2000).

601 51 Wei, W. *et al.* Missense mutations near the N-glycosylation site of the A2 domain lead to various  
602 intracellular trafficking defects in coagulation factor VIII. *Sci Rep* **7**, 45033,(2017).

603 52 Loriol, C., Audfray, A., Dupuy, F., Germot, A. & Maftah, A. The two N-glycans present on bovine  
604 Pofut1 are differently involved in its solubility and activity. *FEBS J* **274**, 1202-1211,(2007).

605 53 Molinari, M. N-glycan structure dictates extension of protein folding or onset of disposal. *Nat*  
606 *Chem Biol* **3**, 313-320,(2007).

607 54 Lerouge, P. *et al.* N-glycoprotein biosynthesis in plants: recent developments and future trends.  
608 *Plant Mol Biol* **38**, 31-48,(1998).

609 55 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis.  
610 *J Comput Chem* **25**, 1605-1612,(2004).

611 56 Okajima, T., Xu, A., Lei, L. & Irvine, K. D. Chaperone activity of protein O-fucosyltransferase 1  
612 promotes notch receptor folding. *Science* **307**, 1599-1603,(2005).

613 57 Ricketts, L. M., Dlugosz, M., Luther, K. B., Haltiwanger, R. S. & Majerus, E. M. O-fucosylation is  
614 required for ADAMTS13 secretion. *J Biol Chem* **282**, 17014-17023,(2007).

615 58 Wang, L. W. *et al.* O-fucosylation of thrombospondin type 1 repeats in ADAMTS-like-1/punctin-1  
616 regulates secretion: implications for the ADAMTS superfamily. *J Biol Chem* **282**, 17024-  
617 17031,(2007).

618 59 Bruckner, K., Perez, L., Clausen, H. & Cohen, S. Glycosyltransferase activity of Fringe modulates  
619 Notch-Delta interactions. *Nature* **406**, 411-415,(2000).

620 60 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput.  
621 *Nucleic Acids Res* **32**, 1792-1797,(2004).

622 61 Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface  
623 for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**, 221-224,(2010).

624 62 Castresana, J. Selection of conserved blocks from multiple alignments for their use in  
625 phylogenetic analysis. *Mol Biol Evol* **17**, 540-552,(2000).



- 626 63 Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies:  
627 assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-321,(2010).
- 628 64 Le, S. Q. & Gascuel, O. An improved general amino acid replacement matrix. *Mol Biol Evol* **25**,  
629 1307-1320,(2008).
- 630 65 Yang, Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates  
631 over sites: approximate methods. *J Mol Evol* **39**, 306-314,(1994).
- 632 66 Lartillot, N., Lepage, T. & Blanquart, S. PhyloBayes 3: a Bayesian software package for  
633 phylogenetic reconstruction and molecular dating. *Bioinformatics* **25**, 2286-2288,(2009).
- 634 67 Lartillot, N. & Philippe, H. A Bayesian mixture model for across-site heterogeneities in the amino-  
635 acid replacement process. *Mol Biol Evol* **21**, 1095-1109,(2004).
- 636 68 Kapustin, Y., Souvorov, A., Tatusova, T. & Lipman, D. Splign: algorithms for computing spliced  
637 alignments with identification of paralogs. *Biol Direct* **3**, 20,(2008).
- 638 69 Beitz, E. Subfamily logos: visualization of sequence deviations at alignment positions with high  
639 information content. *BMC Bioinformatics* **7**, 313,(2006).
- 640 70 Beitz, E. TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon.  
641 *Bioinformatics* **16**, 135-139,(2000).
- 642 71 Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary  
643 Genetics Analysis version 6.0. *Mol Biol Evol* **30**, 2725-2729,(2013).
- 644 72 Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for  
645 Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* **41**, 95-98,(1999).
- 646 73 Brun, C., Monestier, O., Legardinier, S., Maftah, A. & Blanquet, V. Murine GASP-1 N-glycosylation  
647 is not essential for its activity on C2C12 myogenic cells but alters its secretion. *Cell Physiol*  
648 *Biochem* **30**, 791-804,(2012).

649

#### 650 **Author contributions statement**

651 AG, AM, FP and SL conceived and designed the experiments. FP and SL performed molecular biology,  
652 biochemical experiments and homology modeling. AG initiated the project and performed molecular  
653 comparisons and phylogenetic analyses. BAJ participated in plasmid construction, sequencing and stable  
654 cell line generation. EP performed mass spectrometry analyses. CEB contributed to molecular cloning and  
655 preliminary phylogenetic studies. AG, FP and SL drafted the manuscript. AM, BAJ and CEB critically  
656 revised the manuscript. All the authors have corrected and approved the manuscript for publication.

657

#### 658 **Competing interests**

659 The authors declare no competing interests.

660

#### 661 **Acknowledgements**



662 This work was supported by a French Ministry of Higher Education and Research doctoral fellowship to  
663 FP. We thank Dr Fabrice Dupuy for his help in site-directed mutagenesis. We also gratefully thank  
664 Pr Willy Morelle for MALDI analyses of WIF1 and Sarah Hanache, Siham Hedir and Aurélie Lagarde for  
665 their past contributions during their respective internships. We thank Dr J Cook-Moreau, CNRS  
666 UMR7276, INSERM U1262, Limoges, France, for English editing. The work was funded by the  
667 GlyCanColor projet within the CORC program.

## 668 **Figure legends**

669 **Figure 1: Phylogenetic tree of bilaterians based on Wif1 comparison.** The tree was reconstructed from  
670 47 species and 301 aligned positions using the PhyML method with LG+ $\Gamma$  (4 rates) evolution model. The  
671 best ML tree had a log-likelihood (LnL) value of -10048.71 and the estimated value of  $\alpha$  shape parameter  
672 of the discrete  $\Gamma$  distribution was 1.35. The tree was drawn to scale and mid-point rooted. The scale bar  
673 represents the number of substitutions per site. Non-parametric bootstrap percentages from ML analysis  
674 (500 replicates) appeared for nodes when  $\geq 50\%$ . Filled circles indicated nodes with estimated posterior  
675 probabilities  $> 0.95$  in Bayesian inference. Genbank accession numbers are indicated after the species  
676 name. Branch colors represent monophyletic groups or species which differed in the number and position  
677 of Wif1 potential *O*-fucosylation sites.

678  
679 **Figure 2: Conservation of potential Wif1 *O*-fucosylation sites in bilaterians. (A)** Presence of the  
680 consensus sequence, C<sup>2</sup>X4(S/T)C<sup>3</sup>, for *O*-fucosylation in deuterostomes and protostomes is indicated by a  
681 black dot. The different font colors of species correspond to those of branches shown in figure 1. Group  
682 and species in black font were not included in the phylogenetic reconstruction. **(B)** Gene structure of *Mus*  
683 *musculus* *Wif1* on chromosome 10 with above numbers corresponding to exon sizes (E1 to E10) and  
684 below, to intron sizes. Only exon 10, partially bordered by a dotted line, was not drawn to scale. The  
685 protein sequence contained a signal peptide, a Wif domain (WD), five EGF-like (I to V) regions and a  
686 C-terminal hydrophilic domain. WD is also found in metazoan Ryk receptor tyrosine kinases. Fill colors  
687 of the different domains are indicated on their corresponding coding exons. Consensus sequence of  
688 gnathostomes for predictive *N*-glycosylation (N<sup>88</sup>, N<sup>245</sup>) and *O*-fucosylation (T<sup>255</sup>, T<sup>319</sup>) sites, homologous  
689 to those present in mouse WIF1, are distinguished according to a HotCold color variation and in uppercase  
690 for identity  $\geq 80\%$ , in lowercase, between 50 and 80 %, and with a dot, when  $< 50\%$ . Logos for Wif1  
691 orthologs and subfamily logos comparing gnathostomes to protostomes were created. Numbering is  
692 according to the mouse WIF1 sequence. The height of the letters represents the amino acid relative



693 frequency at each position. Sub-family logos display relevant deviations (\*) of a sub-family compared to  
694 the other.

695  
696 **Figure 3: N-glycosylation of recombinant mouse WIF1.** Immunoblots with anti-V5-HRP antibody for  
697 recombinant WT and mutant WIF1 proteins, untreated or digested with endoglycosidase F (or PNGase F)  
698 (A) Full-length WT WIF1, secreted by stable CHO cells, was partially digested by PNGase F to highlight  
699 the presence of its two N-glycans (B) WIF1 WT, single (N88Q and N245Q) and double (N88/245Q)  
700 mutant WIF1 proteins were digested with 1U PNGase F and compared to untreated samples by Western  
701 blot.

702  
703 **Figure 4: O-fucosylation of recombinant isolated EGF-III and EGF-V from mouse WIF1.** Isolated  
704 WT EGF-III and EGF-V, as well as their mutant counterparts (T255A and T319A), were produced and  
705 purified from *E. coli*. Mouse NOTCH1 EGF26 (N1-EGF26) was used as a positive control. (A) Proteins  
706 were first incubated with mouse POFUT1 and GDP-azido-fucose. After CuAAC with alkynyl biotin,  
707 EGFs were revealed by blotting technique using HRP-streptavidin. Positive signals correspond to  
708 successful *in vitro* POFUT1-mediated O-fucosylation to EGFs (upper panel), for which quantity and  
709 purity were checked by Coomassie blue-stained polyacrylamide gels (lower panel). (B) After incubation  
710 of WT and mutated EGFs with mouse POFUT1 and unlabeled GDP-fucose, reactions were reduced,  
711 alkylated and digested by trypsin and resulting peptides analyzed by MRM-MS. Peaks correspond to the  
712 presence of non-fucosylated peptides, which harbor the sequence of the C<sup>2</sup>-C<sup>3</sup> subdomain with T or A at  
713 position C<sup>2</sup>+5 (upper panels). A peak corresponding to the presence of O-fucosylated peptides (lower  
714 panels) was seen only for WT EGF-III.

715  
716 **Figure 5: O-fucosylation of recombinant full-length mouse WIF1.** Full-length WT WIF1 and its  
717 counterparts mutated on O-fucosylation sites (T255A, T319A, T255/T319A) were produced and purified  
718 from stably transfected CHO cell lines. Using the same CuAAC technique as for isolated EGFs, proteins  
719 incubated *in vitro* with POFUT1 and GDP-azido-fucose were analyzed by blotting technique. (A) Positive  
720 signals correspond to successful *in vitro* POFUT1-mediated O-fucosylation of full-length WIF1 proteins,  
721 for which quantity and purity were checked by silver-stained polyacrylamide gels. (B) Similarly, the same  
722 proteins were incubated *in vitro* with mouse LFNG and UDP-azido-GlcNAc before CuAAC and blotting  
723 technique. Positive signals indicate successful cycloaddition of alkynyl biotin to azido-GlcNAc linked to  
724 O-fucose of full-length WIF1 proteins by LFNG.

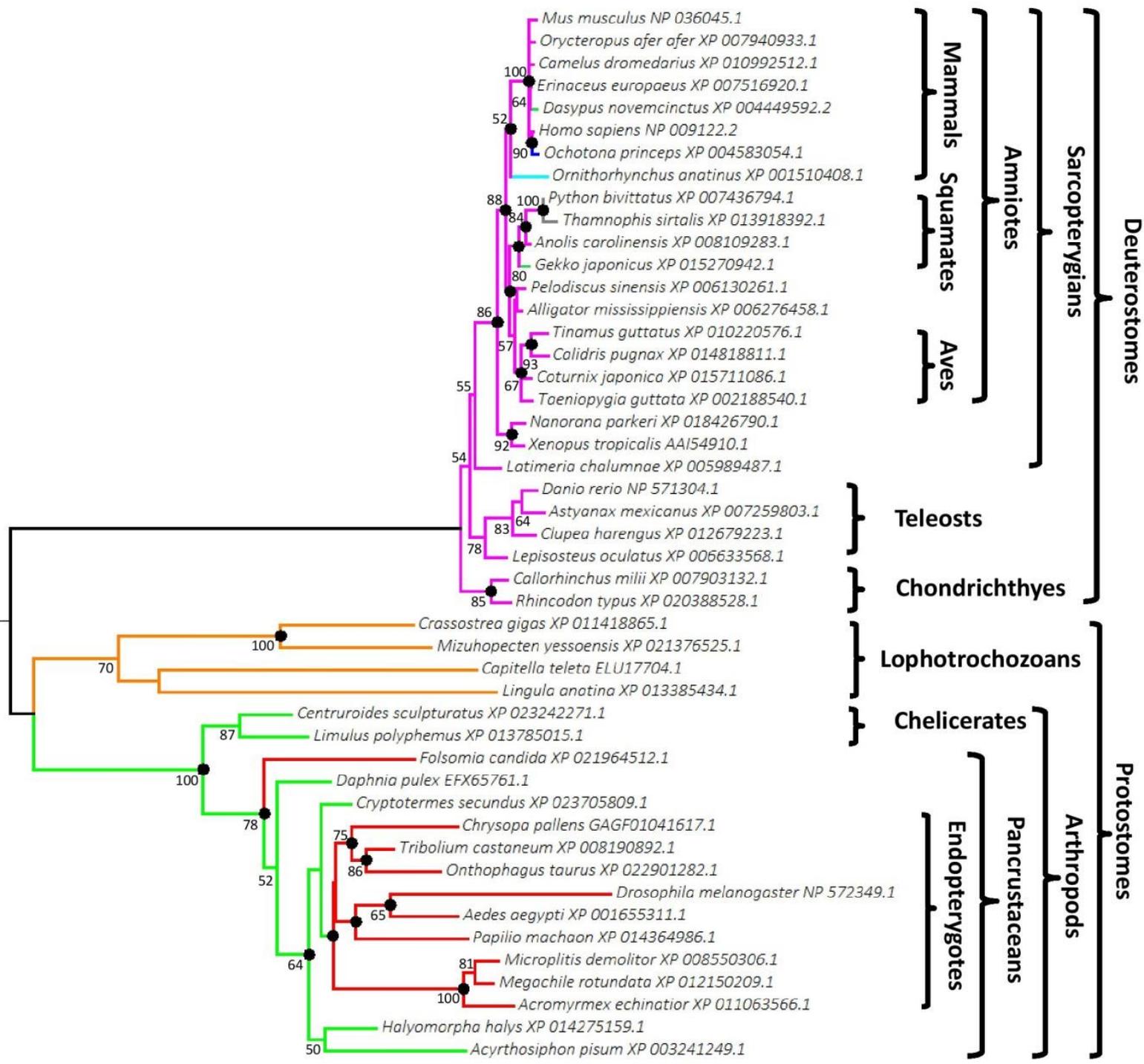
725



726 **Figure 6: Secretion of recombinant full-length WT WIF1 protein and its counterparts mutated at**  
727 **the *N*-glycosylation and *O*-fucosylation consensus sites. (A)** Cell culture supernatants and **(B)** cell  
728 pellets from stable Flp-In<sup>TM</sup> CHO cells were analyzed by Western blot using anti-V5-HRP antibody to  
729 reveal WT WIF1 and WIF1 proteins mutated on sites for *N*-glycosylation (N88Q, N245Q and N88/245Q)  
730 and *O*-fucosylation (T245A, T319A and T255/319A). Histograms represent **(A)** either the % secretion of  
731 mutant proteins compared to WT WIF1 or **(B)** the % protein retention in cells compared to WT WIF1. V5  
732 to GAPDH signal ratios were used for protein retention in cell pellets. Mean  $\pm$  SEM (n>3) are shown  
733 (t-test two-tailed, \*: p<0.05, \*\*: p<0.01, \*\*\*: p<0.001).

734  
735 **Figure 7: Automated homology models for mouse WIF1 EGF-III and EGF-V and their interaction**  
736 **with mouse POFUT1.** Homology models were generated by using Swiss-model Server for mouse WIF1  
737 EGF-III (blue) and EGF-V (red) using X-ray structure of human WIF1 (PDB 2YGQ) as a reference  
738 template. **(A)** Using Matchmaker of CHIMERA, these models were superimposed with mouse NOTCH1  
739 EGF26 (pink) alone (*left*) or in complex with mouse POFUT1 (PDB 5KY4) and GDP-fucose  
740 (black)(*right*). The comparison of these three structures (*left*), partly based on the conservation of the three  
741 disulfide bonds (C<sup>1</sup>-C<sup>3</sup>, C<sup>2</sup>-C<sup>4</sup>, C<sup>5</sup>-C<sup>6</sup>)(yellow), shows similar C<sup>2</sup>-C<sup>3</sup> and C<sup>5</sup>-C<sup>6</sup> subdomains and a more  
742 divergent C<sup>1</sup>-C<sup>2</sup> loop. The threonine side chain of the *O*-fucosylation consensus motifs is shown in cyan.  
743 The inset shows the zoomed interaction region, with Pofut1 residues in green interacting with residues of  
744 the EGF C<sup>2</sup>-C<sup>3</sup> subdomains (*right*). **(B)** Substitution of NOTCH1 EGF26 either by EGF-III (*left*) or  
745 EGF-V (*right*) in mouse POFUT1 (PDB 5KY4) is shown. Hydrogen and sulphur-hydrogen bonds  
746 between key residues (green) of mouse Pofut1<sup>28</sup> and those of the C<sup>2</sup>-C<sup>3</sup> subdomain might be unaffected in  
747 EGF-III since this EGF is *O*-fucosylated. However, steric clash and/or charge repulsion between EGF-V  
748 and residues in orange (R<sup>48</sup>, Y<sup>78</sup>) could explain the inability of Pofut1 to add *O*-fucose to this EGF.





0.2



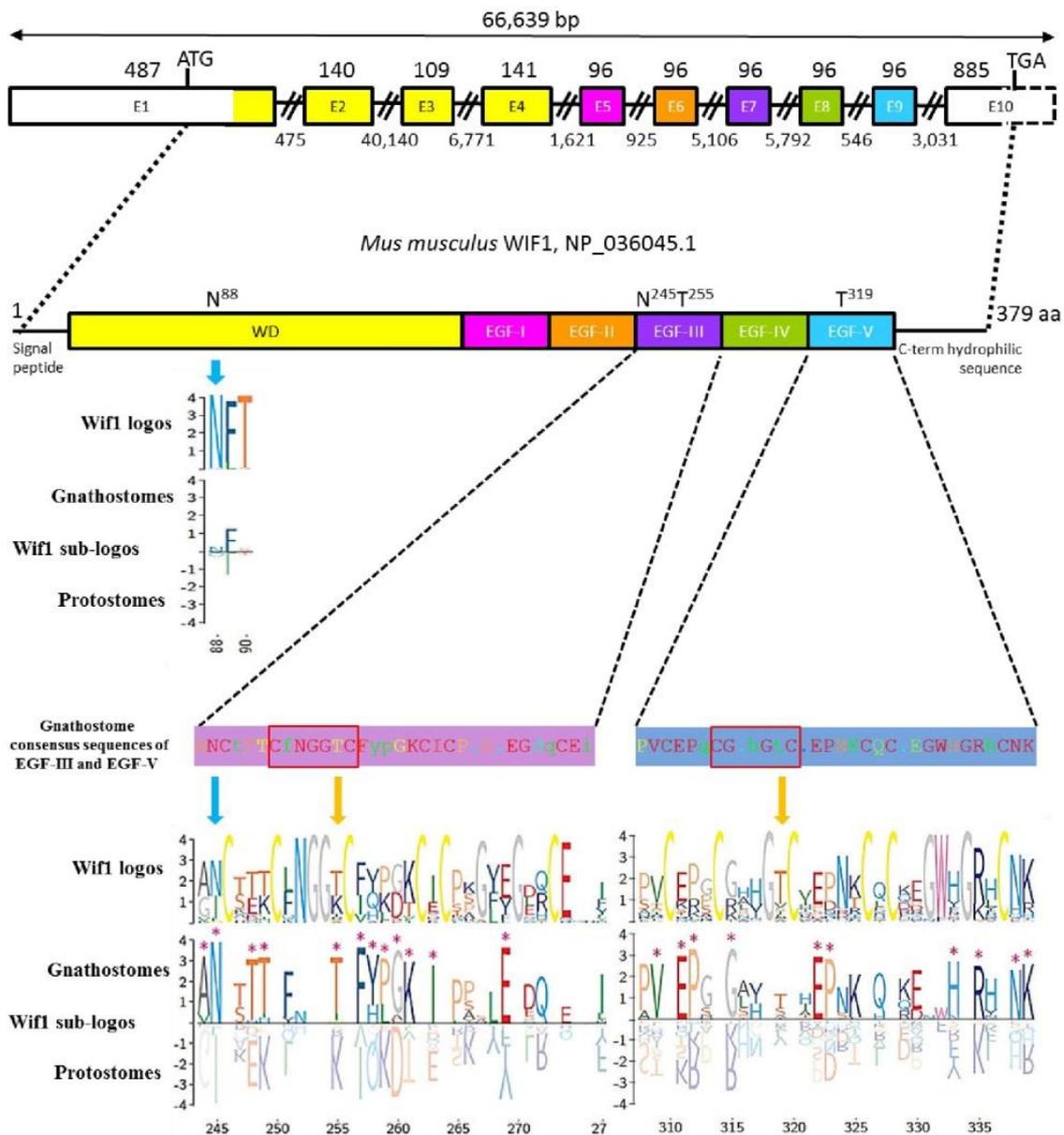
A.

Taxons		EGF-like				
		I	II	III	IV	V
Deuterostomes	Deuterostomes, except:			•		•
	<i>Dasyypus novemcinctus</i> , <i>Gekko japonicus</i> , <i>Sorex araneus</i> XP_004602698.1			•		
	<i>Ochotona princeps</i>					•
	<i>Ornithorhynchus anatinus</i>	•				•
	<i>Python bivittatus</i> , <i>Thamnophis sirtalis</i>					
Protostomes	Ambulacrarians (partial sequences) <i>S. kowalevskii</i> NP_001161492.1, <i>L. variegatus</i> JI441084, <i>S. purpuratus</i> XP_003724586.1, XP_011670169.1, XP_783155.2		•		?	?
	Lophotrochozoans		•		•	(a)
	<i>Capitella teleta</i> , <i>Crassostrea gigas</i> , <i>Crassostrea virginica</i> XP_022296460.1		•			
	<i>Platynereis dumerilii</i> ANS60425.1		•	•		
	<i>Mizuhopecten yessoensis</i>		•			
	<i>Lingula anatina</i>			•	•	
	Arthropods		•			
Chelicerates						
Hexapod non insect <i>Folsomia candida</i>						
Crustacean <i>Daphnia pulex</i> , Insects Polyneoptera and Paraneoptera		•				
Insects Endopterygotes	(b)					

- (a) EGF-V absent in *Crassostrea* spp.
- (b) EGF-I absent in Hymenoptera
- ? EGF-IV and V absent due to partial sequences

B.

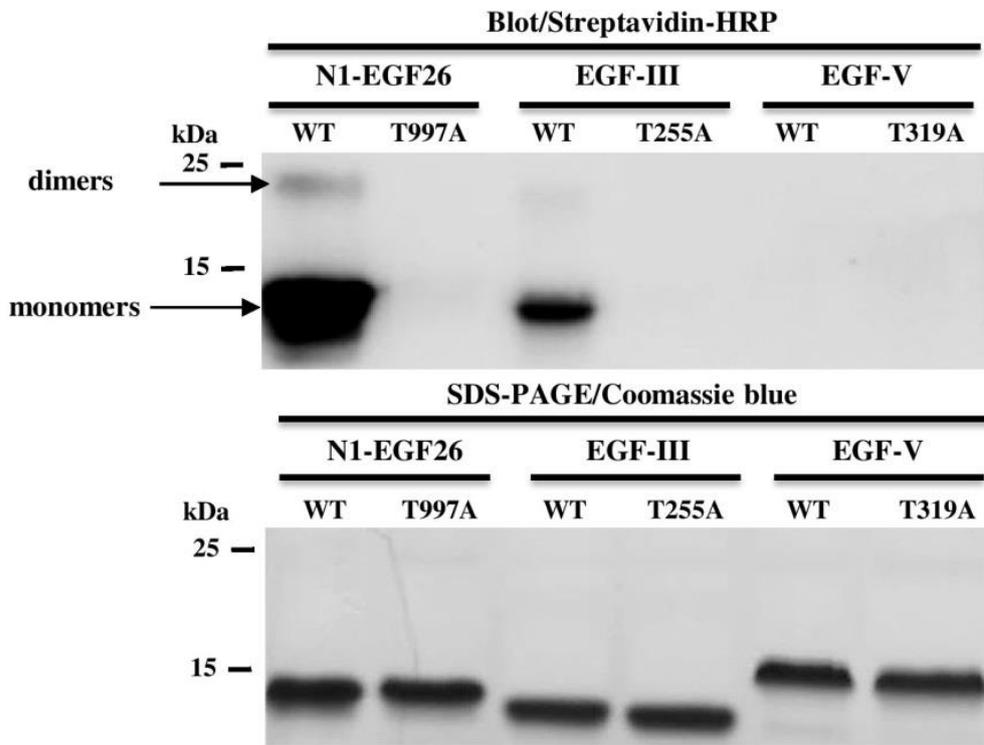
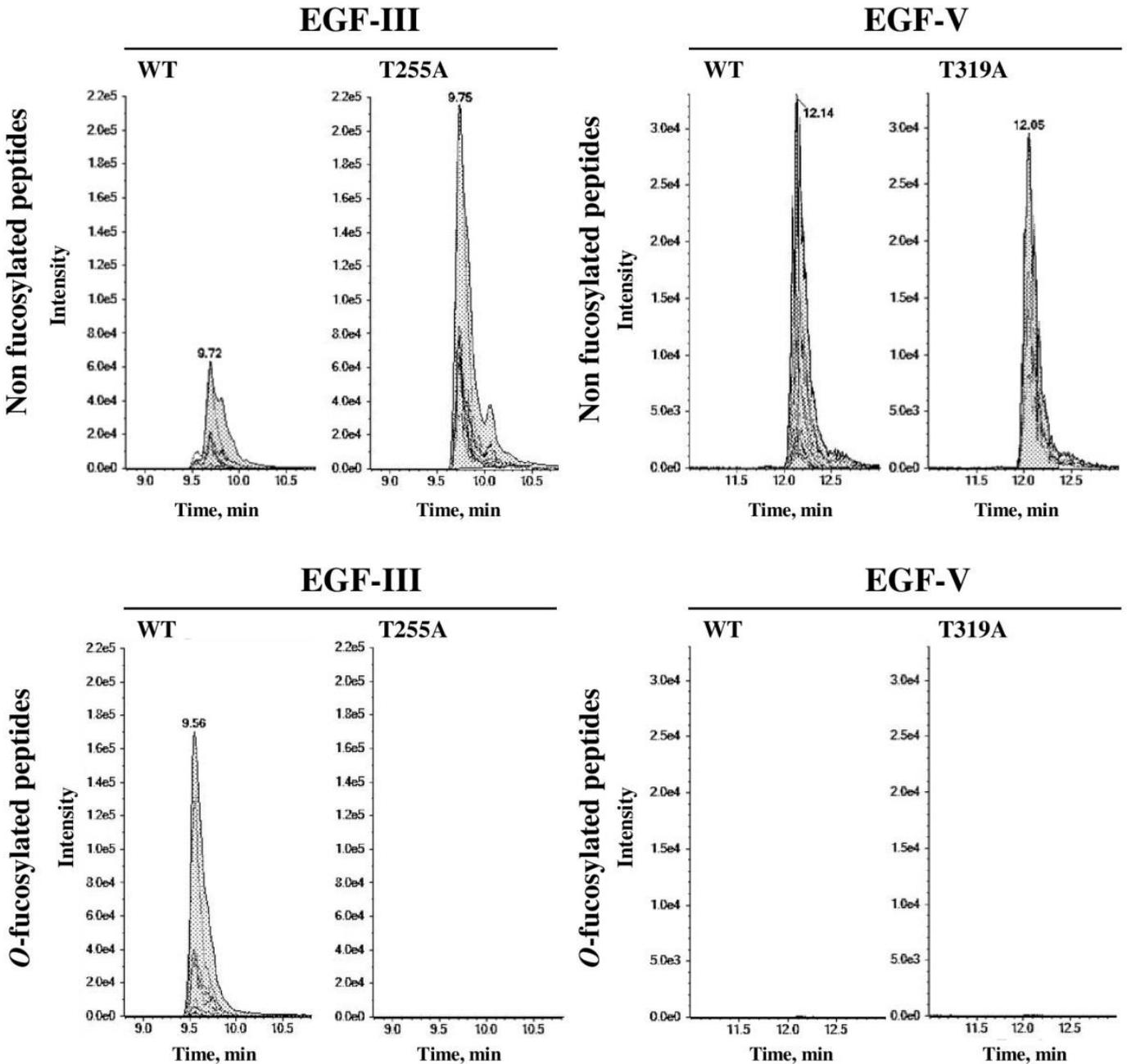
*Mus musculus* *Wif1* on Chr. 10 D2-D3, NC\_000076.6



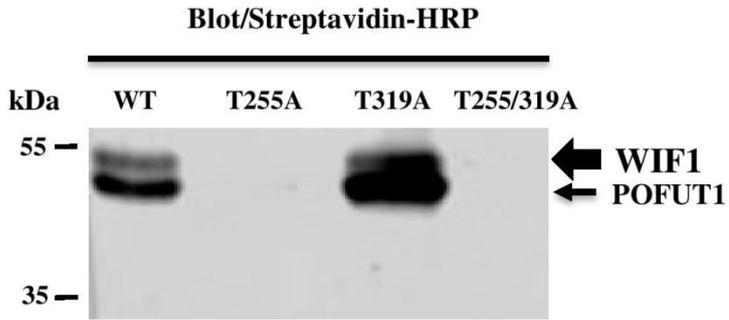
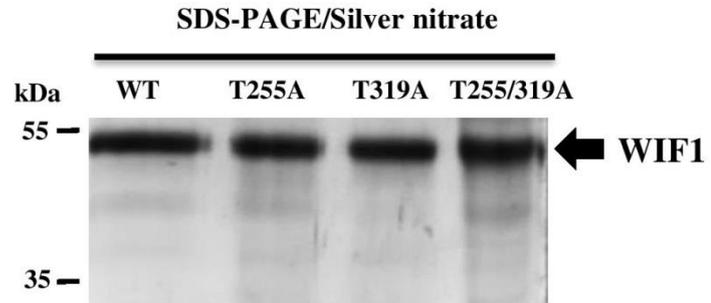
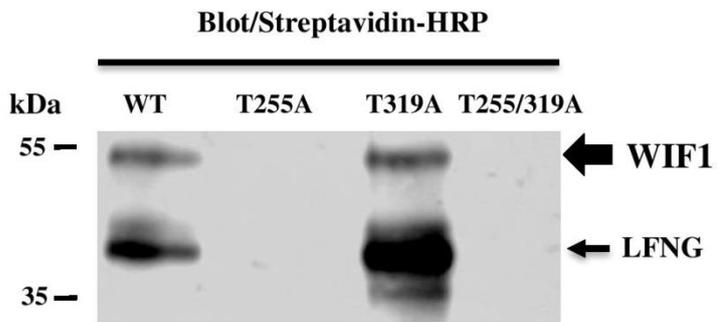






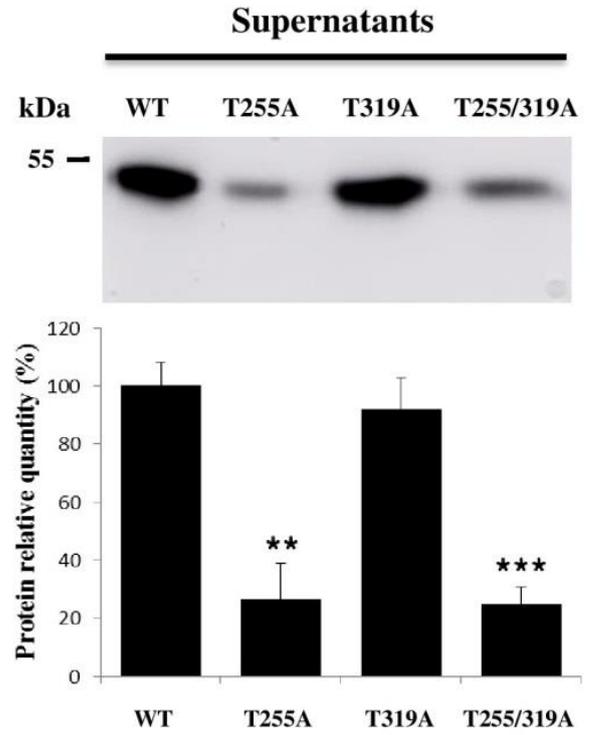
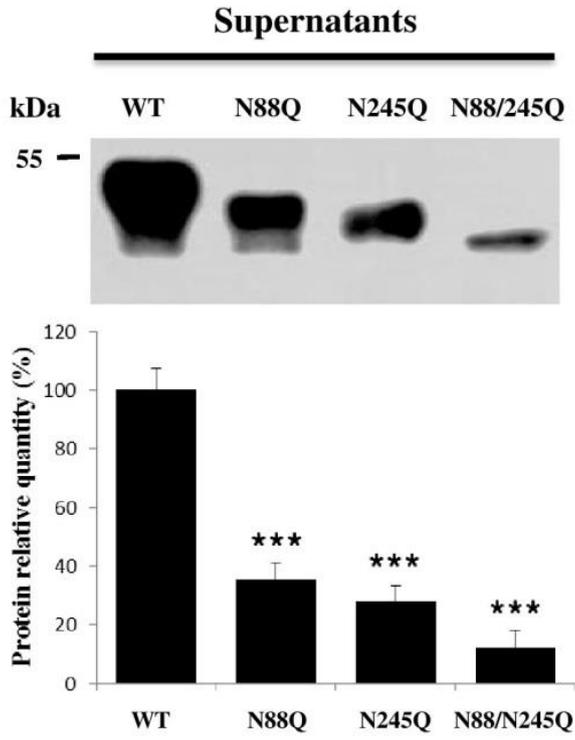
**A.****B.**



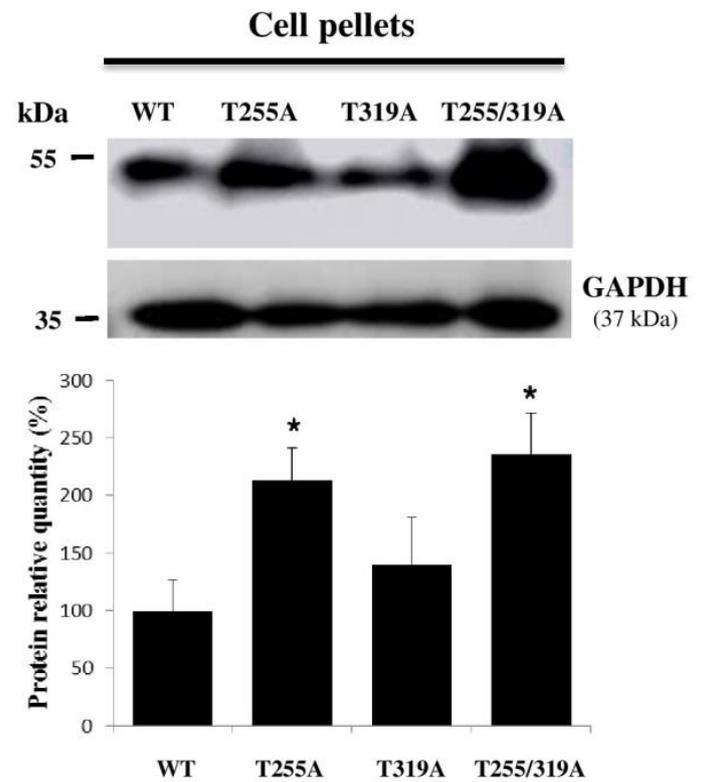
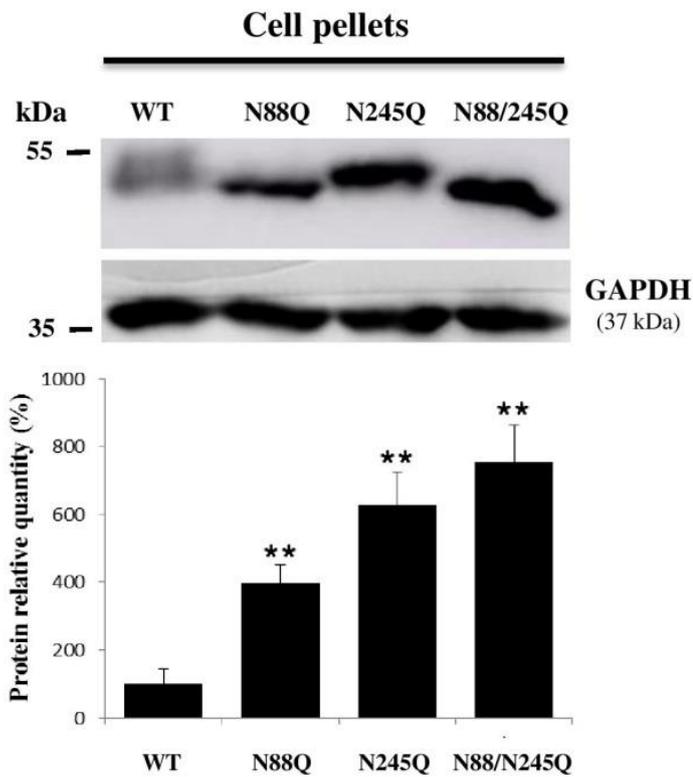
**A.****B.****C.**



**A.**

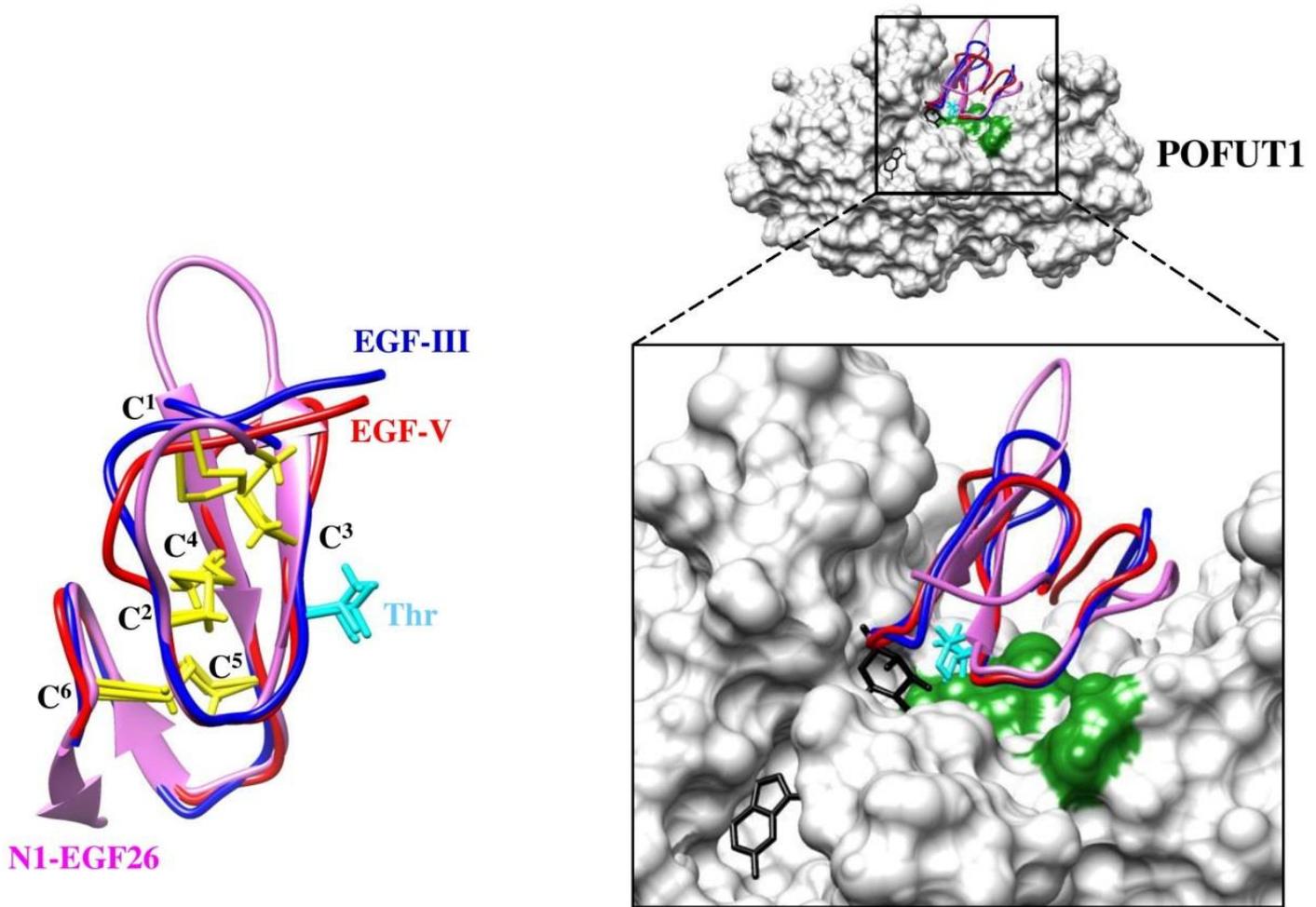


**B.**

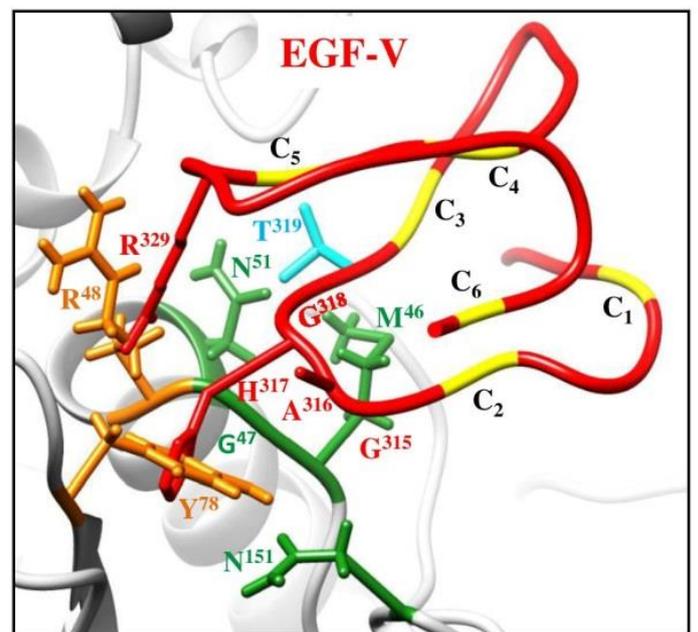
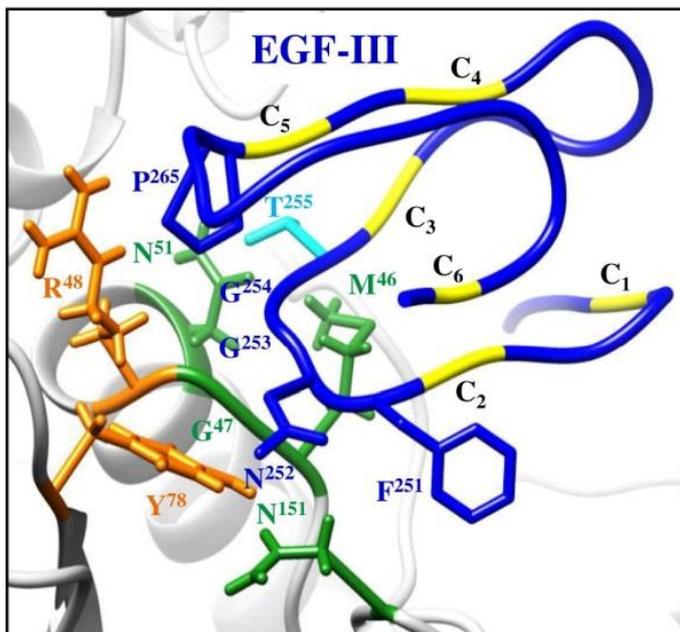




A.



B.





# **SUPPLEMENTARY DATA**

**Mouse WIF1 is only *O*-fucosylated on EGF-III despite two evolutionarily conserved consensus sites**

**Pennarubia Florian<sup>1</sup>, Al Jaam Bilal<sup>1</sup>, Pinault Emilie<sup>1,2</sup>, Brun Caroline E.<sup>1,3</sup>, Maftah Abderrahman<sup>1,#,\*</sup>,  
Germot Agnès<sup>1,#</sup> and Legardinier Sébastien<sup>1,#</sup>**

<sup>1</sup> Univ. Limoges, PEIRENE, EA 7500, Glycosylation and cell differentiation, F-87000 Limoges, France.

<sup>2</sup> Univ. Limoges, BISCEM Mass Spectrometry Platform, F-87025 Limoges, France

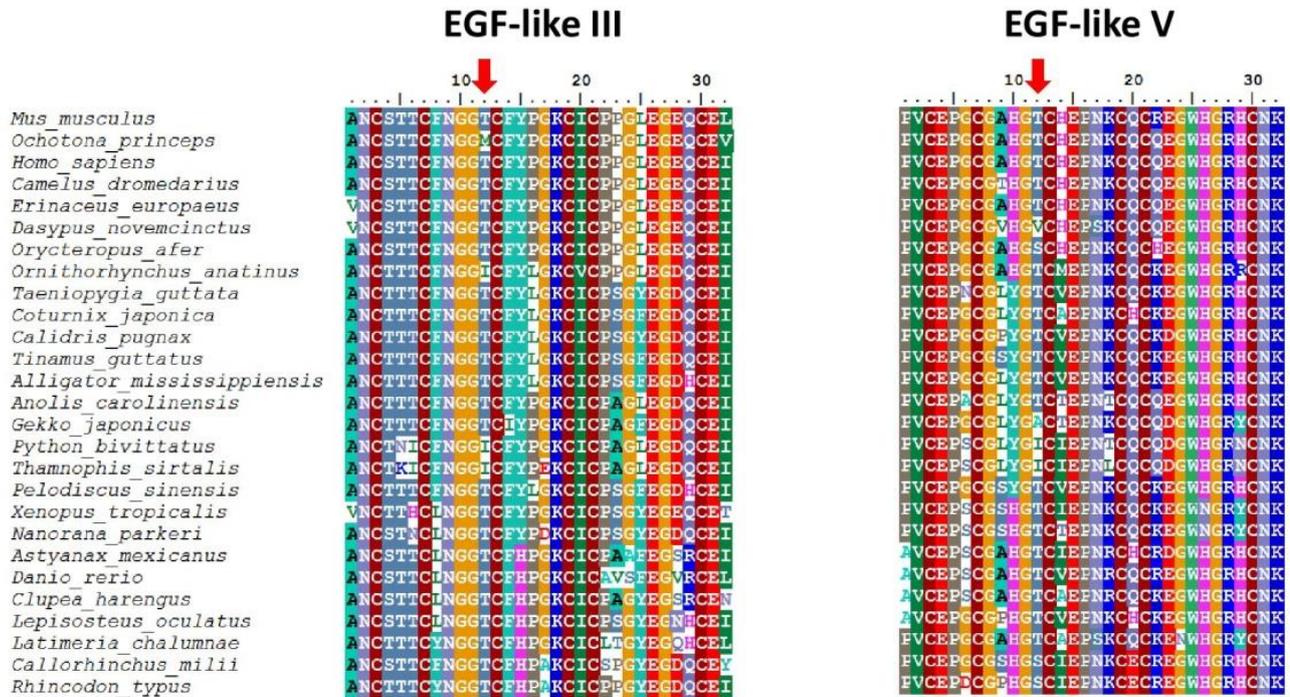
<sup>3</sup> Regenerative Medicine Program, Department of Cellular and Molecular Medicine, Faculty of Medicine, Sprott Center for Stem Cell Research, Ottawa Hospital Research Institute, University of Ottawa, Canada.

#A. Maftah, A. Germot and S. Legardinier are considered co-last authors and contributed equally to this work

\*corresponding author: Abderrahman Maftah, [abderrahman.maftah@unilim.fr](mailto:abderrahman.maftah@unilim.fr)



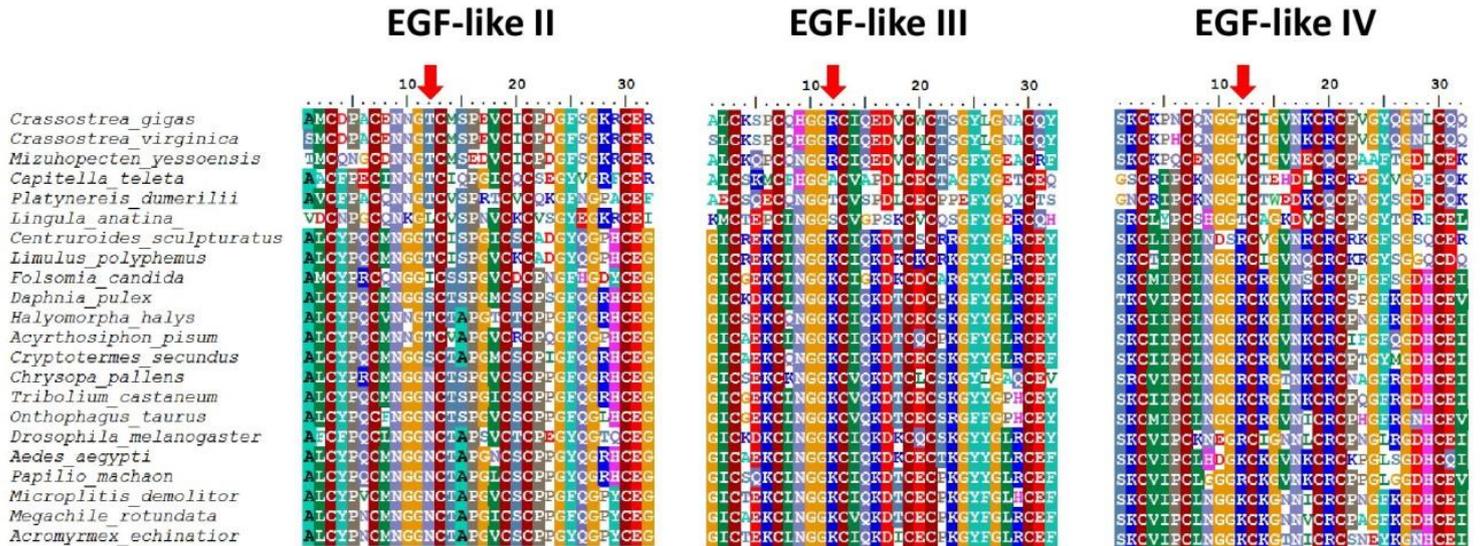
# SUPPLEMENTARY DATA



**Figure 1.** Alignment of the Wif1 EGF-like III and V domains among Gnathostomes. Red arrows point the potential *O*-fucosylated threonine or serine included in the consensus sequence C<sup>2</sup>X<sub>4</sub>(S/T)C<sup>3</sup>.



# SUPPLEMENTARY DATA

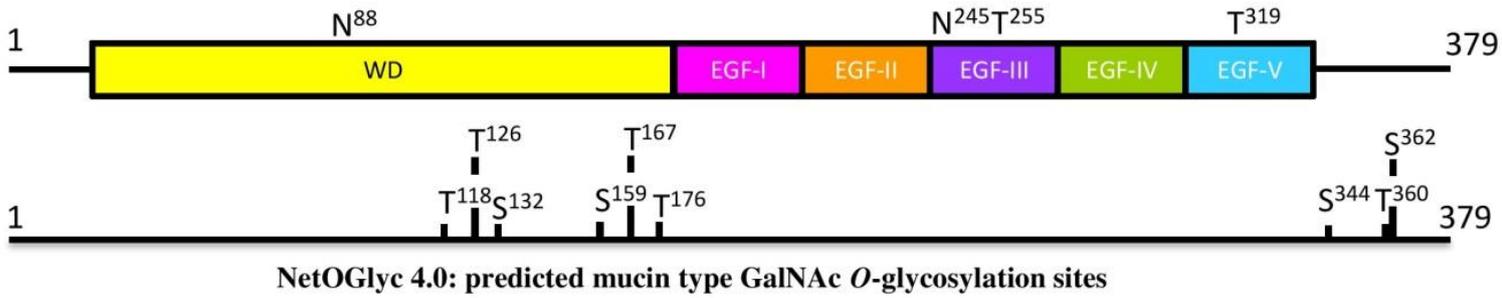


**Figure 2.** Alignment of the Wif1 EGF-like II, III and IV domains among Protostomes. Red arrows point the potential *O*-fucosylated threonine or serine included in the consensus sequence C<sup>2</sup>X<sub>4</sub>(S/T)C<sup>3</sup>.

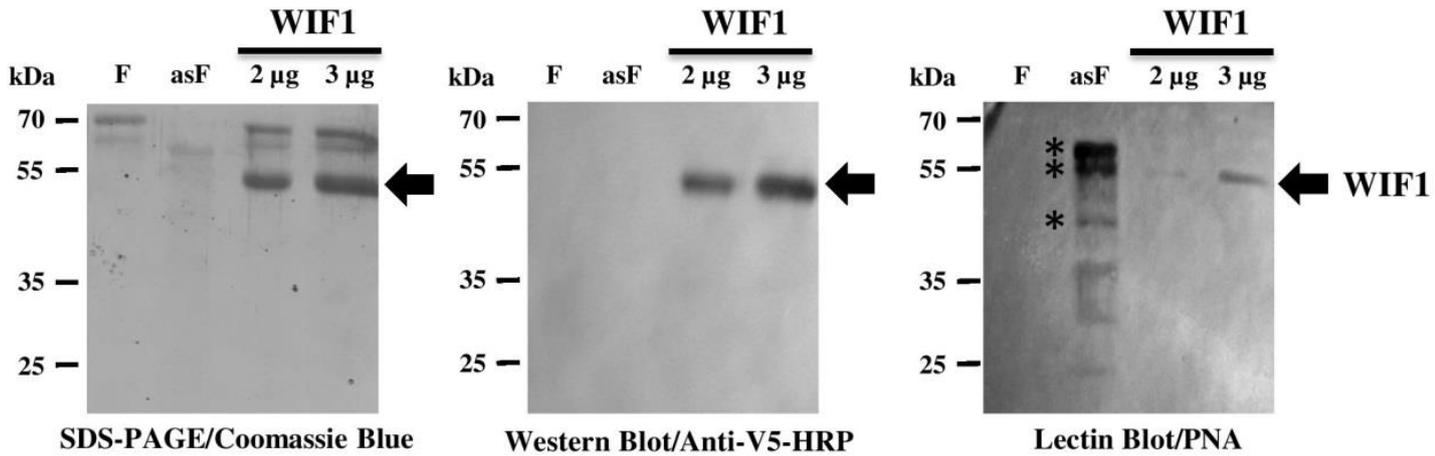


# SUPPLEMENTARY DATA

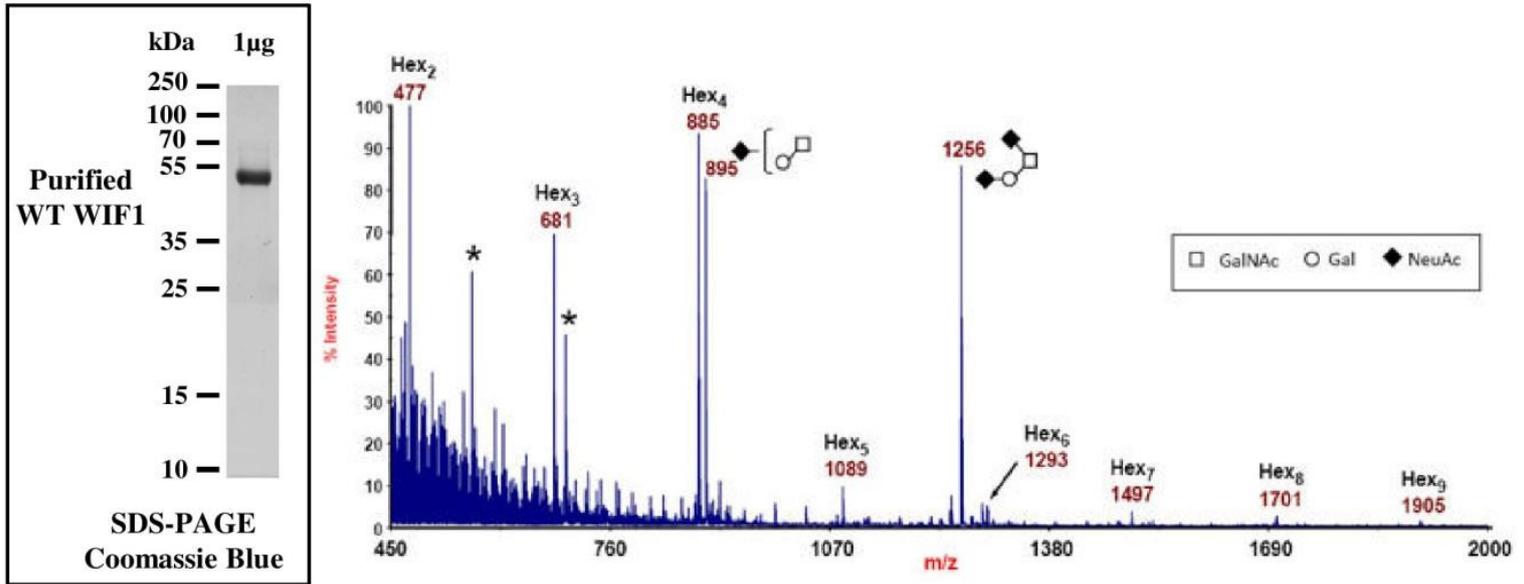
A.



B.



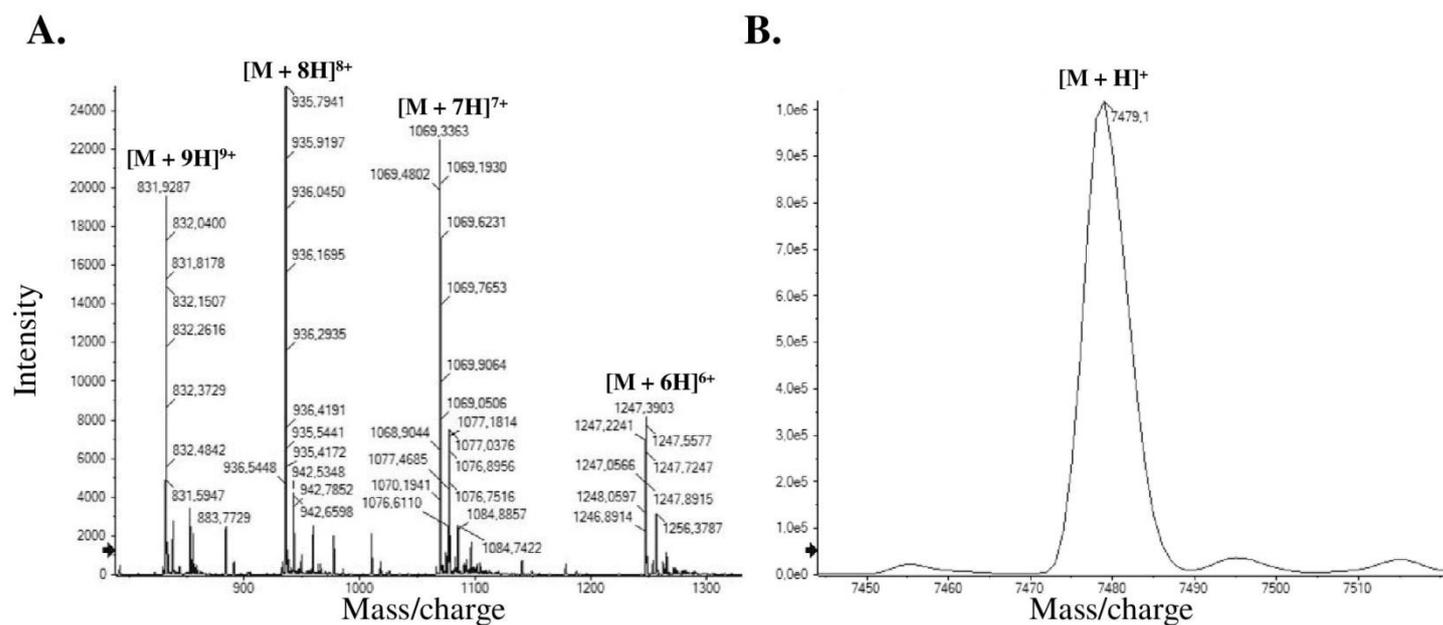
C.



**Figure 3.** Mucin-type *N*-acetylgalactosamine (GalNAc) *O*-glycosylation of mouse WIF1: (A) Nine mucin-type *O*-GalNAc *O*-glycosylation sites were predicted by the NetOglyc Server (<http://www.cbs.dtu.dk/services/NetOglyc/>). Six sites were located in the WIF domain (WD), three sites in the hydrophilic C-terminus and none within the EGF-like domains. (B) 500 ng of Fetuin (F), asialofetuin (asF) from the DIG Glycan Differentiation Kit (Roche Applied Science, Mannheim, Germany) and partially purified WT WIF1 (2 and 3 μg) were analysed by Coomassie blue stained SDS-PAGE, Western blot and lectin blot. Western blot using anti-V5-HRP antibody allowed identification of the recombinant WIF1 protein only and the Peanut agglutinin (PNA) lectin specifically recognized unsialylated mucin-type core disaccharide Galβ(1-3)GalNAc on recombinant WT WIF1 and the glycoprotein control asF (asterisks show the three main bands for commercial asF) (C) After β-elimination, the carbohydrate moiety of recombinant mouse WIF1, purified from CHO cell culture medium (see inset), was analyzed by mass spectrometry (MALDI). Two types of mucin-like GalNAc *O*-glycans were detected, as shown by the spectrum.

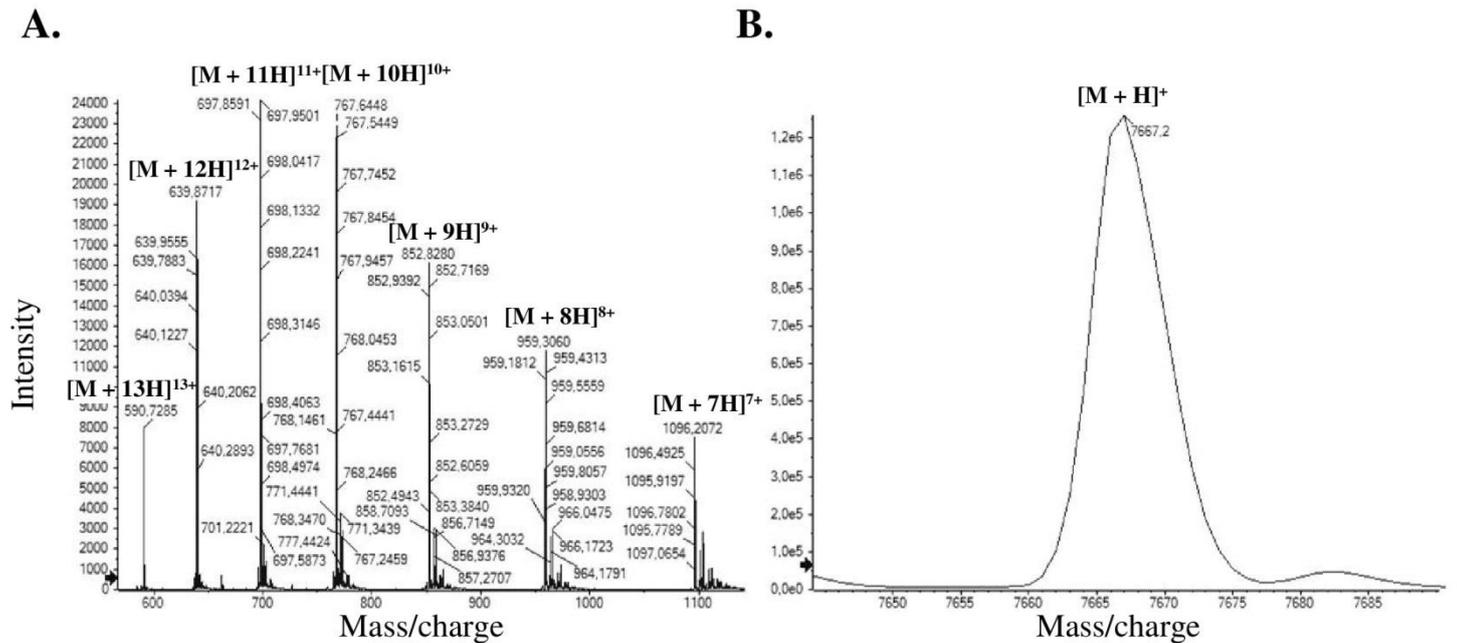


# SUPPLEMENTARY DATA





# SUPPLEMENTARY DATA

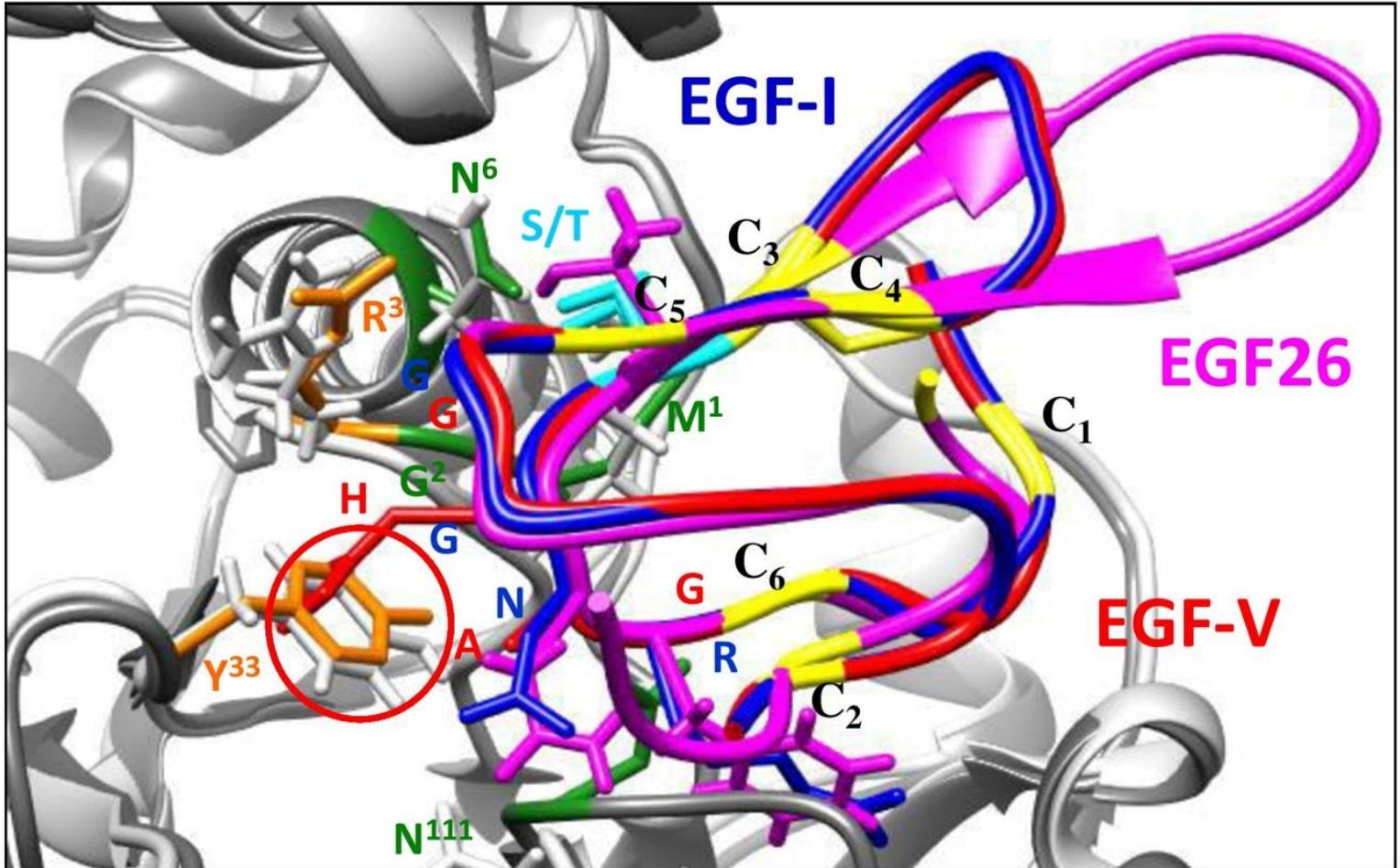


**Figure 5.** ESI-TOFMS spectrum of purified WIF1 EGF-V: (A) The signals corresponding to multi-charged forms ( $[M + 13H]^{13+}$  to  $[M + 7H]^{7+}$ ) of EGF-V were detected. (B) Deconvoluted ESI mass spectra of purified WIF1 EGF-V suggesting the presence of 3 disulfide bonds comparing to the theoretical molecular weight.



## SUPPLEMENTARY DATA

# *Ornithorhynchus anatinus*



```
>Ornithorhynchus_anatinus_XP_001510408.1_EGF-I  
AECPPGGCRNGGSCNERRVCDCPDGFYGPHCEK
```

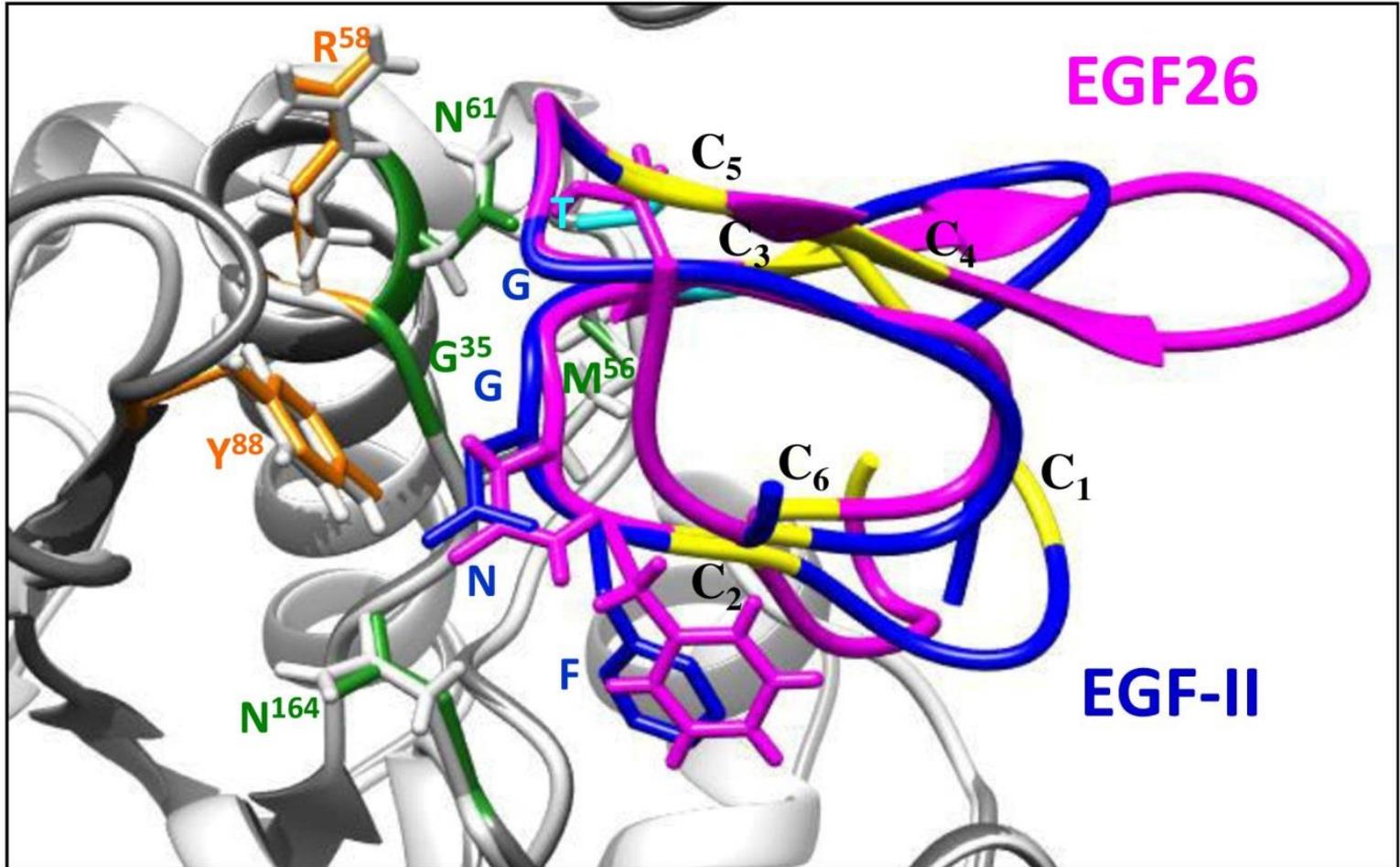
```
>Ornithorhynchus_anatinus_XP_001510408.1_EGF-V  
PVCEPGCGAHGTCMEPNKCQCKEGWHGRRCNK
```

**Figure 6:** Using Matchmaker, structural automated models for *Ornithorhynchus anatinus* (*Oa*) Pofut1 (XP\_016081535.1, 353 aa, with partial N-terminal sequence) and for *Oa* Wif1 EGF-I (XP\_001510408.1\_EGF1)(blue) and EGF-V (XP\_001510408.1\_EGF5)(red) were superimposed to mouse POFUT1 (light gray)/N1-EGF26 (magenta) complexes (PDB5KY4). Except N<sup>111</sup>, key residues (green) of *Oa* Pofut1 matched with those of mouse Pofut1, known to be involved in interaction with the C<sup>2</sup>-C<sup>3</sup> subdomain of N1-EGF26. Side chains of residues are shown for the C<sup>2</sup>-C<sup>3</sup> subdomains and residues of interest. A steric clash is shown (red circle) between Y<sup>33</sup> of *Oa* Pofut1 and the histidine residue at the position C<sup>2</sup>+3 in EGF-V.



## SUPPLEMENTARY DATA

# *Saccoglossus kowalevskii*



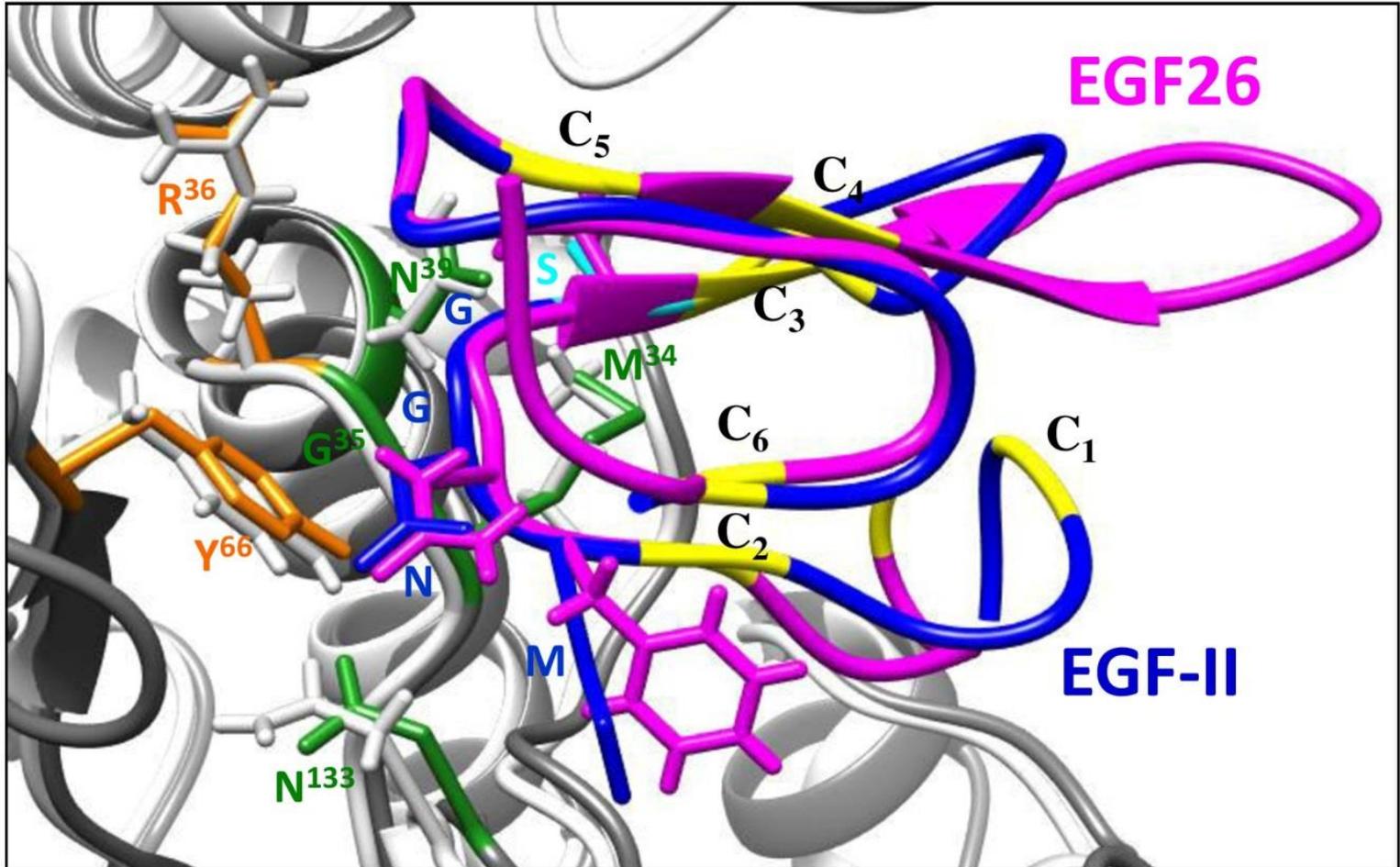
```
>Saccoglossus_kowalevskii_NP_001161492.1_EGF-II  
ALCSPRCFNGGTCIAPGICVCADDF'TGRSCEK
```

**Figure 7:** Using Matchmaker, structural automated models for *Saccoglossus kowalevskii* (*Sk*) Pofut1 (XP\_006823249.1, 407 aa)(dim gray) and for *Sk* Wif1 EGF-II (NP\_001161492.1\_EGF2)(blue) were superimposed to mouse POFUT1 (light gray)/N1-EGF26 (magenta) complexes (PDB5KY4). Key residues (green) of *Sk* Pofut1 matched with those of mouse Pofut1, known to be involved in interaction with the C<sup>2</sup>-C<sup>3</sup> subdomain of N1-EGF26. Side chains of residues are shown for the C<sup>2</sup>-C<sup>3</sup> subdomains and residues of interest. No steric clash was found.



## SUPPLEMENTARY DATA

# *Daphnia pulex*



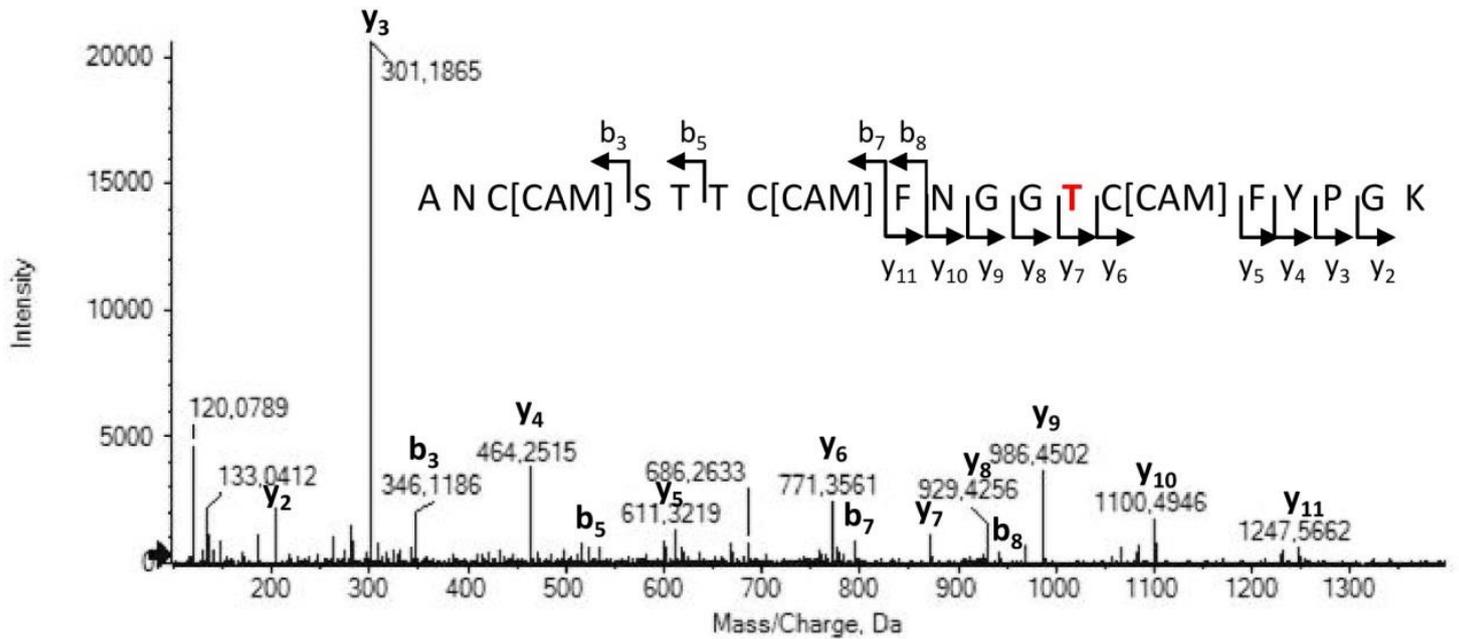
```
>Daphnia_pulex_EFX65761.1_EGF-II  
ALCYPQCMNGGSCTSPGMCSCPSGFQGRHCEG
```

**Figure 8:** Using Matchmaker, structural automated models for *Daphnia pulex* (*Dp*) Pofut1 (EFX81920.1, 381 aa)(dim gray) and for *Dp* Wif1 EGF-II (EFX65761.1\_EGFII)(blue) were superimposed to mouse POFUT1 (light gray)/N1-EGF26 (magenta) complexes (PDB5KY4). Key residues (green) of *Dp* Pofut1 matched with those of mouse Pofut1, known to be involved in interaction with the C<sup>2</sup>-C<sup>3</sup> subdomain of N1-EGF26. Side chains of residues are shown for the C<sup>2</sup>-C<sup>3</sup> subdomains and residues of interest. No steric clash was found.

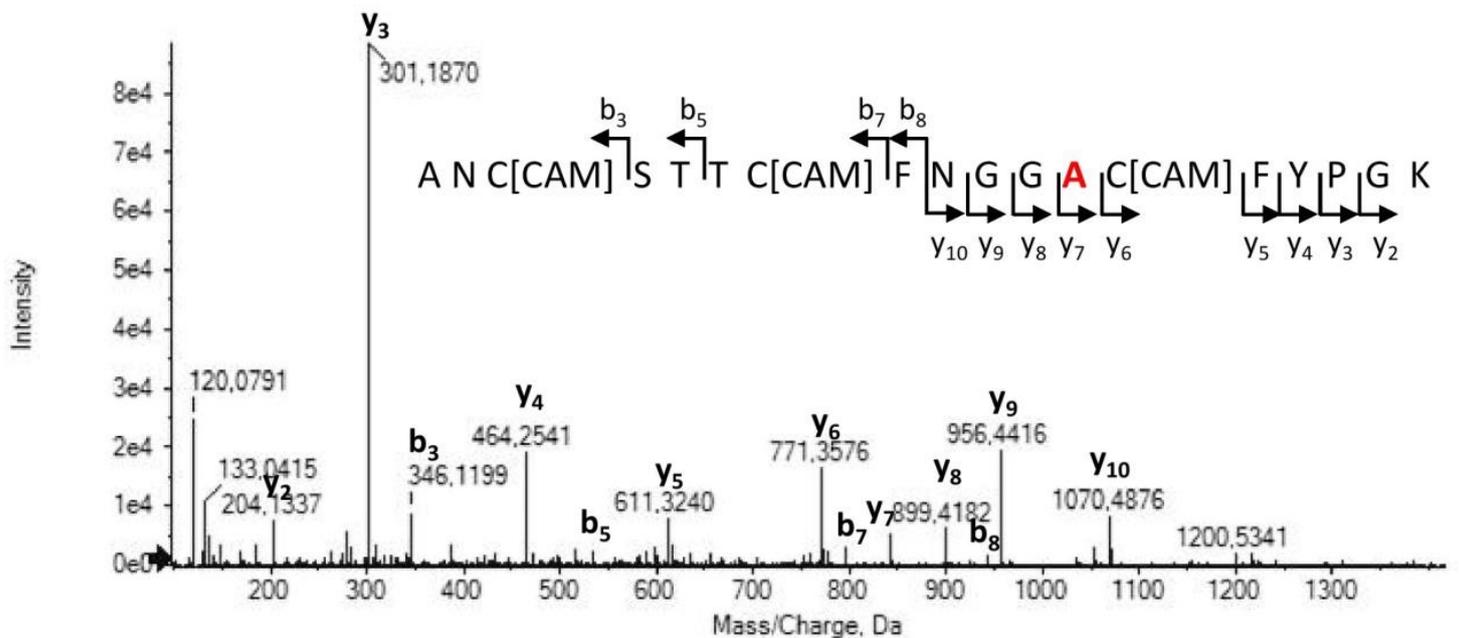


# SUPPLEMENTARY DATA

A.



B.

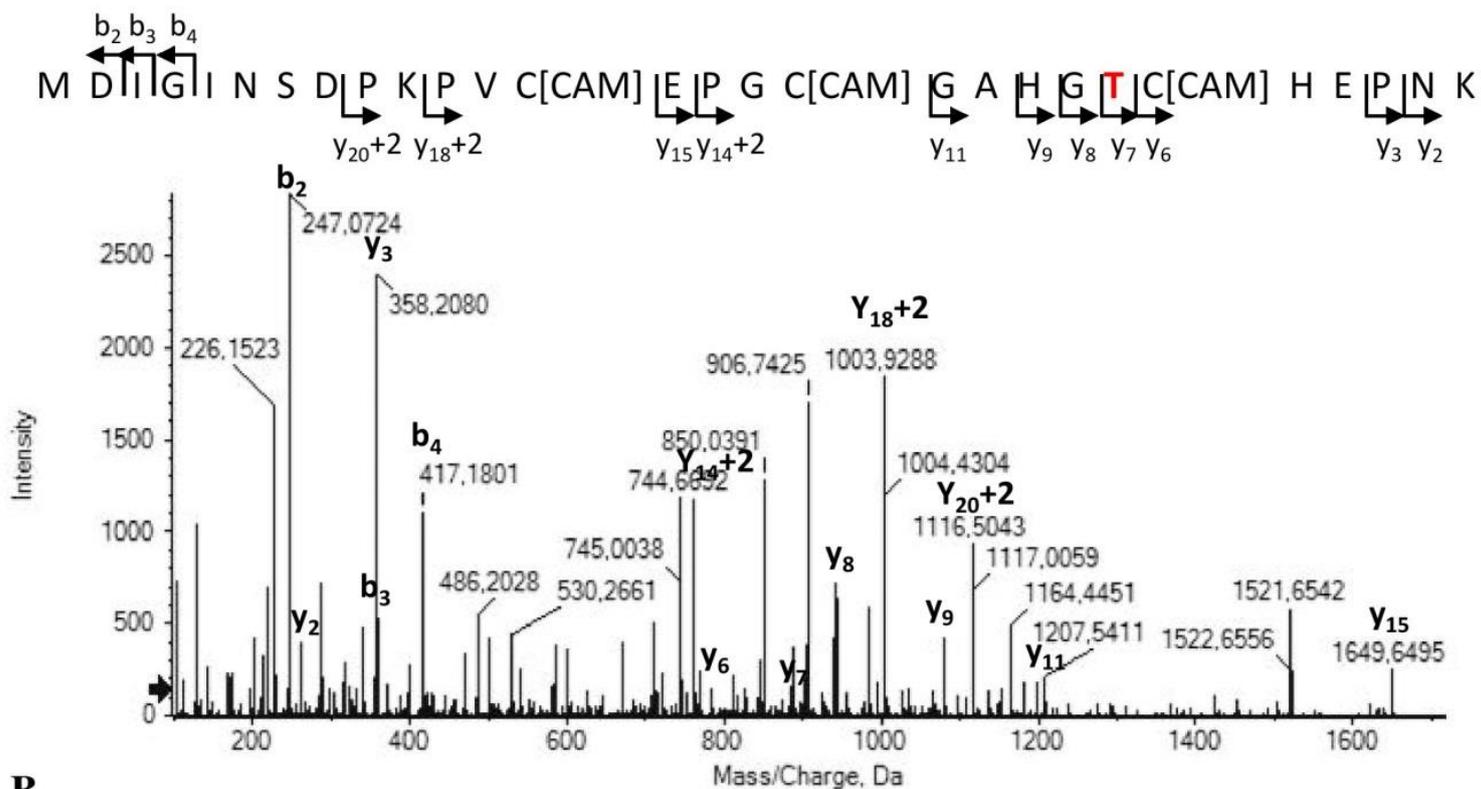


**Figure 9:** Representative MS/MS spectra for WT (A) and mutant (B) WIF1 EGF-III tryptic peptide containing the *O*-fucosylation consensus site (in red bold). Spectra were acquired with a micro-LC TripleTOF system in IDA mode. Detected b/y fragments are annotated on each spectrum confirming the identified peptidic sequence.

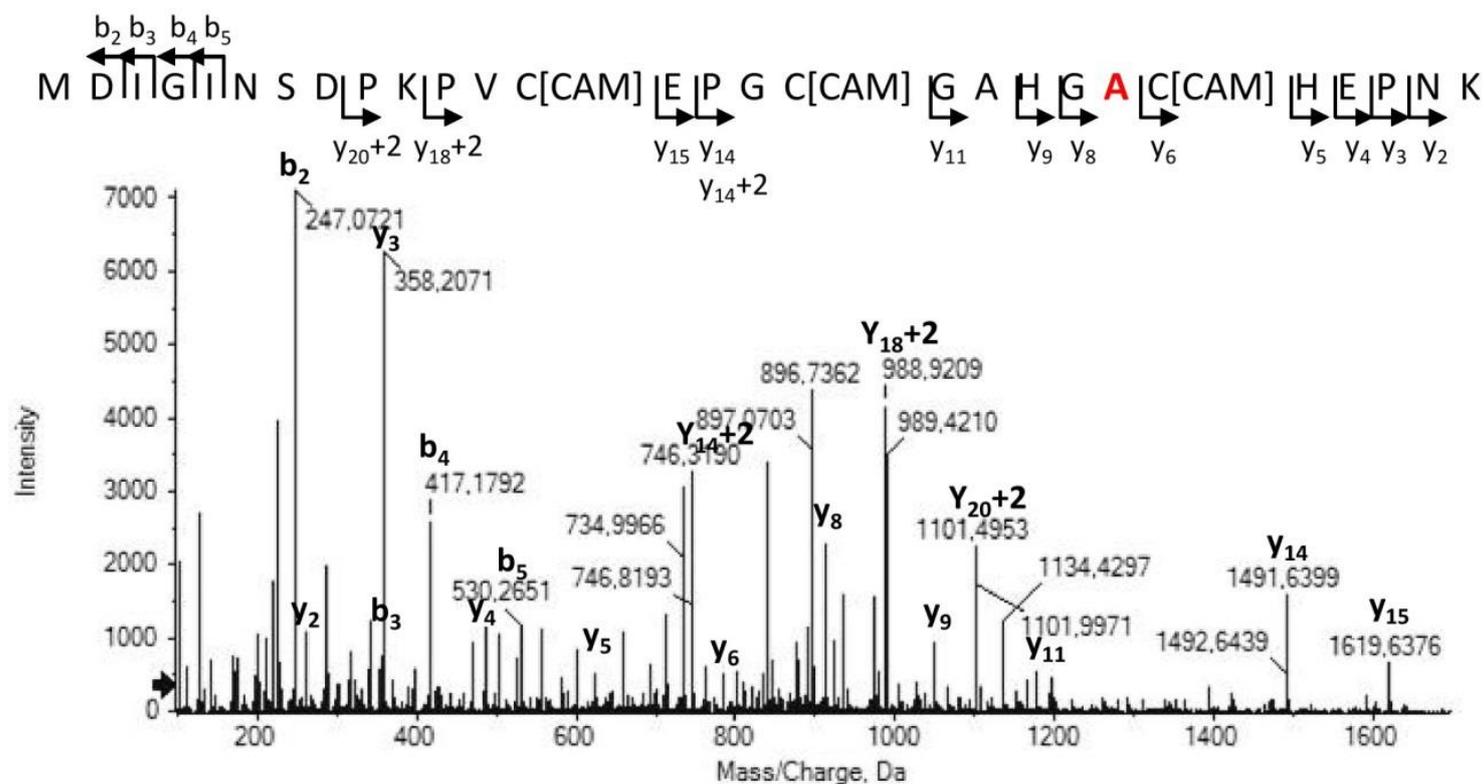


# SUPPLEMENTARY DATA

A.



B.



**Figure 10:** Representative MS/MS spectra for WT (A) and mutant (B) WIF1 EGF-V tryptic peptide containing the *O*-fucosylation consensus site (in red bold). Spectra were acquired with a micro-LC TripleTOF system in IDA mode. Detected b/y fragments are annotated on each spectrum confirming the identified peptidic sequence.



# SUPPLEMENTARY DATA

Targeted sequence <sup>a</sup>	charge state	Non fucosylated peptide		Fucosylated peptide		DP (V)	CE (V)			
		MS1	MS2	MS1	MS2					
WT EGF III ANC[CAM]STTC[CAM]FNGG <u>T</u> [CAM]FYPGK	+2	1021.4137	+2y13	1508.6296	1094.4426	+2y13	1654.68760	105.6	45.6	
			+2y11	1247.5513		+2y11	1393.6092			
			+2y9	986.4400		+2y9	1132.4979			
			+2y4	464.2503		+2y4	464.2503			
			+2y3	301.1870		+2y3	301.1870			
			+2b3	346.1179		+2b3	346.1179			
	+3	681.2782	+3y10	1100.4829	729.9642	+3y10	1246.5408	80.8	34.7	
			+3y9	986.4400		+3y9	1132.4979			
			+3y6	771.34942		+3y6	771.3494			
			+3y4	464.2503		+3y4	464.2503			
			<b>+3y3</b>	<b>301.1870</b>		<b>+3y3</b>	<b>301.1870</b>			
			+3b4	433.1499		+3b4	433.1499			
	Mutant EGF III ANC[CAM]STTC[CAM]FNGG <u>A</u> [CAM]FYPGK	+2	1006.4084	+2y12	1377.5714	1079.4378	+2y12	1523.6293	104.5	45.1
				+2y11	1217.5408		+2y11	1363.5987		
+2y10				1070.4723	+2y10		1216.5303			
+2y9				956.4294	+2y9		1102.4873			
+2y4				464.2503	+2y4		464.2503			
+2y3				301.1870	+2y3		301.1870			
+3		671.2747	+3y10	1070.4723	719.9607	+3y10	1216.5303	80.1	34.1	
			+3y9	956.42946		+3y9	1102.4873			
			+3y8	899.4080		+3y8	1045.4659			
			+3y6	771.3494		+3y6	771.3494			
			+3y4	464.2503		+3y4	464.2503			
			<b>+3y3</b>	<b>301.1870</b>		<b>+3y3</b>	<b>301.1870</b>			

**Table 1:** HR-MRM parameters for the detection of WIF1 EGF-III WT and mutant peptides. The *O*-fucosylation site is indicated in bold and underlined. Each peptide is detected by multiple charge states corresponding to different MS1 (with optimized Declustering Potential (DP) and collision energy (CE)) recorded during the 30 min analysis. Peak detection with MultiQuant was achieved with the six most abundant fragments (MS2) with a 10,000 resolution. Quantification fragment is indicated in bold. The 5 other MS2 were used as confirmation fragments.



# SUPPLEMENTARY DATA

Targeted sequence	charge state	Non fucosylated peptide		Fucosylated peptide		DP (V)	CE (V)								
		MS1	MS2	MS1	MS2										
<b>WT EGF V</b> MDIGINSDPKPVC[CAM]EPGC[CAM]GAHG <u>T</u> C[CAM]HEPNK	+3	1026.4492	+3y8	942.4097	1075.1352	+3y8	1088.4677	105.9	53.4						
			+3y3	358.2084		+3y3	358.2084								
			+3y18+2	1003.9169		+3y18+2	1076.9458								
			+3b2	247.0747		+3b2	247.0747								
			+3b4	417.1802		+3b4	417.1802								
			+3b8	846.3661		+3b8	846.36619								
	+4	770.0887	+4y8	942.4097	806.6032	+4y8	1088.4677	87.3	39.5						
			<b>+4y3</b>	<b>358.2084</b>		<b>+4y3</b>	<b>358.2084</b>								
			+4y20+2	1116.4907		+4y20+2	1189.5197								
			+4y18+2	1003.9169		+4y18+2	1076.9458								
			+4y14+2	761.3197		+4y14+2	834.3486								
			+4b2	247.0747		+4b2	247.0747								
			<b>Mutant EGF V</b> MDIGINSDPKPVC[CAM]EPGC[CAM]GAHG <u>A</u> C[CAM]HEPNK	+4		762.5861	+4y8			912.3992	799.1006	+4y8	1058.4571	86.7	39.1
							<b>+4y3</b>			<b>358.2084</b>		<b>+4y3</b>	<b>358.2084</b>		
+4y20+2	1101.4855	+4y20+2			1174.5144										
+4y18+2	988.9116	+4y18+2			1061.9406										
+4y14+2	746.3144	+4y14+2			819.3433										
+4b2	247.0747	+4b2			247.0747										

**Table 2:** HR-MRM parameters for the detection of WIF1 EGF-V WT and mutant peptides. The *O*-fucosylation site is indicated in bold and underlined. Each peptide is detected by multiple charge states corresponding to different MS1 (with optimized Declustering Potential (DP) and collision energy (CE)) recorded during the 30 min analysis. Peak detection with MultiQuant was achieved with the six most abundant fragments (MS2) with a 10,000 resolution. Quantification fragment is indicated in bold. The 5 other MS2 were used as confirmation fragments.



# **SUPPLEMENTARY INFORMATION**

## **Full length gel/blot images**

**Mouse WIF1 is only *O*-fucosylated on EGF-III despite two evolutionarily conserved consensus sites**

**Pennarubia Florian<sup>1</sup>, Al Jaam Bilal<sup>1</sup>, Pinault Emilie<sup>1,2</sup>, Brun Caroline E.<sup>1,3</sup>, Maftah Abderrahman<sup>1,#,\*</sup>,  
Germot Agnès<sup>1,#</sup> and Legardinier Sébastien<sup>1,#</sup>**

<sup>1</sup> Univ. Limoges, PEIRENE, EA 7500, Glycosylation and cell differentiation, F-87000 Limoges, France.

<sup>2</sup> Univ. Limoges, BISCEM Mass Spectrometry Platform, F-87025 Limoges, France

<sup>3</sup> Regenerative Medicine Program, Department of Cellular and Molecular Medicine, Faculty of Medicine, Sprott Center for Stem Cell Research, Ottawa Hospital Research Institute, University of Ottawa, Canada.

#A. Maftah, A. Germot and S. Legardinier are considered co-last authors and contributed equally to this work

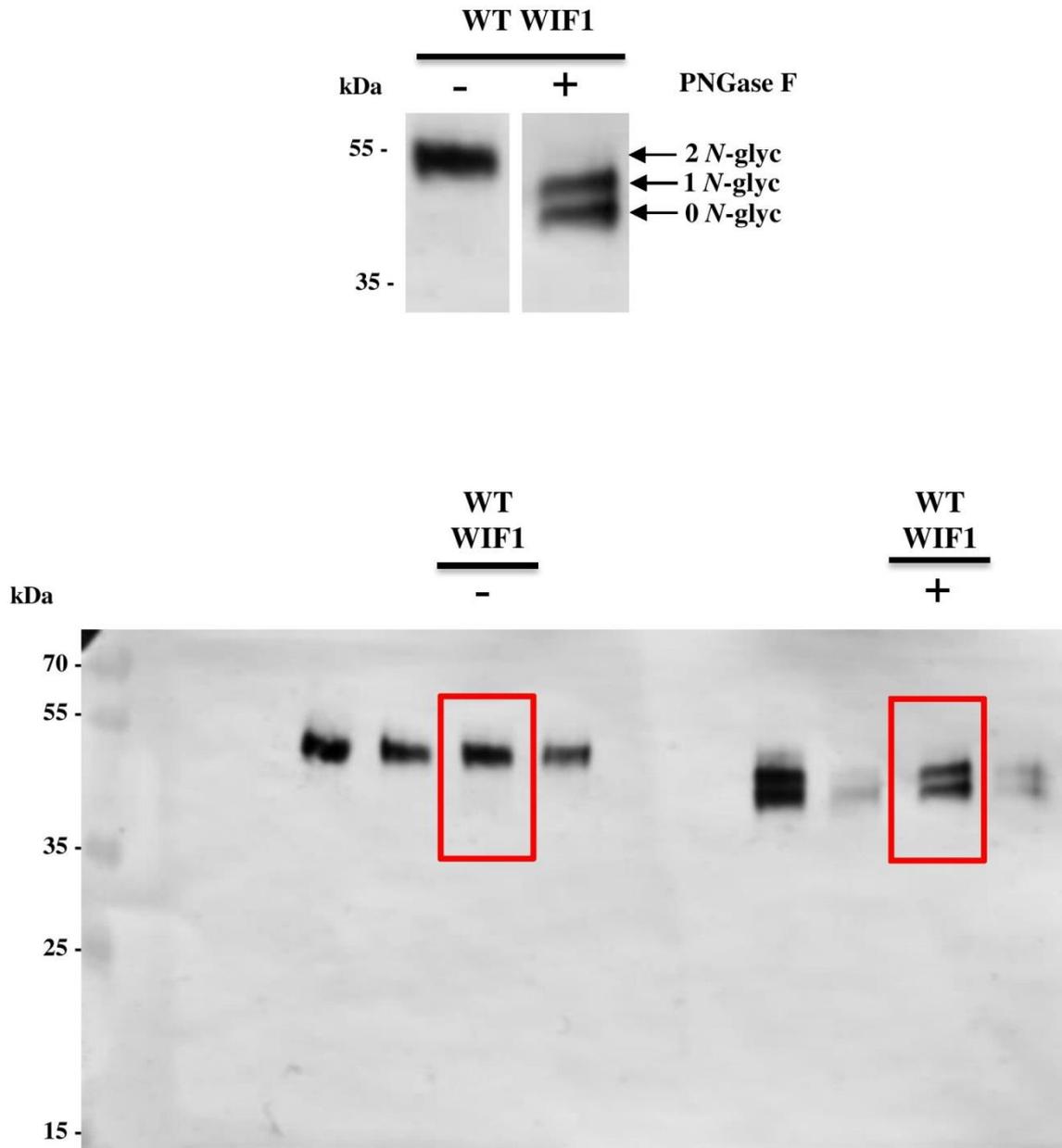
\*corresponding author: Abderrahman Maftah, [abderrahman.maftah@unilim.fr](mailto:abderrahman.maftah@unilim.fr)



# SUPPLEMENTARY INFORMATION

## FIGURE 3

A.

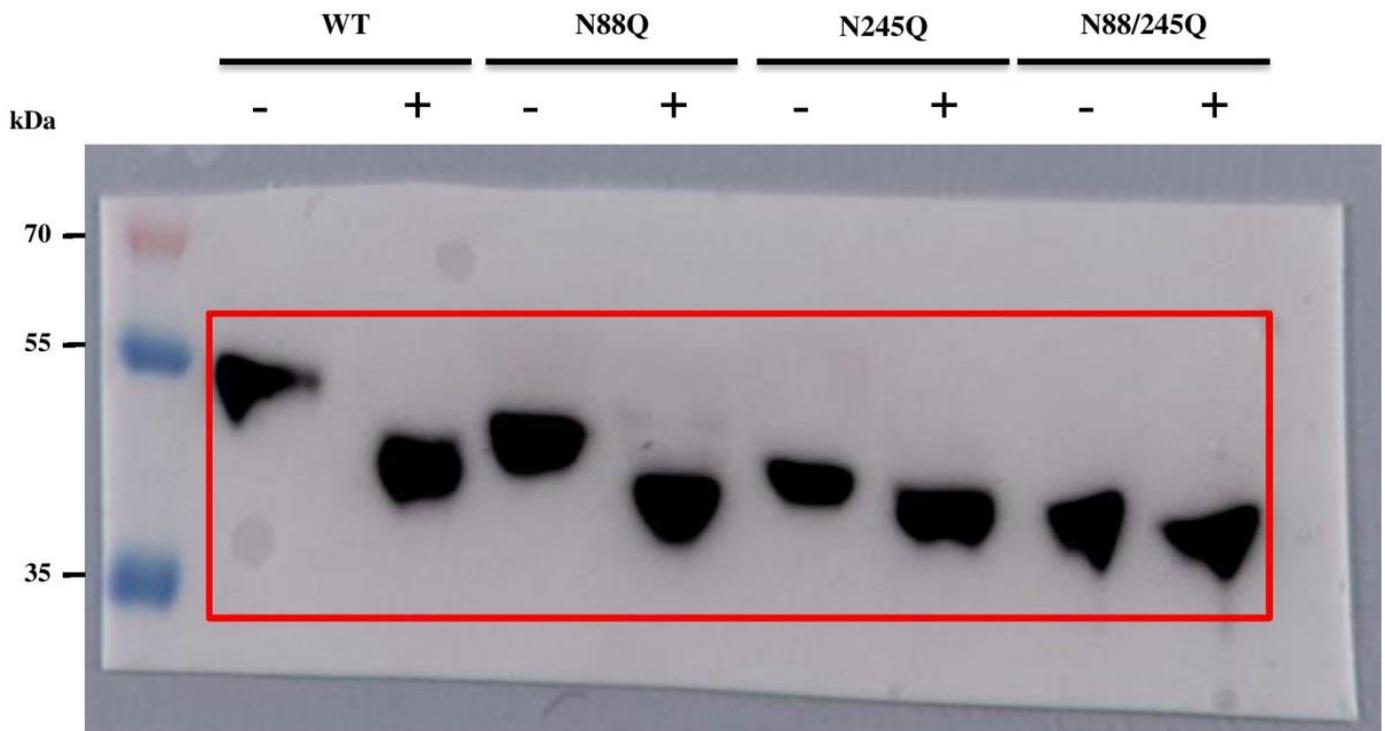
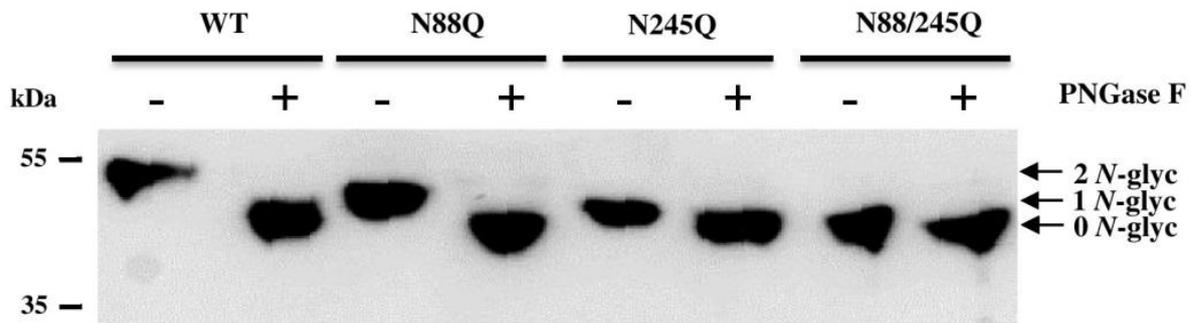




# SUPPLEMENTARY INFORMATION

## FIGURE 3

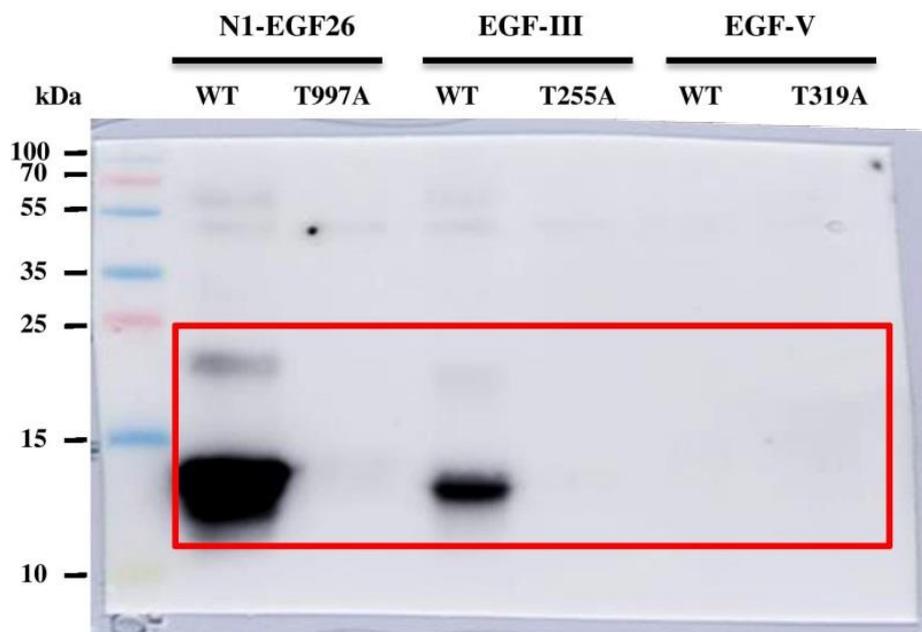
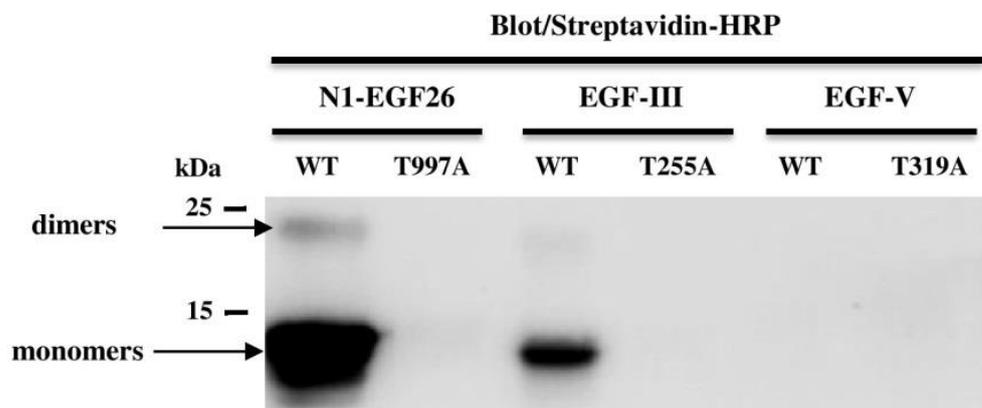
**B.**





# SUPPLEMENTARY INFORMATION

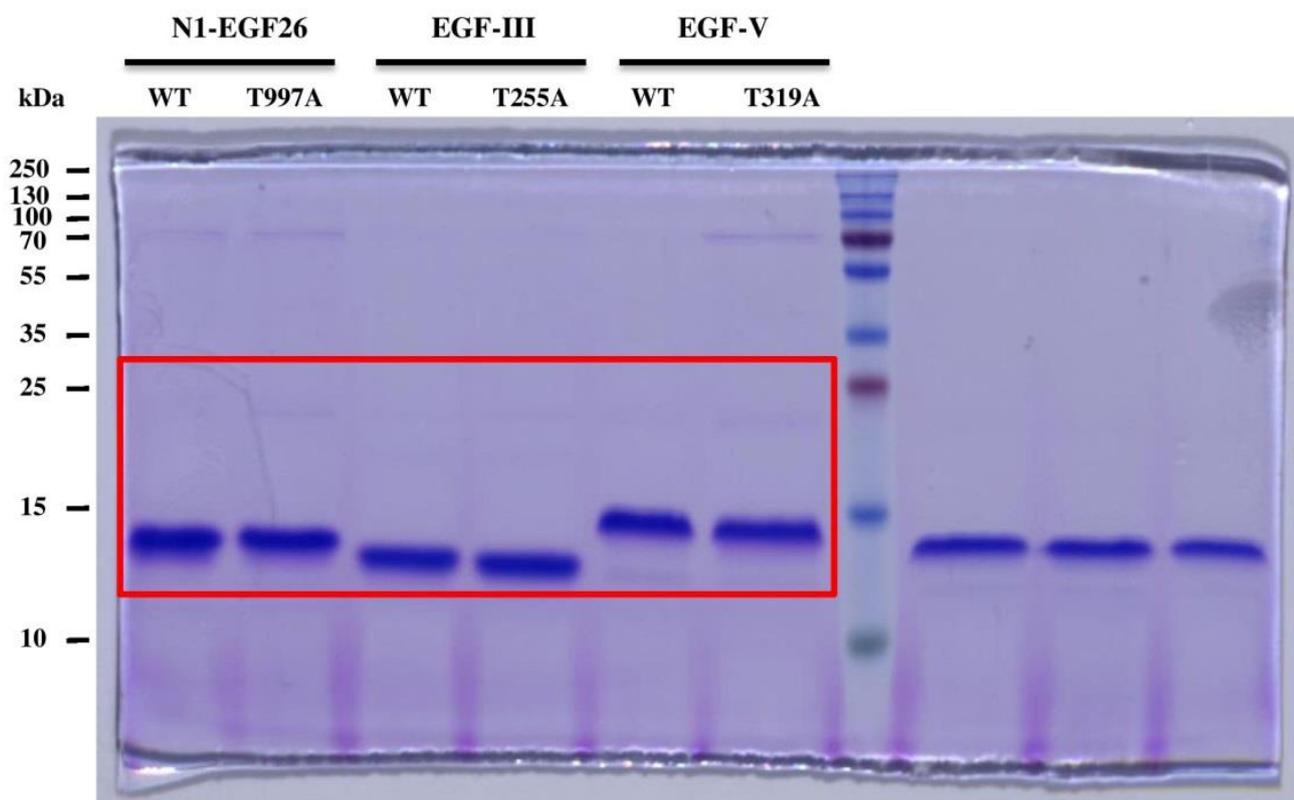
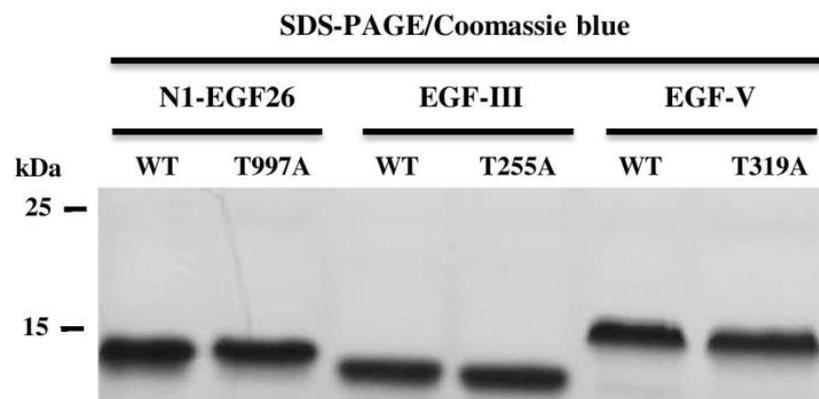
## FIGURE 4.A





# SUPPLEMENTARY INFORMATION

## FIGURE 4.A

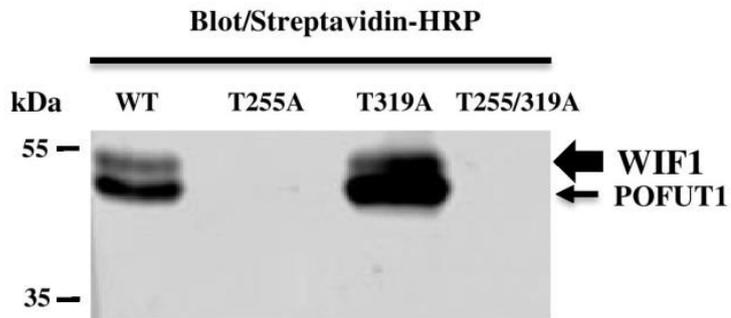




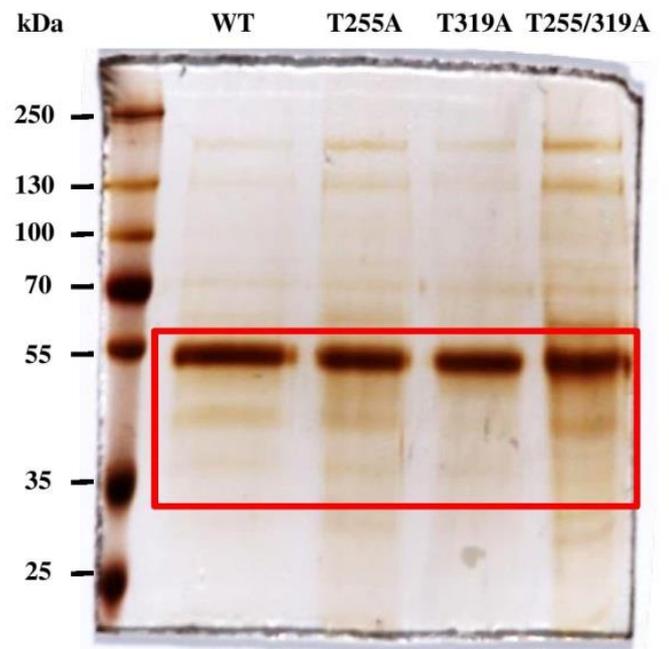
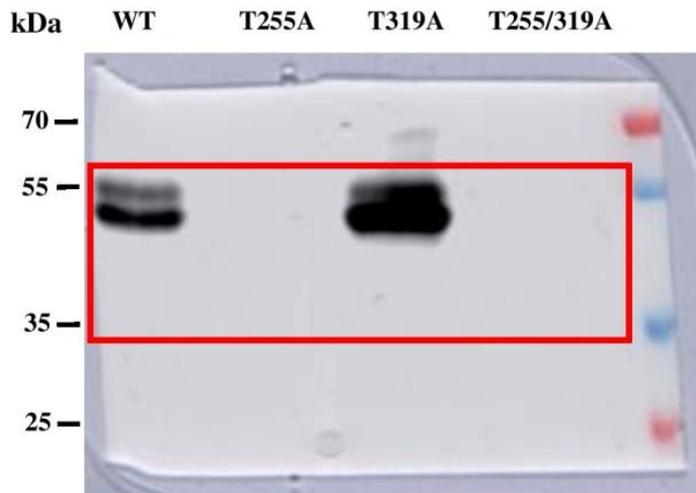
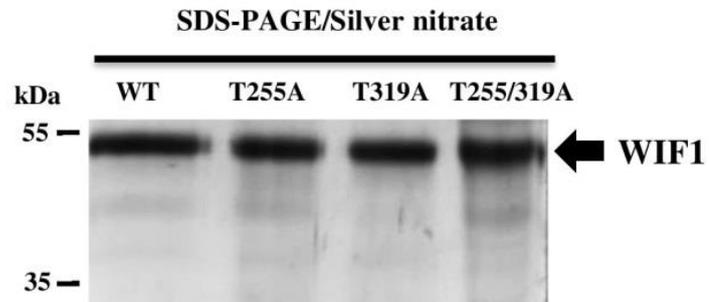
# SUPPLEMENTARY INFORMATION

## FIGURE 5

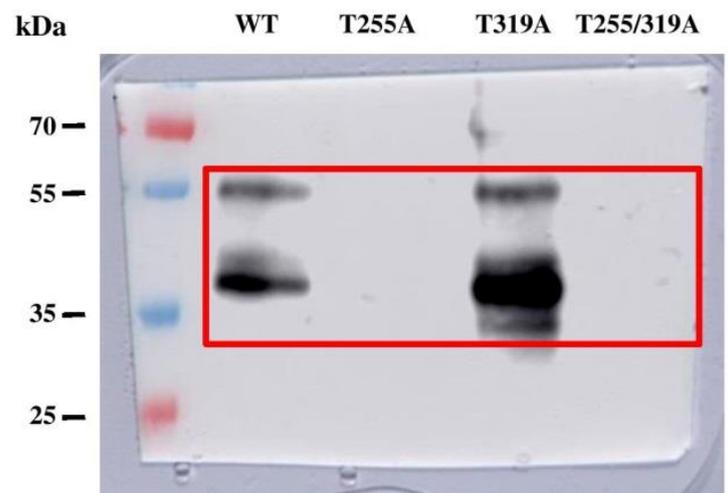
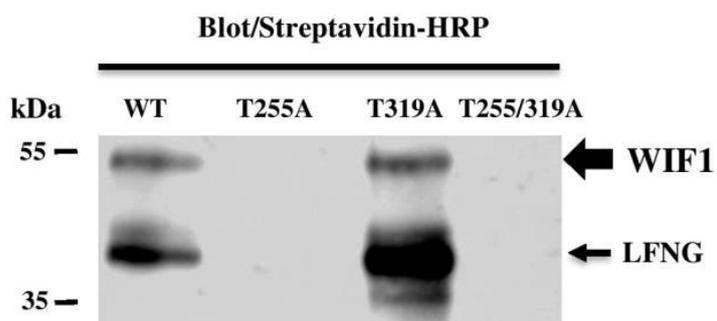
A.



B.



C.

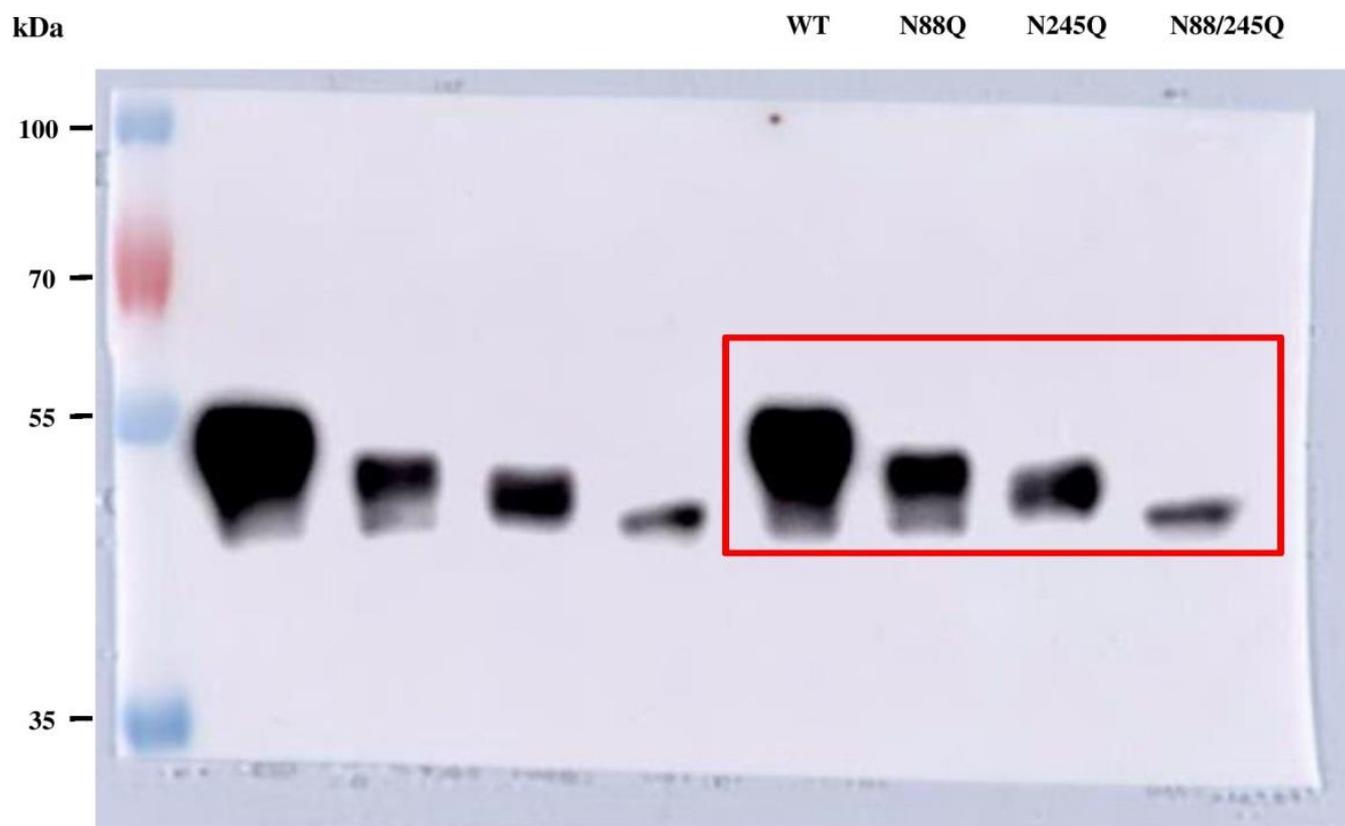
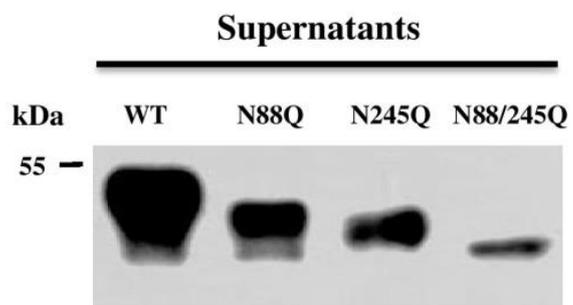




# SUPPLEMENTARY INFORMATION

## FIGURE 6

A.

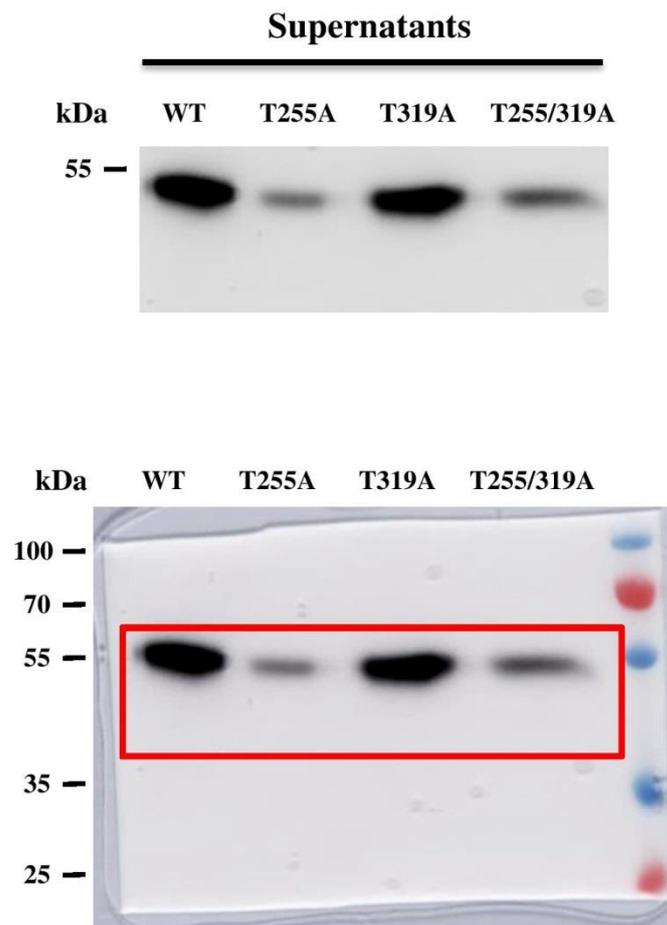




# SUPPLEMENTARY INFORMATION

## FIGURE 6

A.

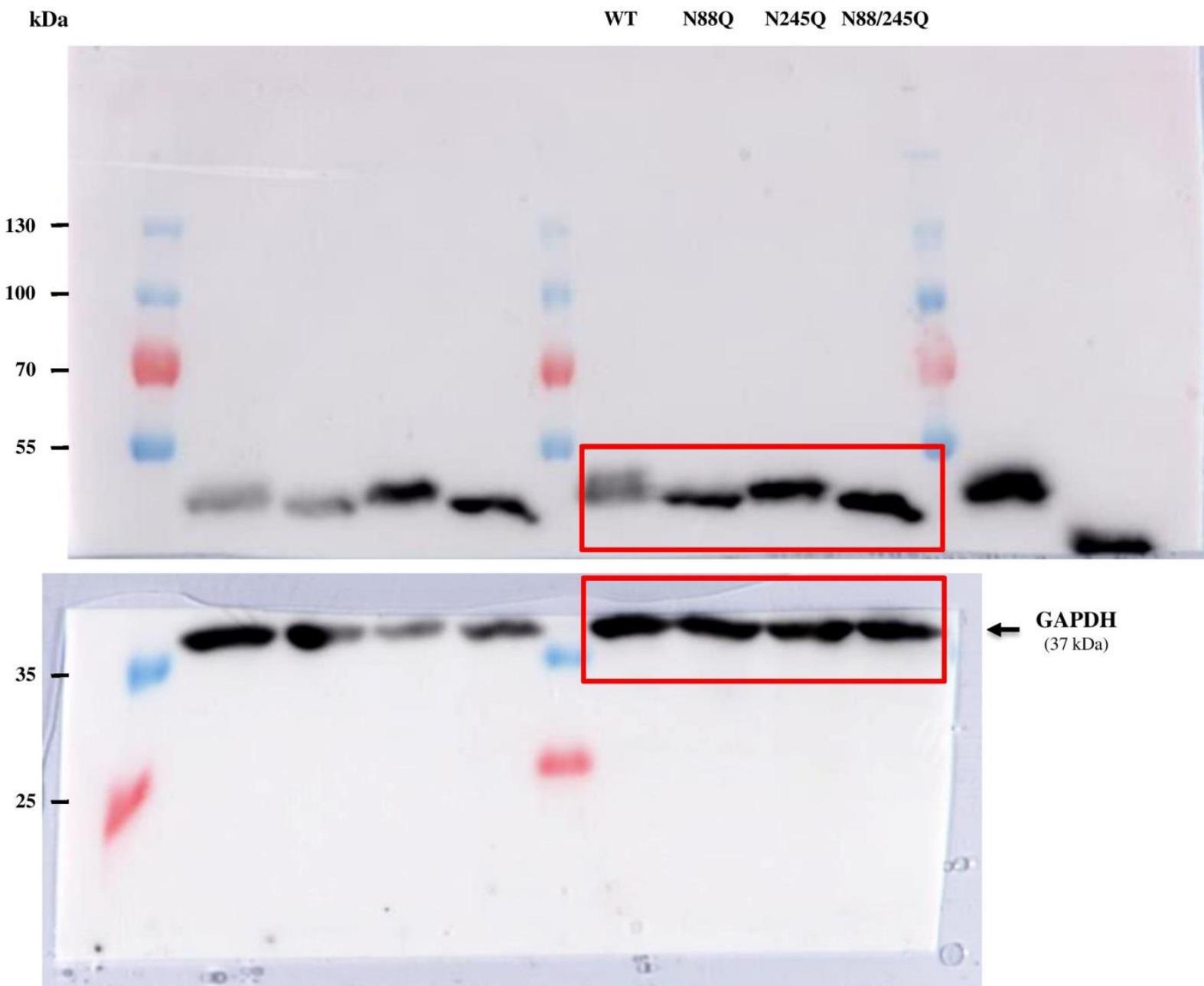
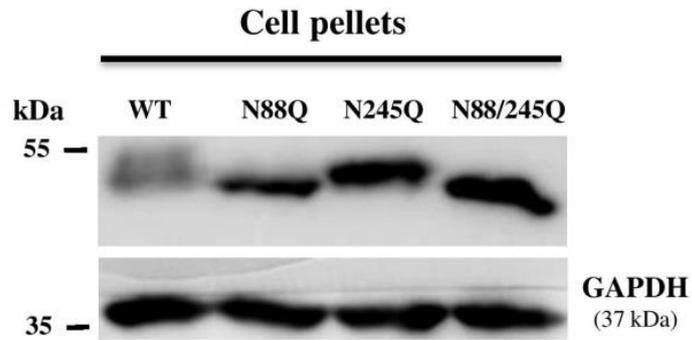




# SUPPLEMENTARY INFORMATION

## FIGURE 6

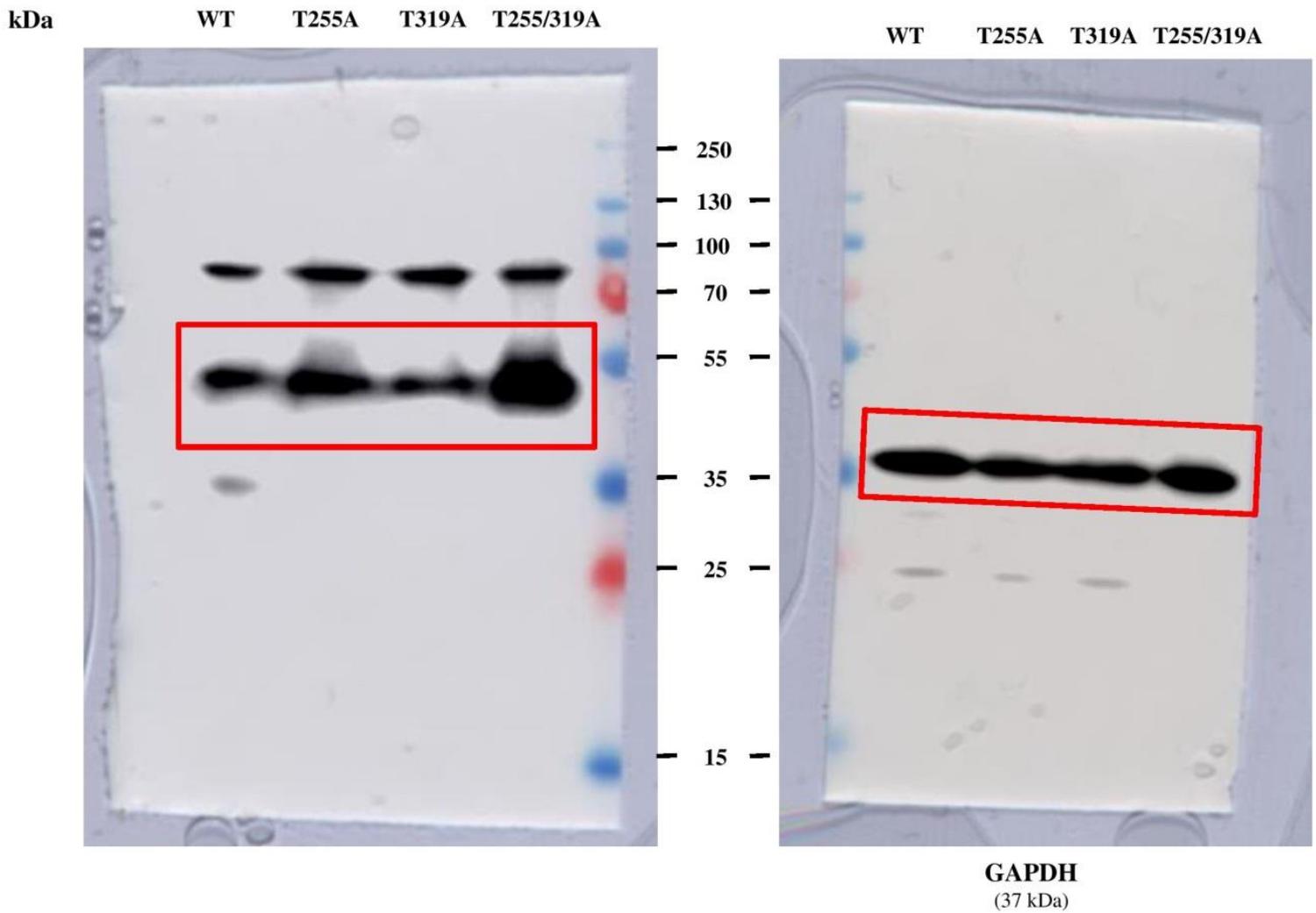
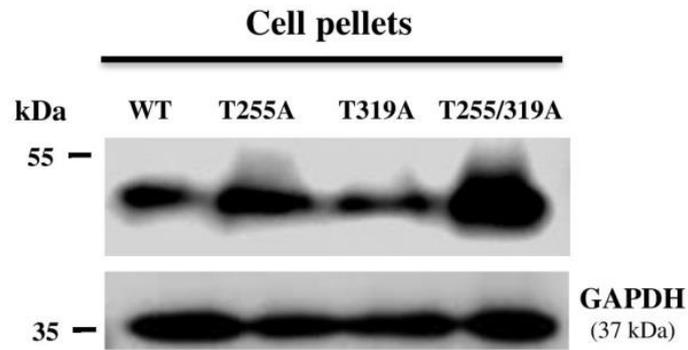
B.





# SUPPLEMENTARY INFORMATION

## FIGURE 6





***In vitro* acellular method to reveal *O*-fucosylation on EGF-like domains**

**Pennarubia Florian**, Pinault Emilie, Maftah Abderrahman and Legardinier Sébastien

**Glycobiology**. 2018 Nov 29. doi: 10.1093/glycob/cwy106.



L'étude de la *O*-fucosylation de WIF1 nous a confronté à la difficulté de mettre en évidence cette glycosylation rare. Cela est principalement dû à l'absence d'anticorps ou de lectines reconnaissant spécifiquement les *O*-fucoses, associé au fait qu'il n'existe aucune méthode chimique ou enzymatique clivant spécifiquement les fucoses *O*-liés. A ce jour, les techniques les plus couramment utilisées sont la spectrométrie de masse (Kentzer et al., 1990) associée ou non au radiomarquage (Rampal et al., 2005).

Bien qu'étant la plus pertinente, l'étude de la *O*-fucosylation sur une protéine dans sa conformation native et naturelle est souvent inenvisageable. Des contraintes liées à la disponibilité, la solubilité et la purification de la protéine d'intérêt imposent fréquemment de recourir à l'étude de protéines recombinantes. En effet, l'étude de la *O*-fucosylation de protéines naturelles par spectrométrie de masse n'a été possible que sur des protéines solubles et disponibles en grande quantité ; c'est le cas des facteurs de coagulation VII, IX, XII (Bjoern et al., 1991; Harris et al., 1992; Nishimura et al., 1992) et de l'UROKINASE (Kentzer et al., 1990). Du fait de leurs faibles niveaux d'expression endogène et de leur caractère membranaire, le récepteur NOTCH1 et ses principaux ligands DLL1 et JAG1 (Moloney et al., 2000b; Panin et al., 2002) ont dû être surexprimés dans des systèmes hétérologues pour les purifier et étudier leur *O*-fucosylation. Ainsi, selon la nature de la protéine (sécrétée ou membranaire), sa biodisponibilité et la facilité de sa purification, la démarche expérimentale doit être adaptée. Une alternative intéressante consiste à produire un ELD d'une protéine donnée de manière isolée (Takeuchi et al., 2017) dans un système procaryote et de montrer sa capacité ou non à être *O*-fucosylé *in vitro* par POFUT1. Si la simplicité et la rapidité de production d'une telle protéine recombinante (> 10 kDa) facilitent grandement l'analyse de sa capacité à être *O*-fucosylée ou non, l'amplification par PCR de la séquence codante d'un ELD donné est difficile du fait de la grande similarité des ELDs entre eux. Nous proposons dans l'article suivant de contourner ces problèmes d'amplification en utilisant une méthode d'hybridation d'oligonucléotides, ne nécessitant ni PCR ni de disposer de l'ADNc codant la protéine d'intérêt.

Dans notre étude sur la *O*-fucosylation de la protéine WIF1 murine (chapitre 2), nous avons testé rapidement la capacité des ELDs isolés III et V à être *O*-fucosylés ou non par POFUT1 avant de poursuivre par des productions et analyses plus complexes sur protéine entière. Comme évoqué plus haut, nous avons hybridé 4 paires d'oligonucléotides



complémentaires pour obtenir un fragment d'ADN double brin à extrémités cohésives portant la séquence codante des ELDs III et V de WIF1. Le clonage orienté de ces hybridés est finalement réalisé dans un vecteur d'expression procaryote. Après expression dans le périplasme, les ELDs sont purifiés et soumis à une incubation avec l'enzyme POFUT1 et un GDP-fucose porteur d'un groupement azido. La chimie click, qui permet de greffer sur le groupement azido une molécule détectable par blot (comme la biotine alcyne), est une technique rapide et peu coûteuse, que nous avons mis en œuvre pour mettre en évidence une nouvelle protéine *O*-fucosylée. En effet, les résultats obtenus précédemment sur la protéine WIF1, nous ont amenés à nous interroger sur la possibilité d'étendre cette approche globale à l'étude rapide de n'importe quel ELD de n'importe quelle protéine. De plus, plusieurs questionnements sur la sensibilité de la technique de chimie click méritaient d'être éclaircis.

Le but de l'article qui suit a donc été de mieux appréhender les avantages et limites de cette nouvelle approche, destinée à évaluer la capacité d'un ELD d'intérêt à être *O*-fucosylé ou non. Pour cela, nous avons opté pour la production de deux ELDs connus pour être *O*-fucosylés, à savoir les ELDs 12 et 26 du récepteur NOTCH1 murin. Ce choix a été motivé par la différence d'affinité de l'enzyme POFUT1 pour ces deux ELDs; l'ELD 12 présentant une affinité 100 fois inférieure à celle de l'ELD 26 (Li et al., 2017). Pourtant, la superposition d'un modèle structural de l'ELD 12 de souris avec la structure 3D de l'ELD 26 murin montre une grande similarité. En accord avec la littérature (Li et al., 2017), nos résultats utilisant cette nouvelle approche de chimie click montrent bien une propension plus faible de l'ELD 12 à être *O*-fucosylé par rapport à l'ELD 26. Ces analyses ont été confirmées par spectrométrie de masse, une technique plus sensible qui nous a permis aussi de quantifier avec précision la proportion d'ELDs réellement *O*-fucosylés.

Pour conclure, cette étude montre une nouvelle approche permettant de générer rapidement et efficacement n'importe quel ELD d'une glycoprotéine issue de n'importe quelle espèce afin d'en étudier le caractère *O*-fucosylable *in vitro*. L'analyse par chimie click associée à des techniques de blot a montré que cette approche est suffisamment sensible pour mettre en évidence des différences d'efficacité de POFUT1 à *O*-fucosyler les ELDs 12 et 26 de NOTCH1, connus pour se lier à l'enzyme avec des affinités différentes. Enfin, même si cette approche demeure moins sensible et moins quantitative que la spectrométrie de masse, elle permet néanmoins de discriminer un ELD facilement *O*-fucosylable d'un autre qui le serait beaucoup moins ou pas du tout.



## ***In vitro* acellular method to reveal *O*-fucosylation on EGF-like domains**

**Pennarubia Florian<sup>1</sup>, Pinault Emilie<sup>1,2</sup>, Maftah Abderrahman<sup>1,\*</sup> and Legardinier Sébastien<sup>1,\*,#</sup>**

<sup>1</sup> Univ. Limoges, PEIRENE, EA 7500, Glycosylation and cell differentiation, F-87000 Limoges, France.

<sup>2</sup> Univ. Limoges, BISCEM Mass Spectrometry Platform, F-87025 Limoges, France

\* A. Maftah and S. Legardinier are considered co-last authors and contributed equally to this work

#To whom correspondence should be addressed: Tel: +33555457792; Fax: +33555457653;

e-mail: [sebastien.legardinier@unilim.fr](mailto:sebastien.legardinier@unilim.fr)

**Running head:** *O*-fucosylation of EGF-like domains (EGF-LD)

**Keywords:** Click chemistry / EGF-like / MRM-MS / *O*-fucosylation / POFUT1

**Supplementary Data Included:** Figures S1-S9 and Tables 1 and 2



## ABSTRACT

A hundred of human proteins have one or more EGF-like domains (EGF-LD) bearing the *O*-fucosylation consensus motif C<sup>2</sup>X4(S/T)C<sup>3</sup> but to date, only a few of them have been shown to be *O*-fucosylated. The protein *O*-fucosyltransferase (POFUT1) specifically recognizes correctly folded EGF-LD of the human EGF (hEGF) type and transfers fucose on serine or threonine residue within the *O*-fucosylation motif. Here, we propose a strategy for a rapid screening for ability of any EGF-LD to be *O*-fucosylated, using copper-catalyzed azide-alkyne cycloaddition (CuAAC). By an oligonucleotide hybridization approach, double-stranded fragments encoding any EGF-LD can be first rapidly cloned into the prokaryotic vector pET-25b to promote its targeting to periplasm and formation of the three conserved disulfide bonds. After protein production and purification, an *in vitro* POFUT1-mediated *O*-fucosylation can be performed with azido GDP-fucose. Successful transfer of *O*-fucose is finally revealed by blotting technique after CuAAC. In this study, we specially focused on mouse NOTCH1 EGF12 and EGF26, which are both known to be *O*-fucosylated although having different binding affinities towards POFUT1. Indeed, we clearly showed here that addition of *O*-fucose by POFUT1 was much more efficient for EGF26 than for EGF12. This experimental approach is rapid and sufficiently sensitive to reveal propensity of any EGF-LD to be *O*-fucosylated; it is thus useful prior to perform structure-function studies on target proteins containing one or several EGF-LD.



## INTRODUCTION

The *O*-fucosylation is a rare post-translational modification widely found in EGF-LD (Harris and Spellman 1993) and in thrombospondin type 1 repeats (TSR) (Hofsteenge et al. 2001) of secreted or transmembrane glycoproteins. The addition of *O*-linked fucose on EGF-LD and TSR is catalyzed by protein *O*-fucosyltransferase 1 (POFUT1)(Lira-Navarrete et al. 2011) and 2 (POFUT2)(Chen et al. 2012) respectively, which are both soluble endoplasmic reticulum (ER)-resident proteins (Luo and Haltiwanger 2005). The study of EGF-LD is of great interest due to their participation in numerous protein-protein interactions (Kansas et al. 1994, Moloney et al. 2000, Rao et al. 1995). They are small domains of 30-40 residues including six highly conserved cysteine residues forming three disulfide bridges (C<sup>1</sup>-C<sup>3</sup>, C<sup>2</sup>-C<sup>4</sup> and C<sup>5</sup>-C<sup>6</sup>) (Wouters et al. 2005), some of them can be *O*-fucosylated by POFUT1 (Wang et al. 2001). This enzyme can allow hydrolysis of the activated nucleotide sugar GDP-fucose in absence of EGF-LD. The presence of EGF-LD can allow fucose transfer to serine or threonine of *O*-fucosylation consensus motif C<sup>2</sup>-X4-(S/T)-C<sup>3</sup> (where C<sup>2</sup> and C<sup>3</sup> correspond to conserved cysteines 2 and 3, respectively). This addition of *O*-fucose by POFUT1 seems to be restricted to correctly folded EGF-LD (Wang and Spellman 1998) and particularly to those of the human EGF-like (hEGF) type (Li et al. 2017), defined as containing 8-9 residues between C<sup>5</sup> and C<sup>6</sup> (Wouters et al. 2005). Three main domains in EGF-LD are involved in interaction with residues of POFUT1, namely the C<sup>2</sup>-C<sup>3</sup> and C<sup>5</sup>-C<sup>6</sup> subdomains and at a lesser extent the C<sup>1</sup>-C<sup>2</sup> loop (Li et al. 2017). Except at the position C<sup>2</sup>+3, X residues of the *O*-fucosylation motif were shown to establish hydrogen bonds with highly conserved residues of mouse POFUT1, which are deeply buried between its two Rossmann-fold domains (Li et al. 2017, Lira-Navarrete et al. 2011). In addition, the cysteine C<sup>3</sup> of interacting EGF-LD is involved in a stabilizing sulphur-hydrogen bond. The X residue in position C<sup>2</sup>+3 is often a glycine as found in the *O*-fucosylation motif C<sup>2</sup>FNGGTC<sup>3</sup> of Notch1 EGF26 and is known to participate in van der Waals interactions with Y<sup>78</sup> of mouse POFUT1. However, other residues with a longer side chain were shown to induce a steric clash leading to a decreased binding to POFUT1 as it is the case for EGF12 (C<sup>2</sup>QNDATC<sup>3</sup>)(Li et al. 2017). Even if they bind to POFUT1 with



different affinities, EGF12 and 26 are both known to be *O*-fucosylated and *O*-linked fucose is critical for Notch signaling (Rampal et al. 2005). Except for Notch receptor and its Delta-like and Jagged ligands, the presence of *O*-fucose on EGF-LD has been experimentally demonstrated only for a limited number of glycoproteins. Four natural glycoproteins in human were demonstrated to be *O*-fucosylated namely human blood coagulation factor VII, IX and XII (Harris et al. 1992, Hofsteenge et al. 2001, Nishimura et al. 1992) and urokinase (Kentzer et al. 1990). In all other cases, the *O*-fucosylation was revealed on full-length or truncated multi-EGF-LD recombinant proteins such as for Notch1 (Moloney et al. 2000), its ligands Delta and Jagged (Panin et al. 2002), Notch3 (Arboleda-Velasquez et al. 2005), Cripto (Schiffer et al. 2001), Agrin (Kim et al. 2008) and Amaco (Gebauer et al. 2008). Mass spectrometry is the most widely used technique to study protein *O*-fucosylation and overcomes lack of specific antibodies or commercial lectins recognizing *O*-fucose. No enzymatic method is available to specifically remove *O*-fucose and  $\beta$ -elimination removes *O*-glycans (Stults and Cummings 1993) without specificity towards *O*-fucosylglycans. In numerous studies, metabolic labeling with radioactive  $^{14}\text{C}$ -fucose (Bjoern et al. 1991) or  $^3\text{H}$ -Fucose (Rampal et al. 2005) was used to highlight protein *O*-fucosylation. However, another method using peracetylated 6-Alkynyl fucose and click chemistry was proposed as an alternative method avoiding radioactivity (Al-Shareffi et al. 2013). In most cases, GlcNAc transferase I-deficient (Lec1) CHO cells (Stanley et al. 1975a, Stanley et al. 1975b), which are unable to synthesize complex- or hybrid-type *N*-glycans, were used to study protein *O*-fucosylation (Al-Shareffi et al. 2013, Kim et al. 2008, Shi et al. 2007).

Based on presence of the *O*-fucosylation motif  $\text{C}^2\text{-X4-(S/T)-C}^3$ , 87 secreted and transmembrane proteins containing from one to more than forty EGF-LD were predicted in humans to be modified by POFUT1 (Schneider et al. 2017). To date, less than 15 natural or recombinant multi-EGF-LD proteins (Schneider et al. 2017) are demonstrated experimentally to be *O*-fucosylated, partly due to technical difficulties. Our present study proposes an original, sensitive and specific strategy to quickly check the *O*-fucosylation status of these EGF-LD. The experimental device includes *in vitro* POFUT1-mediated *O*-fucosylation and CuAAC. The relevance of results was confirmed by multiple reactions monitoring-mass



spectrometry (MRM-MS). To test efficiency, specificity and sensibility of this approach, we compared ability of mouse POFUT1 to add *O*-fucose to mouse Notch1 EGF12 and EGF26, known to be *O*-fucosylated (Rampal et al. 2005) and having different binding affinities towards POFUT1 (Li et al. 2017). This whole approach is particularly useful to reveal EGF-LD *O*-fucosylation but further studies are still required to determine functional consequences of *O*-fucose in structure and/or functions of secreted or transmembrane EGF-LD-containing proteins.

## RESULTS AND DISCUSSION

**Strategies to reveal protein *O*-fucosylation.** The two multi-step complementary approaches to reveal EGF-LD *O*-fucosylation are presented in Fig. 1. First, double-stranded DNA fragments encoding any EGF-LD can be rapidly produced using four complementary and overlapping pairs of oligonucleotides (A-B, C-D, E-F, G-H) with cohesive ends (step 1, Fig. 1). This oligonucleotide hybridization technique is an efficient alternative to classical cloning if total cDNAs are not available. This avoids PCR that is not effective in amplifying specifically coding sequence of a targeted EGF-LD rather than another EGF-LD. After overnight hybridization, double-stranded DNA fragments encoding an EGF-LD is cloned into the prokaryotic expression vector pET-25b vector using unique cloning sites (step 2, Fig. 1). This vector allows protein expression and targeting to *E.coli* periplasm, where disulfide bonds can be formed (step 3, Fig. 1). After protein extraction from induced BL21 cells, the EGF-LD extended with a C-terminal polyhistidine tag can be purified by nickel affinity chromatography (step 4, Fig. 1). These purified proteins, namely wild-type EGF-LD and its counterpart mutated on the *O*-fucosylation site (T/A mutation), are incubated with recombinant POFUT1 and unlabeled or azido (N3) GDP-fucose (step 5, Fig. 1). Then, successful addition of *O*-fucose is revealed after CuAAC of alkynyl biotin to the azido-*O*-fucosylated EGF-LD (step 6a, Fig. 1) by blotting technique (step 7a, Fig. 1). If unlabeled GDP-fucose is alternatively used, glycosyltransferase reactions are subjected to reduction, alkylation and in-solution digestion (step 6b, Fig. 1) and analyzed by MRM-MS (step 7b, Fig. 1).



**Comparison of biochemical properties of Notch1 EGF12 and 26 purified from *E.coli*.** Theoretically, any EGF-LD of any species can be cloned, produced and purified in a week using our strategy. To test efficiency, specificity and sensitivity of CuAAC approach, we chose to produce mouse NOTCH1 EGF12 and 26, known to be *O*-fucosylated on T<sup>466</sup> and T<sup>997</sup> respectively, and their T/A mutants. Both were cloned using the oligonucleotide hybridization technique (Supplementary Fig.1). Mouse EGF12 and EGF26 are 37-residue long EGF-LD sharing 37.84 % identity and possessing the *O*-fucosylation motif C<sup>2</sup>X4(S/T)C<sup>3</sup> (Fig. 2a). An automated homology model was generated for mouse EGF12 using the Swiss-model server. The relevance of the model is supported by a Global Model Quality Estimation (GMQE) of 0.99, using as a template the 3D structure of human EGF12 (McMillan et al. 2017) which shares 94.49% identity with the mouse EGF12. Using Matchmaker of CHIMERA, a structural model for mouse EGF12 was superimposed to X-ray structure of mouse EGF26 (Fig. 2b). The C<sup>2</sup>-C<sup>3</sup> subdomains, known to interact with key residues of mouse POFUT1 (Li et al. 2017), as well as the C<sup>5</sup>-C<sup>6</sup> subdomains were perfectly superimposed. However, the flexible hydrophilic C<sup>1</sup>-C<sup>2</sup> loop, known to “accommodate” numerous different structures, and the C<sup>3</sup>-C<sup>4</sup> subdomain, known for its lack of interaction with POFUT1, were a little bit more divergent between these two EGF-LD.

**Production, purification and characterization of soluble mouse POFUT1.** The co-crystallization of mouse POFUT1 with Notch1 EGF26 (Li et al. 2017) shows that this EGF-LD mainly interacts by establishing hydrogen bonds (Li et al. 2017) through its C<sup>2</sup>-C<sup>3</sup> subdomain (C<sup>2</sup>FNGGTC<sup>3</sup>) with key residues (G<sup>47</sup>, M<sup>46</sup>, N<sup>51</sup> and N<sup>151</sup>) located in the deep groove of POFUT1 (Fig. 2c). Alignment and superimposition of known X-ray structures of POFUT1 from different species (human, mouse and *C.elegans*) showed that these key residues were conserved at the same spatial positions (Fig. 2d). We also report high conservation of these key residues among widely studied species (Supplementary Fig. 2). This suggests that any POFUT1 from these species could add *O*-fucose on mouse N1 EGF12 and 26, as previously demonstrated by successful addition of *O*-fucose on mouse EGF12 by *C.elegans* POFUT1 (Lira-Navarrete et al. 2011). Although commercial human POFUT1 was available, we chose to produce



recombinant mouse POFUT1 as a secreted polyhistidine-V5-tagged protein by stably transfected Flp-In™ CHO cells. The cDNA encoding mouse POFUT1 without its endogenous signal peptide and KDEL-like endoplasmic reticulum (ER)-retention signal was first cloned into homemade pSec-NtermHis6 vector. Then, a V5 epitope tag was added after oligonucleotide hybridization downstream from the N-terminal polyhistidine tag (Supplementary Fig. 3). Recombinant mouse POFUT1 was purified from supernatants of stable CHO cells culture media using Ni-NTA purification and the purity of the protein has been confirmed by Coomassie-blue stained polyacrylamide gels (Fig. 2e). As expected, recombinant mouse POFUT1 was specifically recognized by both anti-POFUT1 and anti-V5-HRP antibodies (Fig. 2e). After 20 hours of incubation with GDP-fucose at 37°C, hydrolytic activity of the enzyme towards this nucleotide-sugar was indirectly quantified. Indeed, inorganic phosphate, which was removed from the leaving GDP by a specific phosphatase, was quantified using Malachite reagents by spectrophotometry. The specific activity of  $0.85 \text{ pmol} \cdot \text{min}^{-1} \cdot \mu\text{g}^{-1}$  that we determined for recombinant mouse POFUT1 (Fig. 2f) was consistent with that of commercial human POFUT1 (R&D Systems Inc., Minneapolis, MN).

**Revelation of *O*-fucosylation of Notch1 EGF12 and 26 by CuAAC approach.** As an alternative to radioactivity, metabolic labeling and mass spectrometry, we first chose to use CuAAC approach to reveal *O*-fucosylation. This approach consists in a copper-catalyzed cycloaddition of alkynyl biotin to azido-fucose anchored to EGF-LD, after incubation with recombinant POFUT1 and GDP-azido-fucose. Revelation of an *O*-linked azido-fucosylated EGF-LD in complex with biotin was performed after blotting with streptavidin-HRP. First, purified WT and T/A mutant EGF12 and 26 were analyzed by Coomassie-blue stained polyacrylamide gels (Fig. 3a, upper panel). The four purified proteins exhibited equivalent purity but migrated at higher molecular weights (>10kDa) than expected (around 8 kDa). More precisely, higher apparent MW were observed for WT and T997A EGF26, which are less negatively charged than WT and T466A EGF12 (12 *versus* 16 negative charges). In a first attempt of 1 hour incubation with recombinant POFUT1, WT EGF26 was revealed after 30 seconds of light exposure. As a negative control, no signal was obtained with T997A mutant. More surprisingly, no signal was detected



for WT EGF12 (Fig. 3a), even with longer light exposure. Considering the lower affinity of EGF12 for POFUT1 (Li et al. 2017), incubation time with recombinant enzyme was increased up to 20 hours before performing CuAAC and analyses by blotting technique. In these experimental conditions, signals were finally revealed for both WT EGF12 and EGF26. However, signal appeared much more rapidly for EGF26 than for EGF12, after 5 and 90 seconds light exposure respectively. Even with longer exposures (10 minutes), no signals were detected for the T/A mutants. In addition, this experimental approach was sufficiently sensitive since different EGF-LD known to bind to POFUT1 with different affinities were discriminated. These data clearly gave insight of sensitivity and specificity of CuAAC approach.

**Detection of *O*-fucosylated Notch1 EGF12 and 26 by MRM-MS.** To confirm CuAAC results, MRM-MS analyses were performed. The same amounts of WT or mutated EGF12 and 26 were incubated with POFUT1 and unlabeled GDP-fucose. In all cases, peaks corresponding to the non-fucosylated peptides were detected (Fig. 3b, upper panel). However, peak intensities were lower for WT EGF12 and 26 than for their T/A counterparts, revealing absence of *O*-fucose on a part of GluC WT peptides harboring the *O*-fucosylation site. Peaks corresponding to *O*-fucosylated peptides (Fig.3b, lower panel) were, as expected, only seen for WT EGF12 and 26, confirming the presence of *O*-fucose on both EGF-LD on T<sup>466</sup> and T<sup>997</sup> respectively. Consistent with CuAAC results, a difference between EGF12 and 26 was nevertheless observed. Indeed, the ratio between *O*-fucosylated and non-*O*-fucosylated peptides for EGF12 and 26 was  $0.069 \pm 0.019$  and  $1.258 \pm 0.203$  respectively ( $p$  value = 0.0096), probably because of the lower affinity of EGF12 for POFUT1 compared to EGF26. HPLC-ESI-TOFMS analyses (Supplementary Fig.7) and reverse-phase HPLC chromatograms (Supplementary Fig.8-9) were consistent with the fact that the lower *O*-fucosylation for WT EGF12 did not result from improperly folded isomers since only one elution peak was detected. Consistent with our previous results obtained with CuAAC, MRM-MS confirmed that EGF26 was easier to *O*-fucosylate by POFUT1 than EGF12. Finally, MRM-MS approach is a more accurate but expensive method than CuAAC to reveal specific addition of *O*-fucose on EGF-LD.



## CONCLUSION

Our experimental approach to highlight protein *O*-fucosylation of EGF-LD is a quick and reliable method, whose specificity and sensitivity were shown by studying two EGF-LD, namely NOTCH1 EGF12 and 26, which differ by their binding affinity to POFUT1 and consequently by their propensity to be *O*-fucosylated *in vitro*. This type of approach could be also carried out with POFUT2 which adds *O*-fucose to TSR (Luo et al. 2006), as well as with POGLUT1 (Fernandez-Valdivia et al. 2011, Takeuchi et al. 2012) and EOGT (Sakaidani et al. 2011, Varshney and Stanley 2017), allowing specific anchorage of *O*-linked glucose and *O*-linked GlcNAc respectively on EGF-LD.

## MATERIALS AND METHODS

**Glycosyltransferase reaction.** Before CuAAC experiments, glycosyltransferase reactions were carried out with 1  $\mu$ g of POFUT1, mixed with 2 nmoles of GDP-azido-fucose (R&D Systems Inc., Minneapolis, MN, USA) as recommended by the manufacturer (R&D Systems) and 2  $\mu$ g of purified EGF-LD in 25  $\mu$ l of reaction buffer (25 mM Tris, 5 mM CaCl<sub>2</sub>, 10 mM MnCl<sub>2</sub>, pH 7.5) and incubated for 1 h or 20 h at 37°C. For mass spectrometry analysis, 2.5  $\mu$ g of POFUT1 were incubated with 5  $\mu$ g of purified EGF-LD and 2 nmoles GDP-fucose in 20  $\mu$ l of reaction buffer and incubated 20 h at 37°C.

**Click chemistry reactions.** CuAAC was performed using 1.25 mM CuCl<sub>2</sub>, 2.5 mM ascorbic acid and 0.125 mM alkynyl biotin (R&D Systems Inc., Minneapolis, MN, USA), directly added to glycosyltransferase reactions. The mixture was incubated in the dark for 1 h at room temperature.

**Targeted Mass Spectrometry.** Proteins were reduced, alkylated and digested by Glu-C. Resulting peptides were analyzed with a nanoLC 425 in micro-flow mode (Eksigent, Dublin, CA, USA) coupled to a TTOF5600+ mass spectrometer (SCIEX, Framingham, USA) in Information-Dependent Acquisition mode. ProteinPilot 5.0 (SCIEX) was applied to search against the recombinant protein sequence database, and the MRM transition list was established using Skyline 3.5.0 (MacCoss Lab, University of



Washington, Seattle, WA, USA). *O*-fucosylation was added *in silico* at the expected position with PeakView software (SCIEX) and *m/z* of precursor and fragments were calculated. The data were acquired in high-Resolution MRM mode and processed with MultiQuant Software 3.0.1 (SCIEX). Areas were collected for the same most intense fragment of fucosylated and non-fucosylated peptides and a ratio of fucosylation was calculated.

**Supplementary Data.** Refer to Web version on PubMed Central for supplementary material.

### **Funding**

This work was supported by a French Ministry of Higher Education and Research doctoral fellowship to FP. The work was partly funded by the project GlyCanColor (CORC 2016).

### **Acknowledgements**

We would like to thank M. Romain Marcillou for his help and DNA sequencing and Pr Philippe Bressolier for RP-HPLC purification.

### **Conflict of interest statement**

The authors declare no competing interests

### **Abbreviations**

CHO, Chinese Hamster Ovary cells

CuAAC, Copper-catalyzed Azide-Alkyne Cycloaddition

EGF-LD, Epidermal Growth Factor-Like domain

ER, Endoplasmic Reticulum

EOGT, EGF domain specific *O*-linked *N*-Acetylglucosamine Transferase

GlcNAc, *N*-acetylglucosamine

GMQE, Global Model Quality Estimation

MRM-MS, Multiple Reactions Monitoring - Mass Spectrometry



POFUT1, Protein *O*-fucosyltransferase 1

POGLUT1, Protein *O*-glucosyltransferase 1

TSR, ThromboSpondin type 1 Repeats

## **Figure legends**

**Figure 1: Experimental approaches to highlight *O*-fucosylation of EGF-LD.** Using hybridization technique (step1), four pairs of complementary oligonucleotides with cohesive ends encoding any EGF-LD are cloned in a prokaryotic expression vector such as pET-25b (step 2). After induction of protein expression (step 3) and nickel affinity purification of His<sub>6</sub> tagged protein (step 4), incubation of purified WT EGF-LD or its counterpart T/A mutated on the *O*-fucosylation site can be performed with unlabeled or azido (N<sub>3</sub>) GDP-fucose and recombinant POFUT1 (step 5). Successful addition of *O*-fucose can be revealed either after CuAAC of alkynyl biotin to the azido-*O*-fucosylated EGF-LD (step 6a) and blotting technique (step 7a) or after reduction, alkylation and in-solution digestion (step 6b) and MRM-MS (step 7b).

**Figure 2:** (A) Using Clustal Omega, mouse WT EGF12 and 26 were aligned: their *O*-fucosylation consensus motifs are in bold, cysteines in orange and *O*-fucosylable threonine in cyan. (B) Homology model of mouse EGF12 (brown) was generated by using Swiss-model Server with X-ray structure of human EGF12 (PDB 5L0R) as a reference template. Using Matchmaker of CHIMERA, this model was superimposed with mouse NOTCH1 EGF26 (blue) (PDB 5KY4), extracted from the X-ray structure of mouse POFUT1 co-crystallized with mouse Notch1 EGF26 (Li et al. 2017). The comparison of these two structures, partly based on the conservation of the three disulfide bonds (C<sup>1</sup>-C<sup>3</sup>, C<sup>2</sup>-C<sup>4</sup>, C<sup>5</sup>-C<sup>6</sup>) (yellow), shows similar C<sup>2</sup>-C<sup>3</sup> and C<sup>5</sup>-C<sup>6</sup> subdomains and a more divergent C<sup>1</sup>-C<sup>2</sup> loop. The threonine side chain of the *O*-fucosylation consensus motif is shown in cyan. (C) X-ray co-crystallized structure of POFUT1 (white) and EGF26 (blue) was used to show POFUT1 key residues (red) involved in interaction with



EGF26 (M<sup>46</sup>, G<sup>47</sup>, N<sup>51</sup> and N<sup>151</sup>) or able to induce a steric clash (Y<sup>78</sup>). **(D)** Using Matchmaker of CHIMERA, X-ray structures of *Mus musculus* (*Mm*) (white) (PDB 5KY4), *Homo sapiens* (*Hs*) (light grey) (PDB 5UXH) and *Caenorhabditis elegans* (*Ce*) (dark grey) (PDB 3ZY5) POFUT1 were superimposed. Superimposition shows the same spatial location and the same orientation for side chains of key residues cited above for *Mm* (red), *Hs* (orange) and *Ce* (brown). **(E)** Recombinant WT POFUT1 was produced as a secreted N-terminal His6 V5-tagged protein and was purified from stable Flp-In<sup>TM</sup> CHO cells. As revealed by Coomassie blue-stained polyacrylamide gels and Western blot, the main recombinant eluted protein, with more than 90% purity, was specifically recognized by both anti-V5-HRP and anti-POFUT1 antibodies. **(F)** After 20 h of incubation with POFUT1 and GDP-fucose, Malachite Green Phosphate Detection kit (R&D Systems Inc., Minneapolis, MN) was used to determine specific hydrolytic activity of recombinant mouse POFUT1.

**Figure 3: CuAAC and MRM-MS to reveal EGF-LD O-fucosylation.** **(A)** The same amounts of recombinant WT EGF12 and EGF26, as well as their mutant counterparts (T466A and T997A) were used as shown on Coomassie blue-stained polyacrylamide gel (upper panel). After 1h incubation of EGF-LD at 37°C with POFUT1 and GDP-azido-fucose, CuAAC with alkynyl biotin was performed and a blotting technique using HRP-streptavidin was used to reveal EGF-LD with O-linked azido-fucose. The positive signal revealed successful addition of O-fucose on WT EGF26 by POFUT1 in 1h. The same technique was used after 20 h incubation, to highlight positive signals for both WT EGF12 and 26. **(B)** After incubation of WT and mutated EGF12 and 26 with mouse POFUT1 and GDP-fucose, reactions were reduced, alkylated and digested by GluC. Resulting peptides were then analyzed by MRM-MS. Peaks are shown for non-fucosylated peptides (upper panels) and O-fucosylated peptides (lower panels). Nested peaks correspond to the superposition of the MS2 ions signals, obtained during the run time.



## REFERENCES

- Al-Shareffi E, Chaubard JL, Leonhard-Melief C, Wang SK, Wong CH, Haltiwanger RS. 2013. 6-alkynyl fucose is a bioorthogonal analog for O-fucosylation of epidermal growth factor-like repeats and thrombospondin type-1 repeats by protein O-fucosyltransferases 1 and 2. *Glycobiology*, 23:188-198.
- Arboleda-Velasquez JF, Rampal R, Fung E, Darland DC, Liu M, Martinez MC, Donahue CP, Navarro-Gonzalez MF, Libby P, D'Amore PA, *et al.* 2005. CADASIL mutations impair Notch3 glycosylation by Fringe. *Hum Mol Genet*, 14:1631-1639.
- Bjoern S, Foster DC, Thim L, Wiberg FC, Christensen M, Komiyama Y, Pedersen AH, Kisiel W. 1991. Human plasma and recombinant factor VII. Characterization of O-glycosylations at serine residues 52 and 60 and effects of site-directed mutagenesis of serine 52 to alanine. *J Biol Chem*, 266:11051-11057.
- Chen CI, Keusch JJ, Klein D, Hess D, Hofsteenge J, Gut H. 2012. Structure of human POFUT2: insights into thrombospondin type 1 repeat fold and O-fucosylation. *EMBO J*, 31:3183-3197.
- Fernandez-Valdivia R, Takeuchi H, Samarghandi A, Lopez M, Leonardi J, Haltiwanger RS, Jafar-Nejad H. 2011. Regulation of mammalian Notch signaling and embryonic development by the protein O-glycosyltransferase Rumi. *Development*, 138:1925-1934.
- Gebauer JM, Muller S, Hanisch FG, Paulsson M, Wagener R. 2008. O-glycosylation and O-fucosylation occur together in close proximity on the first epidermal growth factor repeat of AMACO (VWA2 protein). *J Biol Chem*, 283:17846-17854.
- Harris RJ, Ling VT, Spellman MW. 1992. O-linked fucose is present in the first epidermal growth factor domain of factor XII but not protein C. *J Biol Chem*, 267:5102-5107.
- Harris RJ, Spellman MW. 1993. O-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology*, 3:219-224.
- Hofsteenge J, Huwiler KG, Macek B, Hess D, Lawler J, Mosher DF, Peter-Katalinic J. 2001. C-mannosylation and O-fucosylation of the thrombospondin type 1 module. *J Biol Chem*, 276:6485-6498.



- Kansas GS, Saunders KB, Ley K, Zakrzewicz A, Gibson RM, Furie BC, Furie B, Tedder TF. 1994. A role for the epidermal growth factor-like domain of P-selectin in ligand recognition and cell adhesion. *J Cell Biol*, 124:609-618.
- Kentzer EJ, Buko A, Menon G, Sarin VK. 1990. Carbohydrate composition and presence of a fucose-protein linkage in recombinant human pro-urokinase. *Biochem Biophys Res Commun*, 171:401-406.
- Kim ML, Chandrasekharan K, Glass M, Shi S, Stahl MC, Kaspar B, Stanley P, Martin PT. 2008. O-fucosylation of muscle agrin determines its ability to cluster acetylcholine receptors. *Mol Cell Neurosci*, 39:452-464.
- Li Z, Han K, Pak JE, Satkunarajah M, Zhou D, Rini JM. 2017. Recognition of EGF-like domains by the Notch-modifying O-fucosyltransferase POFUT1. *Nat Chem Biol*, 13:757-763.
- Lira-Navarrete E, Valero-Gonzalez J, Villanueva R, Martinez-Julvez M, Tejero T, Merino P, Panjekar S, Hurtado-Guerrero R. 2011. Structural insights into the mechanism of protein O-fucosylation. *PLoS One*, 6:e25365.
- Luo Y, Haltiwanger RS. 2005. O-fucosylation of notch occurs in the endoplasmic reticulum. *J Biol Chem*, 280:11289-11294.
- Luo Y, Koles K, Vorndam W, Haltiwanger RS, Panin VM. 2006. Protein O-fucosyltransferase 2 adds O-fucose to thrombospondin type 1 repeats. *J Biol Chem*, 281:9393-9399.
- McMillan BJ, Zimmerman B, Egan ED, Lofgren M, Xu X, Hesser A, Blacklow SC. 2017. Structure of human POFUT1, its requirement in ligand-independent oncogenic Notch signaling, and functional effects of Dowling-Degos mutations. *Glycobiology*, 27:777-786.
- Moloney DJ, Shair LH, Lu FM, Xia J, Locke R, Matta KL, Haltiwanger RS. 2000. Mammalian Notch1 is modified with two unusual forms of O-linked glycosylation found on epidermal growth factor-like modules. *J Biol Chem*, 275:9604-9611.
- Nishimura H, Takao T, Hase S, Shimonishi Y, Iwanaga S. 1992. Human factor IX has a tetrasaccharide O-glycosidically linked to serine 61 through the fucose residue. *J Biol Chem*, 267:17520-17525.



- Panin VM, Shao L, Lei L, Moloney DJ, Irvine KD, Haltiwanger RS. 2002. Notch ligands are substrates for protein O-fucosyltransferase-1 and Fringe. *J Biol Chem*, 277:29945-29952.
- Rampal R, Arboleda-Velasquez JF, Nita-Lazar A, Kosik KS, Haltiwanger RS. 2005. Highly conserved O-fucose sites have distinct effects on Notch1 function. *J Biol Chem*, 280:32133-32140.
- Rao Z, Handford P, Mayhew M, Knott V, Brownlee GG, Stuart D. 1995. The structure of a Ca(2+)-binding epidermal growth factor-like domain: its role in protein-protein interactions. *Cell*, 82:131-141.
- Sakaidani Y, Nomura T, Matsuura A, Ito M, Suzuki E, Murakami K, Nadano D, Matsuda T, Furukawa K, Okajima T. 2011. O-linked-N-acetylglucosamine on extracellular protein domains mediates epithelial cell-matrix interactions. *Nat Commun*, 2:583.
- Schiffer SG, Foley S, Kaffashan A, Hronowski X, Zichittella AE, Yeo CY, Miatkowski K, Adkins HB, Damon B, Whitman M, *et al.* 2001. Fucosylation of Cripto is required for its ability to facilitate nodal signaling. *J Biol Chem*, 276:37769-37778.
- Schneider M, Al-Shareffi E, Haltiwanger RS. 2017. Biological functions of fucose in mammals. *Glycobiology*, 27:601-618.
- Shi S, Ge C, Luo Y, Hou X, Haltiwanger RS, Stanley P. 2007. The threonine that carries fucose, but not fucose, is required for Cripto to facilitate Nodal signaling. *J Biol Chem*, 282:20133-20141.
- Stanley P, Caillibot V, Siminovitch L. 1975a. Selection and characterization of eight phenotypically distinct lines of lectin-resistant Chinese hamster ovary cell. *Cell*, 6:121-128.
- Stanley P, Narasimhan S, Siminovitch L, Schachter H. 1975b. Chinese hamster ovary cells selected for resistance to the cytotoxicity of phytohemagglutinin are deficient in a UDP-N-acetylglucosamine--glycoprotein N-acetylglucosaminyltransferase activity. *Proc Natl Acad Sci U S A*, 72:3323-3327.
- Stults NL, Cummings RD. 1993. O-linked fucose in glycoproteins from Chinese hamster ovary cells. *Glycobiology*, 3:589-596.
- Takeuchi H, Kantharia J, Sethi MK, Bakker H, Haltiwanger RS. 2012. Site-specific O-glucosylation of the epidermal growth factor-like (EGF) repeats of notch: efficiency of glycosylation is affected by proper folding and amino acid sequence of individual EGF repeats. *J Biol Chem*, 287:33934-33944.



Varshney S, Stanley P. 2017. EOGT and O-GlcNAc on secreted and membrane proteins. *Biochem Soc Trans*, 45:401-408.

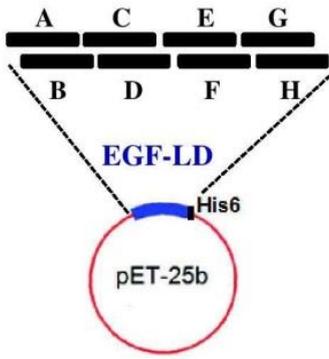
Wang Y, Shao L, Shi S, Harris RJ, Spellman MW, Stanley P, Haltiwanger RS. 2001. Modification of epidermal growth factor-like repeats with O-fucose. Molecular cloning and expression of a novel GDP-fucose protein O-fucosyltransferase. *J Biol Chem*, 276:40338-40345.

Wang Y, Spellman MW. 1998. Purification and characterization of a GDP-fucose:polypeptide fucosyltransferase from Chinese hamster ovary cells. *J Biol Chem*, 273:8112-8118.

Wouters MA, Rigoutsos I, Chu CK, Feng LL, Sparrow DB, Dunwoodie SL. 2005. Evolution of distinct EGF domains with specific functions. *Protein Sci*, 14:1091-1103.

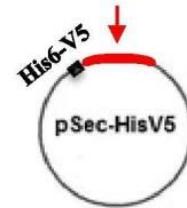


# 1. Oligonucleotide hybridization



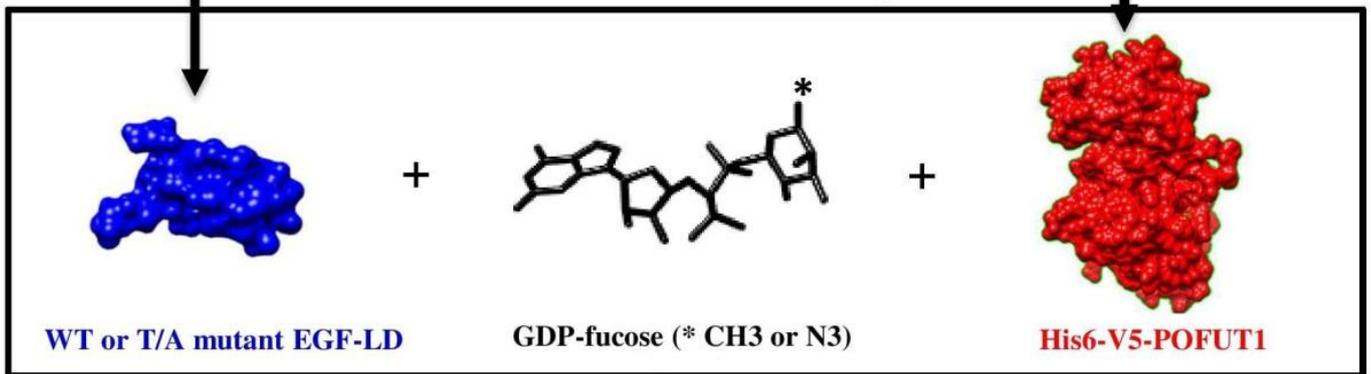
- 2. Cloning into expression vectors
- 3. Protein expression
- 4. Ni-NTA protein purification

Pofut1<sup>w/o</sup> RDEF cDNA



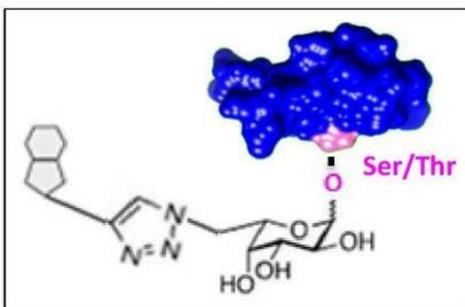
Stable Flp-In™ CHO cells  
Secreted protein

# 5. *In vitro* POFUT1-mediated *O*-fucosylation



# 6a. Copper-catalyzed azide-alkyne Cycloaddition (CuAAC)

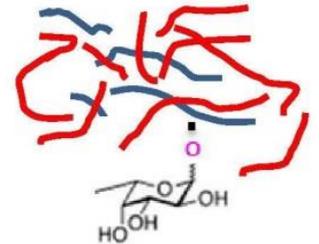
Biotin alkyne + CuCl<sub>2</sub> + Ascorbic Acid



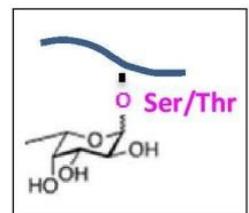
Biotinylated *O*-fucosylated EGF-LD

# 7a. HRP-streptavidin/BLOT

# 6b. Reduction, alkylation and in-solution digestion



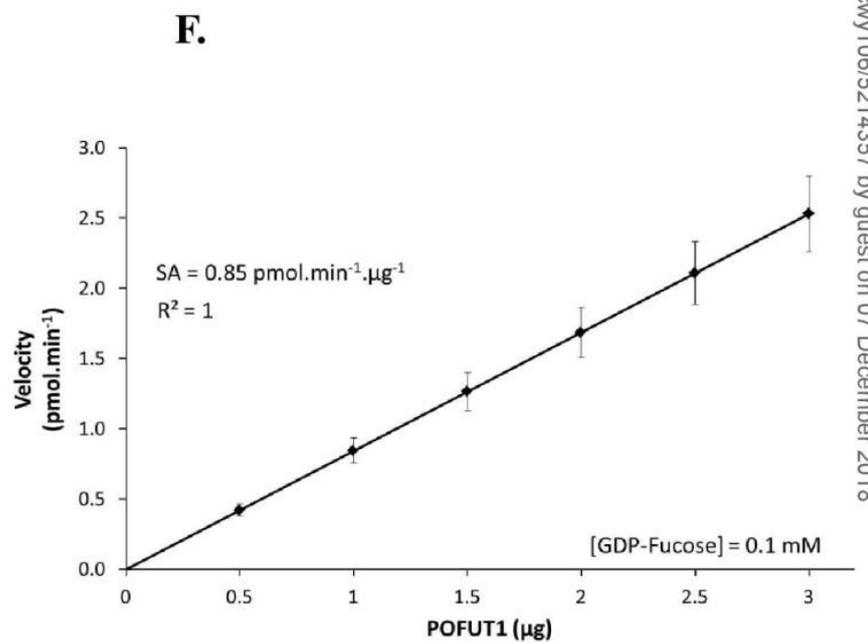
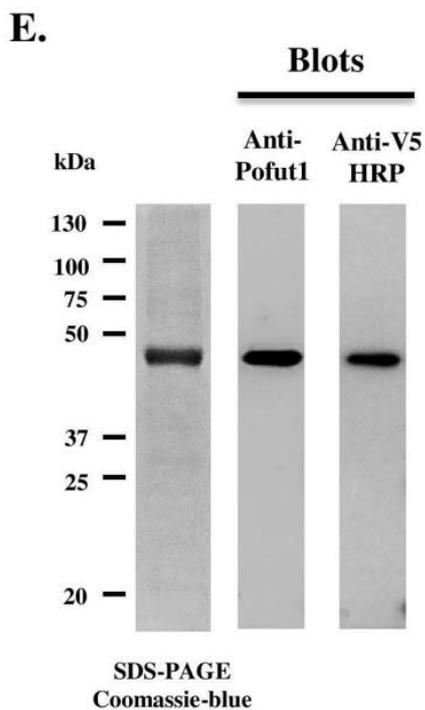
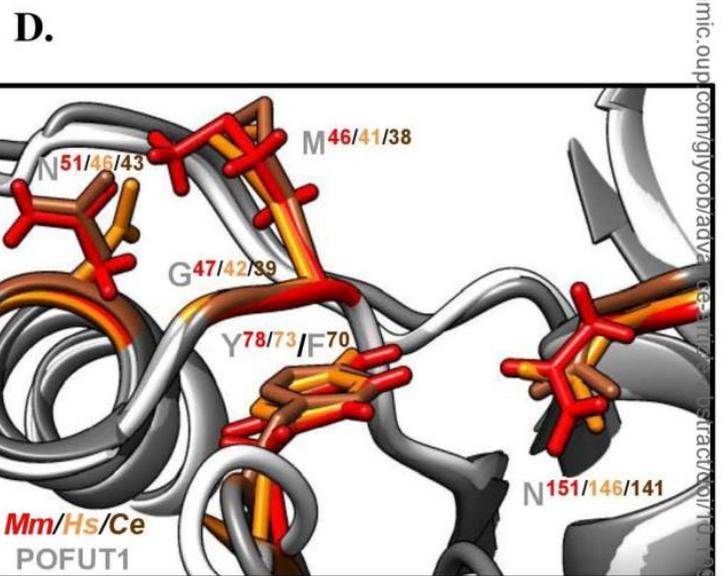
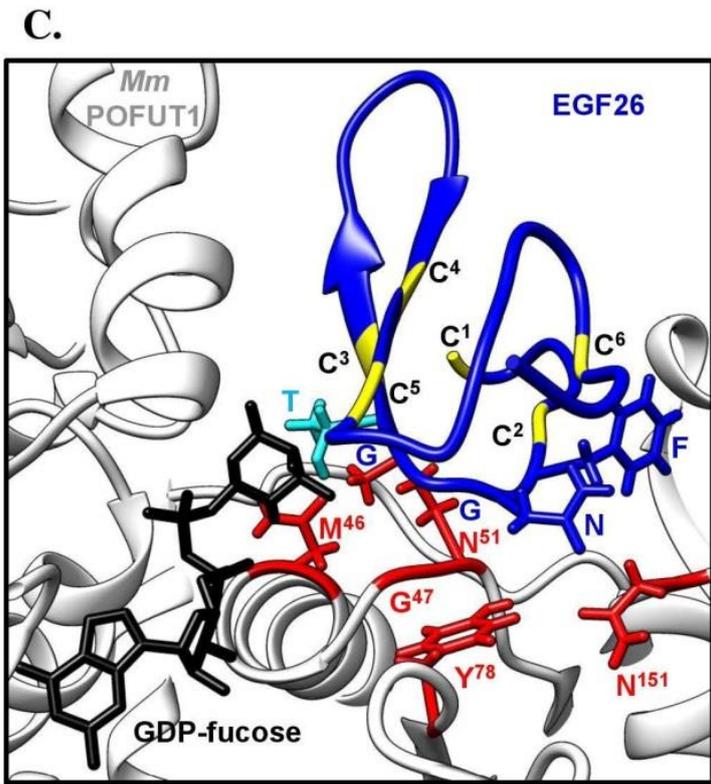
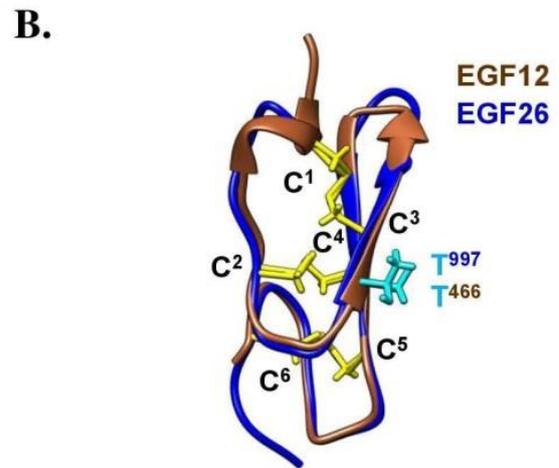
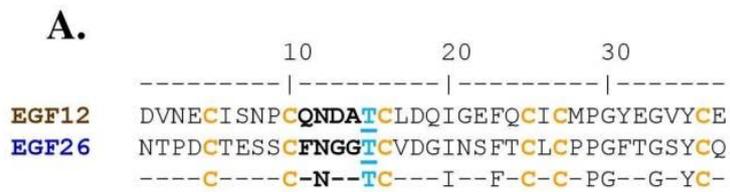
*O*-fucosylated peptide



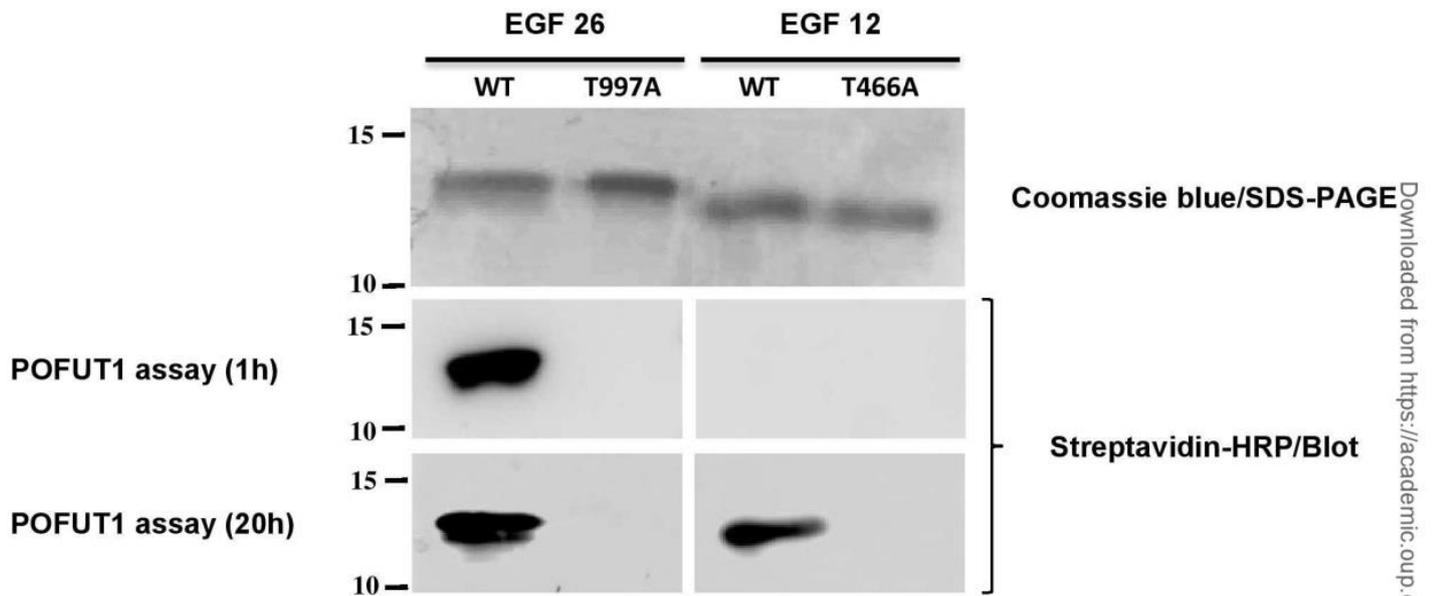
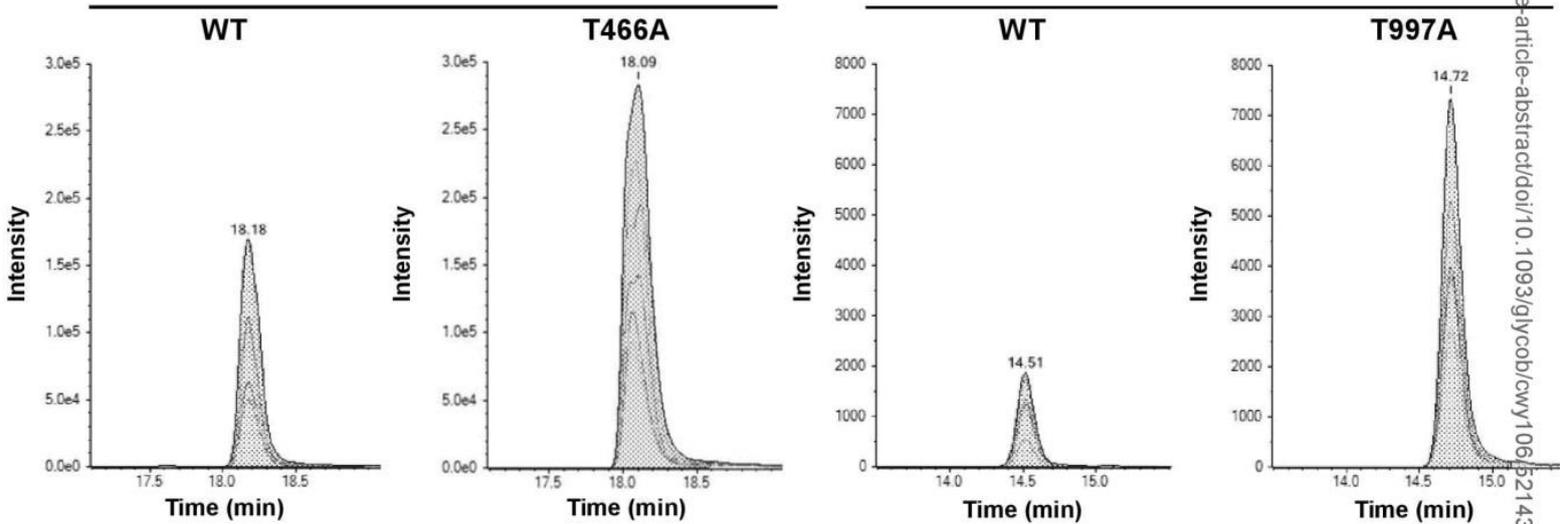
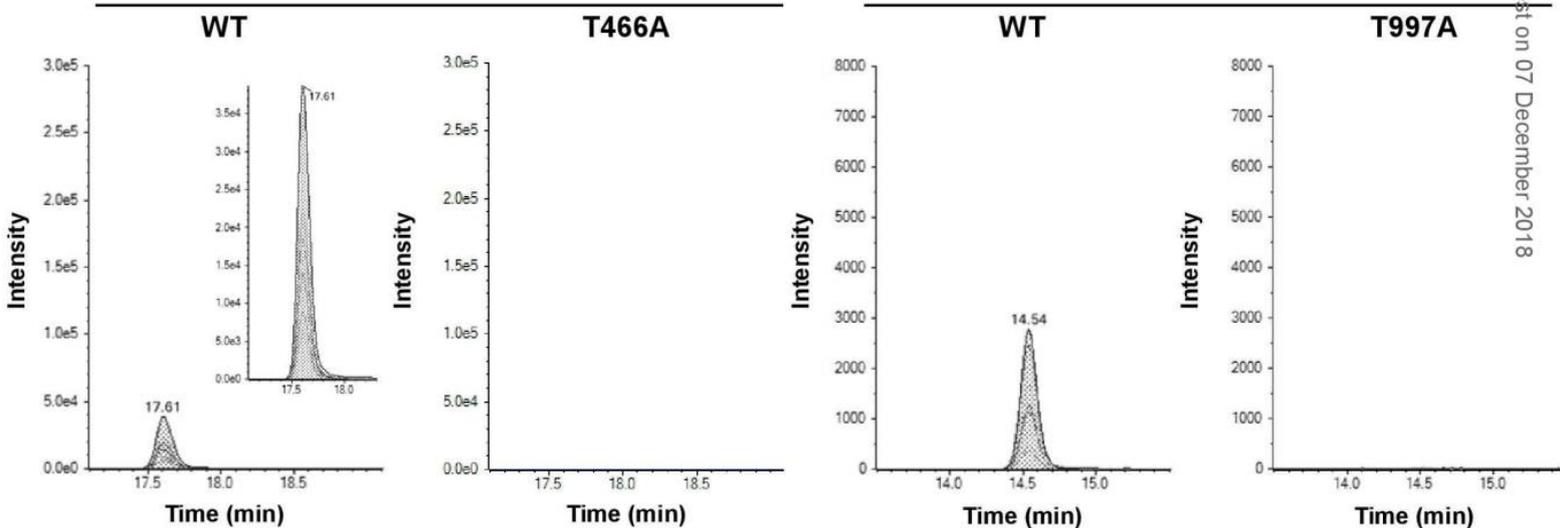
# 7b. MRM-MS

(targeted mass spectrometry)







**A.****B.****Non-fucosylated peptides****EGF 12****EGF 26****O-fucosylated peptides****EGF 12****EGF 26**



## ***In vitro* acellular method to reveal *O*-fucosylation on EGF-like domains**

**Pennarubia Florian<sup>1</sup>, Pinault Emilie<sup>1,2</sup>, Maftah Abderrahman<sup>1,\*</sup> and Legardinier Sébastien<sup>1,\*,#</sup>**

<sup>1</sup> Univ. Limoges, PEIRENE, EA 7500, Glycosylation and cell differentiation, F-87000 Limoges, France.

<sup>2</sup> Univ. Limoges, BISCEM Mass Spectrometry Platform, F-87025 Limoges, France

### **SUPPLEMENTARY MATERIALS AND METHODS**

***Plasmid constructs to produce recombinant EGF-LD.*** Eight oligonucleotides A to H were designed to form four complementary pairs of double-stranded fragments with cohesive ends, encoding EGF12 (amino acids 452-488) and EGF26 (amino acids 983-1019) of mouse NOTCH1 (NP\_032740.3) (Supplementary Fig.1). 1 µg of each internal oligonucleotide B-G was first incubated with 0.5 mM ATP and 4 units of polynucleotide kinase in its reaction buffer (Sigma-Aldrich, Saint Louis, USA) for 1 h at 37°C to phosphorylate 5'ends. After 10 min enzyme inactivation at 65°C, phosphorylated oligonucleotides B-G were separately heated for 10 min at 80°C as well as non-phosphorylated oligonucleotides A and H. For each EGF-LD, the eight denatured oligonucleotides were equally mixed (1µg each) and incubated for 10 min at 80°C just before hybridization that took place overnight during progressive decrease of temperature down to room temperature. Same procedure with appropriate oligonucleotides (3' and 4' instead of 3 and 4 oligonucleotides) was carried out to construct their mutated counterparts, namely T466A EGF12 and T997A EGF26 (Supplementary Fig.1). Hybridized oligonucleotides were cloned between *Bam*HI and *Xho*I into the pET-25b(+) vector (Novagen, Millipore, MA, USA), upstream from the C-terminal HSV and polyhistidine tags. All the sequences of pET-derived constructs were verified.

***Plasmid constructs to produce recombinant mouse POFUT1.*** By the same oligonucleotide hybridization technique as described above, commercial pSecTag/FRT/V5-His-TOPO<sup>R</sup> vector (Thermo Fisher



Scientific, Waltham, MA, USA) was first modified between *NheI* and *KpnI* restriction unique sites with eight hybridized oligonucleotides (A-H) to introduce new DNA fragment encoding the same secretory signal peptide as found in commercial vector fused to a polyhistidine tag (Supplementary Fig.2A). Mouse POFUT1 (NP\_536711.3, residues 31-389) cDNA without sequences encoding signal peptide (SP), KDEL-like motif RDEF but with a new stop codon added, was cloned after PCR amplification with I and J primers between *KpnI* and *BamHI* in the modified vector named pSec-NtermHis6 (Supplementary Fig.2B). We then chose to add a sequence encoding V5 epitope using *KpnI* unique restriction site downstream from sequence of His6 tag by oligonucleotide hybridization technique (Supplementary Fig.2C). This plasmid construct referred to as pSec-NtermHis6V5Pofut1 was used after sequence verification to produce stably transfected Flp-In<sup>TM</sup> CHO cells (Thermo Fisher Scientific) after cotransfection with the Flp recombinase vector pOG44 (Thermo Fisher Scientific).

**Protein expression and purification.** Recombinant mouse POFUT1 with its N-terminal polyhistidine-V5 tag was produced as a secreted protein by stable Flp-In<sup>TM</sup> CHO cells. After production during 96 h in serum-free Opti-MEM I medium (Thermo Fisher Scientific), proteins were recovered by centrifugation from cell culture supernatants, concentrated in binding buffer (25 mM Tris-HCl, 500 mM NaCl, 5 mM CaCl<sub>2</sub>, 20 mM imidazole, pH 7.5) and purified on Ni-NTA column by imidazole gradient using AKTA prime system (GE Healthcare, Piscataway, NJ, USA). Recombinant EGF12 and 26 of mouse NOTCH1 were produced in BL21 after 4 h incubation at 37°C in LB broth supplemented with 100 µg.mL<sup>-1</sup> ampicillin and 1 mM IPTG. After lysis with 0.5 mg.mL<sup>-1</sup> lysozyme and sonication, soluble proteins were recovered in supernatants after centrifugation at 10,000 g for 20 min at 4°C and diluted in binding buffer before being purified in the same way and concentrated by ultrafiltration.

**Hydrolysis of GDP-fucose of recombinant Pofut1-** Hydrolysis of GDP-fucose by recombinant POFUT1 was carried out in 50 µl of reaction buffer (25 mM Tris, 5 mM CaCl<sub>2</sub>, 10 mM MnCl<sub>2</sub>, pH 7.5) and incubated 20 h at 37°C. Different concentrations of recombinant POFUT1 and GDP-fucose were mixed with 25 ng of coupling phosphatase 1. Then, release of inorganic phosphate was quantified with the



Malachite Green Phosphate Detection kit (R&D Systems Inc., Minneapolis, MN) and absorbance was read at 620 nm with FLUOstar Omega Microplate Reader (BMG LABTECH).

**SDS-PAGE and gel blotting.** Purified proteins were separated on 12 or 15% acrylamide gels (SDS-PAGE) and blotted to nitrocellulose membranes. For Western blot, membranes were blocked with 5% fat free milk in TBST (50mM Tris-HCl, 150 mM NaCl, 0.1% Tween-20, pH 7.4) for 1 h at room temperature and incubated overnight at 4°C with anti-V5-HRP (Thermo Fisher Scientific) or anti-Pofut1 (Der Vartanian et al. 2015) diluted to 1:2,000 in 2.5% fat free milk-TBST. For anti-Pofut1, membranes were incubated after three washes in TBST for 1 h at room temperature with secondary HRP conjugate antibodies diluted to 1:2,000 (Dako, Glostrup, Denmark) in 2.5% fat free milk-TBST. After CuAAC, gel separation and transfer, membranes were blocked with 10% fat free milk-TBST for 10 min, before being incubated with streptavidin-HRP in TBST at 25 ng/ml for 30 min. The membranes were washed three-times before and after streptavidin-HRP incubation with TBST, for 15 min per wash. In both cases, membranes were revealed using enhanced chemiluminescence peroxidase substrate. Signals were visualized and quantified using Amersham Imager 600 (GE Healthcare).

**Alignments and superimposition of X-ray structures.** An automatic homology model was generated for mouse NOTCH1 (N1) EGF12 on the Swiss-model server (<https://swissmodel.expasy.org>) using X-ray structure of human one (PDB 5L0R) (36) as a reference template, since they share more than 94% identity. This structural model was considered as relevant on the basis of selected template and the Global Model Quality Estimation (GMQE) score. Using MatchMaker of UCSF CHIMERA (Pettersen et al. 2004), mouse N1-EGF26, co-crystallized with mouse POFUT1 (PDB 5KY4)(Li et al. 2017), was superimposed with obtained model for EGF12 . Same method was also used to superimpose crystalized structures of Pofut1 from *Mus Musculus* (PDB 5KY4)(10), *Homo sapiens* (PDB 5UXH) (McMillan et al. 2017) and *Caenorhabditis elegans* (PDB 3ZY5)(Lira-Navarrete et al. 2011).



***Protein digestion and peptide purification before mass spectrometry (MS) analyses.*** Proteins in 25 mM phosphate buffer pH 7.8 were reduced in 5mM dithiothreitol, alkylated in 10 mM iodoacetamide and digested overnight at 37°C using 0.5 µg of endoproteinase Glu-C (Sigma-Aldrich, Saint Louis, USA). Peptide samples were then purified on 1 cc 30mg HLB cartridge (Waters Corporation, Milford, USA) with following steps: conditioning with 1 ml methanol, equilibration with 0.5 % formic acid in water, loading of sample diluted in 0.5 % formic acid in water, washing twice with 0.5 % formic acid in water and elution with 1 ml methanol. After evaporation under nitrogen, purified peptides were resolubilized in 50 µl of loading mobile phase (water/acetonitrile/trifluoroacetic acid (98/2/0.05%)) and filtered on 0.22µm spin column (Agilent Technologies, Santa Clara, CA).

***MicroLC MS/MS analysis of Glu-C peptides.*** Purified Glu-C peptides were analysed by microLC-MS/MS using a nanoLC 425 in micro-flow mode (Eksigent, Dublin, CA) system coupled with a TripleTOF 5600+ (Sciex, Framingham, MA). 5µl of each sample were trapped on C18 Pepmap100 cartridge (300 µm Id x 5 mm, 5µm; Thermo Scientific), and desalting was carried out at 10 µL/min with loading mobile phase for 5 minutes. The chromatographic separation was performed on a ChromXP C18 column (150 x 0.3 mm i.d., 120Å, 3 µm; Sciex) at a flow rate of 3 µl/min. The mobile phase was a gradient of water/acetonitrile/formic acid 100/0/0.1% (A) and 5/95/0.1% (B) programmed as follows: initial, 5% B, increased to 25% in 20 min, then increased to 95% B in 2 min, maintained at 95% for 4 min, and finally, decreased to 5% B for reequilibration. The TTOF 5600+ was operated in information-dependent acquisition (IDA) mode with Analyst 1.7TF software (Sciex). MS and MS/MS data were continuously recorded with 1.3s cycle time with up to 20 precursors selected for fragmentation from each MS survey scan. Precursor selection was based upon ion intensity and whether or not the precursor has been previously selected for fragmentation (dynamic exclusion). Collision energies were automatically adjusted to the charge state and  $m/z$  value of the precursor ions.

***Data processing and creation of the targeted MRM-MS method.*** ProteinPilot 5.0 (SCIEX) was applied to search against the recombinant protein sequence database (Supplementary Figures 4-6), and the MRM



transition list was established using Skyline 3.5.0 (MacCoss Lab, University of Washington, Seattle, WA, USA) for the WT and the mutant non-fucosylated peptides. O-fucosylation (coded by [dHx] amino acid modification) was added *in silico* at expected position with PeakView software (SCIEX) and m/z of precursor and fragments were calculated. Data were acquired in high-Resolution MRM mode: product ion scans were collected for the m/z corresponding to fucosylated or non-fucosylated WT and mutant peptides during 30 min using the same parameters as previously described in the IDA method. Data were processed with MultiQuant Software 3.0.1 (SCIEX), considering the 8 most abundant fragments for each peptide with a resolution of 10,000. Same fragments were used for non-fucosylated and fucosylated peptides (Supplementary Tables 1 and 2). Areas were collected for the same most intense fragment of fucosylated and non-fucosylated peptides and a ratio of fucosylation was calculated.

**Statistical analysis.** All experiments were performed in biological triplicates and results were reported as the means  $\pm$  SEMs. Statistical comparisons were performed using two-tailed *t* tests implemented in Prism, version 5.03 (GraphPad Software, Inc., San Diego, CA, USA). A *p* value of 0.05 or less was considered statistically significant.

## REFERENCES

- Der Vartanian A, Audfray A, Al Jaam B, Janot M, Legardinier S, Maftah A, Germot A. 2015. Protein O-fucosyltransferase 1 expression impacts myogenic C2C12 cell commitment via the Notch signaling pathway. *Mol Cell Biol*, 35:391-405.
- Li Z, Han K, Pak JE, Satkunarajah M, Zhou D, Rini JM. 2017. Recognition of EGF-like domains by the Notch-modifying O-fucosyltransferase POFUT1. *Nat Chem Biol*, 13:757-763.
- Lira-Navarrete E, Valero-Gonzalez J, Villanueva R, Martinez-Julvez M, Tejero T, Merino P, Panjekar S, Hurtado-Guerrero R. 2011. Structural insights into the mechanism of protein O-fucosylation. *PLoS One*, 6:e25365.
- McMillan BJ, Zimmerman B, Egan ED, Lofgren M, Xu X, Hesser A, Blacklow SC. 2017. Structure of human POFUT1, its requirement in ligand-independent oncogenic Notch signaling, and functional effects of Dowling-Degos mutations. *Glycobiology*, 27:777-786.
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, 25:1605-1612.



# SUPPLEMENTARY DATA

## WT EGF12

<b>BamHI</b> G/GATC CCGATGTTAAATGAGTGCATCTCCAACC/CATGTCA GAATGACGCCACTTGCCTGGACCA/GAATTGGG GAGTTCCAAATGCATATGATGCCC/AGGTTAT GAAGGTGTATACTGTGAAC/TCGA G D P D V N E C I S N P C Q N D A T C L D Q I G E F Q C I C M P G Y E G V Y C E L E C CTAG/GCCTACAATTACTCAGTAGAGTTGG GTACAGT/CTTACTGGGGTGAACGGACTGGT CTAACCC/CTCAAGGTTACGTATACATACAGG TCCAATA/CTTCCACATATGACACTTG AGCT/C	C .../CATGTCA GAATGACGCCACTTGCCTGGACCA/... .../CTTACTGGGGTGAACGGACTGGT CTAACCC/...	E G H	XhoI
--	--	-------------	------

## T466A EGF12

D P D V N E C I S N P C Q N D A A C L D Q I G E F Q C I C M P G Y E G V Y C E L E .../CATGTCA GAATGACGCCACTTGCCTGGACCA/... .../CTTACTGGGGTGAACGGACTGGT CTAACCC/...	C' .../CATGTCA GAATGACGCCACTTGCCTGGACCA/... .../CTTACTGGGGTGAACGGACTGGT CTAACCC/...	D' H	XhoI
--	---	---------	------

## WT EGF26

<b>BamHI</b> G/GATC CGAACACACCTGACTGTACTGAGAGCT/CCTGCTT CAATGGTGGTACCTGTGTGGATGG/TATCAAC TCCTTACCTGTCTGTCCACC/TGGCTTC ACGGCAGCTACTGTGAGC/TCGA G D P N T P D C T E S S C F N G G T C V D G I N S F T C L C P P G F T G S Y C Q L E C CTAG/GCTTGTGGACTGACATGACTCTCGA GGAAGAA/GTTACCACCACCTTACC ATAGTTG/AGGAAGTGGACAGACAGAGTGG ACCGAAG/TGCCCGTCCGATGACAGTGG AGCT/C	C .../CATGTCA GAATGACGCCACTTGCCTGGACCA/... .../CTTACTGGGGTGAACGGACTGGT CTAACCC/...	E G H	XhoI
--	--	-------------	------

## T997A EGF26

D P N T P D C T E S S C F N G G A C V D G I N S F T C L C P P G F T G S Y C Q L E .../CATGTCA GAATGACGCCACTTGCCTGGACCA/... .../CTTACTGGGGTGAACGGACTGGT CTAACCC/...	C' .../CATGTCA GAATGACGCCACTTGCCTGGACCA/... .../CTTACTGGGGTGAACGGACTGGT CTAACCC/...	D' H	XhoI
--	---	---------	------

**Figure 1:** For each EGF-LD (WT and T/A mutated EGF12 and 26), four pairs of complementary oligonucleotides (A-B, C-D or C'-D', E-F, G-H) were designed with cohesive ends. After heat-denaturation at 80°C for 10 min, the eight oligonucleotides were equally mixed and heated again at 80°C for 10 min. Hybridization occurred by a progressive decrease of temperature from 80°C to room temperature resulting in double-stranded DNA fragments. These fragments were then cloned into the prokaryotic expression vector pET-25b(+) using *Bam*HI and *Xho*I cloning sites (in bold). The *O*-fucosylation sites and T/A mutation of these sites are in bold and underlined.



# SUPPLEMENTARY DATA

<u>Homo sapiens</u>	MGAAAWAREPLS-----VSFLLLLPLPGMPAGSNDPAGYLLYCPGMRFGNQADHFLGSLAFAKLLNRTLAVPPWIEYOHKPKPPFTNLHVSQYKFKLEPQAVHRVISEDLEFMKLL	112
<u>Mus musculus</u>	MGAAAWAPPHLL----LRASFLLLLPLRGRSAGSNDLAGYLLYCPGMRFGNQADHFLGSLAFAKLLNRTLAVPPWIEYOHKPKPPFTNLHVSQYKFKLEPQAVHRVISEDLEFMENL	117
<u>Caenorhabditis elegans</u>	-----MR-VSKVLTASFTVCSYSEAKSNETDPNGYIVFCPCMRFGNQVDFLGLVAFAKALDRTLVLPNFIYKHEE----TKMIFEFLLPQVGTAVKTRVVMQEFKKI	105
<u>Gallus gallus</u>	-----MVL-----L-----SQAALLLATVTVAAAGSNDTAGYLLYCPGMRFGNQADHFLGSLAFAKALNRTLAVPPWIEYRHRHPPFTNLHVPYEEYFKLEPQQHVKVLSLEQFMQL	104
<u>Xenopus laevis</u>	-----MERGVWLLLLLPV--GCEGLYNSNGYIVYCPGMRFGNQADHFLGSLAFAKMNRTLVVPPWIVNHRHPPFTNLHVPYEEYFKLEPQQHVIISHEEFMBQL	104
<u>Danio rerio</u>	MA-----ARPKHLKPIVTVLHTSLIFLNVFIDATDVTWDENGYILYCPGMRFGNQVDFHFLGSLAFAKMLNRTLAVPPWIEYRHSPPFTNLHVPYSEYFQLEPKQHRVSVLEDEHRL	117
<u>Drosophila melanogaster</u>	-----MQWLKMKLRFVNLILLLSSTCAQLGGDPNGYILYCPGMRFGNQADHFLGSLAFAKALNRTLILPPWIEYRGRGE--LRSRQVFNFTYFEVEPLKEHRVITMADFMHLL	108
<u>Homo sapiens</u>	APTWPPPKRVA YCFEVA-----AQRSPDKKTCPMKGNPFQFWDQFHVSNKSELFTGIS--FSAS----YREQWSQRFSPKHPVIALPGAPAFPPVLEHRPLQKXVMWSDEMVKTG	222
<u>Mus musculus</u>	APSHWPPKRVAYCFEVA-----AQRSPDKKTCPMKGNPFQFWDQFHVSNKSELFTGIS--FSAS----YKEQWTQRFPAKEHPVIALPGAPAFPPVLEHRPLQKXVMWSDEMVRTG	227
<u>Caenorhabditis elegans</u>	MPTHVGTQRQA IYDK-----SAEPGSHKSGNPFQFWDQIDVDFVGDYFGLDFGGDFLNQMSGRKKWLEKFPSEYFVIAFSSAPAFPPSKGKVMYSIQKYLKWSRRIETQA	214
<u>Gallus gallus</u>	APQHWPGRVAYCFEAA-----AQRSDAKSTCPMKGNPFQFWDQFKVDFKSELFTGIS--FSAS----YRDHWIQRFPPSEHPVIALPGAPAFPPVLEHRPLHKKVMWSDEMVKKG	214
<u>Xenopus laevis</u>	APVLPFPTKRLAYCFATA-----AQRSPDKKSCPMKGNPFQFWDHFDYDFQSELFDGIT--FSAY----YQDHWISRFPPSQHPVIALPGAPAFPPVLEHRSLQRYIIVWSEKIVREG	214
<u>Danio rerio</u>	APKHWPHGQRIAYCFESA-----AHRSDQKKSCPMKGNPFQFWDHIGVDFDRSVLFTGIS--FSSY----YQPHWIKRFPKHPVIALPGAPAFPPVLEHIGLQGFVWSDKIVQEG	227
<u>Drosophila melanogaster</u>	ADDIWPESERVSYFCYKERYSLQEQEKNDPDKPNCCHAKGNPFQFWDHFDVVRSEFYAPLH--FDVHHSNEAAKQWTKYPAESYFVIAFTGAPAFPPVQLENCIKLQRYLQWSQRVREAS	226
<u>Homo sapiens</u>	EAQIHHLVR-PYVGIHLRIGSDWKNAKMLKDGTAGSHFVWASPOCVGYSRSTAAFLTMTCLPDLKEIQRAVKLWVRS-----LDAQSVIVVATDSESYVPELQQLFK-GKVKVYSLK	333
<u>Mus musculus</u>	EALISAHLYR-PYVGIHLRIGSDWKNAKMLKDGTAGSHFVWASPOCVGYSRSTATLMTCLPDLKEIQRAVTLWVRA-----LNARSVVIATDSESYVSEIQQLFK-DKVRVYSLK	338
<u>Caenorhabditis elegans</u>	KKFISANLAK-PFVAHLRNDADWVRCHEIDTT-TNRPLEASEQCUGEGHHL-GTLTKEICSPSKQQLIQLIIVEKVG-----IGAKSVFVADSDRDMIDEINEALKPYEIAHRQE	324
<u>Gallus gallus</u>	EAYTHSLVR-PYVGIHLRIGSDWKNAKMLKDGTAGSHFVWASPOCVGYSRSTAVPLTMDCLPDLKEIKRALKVVWKK-----TVAKSIIATDSEPYTKIQQFQ-GKIKVYVSLQ	325
<u>Xenopus laevis</u>	EEQFSLVR-PYVGIHLRIGSDWKNAKMLKDGTAGSHVWASAOVCYDRYPAGLTMDCPLDKEMRALTLWVER-----SKARSVIATDSESYTAEIOALIG-NEVRVCLQ	325
<u>Danio rerio</u>	EGHFRNLNR-PYVGIHLRIGSDWQNAKMLKDGTDGFHFVWASPOCVGYDRQATLPTMNCPLDTEIRAVKLMVRN-----TGAQSVIATDSEHTTBIQKLFK-GKVKVYVSLQ	338
<u>Drosophila melanogaster</u>	KDFIREQLPFGATGIHLRNGIDWVRACEHVKDS---QHLFASPOCVGYKNER-GALYPELCMPKSEAITLQIKKTIKNVROTDPDNETKSVFVADSNSHMIQELNTALSRMGI SVHKLP	342
<u>Homo sapiens</u>	PEVAQVDLYILGQADHFI GNCVSSFTAFVKRRDLQG---RFSFFGMDRPPKLRDEF-----388	388
<u>Mus musculus</u>	PEVAQIDLILGQADHFI GNCVSSFTAFVKRRDLHG---RQSSFFGMDRPSQLRDEF-----393	393
<u>Caenorhabditis elegans</u>	PDMYTSLAIMGADLFI GNCVSTFTHVKRRDRHAGQSPRPSAFFGIRAVKRHIDL-----381	381
<u>Gallus gallus</u>	PEVAQIDLILGQSDHFI GNCVSSFTAFVKRRDVHG---KFSFFGMDHPRLRDEF-----380	380
<u>Xenopus laevis</u>	PEVAQIDLILGQADHFI GNCVSSFTAFVKRRDVHG---RFSFFGMDSESPVQDEL-----380	380
<u>Danio rerio</u>	PDTAQVDLYILGQADHFI GNCVSSFSFAVKRRNIHG---RFSFFGLDYFGKKNREE---L	395
<u>Drosophila melanogaster</u>	EDDPYLDLAILGQSNHFI GNCISYSYSAFVKRRDVHG---FPSYFDFGPKKDRKHTNVHEEL	402

\* Single, fully conserved residue

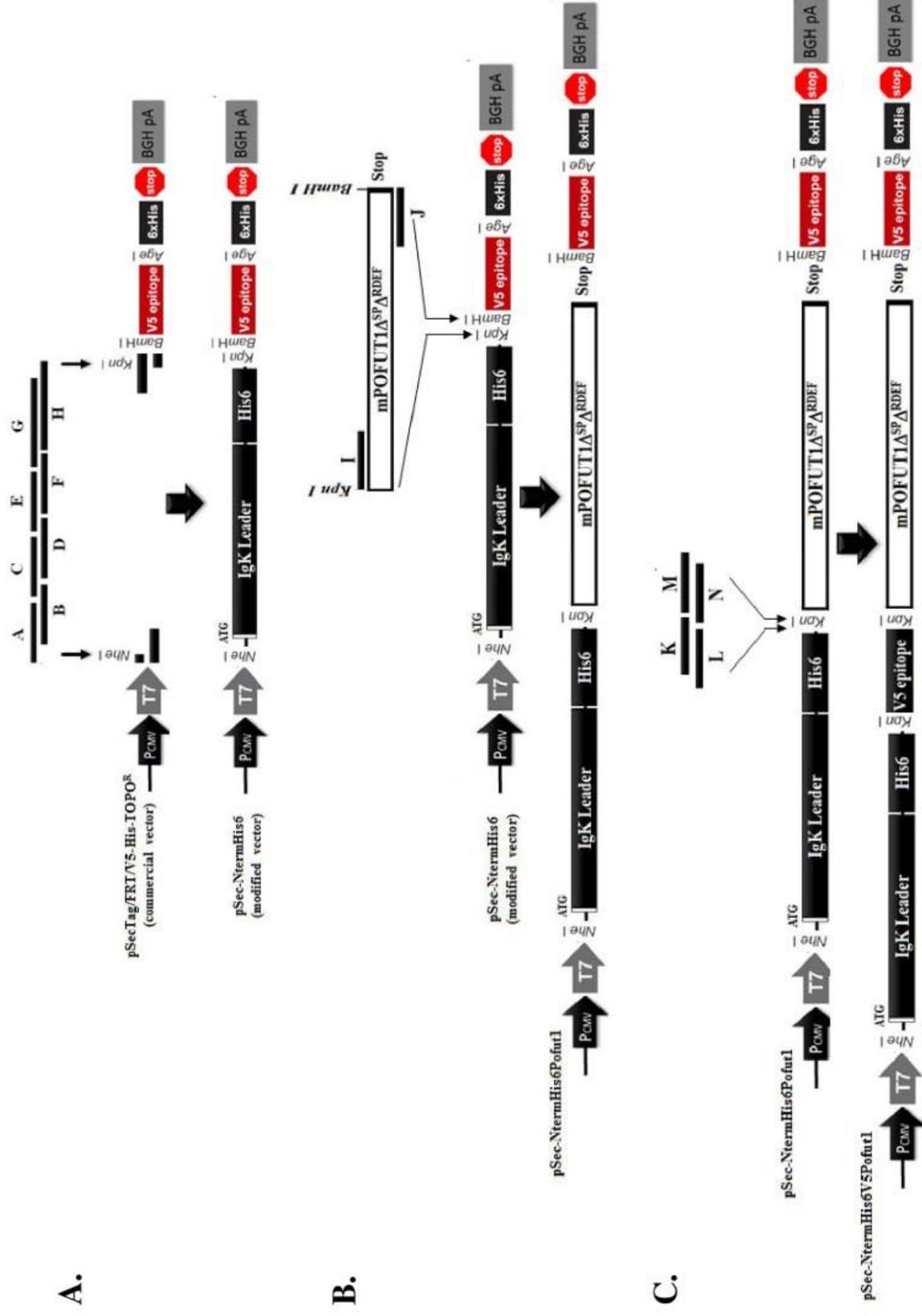
• Conservation between groups of strongly similar properties

• Conservation between groups of weakly similar properties

**Figure 2:** Protein multiple sequence alignment (using Clustal Omega) for Pofut1 of widely studied species, which is composed of a signal peptide and a C-terminal KDEL-like motif. Underlined species names are those for which POFUT1 X-ray structure is available in Protein Data Bank (PDB). Residues highlighted in cyan are key amino acids involved in interaction with EGF-LD, found in highly conserved regions. Aromatic Residues (Y or F) highlighted in black are potentially involved in a steric clash with EGF-LD.



# SUPPLEMENTARY DATA

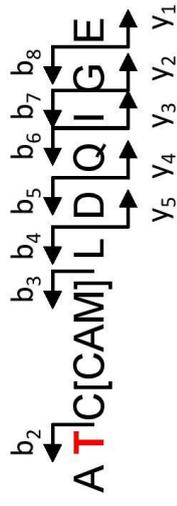
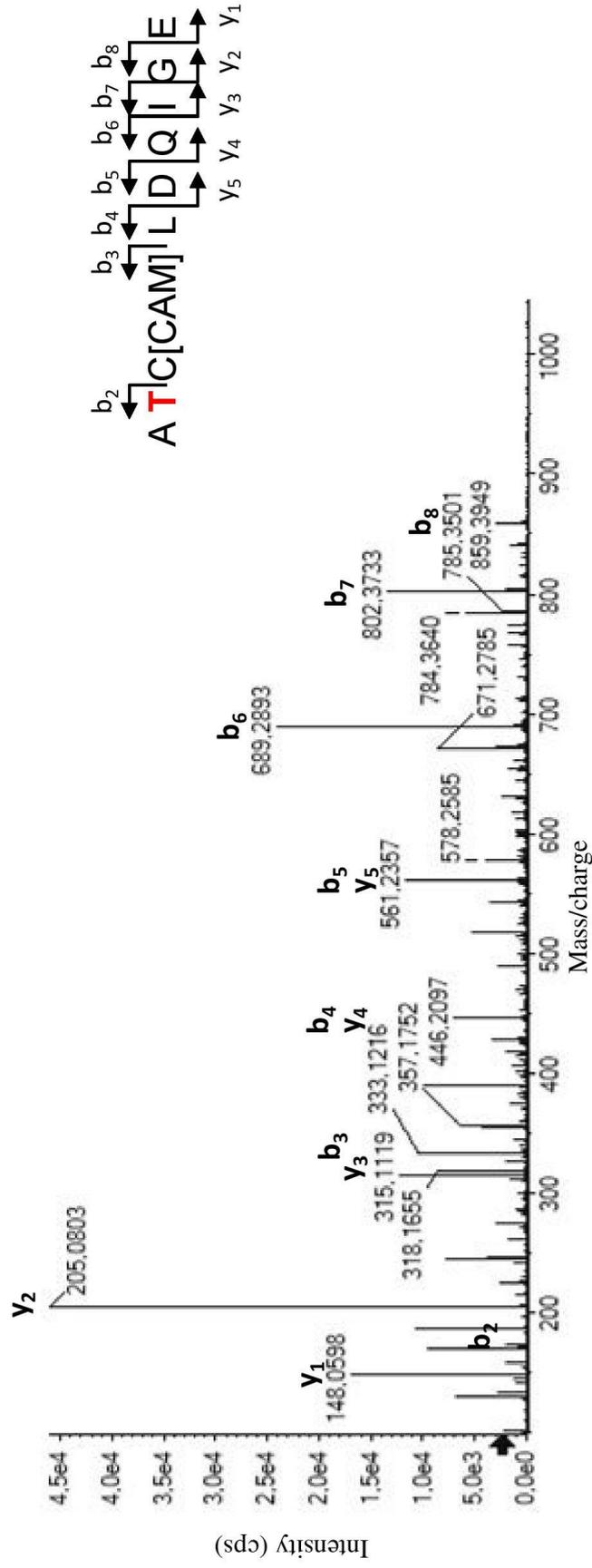


**Figure 3:** (A) Eight oligonucleotides A-H were hybridized and cloned into commercial pSecTag/FRT/V5-His-TOPO<sup>R</sup>, previously digested with *NheI* and *KpnI* restriction enzymes, to introduce a new secretory signal peptide (IgK Leader) as found originally in commercial vector, but fused to six histidine residues (His6). (B) In this modified pSec-NtermHis6 vector, cDNA encoding mouse POFUT1 without endogenous signal peptide (SP) and C-terminal KDEL-like motif (RDEF) but with a new stop codon was cloned between *KpnI* and *BamHI* restriction sites after PCR amplification with I and J primers. (C) As an optional step, a V5 epitope was added downstream the His6 tag, by cloning four hybridized oligonucleotides (K-N) at the *KpnI* cloning site. This construct was used to produce a secreted POFUT1 with N-terminal polyhistidine and V5 tags in CHO cells.

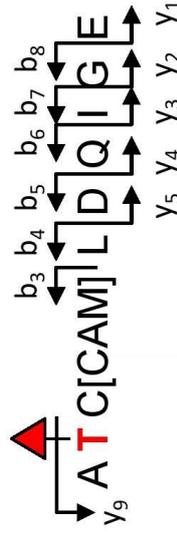
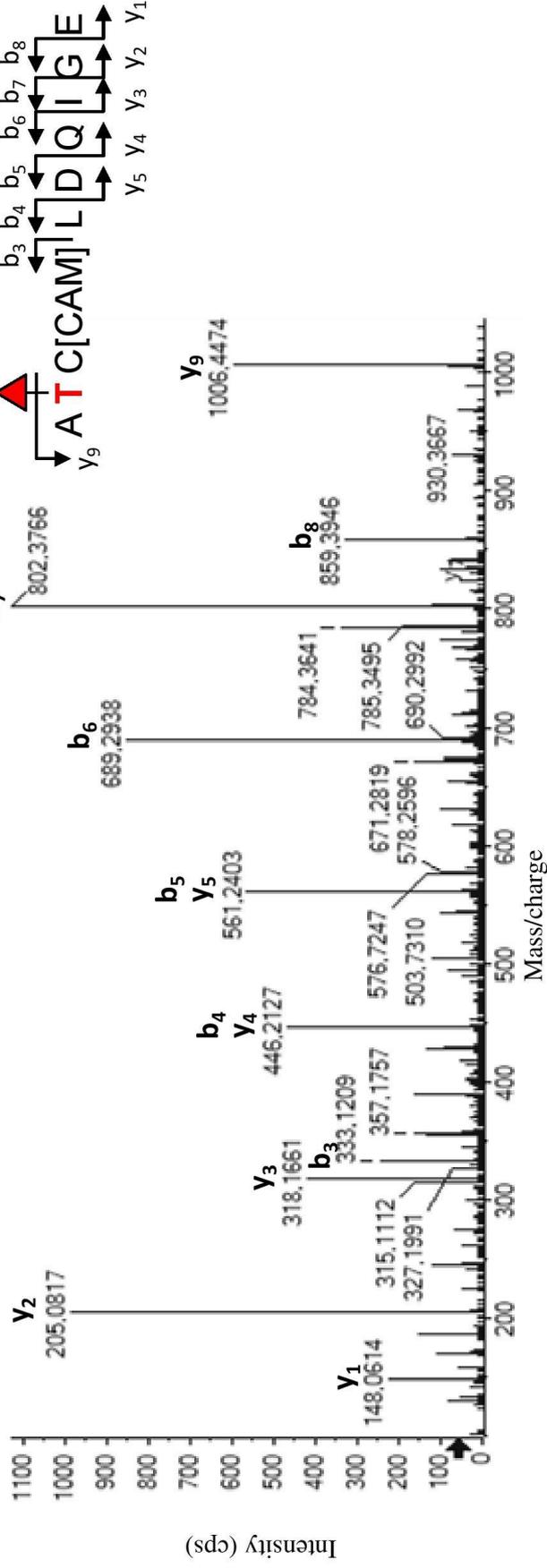


# SUPPLEMENTARY DATA

A.



B.

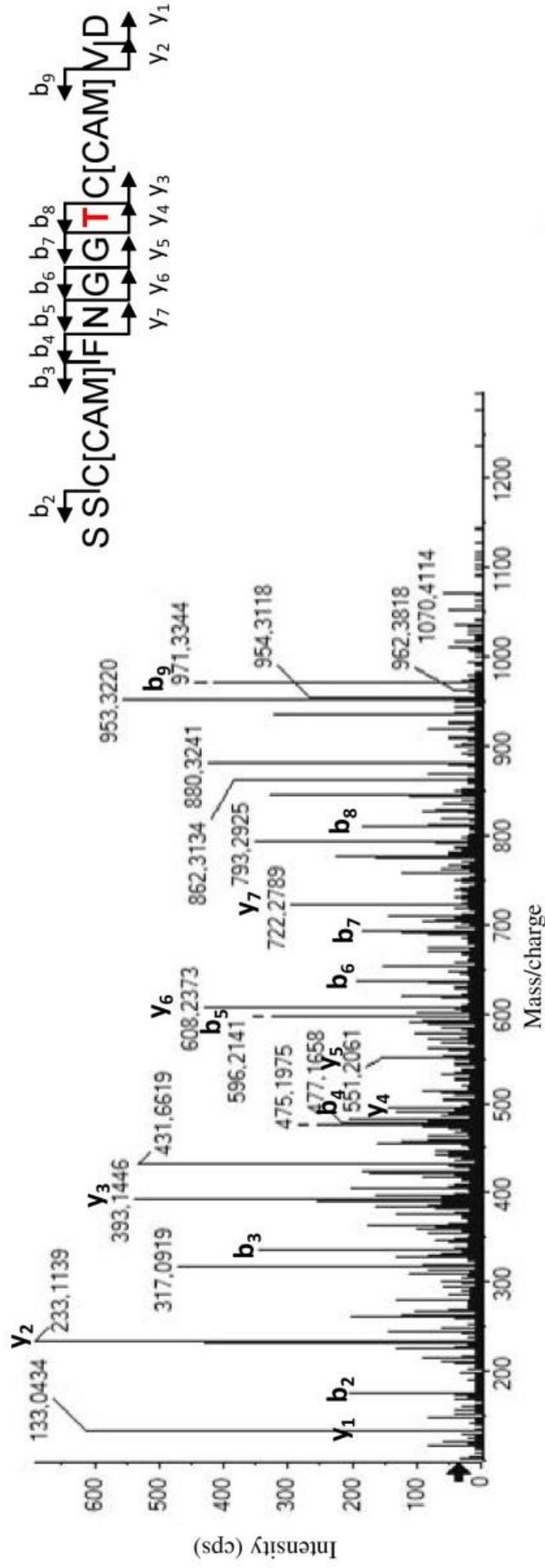


**Figure 4:** Representative MS/MS spectra for non fucosylated (A) and fucosylated (B) WT NOTCH1 EGF12 peptide containing the *O*-fucosylation site (in red) after Glu-C digestion. Spectra were acquired with a micro-LC TripleTOF system in IDA mode. Detected b/y fragments are annotated on each spectrum confirming the identified peptidic sequence. For fucosylated peptide, the loss of fucose occurs first, leading to the presence of y<sub>9</sub> fragment and other non fucosylated b/y fragments.

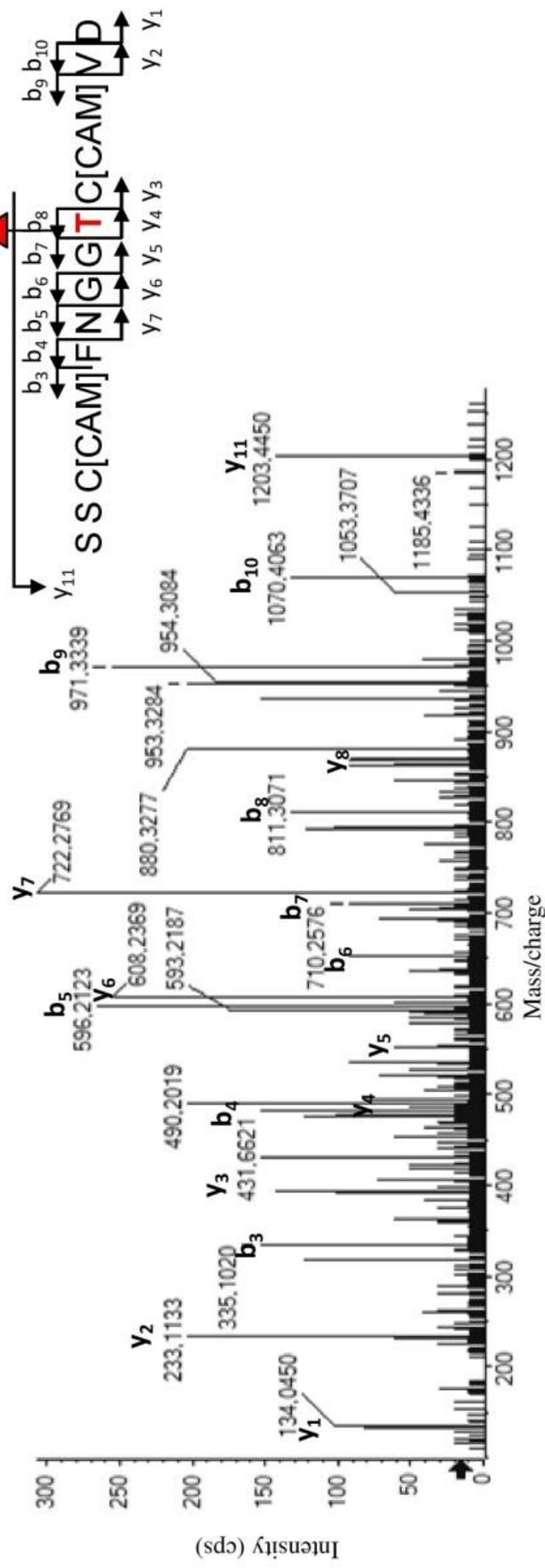


# SUPPLEMENTARY DATA

A.



B.

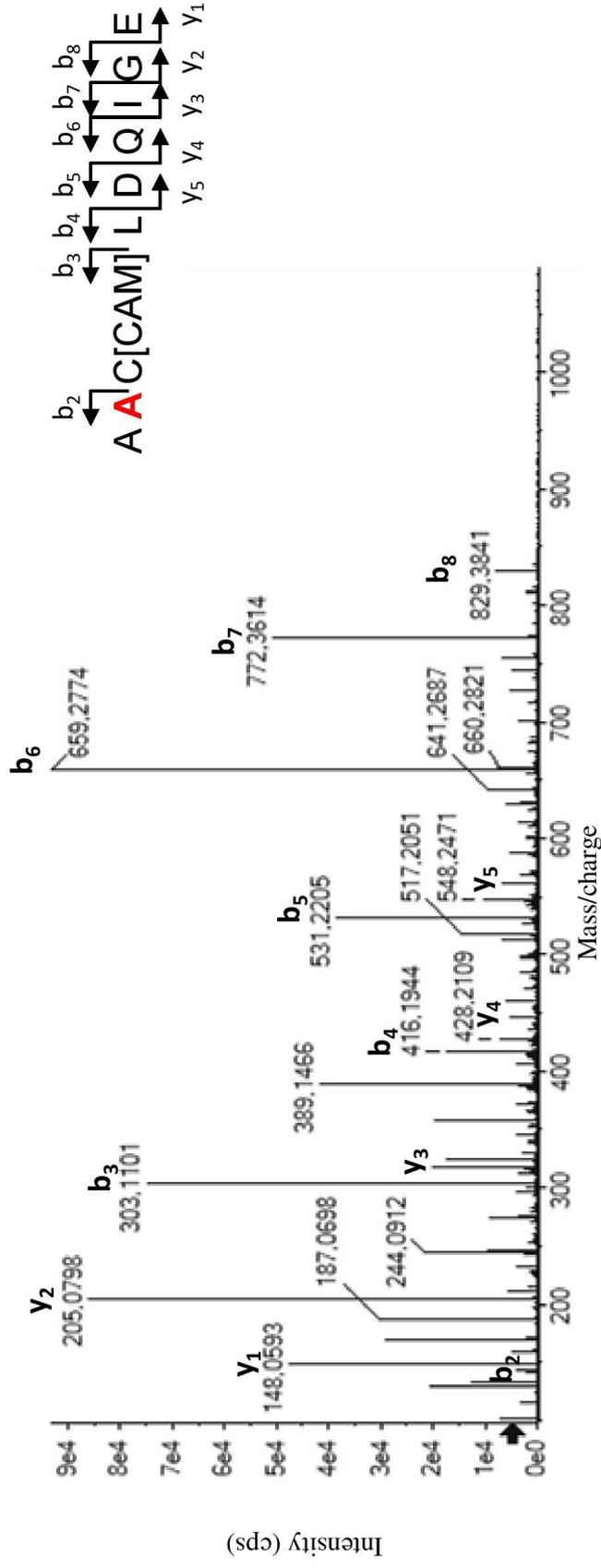


**Figure 5:** Representative MS/MS spectra for non fucosylated (A) and fucosylated (B) WT NOTCH1 EGF26 peptide containing the O-fucosylation site (in red) after Glu-C digestion. Spectra were acquired with a micro-LC TripleTOF system in IDA mode. Detected b/y fragments are annotated on each spectrum confirming the identified peptic sequence. For fucosylated peptide, the loss of fucose occurs first, leading to the presence of y<sub>11</sub> fragment and other non fucosylated b/y fragments.

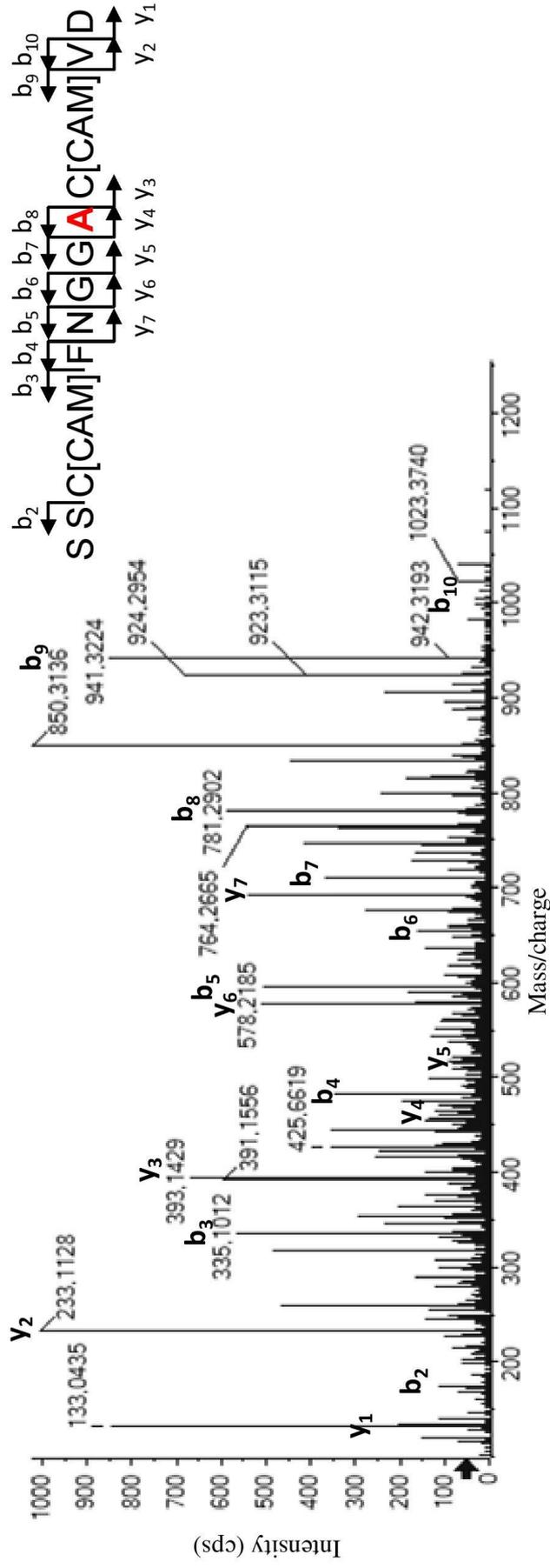


# SUPPLEMENTARY DATA

A.



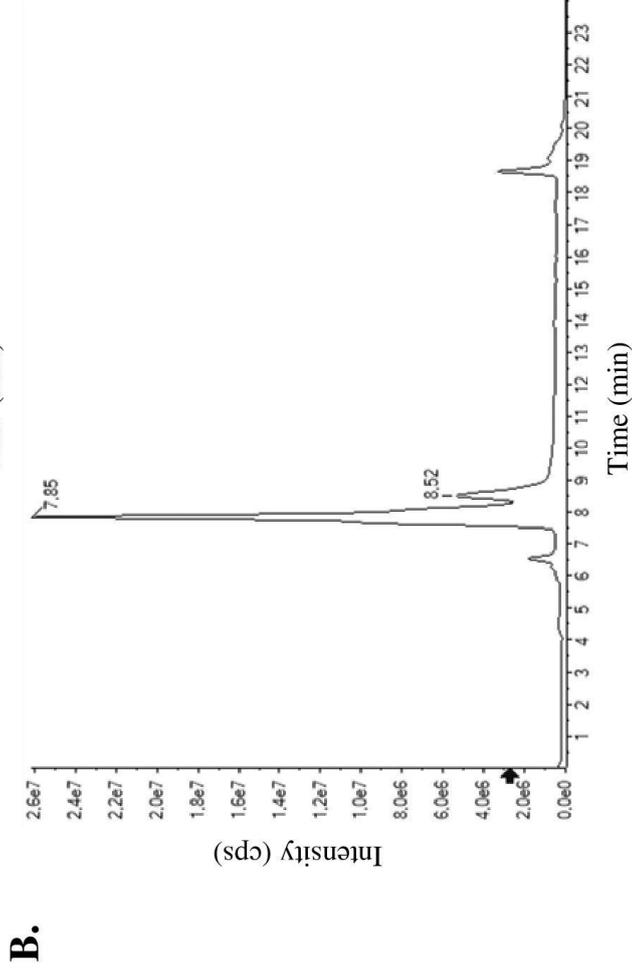
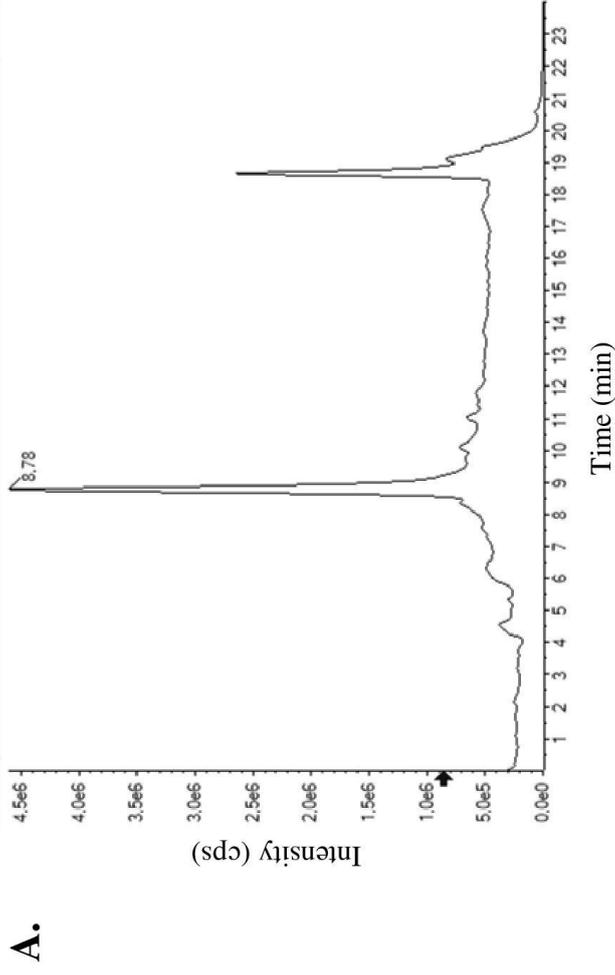
B.



**Figure 6:** Representative MS/MS spectra for mutant NOTCH1 EGF12 (A) and EGF26 (B) peptide containing the modified *O*-fucosylation site (in red) after Glu-C digestion. Spectra were acquired with a micro-LC TripleTOF system in IDA mode. Detected b/y fragments are annotated on each spectrum confirming the identified peptidic sequence.



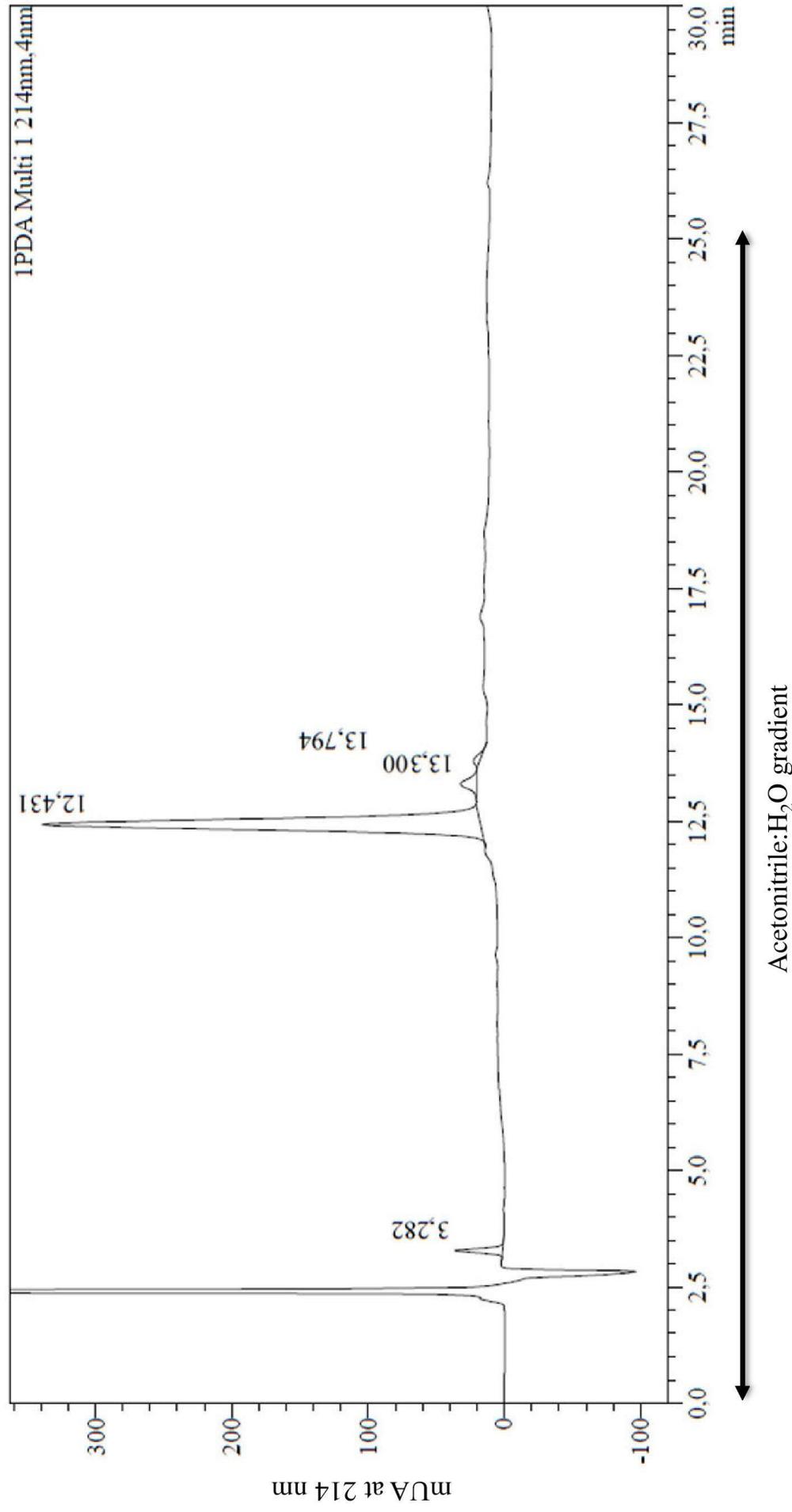
# SUPPLEMENTARY DATA



**Figure 7:** Reverse-phase HPLC of WT NOTCH1 EGF12 (A) and EGF26 (B). Chromatography was carried out as described by Wang et al., 1998 with minor modifications. HPLC-MS analyses were performed with the same system used for the analyses of protein tryptic digests except the gradient was programmed as follows: initial, 5% B, increased to 25% in 3 min, then increased to 30% B in 11 min, then to 98% B in 2min, maintained at 98% for 5 min, and finally, decreased to 5% B for reequilibration. Mass scans were performed in ESI-TOFMS mode from 400 to 2500  $m/z$  with a step size of 0.08 amu and a dwell time of 0.08 ms.



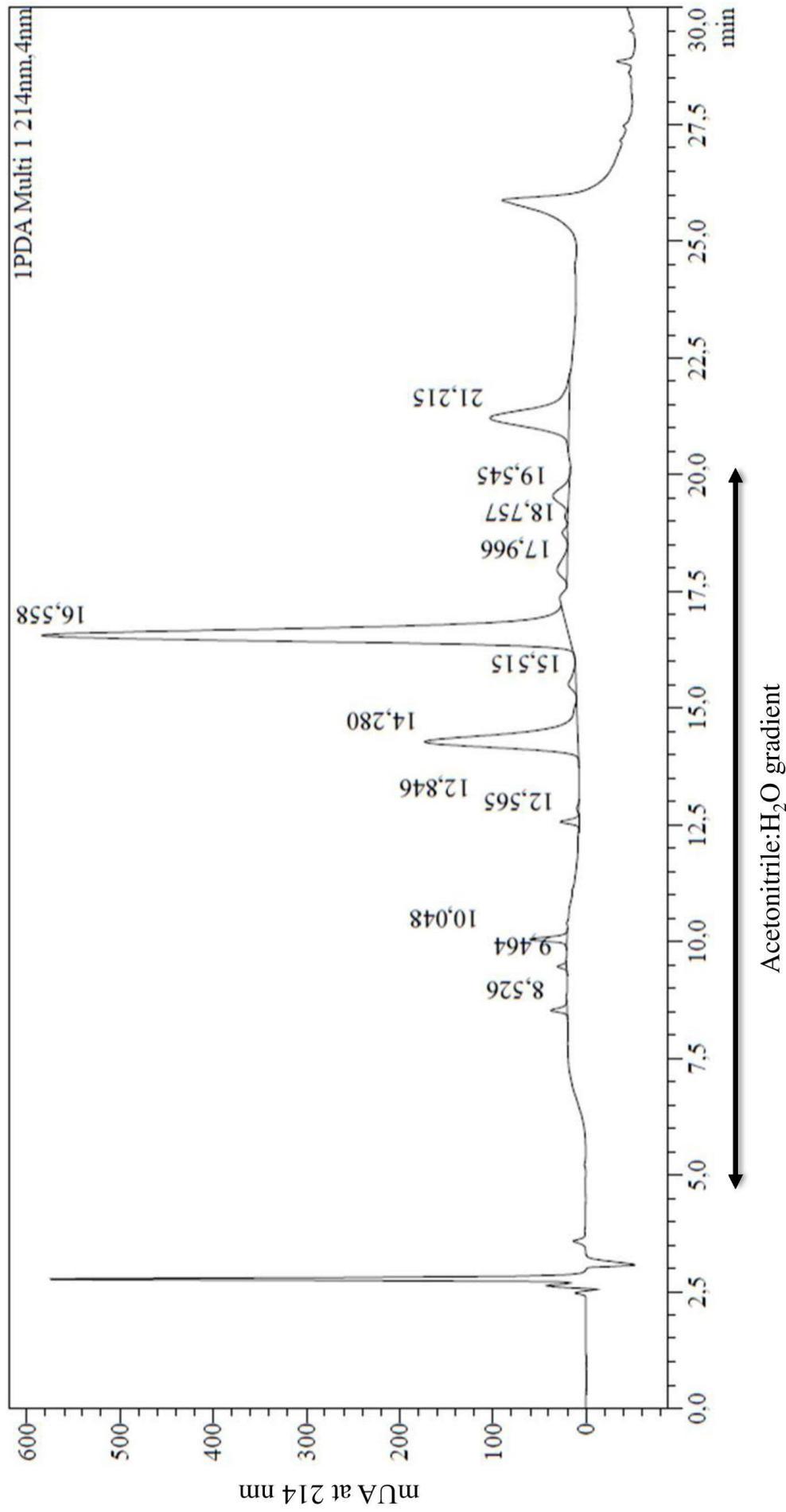
# SUPPLEMENTARY DATA



**Figure 8.** Reverse-phase HPLC of WT NOTCH1 EGF12. Chromatography was carried out as described by Li et al, 2018. HPLC analyses were performed on a C18 column with an acetonitrile:H<sub>2</sub>O gradient running from 20% to 50% (v/v) in the presence of 0.06% (v/v) trifluoroacetic acid.



# SUPPLEMENTARY DATA



**Figure 9.** Reverse-phase HPLC of WT NOTCH1 EGF26. Chromatography was carried out as described by Li et al, 2018. HPLC analyses were performed on a C18 column with an acetonitrile:H<sub>2</sub>O gradient running from 20% to 30% (v/v) in the presence of 0.06% (v/v) trifluoroacetic acid.



Targeted sequence	charge state	Non fucosylated peptide				Fucosylated peptide				DP (V)	CE (V)
		MS1		MS2		MS1		MS2			
		observed	theoretical	observed	theoretical	observed	theoretical	observed	theoretical		
<b>WT EGF 12</b>	+2	503.7291		561.2357	576.7581	+2y5	561.2514	561.2403	561.2403	67.8	27.2
<u>ATC</u> [CAM]LDQIGE		detected		446.2097	detected	+2y4	446.2245	446.2127	446.2127		
				318.1655		+2y3	318.1659	318.1661	318.1661		
				<b>205.0803</b>		<b>+2y2</b>	205.0818	<b>205.0817</b>	<b>205.0817</b>		
				148.0598		+2y1	148.0604	148.0614	148.0614		
				333.1216		+2b3	333.1227	333.1209	333.1209		
				689.2893		+2b6	689.2923	689.2938	689.2938		
				<b>802.3733</b>		<b>+2b7</b>	<b>802.3763</b>	<b>802.3766</b>	<b>802.3766</b>		
<b>Mutant EGF 12</b>	+2	488.7238		318.1663	561.7528	+2y3	318.1659	ND	318.1659	66.7	26.5
<u>AA</u> C[CAM]LDQIGE		detected		205.0798	not	+2y2	205.0818	ND	205.0818		
				148.0593	detected	+2y1	148.0604	ND	148.0604		
				303.1101		+2b3	303.1121	ND	449.1700		
				416.1944		+2b4	416.1962	ND	562.2541		
				531.2205		+2b5	531.2231	ND	677.2810		
				659.2774		+2b6	659.2817	ND	805.3396		
				<b>772.3614</b>		<b>+2b7</b>	<b>772.3658</b>	ND	<b>918.4237</b>		

**Table 1:** HR-MRM parameters for the detection of NOTCH1 EGF12 WT and mutant peptides. The *O*-fucosylation site is indicated in bold and underlined. Each peptide is detected by a single charge state corresponding to a MS1 value (with optimized Declustering Potential (DP) and collision energy (CE)) recorded during the 30 min analysis. Peak detection with MultiQuant was achieved with the eight most abundant fragments (MS2) with a 10,000 resolution. Observed fragments for fucosylated peptides didn't correspond to theoretical fragments due to the loss of fucose occurring first and leading to non fucosylated b/y fragments. Quantification fragment is indicated in bold. The 7 other MS2 were used as confirmation fragments. Fucosylated mutant peptide and its related fragments were not detected (ND).



Targeted sequence	charge state	Non fucosylated peptide				Fucosylated peptide				DP (V)	CE (V)
		MS1	MS2		MS1	MS2					
			theoretical	observed		theoretical	observed				
<b>WT EGF 26</b>	+2	602.2239	+2y6	608.2344	608.2373	675.2528	+2y6	754.2923	608.2369	75	31.5
SSC[CAM]FNGGTC[CAM]VD		detected	+2y3	393.1438	393.1446	detected	+2y3	393.1438	393.1460		
			+2y2	233.1131	233.1139		+2y2	233.1131	233.1133		
			+2b3	335.1019	335.1030		+2b3	335.1019	335.1020		
			<b>+2b5</b>	596.2133	<b>596.2141</b>		<b>+2b5</b>	596.2133	<b>596.2123</b>		
			+2b6	653.2347	653.2360		+2b6	653.2347	653.2362		
			+2b7	710.2560	710.2530		+2b7	710.256	710.2576		
			+2b9	971.3345	971.3344		+2b9	1117.3924	971.3339		
<b>Mutant EGF 26</b>	+2	587.2186	+2y7	692.2668	692.2624	660.2476	+2y7	838.3247	ND	73.9	30.8
SSC[CAM]FNGGAC[CAM]VD		detected	+2y3	393.1438	393.1429	not	+2y3	393.1438	ND		
			+2y2	233.1131	233.1128	detected	+2y2	233.1131	ND		
			+2b3	335.1019	335.1012		+2b3	335.1019	ND		
			+2b5	596.2133	596.2141		+2b5	596.2133	ND		
			+2b8	781.2933	781.2902		+2b8	927.3512	ND		
			+2b9	941.3240	941.3224		+2b9	1087.3819	ND		
			+2b8+2	391.1503	391.1556		+2b8+2	464.1792	ND		

**Table 2.** HR-MRM parameters for the detection of NOTCH1 EGF26 WT and mutant peptides. The *O*-fucosylation site is indicated in bold and underlined. Each peptide is detected by a single charge state corresponding to a MS1 value (with optimized Declustering Potential (DP) and collision energy (CE)) recorded during the 30 min analysis. Peak detection with MultiQuant was achieved with the eight most abundant fragments (MS2) with a 10,000 resolution. Observed fragments for fucosylated peptides didn't correspond to theoretical fragments due to the loss of fucose occurring first and leading to non fucosylated b/y fragments. Quantification fragment is indicated in bold. The 7 other MS2 were used as confirmation fragments. Fucosylated mutant peptide and its related fragments were not detected (ND).



## DISCUSSION GÉNÉRALE ET PERSPECTIVES

---

### **La sous-expression de POFUT1 chez les souris *Pofut1*<sup>cax/cax</sup> diminue l'activation de la voie Notch, en modulant les interactions récepteurs-ligands**

La sous-expression du gène *Pofut1* chez les souris *Pofut1*<sup>cax/cax</sup> influe sur la *O*-fucosylation globale du domaine extracellulaire (NECD) du récepteur NOTCH et probablement aussi sur celle de ses différents ligands (Panin et al., 2002). Cependant, seule la *O*-fucosylation du récepteur NOTCH, et non celle de ses ligands de type DSL (Müller et al., 2014), est essentielle pour les interactions récepteurs-ligands et donc pour la modulation de l'activation de la voie canonique de Notch (Stahl et al., 2008).

Notre étude sur les souris *Pofut1*<sup>cax/cax</sup> montre que la sous-expression de *Pofut1* conduit en particulier à une diminution d'interaction du ligand DLL1 (apporté de manière exogène) avec les récepteurs NOTCH présents à la surface des MDCS *Pofut1*<sup>cax/cax</sup> (Al Jaam et al., 2016), qui seraient hypo-*O*-fucosylés comparés à ceux des MDCS sauvages C3H. Contrairement à DLL1, le ligand JAG1 ne se lie que très peu aux MDCS. De plus, pour ce ligand, aucune différence de fixation n'a été observée entre les MDCS *Pofut1*<sup>cax/cax</sup> et sauvages, quelle que soit la quantité de ligand testée. Cette absence de variation entre les MDCS *Pofut1*<sup>cax/cax</sup> et sauvages pour JAG1 n'exclut pourtant pas une diminution de la signalisation Notch induite par ce ligand. En effet, une étude récente a suggéré que des modifications du profil de *O*-fucosylation du NECD étaient associées à une inhibition de la voie de signalisation Notch induite par JAG1, sans aucune altération de l'interaction récepteur-ligand (Kakuda and Haltiwanger, 2017). Les auteurs émettent l'hypothèse que la *O*-fucosylation du NECD est impliquée dans un changement de conformation du récepteur NOTCH1 (sans incidence sur sa capacité à interagir avec JAG1) induit par la fixation de JAG1 et qui conduit au clivage du NICD.

Ainsi, la diminution d'activation de la signalisation Notch (moindre clivage du NICD et activation plus faible des gènes cibles) chez les MDCS *Pofut1*<sup>cax/cax</sup> serait due à l'hypo-*O*-fucosylation des récepteurs NOTCH qui entrainerait une diminution de la transduction du signal, induite *a minima* par DLL1 suite à son défaut d'interaction avec NOTCH1. Cette altération de la voie Notch pourrait être renforcée par une plus faible activation induite par JAG1, non dépendante d'une diminution de son interaction avec les récepteurs NOTCH.



Outre son rôle de *O*-fucosyltransférase, *Pofut1* est connue chez la drosophile pour son activité chaperonne (Sasamura et al., 2007). Ce rôle reste néanmoins controversé chez les mammifères (Stahl et al., 2008). Nous avons montré que la quantité de récepteurs NOTCH1 à la surface des MDCS *Pofut1*<sup>cax/cax</sup> est équivalente à celle des MDCS sauvages (Al Jaam et al., 2016). Ces résultats n'étaient donc pas l'hypothèse d'une activité chaperonne de POFUT1 dans ce modèle murin contrairement à ce qui est connu chez la drosophile. Ainsi, la diminution d'interaction de Notch hypo-*O*-fucosylé avec DLL1 ne peut donc s'expliquer par une plus faible expression des récepteurs à la surface cellulaire.

L'étude de la *O*-fucosylation des récepteurs NOTCH exprimés à la surface des MDCS *Pofut1*<sup>cax/cax</sup> n'a pu être effectuée que de manière indirecte. La caractérisation des sites réellement *O*-fucosylés du NECD serait pourtant particulièrement intéressante puisque les différents *O*-fucoses du récepteur NOTCH1 n'ont pas tous la même fonction sur la transduction du signal (Kakuda and Haltiwanger, 2017). Il serait donc intéressant de déterminer, chez les MDCS *Pofut1*<sup>cax/cax</sup>, si tous les sites de *O*-fucosylation du NECD sont impactés de la même manière par la sous-expression de *Pofut1* ou si certains d'entre eux, moins accessibles à l'enzyme, sont plus sévèrement touchés.

De la même façon que pour la *O*-fucosylation, l'extension des *O*-fucoses, catalysée par les différentes enzymes de la famille Fringe, est essentielle au bon fonctionnement de la voie Notch puisqu'elle module la capacité des différents ligands à induire la transduction du signal. Il a également été montré que la présence de Fringe induit un changement conformationnel du récepteur NOTCH1, probablement suite à l'extension d'un ou plusieurs *O*-fucoses qu'il contient. Ceci augmenterait l'accessibilité et donc la *O*-fucosylation de l'ELD 26 indispensable au bon fonctionnement de cette voie Notch (Kakuda and Haltiwanger, 2017). La localisation golgienne de Fringe implique l'existence d'un transport rétrograde de NOTCH1 vers le RE suite à la modification de certains de ses *O*-fucoses par Fringe. Ainsi, le site de *O*-fucosylation de l'ELD 26, libéré de certaines contraintes conformationnelles, peut être *O*-fucosylé. Nous pouvons donc nous poser la question de l'impact de la sous-expression de POFUT1 sur le ratio *O*-fucoses courts/ *O*-fucoses rallongés car l'expression des enzymes de la famille Fringe demeure inchangée. Si ce ratio variait, l'effet des ligands DLL1 et JAG1 sur l'activation de la voie Notch pourrait aussi être modifié, ce qui expliquerait la grande perturbation de cette voie dans notre modèle murin. Malheureusement, de telles études nécessitent l'analyse du récepteur NOTCH1 endogène exprimé à la surface des MDCS, ce qui



n'est pas envisageable avec les techniques actuelles. En effet, la très faible expression du récepteur NOTCH1, qui est une protéine de très grande taille et de nature membranaire, exclut la purification de la protéine endogène. De plus, l'absence d'anticorps et de lectines reconnaissant les *O*-fucoses de manière spécifique ainsi que l'absence de traitement enzymatique ou chimique supprimant exclusivement ces sucres, accentue fortement la difficulté d'analyser la *O*-fucosylation d'une telle protéine.

Comme beaucoup d'études, nous nous sommes particulièrement intéressés au récepteur NOTCH1. Néanmoins, il serait intéressant d'analyser le comportement de NOTCH3 par des techniques similaires à celles utilisées pour NOTCH1. En effet, le récepteur NOTCH3 est très proche structurellement de NOTCH1 et partage un certain nombre de ses caractéristiques, à savoir plusieurs ELDs *O*-fucosylables (20 pour NOTCH1 et 15 pour NOTCH3) et les mêmes ligands DSL. Son étude serait d'autant plus intéressante que la signalisation induite par NOTCH3 semble antagoniste à celle de NOTCH1 (Beatus et al., 1999; Bellavia et al., 2008). Ainsi, l'effet de la sous-expression du gène *Pofut1* sur la *O*-fucosylation de NOTCH3 aurait pu contrebalancer celle de NOTCH1, ce qui ne semble pas être totalement le cas. Il est donc possible que la diminution de la quantité de POFUT1 ait plus d'incidence sur la *O*-fucosylation de NOTCH1 que sur celle de NOTCH3 dans notre modèle murin ou que l'expression de ce récepteur soit bien moindre pour que la diminution de son taux de *O*-fucosylation ait un effet majeur.

Enfin, dans un souci d'analyser l'altération de la voie de signalisation Notch dans sa globalité chez les souris *Pofut1*<sup>cax/cax</sup>, il serait pertinent d'étudier l'effet de la diminution de quantité de l'enzyme POFUT1 sur l'expression en surface des récepteurs NOTCH 2 et 4, leur clivage et l'activation des gènes cibles. Dans le même esprit, l'analyse de l'apport exogène de leurs différents ligands canoniques, non testés dans notre étude (JAG2, DLL3 et DLL4), nous permettrait d'avoir une vue d'ensemble sur le bon fonctionnement de la voie Notch et le caractère essentiel de la *O*-fucosylation induite par POFUT1.

### **Effets induits par la sous-expression du gène *Pofut1* chez les souris *Pofut1*<sup>cax/cax</sup>**

L'analyse des souris *Pofut1*<sup>cax/cax</sup> a révélé plusieurs défauts phénotypiques en relation avec la dérégulation de la voie Notch tels que des altérations de la somitogenèse (corps compact, queue tordue, vertèbres fusionnées...) et du développement musculaire post-natal



(hypertrophie musculaire, diminution du pool de cellules satellites...) (Al Jaam et al., 2016; Schuster-Gossler et al., 2009). Il a également été observé chez ces souris *Pofut1*<sup>cax/cax</sup> une sous-expression de POFUT1 dans de nombreux tissus et organes. Cette sous-expression de POFUT1, montrée de manière ubiquitaire (Heu, 2013), pourrait perturber *via* la dérégulation de la voie Notch, de nombreux processus développementaux autres que le développement musculo-squelettique ainsi que le maintien de l'homéostasie adulte (Dou et al., 2011; Janghorban et al., 2018; Zhang et al., 2018).

Dans le but de révéler d'autres effets potentiels de la sous-expression de POFUT1 chez les souris *Pofut1*<sup>cax/cax</sup>, un crible phénotypique primaire a été réalisé à la clinique de la souris en Allemagne ([www.mouseclinic.de/](http://www.mouseclinic.de/)). Ce crible avait pour but d'analyser sur 10 semaines 19 fonctions biologiques telles que le comportement, le profil neurologique, le métabolisme énergétique, la nociception, la morphologie oculaire, les propriétés hématologiques, la biochimie clinique, l'immunologie, etc... Ce crible phénotypique a confirmé les défauts musculo-squelettiques et mis en évidence un certain nombre de nouvelles fonctions altérées chez les souris *Pofut1*<sup>cax/cax</sup> telles que des défauts d'audition, des perturbations du système digestif et urinaire, une résistance à l'anesthésie, des variations morphologiques (oculaires et squelettiques) mais aussi des problèmes comportementaux (potentiellement liés à l'anxiété, l'hyperphagie et l'hyperactivité), métaboliques (plus faible concentration d'insuline, de glucose, de fer et de LDH) et hématologiques (quantité différente d'hématies, de leucocytes, et de plaquettes) (Annexe). L'autre objectif de ce crible phénotypique était d'évaluer les conséquences de l'hypertrophie musculaire modérée des souris *Pofut1*<sup>cax/cax</sup> sur la force et la résistance musculaire. Cependant, la comparaison avec des souris sauvages C3H n'a pas révélé de différences significatives.

Certaines expériences faisant partie du criblage phénotypique et montrant une altération de fonction ont été effectuées une seconde fois (à quelques semaines d'intervalle) mais les résultats n'ont pas tous été confirmés; c'est notamment le cas pour les tests hématologiques. De plus, certaines perturbations physiologiques (coordination, concentration d'insuline, perte d'eau, température de surface) ne sont retrouvées que chez un des deux sexes. Ces problèmes de reproductibilité des résultats pourraient s'expliquer par des différences individuelles dans le niveau d'expression du gène *Pofut1* au sein des souris *Pofut1*<sup>cax/cax</sup>. En effet, ces souris présentent des défauts squelettiques de sévérité variable (Schuster-Gossler et al., 2009) liés à la nature même de la mutation, à savoir l'insertion d'un élément rétrotransposable de type IAP dans l'intron 4 du gène *Pofut1*. Or, les éléments IAP sont connus pour posséder



leurs propres séquences promotrices de type LTR et pour s'intégrer dans le génome en position anti-sens (Qin et al., 2010). Leur présence pourrait par conséquent induire la synthèse de transcrits chimères potentiellement capables de s'hybrider aux transcrits sauvages de *Pofut1* et de les dégrader par un mécanisme d'interférence. Bien que ce mécanisme d'interférence n'ait pas été démontré chez les souris *Pofut1<sup>cax/cax</sup>*, la présence de transcrits chimères de *Pofut1* a bien été mise en évidence (Heu, 2013). Ainsi, il est tout à fait envisageable que la régulation épigénétique du promoteur de l'élément IAP intégré au génome des souris *Pofut1<sup>cax/cax</sup>* induise une diminution de la quantité de transcrits *Pofut1*, de manière variable en fonction des individus et par conséquent de la quantité de l'enzyme associée. Ceci pourrait entraîner une dérégulation variable de la voie de Notch en fonction des individus, avec comme conséquence des défauts squelettiques de sévérité plus ou moins importante. Cela pourrait aussi expliquer la non-reproductibilité ainsi que la faible significativité de certains résultats obtenus lors du crible phénotypique primaire.

Il serait intéressant d'effectuer ce même crible phénotypique mais en séparant les souris *Pofut1<sup>cax/cax</sup>* selon la sévérité de leur phénotype squelettique et en s'assurant de la corrélation avec leur niveau d'expression de *Pofut1*. Cependant, cela nécessiterait la génération d'un grand nombre d'animaux afin de pouvoir les catégoriser en fonction de leurs sexes et de la sévérité de leur phénotype squelettique. Ceci s'avère assez délicat étant donné la difficulté d'obtenir un grand nombre de souris avec un phénotype sévère, induisant souvent une létalité. Une autre possibilité serait d'effectuer des analyses plus précises sur les différentes fonctions biologiques altérées de façon significative chez les souris *Pofut1<sup>cax/cax</sup>*. Cependant, la plupart de ces fonctions (comportement, neurogenèse, morphologie oculaire, hématologies, métabolisme, etc...) sont beaucoup trop éloignées de nos thématiques actuelles et nécessitent donc de se rapprocher de différents laboratoires ou d'envisager certains cribles phénotypiques secondaires réalisables à la clinique de la souris.

Enfin, nous pourrions étudier l'effet de la sous-expression du gène *Pofut1* sur le processus de régénération musculaire chez les souris *Pofut1<sup>cax/cax</sup>*. En effet, la déplétion en cellules satellites chez ces souris pourrait fortement impacter leur capacité à régénérer un muscle lésé suite à une blessure ou une pathologie. Il serait également intéressant d'observer ces capacités de régénération chez des souris âgées, puisque le vieillissement est associé à une diminution du pool de cellules satellites dans les conditions physiologiques (Brack et al., 2005). Ainsi, les potentiels problèmes de régénération musculaires des souris *Pofut1<sup>cax/cax</sup>* pourraient être accentués chez les individus âgés.



## **Autres protéines potentiellement affectées par la sous-expression du gène *Pofut1* chez les souris *Pofut1<sup>cax/cax</sup>***

L'hypo-*O*-fucosylation du récepteur NOTCH1 entraîne une dérégulation de la voie Notch associée à des défauts de somitogenèse et de myogenèse chez les souris *Pofut1<sup>cax/cax</sup>* ainsi que différentes altérations révélées par le crible phénotypique. Le récepteur NOTCH1 n'est cependant pas la seule protéine à être impactée par la sous-expression du gène *Pofut1*. En effet, plusieurs autres protéines sont également *O*-fucosylées ; c'est notamment le cas du récepteur NOTCH3 (Arboleda-Velasquez et al., 2005) qui possède une fonction antagoniste à celle du récepteur NOTCH1 (Beatus et al., 1999) et dont l'altération du profil de *O*-fucosylation pourrait conduire à une activation atténuée ou exacerbée de la signalisation Notch (selon le rôle des *O*-fucoses de NOTCH3). Les ligands des différents récepteurs NOTCH, DLL1 et JAG1, sont également porteurs de *O*-fucoses (Panin et al., 2002). Il est néanmoins admis que les fucoses *O*-liés portés par ces ligands n'influent pas sur l'activation de la voie Notch (Müller et al., 2014) ; néanmoins d'autres effets peuvent être envisagés concernant l'adressage, la maturation et/ou la régulation de ces ligands.

Des protéines indépendantes de la voie de signalisation Notch possèdent aussi des fucoses *O*-liés. C'est le cas de la protéine CRIPTO dont la présence d'un *O*-fucose sur son unique ELD favorise la signalisation de NODAL (Schiffer et al., 2001), essentielle à la formation du mésoderme et de la structure axiale lors du développement embryonnaire. Etant donné que le mésoderme est le feuillet embryonnaire responsable de la mise en place de somites, il est donc possible que la perturbation de la *O*-fucosylation de CRIPTO puisse induire une altération de la signalisation de Nodal et donc de la somitogenèse. Ainsi, CRIPTO pourrait aussi participer au défaut de mise en place des somites chez les souris *Pofut1<sup>cax/cax</sup>*, se traduisant principalement par les troubles musculo-squelettiques. Néanmoins, une autre étude montre que l'effet de CRIPTO sur Nodal serait plutôt dépendant de la présence de la thréonine *O*-fucosylable et non du fucose *O*-lié lui-même (Shi et al., 2007).

D'autres protéines pourraient être corrélées aux perturbations des différents processus physiologiques révélées par le crible phénotypique primaire des souris *Pofut1<sup>cax/cax</sup>*. La perte de la *O*-fucosylation de l'AGRINE stimule la formation de cluster des récepteurs à l'acétylcholine (AChR) (Kim et al., 2008), ce qui pourrait éventuellement expliquer la résistance à l'anesthésie observée chez les souris *Pofut1<sup>cax/cax</sup>*. Les facteurs de coagulation VII, IX et XII ainsi que l'UROKINASE (impliquée dans la fibrinolyse), également *O*-fucosylés



chez l'Homme (Bjoern et al., 1991; Harris et al., 1992; Kentzer et al., 1990; Nishimura et al., 1992), pourraient quant à eux être associés aux altérations hématologiques (excepté pour le facteur IX, dont la séquence de *O*-fucosylation est absente chez la souris) mises en évidence par la clinique de la souris (Annexe) mais également au sein de notre équipe (plus forte viscosité sanguine chez les souris Pofut1<sup>cax/cax</sup>). Il faut également noter l'existence d'une autre protéine *O*-fucosylée, AMACO (Gebauer et al., 2008), dont le rôle est encore inconnu, ce qui rend impossible d'anticiper un éventuel rôle de sa *O*-fucosylation.

L'étude de la *O*-fucosylation de ces protéines s'avère très délicate pour les mêmes raisons que celles évoquées pour le récepteur NOTCH1. Seuls les facteurs de coagulation VII, IX et XII ainsi que l'UROKINASE sont naturellement présents en quantité suffisante pour pouvoir être étudiés directement (Bjoern et al., 1991; Harris et al., 1992; Kentzer et al., 1990; Nishimura et al., 1992). Au sein de l'équipe, des tentatives ont été initiées pour déterminer les profils de *O*-fucosylation des protéines sériques des souris Pofut1<sup>cax/cax</sup> par spectrométrie de masse (Heu, 2013). Cependant, la faible proportion des facteurs de coagulation par rapport à d'autres protéines sériques comme l'albumine ainsi que la présence de nombreuses glycosylations ont empêché l'obtention de résultats concluants. La déplétion des échantillons sanguins en albumine est une solution envisageable afin de concentrer les autres protéines sériques dont les facteurs de coagulation. Néanmoins, nous nous heurterons toujours aux problèmes liés à la présence de nombreuses glycosylations au sein des protéines sériques restantes. En effet, s'il est possible de supprimer les *N*-glycanes des protéines par traitement enzymatique, il est impossible de cliver les *O*-glycanes sans que les *O*-fucoses soient touchés.

Outre ces protéines *O*-fucosylées, il en existe d'autres susceptibles de porter des *O*-fucoses. Deux listes de protéines (humaines et murines), basées sur la présence d'un ou plusieurs ELDs possédant une séquence consensus de *O*-fucosylation, ont révélé la présence d'un peu moins d'une centaine de protéines potentiellement *O*-fucosylables (Al Jaam, 2016; Schneider et al., 2017). Toutes ces protéines jouent des rôles très divers et l'altération de leur éventuelle *O*-fucosylation pourrait être associée aux phénotypes des souris Pofut1<sup>cax/cax</sup>. Il est cependant nécessaire de démontrer la nature *O*-fucosylée ou non de ces protéines potentiellement cibles de POFUT1, avant de s'intéresser aux rôles de ces sucres dans l'activité de ces protéines et particulièrement dans le contexte des souris Pofut1<sup>cax/cax</sup>.

Récemment, nous avons démontré que la protéine WIF1, connue pour inhiber la voie de signalisation Wnt, pouvait être *O*-fucosylée (Pennarubia et al., 2018a). Chez les souris



*Pofut1*<sup>cax/cax</sup>, la *O*-fucosylation de WIF1 pourrait être diminuée voire absente, ce qui pourrait avoir une incidence sur la modulation de la voie Wnt impliquée notamment dans la myogenèse mais aussi dans d'autres processus développementaux. Par conséquent, l'étude de la *O*-fucosylation de WIF1 chez les souris *Pofut1*<sup>cax/cax</sup> s'avère très intéressante, mais la présence de WIF1 en très faible quantité rend son étude inenvisageable. Il serait néanmoins possible d'étudier la voie de signalisation Wnt dans le contexte de la sous-expression de *Pofut1* de notre modèle murin afin de détecter d'éventuelles dérégulations de cette voie.

### **La protéine WIF1 est une nouvelle protéine *O*-fucosylée : Quels sont les rôles possibles du fucose *O*-lié ?**

Nous avons démontré que la protéine WIF1 murine était porteuse d'un *O*-fucose uniquement sur son ELD III et que ce fucose *O*-lié était favorablement impliqué dans la sécrétion de cette protéine (Pennarubia et al., 2018a). Cependant, nous ignorons encore si ce *O*-fucose peut être impliqué dans les autres fonctions biologiques de WIF1, ce qui offre des perspectives intéressantes pour la poursuite de notre étude.

WIF1 est un inhibiteur de la voie Wnt, qui exerce son activité en se liant directement aux protéines de la famille Wnt; les empêchant ainsi de se fixer à leur récepteur Frizzled et donc d'induire la transduction du signal. Le domaine WIF est indispensable et suffisant pour conférer à WIF1 son activité inhibitrice envers les protéines Wnts (Hsieh et al., 1999). Néanmoins, les ELDs renforcent l'interaction WIF1/Wnt et donc l'inhibition de la voie Wnt (Malinauskas et al., 2011). Un des rôles connus des *O*-fucoses est de moduler les interactions protéines-protéines, comme c'est le cas pour le récepteur NOTCH et ses ligands (Moloney et al., 2000a; Rampal et al., 2005). Il est donc probable que la *O*-fucosylation de WIF1 puisse avoir un effet modulateur sur ces interactions protéiques, c'est d'ailleurs l'objet de plusieurs expérimentations en cours dans l'équipe.

La structure de la protéine WIF1 humaine, obtenue après cristallisation et diffraction aux rayons X, montre une zone d'interaction hydrophobe entre l'ELD III et le domaine WIF, qui procure à la protéine plus de stabilité (Malinauskas et al., 2011). Nous avons révélé la présence d'un *O*-fucose et aussi d'un *N*-glycane sur l'ELD III de la protéine WIF1 murine, qui pourraient être impliqués dans cette interaction et donc influencer sur la conformation de la protéine WIF1. Ainsi, le défaut de sécrétion de la protéine WIF1 murine mutée au niveau du



site de *O*-fucosylation de l'ELD III pourrait être lié à un changement de conformation dû à l'absence du *O*-fucose.

Il a été montré que les ELDs II à V (globalement chargés positivement) de la protéine WIF1 étaient capables d'établir des liaisons ioniques avec des protéoglycanes à héparane sulfate (HSPGs), globalement chargés négativement. Cette liaison de WIF1 au HSPGs permettrait de contrôler le gradient morphogénétique des Wnts, en empêchant une partie des protéines Wnt d'aller se fixer à leur récepteur Frizzled (Malinauskas et al., 2011). La quantification des interactions WIF1 (sauvage ou muté)/HSPGs nous permettrait ainsi de savoir si la *O*-fucosylation de l'ELD III de WIF1 influe sur ces interactions avec les HSPGs de la matrice extracellulaire. Des tests *in cellulo* consistant en l'ajout de WIF1 (sauvage et muté) exogène sur des cultures cellulaires, suivi d'une quantification de leur fixation à la surface cellulaire pourraient nous éclairer sur l'implication des fucoses *O*-liés dans l'interaction WIF1/HSPGs. Cependant, ces méthodes ne nous permettraient pas de savoir si la *O*-fucosylation influe directement sur l'interaction WIF1/Wnt ou uniquement sur la liaison WIF1/HSPGs. En effet, la variation de la liaison de WIF1 aux HSPGs a une grande probabilité de se répercuter sur la disponibilité des Wnts à se fixer sur leur récepteur, sans que l'affinité de liaison entre WIF1 et Wnt ne soit modifiée. Parmi les approches visant à étudier l'interaction WIF1/Wnt, la technique NanoLuc sera expérimentée dans l'équipe. Elle consiste à produire nos deux protéines recombinantes d'intérêt (WIF1 et une protéine Wnt potentiellement partenaire), fusionnées chacune avec une partie tronquée non fonctionnelle de la protéine NanoLuc. Si l'interaction WIF1/Wnt est effective, elle induira la reconstitution de la protéine NanoLuc, ce qui la rendra active pour émettre un signal lumineux quantifiable. Selon le même principe, il est également possible d'utiliser des anticorps secondaires spécifiques fusionnés à ces deux parties tronquées de la protéine NanoLuc, ciblant les anticorps primaires dirigé contre nos protéines d'intérêt. Cette deuxième stratégie, nécessitant de disposer de très bons anticorps primaires, est à privilégier car elle éviterait la production de nouvelles protéines recombinantes, fusionnées à un peptide supplémentaire susceptible d'influer sur l'interaction WIF1/Wnt.

Nous avons aussi montré que le *O*-fucose porté par l'ELD III de WIF1 pouvait être rallongé *in vitro* par un GlcNAc, suite à l'action de Lunatic Fringe produit de manière recombinante. Il est connu que l'allongement des *O*-fucoses du récepteur NOTCH1 module son interaction avec ses différents ligands DSL (Kakuda and Haltiwanger, 2017). Ainsi, l'action de Fringe sur le *O*-fucose de WIF1 pourrait moduler les interactions WIF1/Wnt et/ou



WIF1/HSPGs. Cependant, avant d'étudier le rôle biologique de l'extension du *O*-fucose de WIF1 par Fringe, il est nécessaire de démontrer que ce processus a réellement lieu *in cellulo*. Pour cela, des analyses par spectrométrie de masse visant à mettre en évidence le disaccharide GlcNAc $\beta$ 1-3Fucose sur la protéine WIF1 recombinante ont été entreprises. Malheureusement, nous n'avons pas réussi à détecter le peptide portant le site de *O*-fucosylation sur l'ELD III après digestion trypsique, même après traitement à la PNGase F de la protéine WIF1 entière. Ceci pourrait résulter de problèmes de clivages, d'une mauvaise ionisation ou de peptides trop longs pour être détectés. Plusieurs expériences supplémentaires sont actuellement en cours, elles consistent à digérer la protéine WIF1 avec d'autres enzymes (Glu-C,...) pour générer des peptides différents et potentiellement détectables. Si la présence d'un *O*-fucose allongé est avérée, il serait alors intéressant de savoir quelles enzymes de la famille Fringe (Lunatic, Manic et/ou Radical) sont impliquées. Cette information serait utile pour décider du ou des KO d'enzymes à effectuer sur nos lignées afin d'obtenir une protéine WIF1 portant uniquement des fucoses *O*-liés non allongés. Une fois ces protéines obtenues, nous pourrions par les techniques citées précédemment nous concentrer sur l'effet possible de l'allongement du *O*-fucose de WIF1 sur sa sécrétion et/ou sa capacité à inhiber la voie Wnt.

Outre l'impact possible de la *O*-fucosylation sur l'interaction WIF1/Wnt, il serait intéressant d'évaluer en amont le rôle biologique de WIF1 dans nos modèles cellulaires (C2C12 et cellules cancéreuses colorectales humaines). Pour cela, nous pourrions regarder l'effet de l'apport exogène de WIF1 (sauvage ou muté) sur les capacités de prolifération et de différenciation des C2C12, comme cela a été réalisé pour les protéines sFRP1 et 2 (Descamps et al., 2008), également antagonistes de la voie Wnt. Nous pourrions ainsi évaluer le rôle de WIF1 dans la différenciation myogénique et le rôle éventuel du fucose *O*-lié, qu'il soit allongé ou non. Des expérimentations similaires sur nos lignées cancéreuses colorectales nous permettraient d'en apprendre davantage sur le rôle de WIF1 et éventuellement de sa *O*-fucosylation dans la capacité proliférative et migratoire des cellules cancéreuses. En effet, il est connu que l'expression de WIF1 est très souvent inactivée par hyperméthylation de son promoteur dans de nombreux cancers, ce qui conduit à une sur-activation de la voie Wnt associée à une prolifération accrue (He et al., 2005; Mazieres et al., 2004; Yang et al., 2010).



## Une conservation inégale des sites de *O*-fucosylation chez les gnathostomes

Nous avons montré par des approches phylogénétiques que le site de *O*-fucosylation de l'ELD III de la protéine WIF1 était fortement conservé chez les gnathostomes, ce qui suggère l'occupation réelle de ce site par un *O*-fucose et son maintien au cours de l'évolution. Cependant, dans quelques cas particuliers, comme chez le pika (*Ochotona princeps*), l'ornithorynque (*Ornithorhynchus anatinus*), les serpents et les ambulacraires, ce site de *O*-fucosylation est absent dans l'ELD III (Pennarubia et al., 2018a). Nous avons montré pour la protéine WIF1 murine que son ELD III portait bien un *O*-fucose, nécessaire à une sécrétion optimale de la protéine. Ainsi, il serait intéressant d'étudier, pour les espèces précédemment citées, l'effet de l'absence du site de *O*-fucosylation sur l'ELD III sur la sécrétion de WIF1 et potentiellement sur son activité inhibitrice. Il est cependant possible que, chez ces espèces, l'absence de fucose *O*-lié sur l'ELD III soit compensée par la présence d'un autre site *O*-fucosylable. Cela pourrait être le cas de l'ornithorynque et des ambulacraires, pour lesquels on retrouve un site de *O*-fucosylation respectivement dans l'ELD I et l'ELD II. Nos modèles structuraux suggèrent que ces ELDs seraient de bons substrats pour POFUT1 et donc probablement *O*-fucosylés (Pennarubia et al., 2018a). Il serait donc pertinent de tester biochimiquement la *O*-fucosylation de ces ELDs par notre approche expérimentale *in vitro* (Pennarubia et al., 2018b) puis de se focaliser sur les effets de la présence de fucose *O*-lié (ailleurs que sur l'ELD III) sur la sécrétion de WIF1 pour ces espèces et ses interactions avec les protéines Wnt et les HSPGs. Dans l'hypothèse où ces protéines WIF1 seraient moins sécrétées et/ou moins actives, il serait judicieux de regarder le profil d'expression des protéines Wnt et des autres inhibiteurs de cette voie (Dickkopf, Cerberus et sFRP), afin de vérifier la présence d'un éventuel phénomène de compensation pour contrôler la signalisation Wnt.

Des études similaires, relatives à la détermination du caractère *O*-fucosylable *in vitro* d'ELDs d'intérêt (Pennarubia et al., 2018b), pourraient être envisagées pour les différentes espèces de protostomiens, qui présentent de nombreuses divergences de séquences protéiques et donc de séquences consensus de *O*-fucosylation. Cependant, de telles études nécessitent d'élucider en amont le rôle de la *O*-fucosylation de WIF1 (en dehors de la sécrétion) afin de rendre toute comparaison possible. Cela suggère la production de nombreuses protéines recombinantes pour chaque espèce étudiée (WIF1, Wnts). Néanmoins, les résidus de POFUT1 impliqués dans l'interaction avec les ELDs et donc dans le processus de *O*-fucosylation sont très conservés (Pennarubia et al., 2018a): il semble donc possible de *O*-fucosyler un ELD



d'une espèce donnée avec une *O*-fucosyltransférase d'une autre espèce (Lira-Navarrete et al., 2011). Ceci suggère qu'une seule lignée puisse être utilisée (cellules CHO) pour produire des protéines WIF1 recombinantes correctement *O*-fucosylées et cela peu importe l'espèce dont elles sont issues.

### **L'interaction POFUT1/ELD est à étudier au cas par cas**

Nos deux dernières études ont montré que la présence d'une séquence consensus de *O*-fucosylation, (C<sup>2</sup>-X-X-X-X-(**S/T**)-C<sup>3</sup>) (Müller et al., 2014), était nécessaire mais insuffisante pour prédire le caractère *O*-fucosylable d'un ELD que ce soit *in vitro* ou *in vivo*. En effet, l'affinité de POFUT1 pour les ELDs module l'efficacité de transfert du *O*-fucose puisque l'interaction POFUT1-ELD est étroitement dépendante de la nature des résidus qui composent leur sous-domaine C<sup>2</sup>-C<sup>3</sup> (portant le site de *O*-fucosylation) mais aussi leur sous-domaine C<sup>5</sup>-C<sup>6</sup> et leur boucle C<sup>1</sup>-C<sup>2</sup> (Li et al., 2017; Pennarubia et al., 2018a, 2018b). L'interaction POFUT1-ELD se fait principalement par l'établissement de liaison hydrogène entre le sous-domaine C<sup>2</sup>-C<sup>3</sup> de l'ELD et certains résidus conservés de POFUT1. Ces interactions permettent le positionnement correct de l'ELD dans la cavité acceptrice (au fond de laquelle est logé le GDP-fucose), située entre les deux domaines de Rossman de l'enzyme, pour qu'il soit *O*-fucosylé. Ainsi, chaque ELD devrait être testé individuellement afin de déterminer son caractère *O*-fucosylable ou non.

Cependant, il est important de garder à l'esprit que la *O*-fucosylation de chaque ELD est dépendante de son accessibilité à POFUT1 au sein de la protéine qui le porte. En effet, alors que l'ELD 26 du récepteur NOTCH1, est globalement moins *O*-fucosylé que l'ELD 12 *in cellulo* (Kakuda and Haltiwanger, 2017), nous avons montré *in vitro* que l'ELD 12 était beaucoup plus difficile à *O*-fucosyler par POFUT1 du fait d'une affinité beaucoup plus faible pour cette enzyme (Li et al., 2017; Pennarubia et al., 2018b). Ainsi, la *O*-fucosylation importante d'un ELD isolé *in vitro* ne signifie pas forcément une *O*-fucosylation efficace au sein de la protéine entière. Cependant, l'inverse semble plus révélateur puisque si un ELD isolé n'est pas apte à recevoir un *O*-fucose *in vitro*, il est peu probable qu'il puisse être *O*-fucosylé *in cellulo*. Il est donc indispensable d'étudier la *O*-fucosylation des ELDs dans le contexte de la protéine entière, pour se rapprocher le plus possible des conditions cellulaires.



Même si l'étude des protéines entières paraît plus pertinente, elle est toutefois plus difficile à réaliser du fait de leur disponibilité, de leur solubilité et de la possibilité de les purifier. Ainsi, il serait intéressant de faire un premier criblage des ELDs portés par les protéines cibles de POFUT1, afin de connaître leur aptitude à recevoir un *O*-fucose *in vitro* ou non. Cela permettrait d'éliminer tous les ELDs inaptes à être *O*-fucosylés *in vitro* et donc probablement *in cellulo*. Enfin, des analyses par spectrométrie de masse et/ou chimie click pourraient être effectuées sur les protéines entières portant des ELDs *O*-fucosylables *in vitro*. Cela permettrait de savoir si leur *O*-fucosylation *in cellulo* est également possible ou soumise à diverses contraintes, telles que l'accessibilité.

### Vers une nouvelle séquence consensus de *O*-fucosylation ?

A ce jour, la séquence consensus de *O*-fucosylation d'un ELD est C<sup>2</sup>-X-X-X-X-(S/T)-C<sup>3</sup> (Müller et al., 2014). Cependant, la nature des résidus du sous-domaine C<sup>2</sup>-C<sup>3</sup> influe sur la capacité de POFUT1 à interagir avec un ELD et donc à transférer un fucose sur la sérine ou la thréonine *O*-fucosylable. En effet, il a été montré sur l'ELD 12 du récepteur NOTCH1 que la présence d'un aspartate en position C<sup>2</sup>+3 (C<sup>2</sup>QNDATC<sup>3</sup>) était associée à une faible affinité de POFUT1 pour cet ELD. L'aspartate, du fait de sa longue chaîne latérale, induit un clash stérique avec la tyrosine 78 de POFUT1 (Li et al., 2017). Par contre, la mutation de ce résidu aspartate en glycine provoque une forte augmentation d'affinité de POFUT1 pour l'ELD (Li et al., 2017). Ceci est en accord avec le fait que la glycine, qui ne dispose pas de chaîne latérale capable d'induire un encombrement stérique, est fréquemment retrouvée en position C<sup>2</sup>+3 dans de nombreux ELDs *O*-fucosylés, tels que l'ELD 26 du récepteur NOTCH1 (C<sup>2</sup>FNGGTC<sup>3</sup>) et l'ELD du facteur VII de coagulation (C<sup>2</sup>QNGGSC<sup>3</sup>), tous deux fortement affins pour POFUT1. Nous avons confirmé que l'ELD 12 de NOTCH1 était moins efficacement *O*-fucosylé *in vitro* par POFUT1 que l'ELD 26 (Pennarubia et al., 2018a), ce qui peut être relié à la faible affinité de l'enzyme pour l'ELD 12 (Li et al., 2017). Nous avons également montré que l'ELD V de WIF1 murin, dont la séquence de *O*-fucosylation est largement conservée chez les gnathostomes, n'était pas *O*-fucosylable *in vitro* ni *O*-fucosylé *in cellulo*. A l'aide de modèles par homologie et de la structure cristallographique du complexe murin POFUT1/ELD 26 (Li et al., 2017), nous avons corrélé ce défaut de *O*-fucosylation à la présence dans l'ELD V d'une histidine en position C<sup>2</sup>+3 (C<sup>2</sup>GAHGTC<sup>2</sup>), dont la longue chaîne latérale induirait un clash stérique important avec le résidu Y<sup>78</sup> de



POFUT1. Ce clash, plus important encore que celui occasionné par l'aspartate de l'ELD 12 de NOTCH1, aboutirait à défaut d'interaction entre l'ELD V et POFUT1, prévenant toute *O*-fucosylation. Ainsi, on peut penser que selon l'importance du clash stérique à cette position C<sup>2</sup>+3, l'affinité de POFUT1 pour un ELD peut être plus ou moins diminuée voire nulle, provoquant dans ce dernier cas un défaut d'interaction et donc de *O*-fucosylation. Etant donné que les ELDs sont structurellement très proches entre eux, il est fort probable que la présence d'une histidine en position C<sup>2</sup>+3 dans un ELD l'empêche d'interagir correctement avec POFUT1 et par conséquent d'être *O*-fucosylé. Ainsi, il pourrait être utile de redéfinir la séquence consensus de *O*-fucosylation en précisant, notamment pour la position C<sup>2</sup>+3, les résidus favorables à une interaction avec POFUT1 et ceux qui le sont peu comme l'aspartate ou pas du tout comme l'histidine. Tous les résidus avec une chaîne latérale plus courte que celle de l'aspartate (glycine, alanine, valine, sérine, thréonine, cystéine et proline) ne devraient pas induire de clash stérique : parmi ceux-ci, notons que la cystéine n'est jamais retrouvée dans le sous-domaine C<sup>2</sup>-C<sup>3</sup> d'un ELD. Les résidus avec une chaîne latérale comparable à celle de l'aspartate (asparagine, leucine et isoleucine) ou légèrement plus imposante (glutamate, glutamine, lysine et méthionine) devraient induire un clash stérique suffisant pour diminuer l'interaction avec POFUT1, sans toutefois empêcher totalement la *O*-fucosylation. Par contre, il est fort probable que des résidus possédant une chaîne latérale plus encombrante que celle de l'histidine (tyrosine, arginine, phénylalanine et tryptophane) induisent un clash stérique drastique prévenant toute *O*-fucosylation. Afin de déterminer précisément quels résidus ne peuvent être présents en position C<sup>2</sup>+3 pour qu'un ELD soit *O*-fucosylé, nous pourrions nous focaliser sur l'un d'entre eux connu pour être très favorable à la *O*-fucosylation comme l'ELD 26. L'étude de formes variantes de cet ELDs, où seul le résidu en position C<sup>2</sup>+3 est remplacé par ceux susceptibles d'induire un clash stérique, nous permettrait donc de redéfinir la séquence consensus de *O*-fucosylation par exclusion des résidus défavorables.

Toujours dans l'optique de redéfinir une séquence consensus de *O*-fucosylation plus pertinente, il faudrait également considérer les résidus aux positions C<sup>2</sup>+1, C<sup>2</sup>+2 et C<sup>2</sup>+4, qui sont impliqués dans la formation de liaisons hydrogène avec certains résidus clés de l'enzyme POFUT1 (Li et al., 2017). L'établissement de toutes ces liaisons permet le bon positionnement de l'ELD dans la cavité d'interaction de POFUT1, afin de permettre sa *O*-fucosylation. En se basant sur la liste des ELDs *O*-fucosylables, nous pouvons noter certaines similarités entre les résidus présents à chacune de ces trois positions. Si la présence



d'une glutamine en position C<sup>2</sup>+1 induit la formation d'une liaison hydrogène avec POFUT1, de nombreux autres résidus peuvent la remplacer sans que cela ne perturbe la *O*-fucosylation. Cependant, aucun des ELDs démontrés comme étant *O*-fucosylés ne comporte de résidu polaire chargé négativement (glutamate ou aspartate) à cette position. Concernant les positions C<sup>2</sup>+2 et C<sup>2</sup>+4, la variation est beaucoup plus restreinte avec respectivement trois (asparagine, glutamine et histidine) et deux (alanine et glycine) acides aminés retrouvés chez tous les ELDs cibles de POFUT1. Etant donné que tous ces résidus varient de façon beaucoup plus importante pour les ELDs des protéines potentiellement cibles de POFUT1, la même approche expérimentale que celle citée précédemment pourrait être entreprise pour déterminer les résidus favorables ou non à chacune des positions du sous-domaine C<sup>2</sup>-C<sup>3</sup>.

Enfin, pour aller plus loin, nous pourrions aussi nous intéresser aux résidus présents hors du sous-domaine C<sup>2</sup>-C<sup>3</sup> des ELDs. En effet, certains d'entre eux sont associés à la formation de liaisons hydrogène ou apolaire (Li et al., 2017). De plus, nos modèles structuraux révèlent que le défaut d'interaction de l'ELD V de WIF1 avec POFUT1 pourrait être renforcé par la présence d'un second clash stérique potentiellement associé à une répulsion de charge en position C<sup>5</sup>+1 (Pennarubia et al., 2018a). Cependant, l'effet de ces résidus hors du domaine C<sup>2</sup>-C<sup>3</sup> sur la capacité d'interaction de POFUT1 et donc sur l'efficacité de la *O*-fucosylation semble beaucoup moins important (Li et al., 2017). L'affinage progressif de la séquence consensus de *O*-fucosylation nous permettrait de diminuer de manière non négligeable la liste des ELDs potentiellement aptes à recevoir un *O*-fucose et donc de redéfinir de manière plus pertinente la liste des protéines cibles de POFUT1. A terme, nous pourrions étudier la *O*-fucosylation *in cellulo* des protéines portant ces ELDs et entreprendre des études structure-fonction pour évaluer les différents rôles possibles de cette modification post-traductionnelle rare.

### **Qu'en est-il des autres *O*-glycosylations rares des ELDs, la *O*-glucosylation et la *O*-GlcNAcylation ?**

Les ELDs peuvent être modifiés par deux autres *O*-glycosylations rares : la *O*-glucosylation catalysée par l'enzyme POGLUT1 (Takeuchi et al., 2012) et la *O*-GlcNAcylation induite par l'enzyme EOGT (Sakaidani et al., 2011). Il serait intéressant de dresser une liste de protéines cibles de POGLUT1 et d'EOGT puis de confronter ces listes à celle établies pour POFUT1 afin d'identifier l'ensemble des ELDs portant plusieurs séquences



de *O*-glycosylation. De la même façon que pour POFUT1, nous pourrions utiliser la structure 3D des complexes POGLUT1/ELD 12 de NOTCH1 (Li et al., 2017) et y positionner *in silico*, à la place l'ELD 12, les différents modèles structuraux générés de manière automatique pour chacun des ELDs potentiellement *O*-glucosylables afin d'identifier d'éventuels problèmes d'interaction. Ces problèmes d'interaction potentiels pourraient par la suite être confirmés ou infirmés expérimentalement par des approches *in vitro* similaires à celles utilisées dans nos études (Pennarubia et al., 2018a, 2018b).

En ce qui concerne la *O*-GlcNAcylation, nous ne disposons pas de la structure 3D de l'enzyme EOGT. Nous ignorons donc comment cette enzyme interagit avec un ELD et quels sont les résidus clés impliqués dans l'interaction. Cependant, le manque d'informations sur les mécanismes d'interaction EOGT/ELD ne représente pas un frein à l'étude expérimentale de la capacité d'EOGT à transférer *in vitro* un GlcNAc sur un ELD. En effet, nous avons montré, à l'aide de la même stratégie expérimentale que celle utilisée pour étudier la *O*-fucosylation, la capacité des enzymes POGLUT1 et EOGT à ajouter respectivement un *O*-glucose et un *O*-GlcNAc sur un ELD recombinant produit de manière isolé.

Pour aller plus loin dans l'étude du fonctionnement de ces trois enzymes, il serait très intéressant d'étudier certains ELDs possédant les trois *O*-glycanes (*O*-fucose, *O*-glucose et *O*-GlcNAc) tel que l'ELD 20 de NOTCH1 (Kakuda and Haltiwanger, 2017). Nous pourrions ainsi aborder la question des capacités de transfert de ces trois enzymes réticulaires (POFUT1, POGLUT1 et EOGT), sont-elles dépendantes ou indépendantes les unes des autres ?



## REFERENCES BIBLIOGRAPHIQUES

---

- Aebi, M. (2013). N-linked protein glycosylation in the ER. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* *1833*, 2430–2437.
- Al Jaam, B. (2016). Aspects biochimiques et cellulaires de la dérégulation du processus myogénique chez des souris hypomorphes pour le gène Pofut I. thesis. Limoges.
- Al Jaam, B., Heu, K., Pennarubia, F., Segelle, A., Magnol, L., Germot, A., Legardinier, S., Blanquet, V., and Maftah, A. (2016). Reduced Notch signalling leads to postnatal skeletal muscle hypertrophy in Pofut1<sup>cax/cax</sup> mice. *Open Biol.* *6*.
- Al-Shareffi, E., Chaubard, J.-L., Leonhard-Melief, C., Wang, S.-K., Wong, C.-H., and Haltiwanger, R.S. (2013). 6-Alkynyl fucose is a bioorthogonal analog for O-fucosylation of epidermal growth factor-like repeats and thrombospondin Type-1 repeats by protein O-fucosyltransferases 1 and 2. *Glycobiology* *23*, 188–198.
- Anelli, T., and Sitia, R. (2008). Protein quality control in the early secretory pathway. *EMBO J.* *27*, 315–327.
- Arboleda-Velasquez, J.F., Rampal, R., Fung, E., Darland, D.C., Liu, M., Martinez, M.C., Donahue, C.P., Navarro-Gonzalez, M.F., Libby, P., D'Amore, P.A., et al. (2005). CADASIL mutations impair Notch3 glycosylation by Fringe. *Hum. Mol. Genet.* *14*, 1631–1639.
- Artavanis-Tsakonas, S. (1999). Notch Signaling: Cell Fate Control and Signal Integration in Development. *Science* *284*, 770–776.
- Beatus, P., Lundkvist, J., Oberg, C., and Lendahl, U. (1999). The notch 3 intracellular domain represses notch 1-mediated activation through Hairy/Enhancer of split (HES) promoters. *Development* *126*, 3925–3935.
- Bellavia, D., Checquolo, S., Campese, A.F., Felli, M.P., Gulino, A., and Screpanti, I. (2008). Notch3: from subtle structural differences to functional diversity. *Oncogene* *27*, 5092–5098.
- Bigas, A., Robert-Moreno, I., and Espinosa, L. (2010). The Notch pathway in the developing hematopoietic system. *Int. J. Dev. Biol.* *54*, 1175–1188.
- Bjoern, S., Foster, D.C., Thim, L., Wiberg, F.C., Christensen, M., Komiyama, Y., Pedersen, A.H., and Kisiel, W. (1991). Human plasma and recombinant factor VII. Characterization of O-glycosylations at serine residues 52 and 60 and effects of site-directed mutagenesis of serine 52 to alanine. *J. Biol. Chem.* *266*, 11051–11057.
- Brack, A.S., Bildsoe, H., and Hughes, S.M. (2005). Evidence that satellite cell decrement contributes to preferential decline in nuclear number from large fibres during murine age-related muscle atrophy. *J. Cell Sci.* *118*, 4813–4821.
- Bray, S.J. (2006). Notch signalling: a simple pathway becomes complex. *Nat. Rev. Mol. Cell Biol.* *7*, 678–689.



- Breton, C., Šnajdrová, L., Jeanneau, C., Koča, J., and Imberty, A. (2006). Structures and mechanisms of glycosyltransferases. *Glycobiology* *16*, 29R-37R.
- Breton, C., Fournel-Gigleux, S., and Palcic, M.M. (2012). Recent structures, evolution and mechanisms of glycosyltransferases. *Curr. Opin. Struct. Biol.* *22*, 540–549.
- Brockhausen, I., Schachter, H., and Stanley, P. (2009). O-GalNAc Glycans. In *Essentials of Glycobiology*, A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, and M.E. Etzler, eds. (Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press), p.
- Brückner, K., Perez, L., Clausen, H., and Cohen, S. (2000). Glycosyltransferase activity of Fringe modulates Notch–Delta interactions. *Nature* *406*, 411–415.
- Buettner, F.F.R., Ashikov, A., Tiemann, B., Lehle, L., and Bakker, H. (2013). C. elegans DPY-19 Is a C-Mannosyltransferase Glycosylating Thrombospondin Repeats. *Mol. Cell* *50*, 295–302.
- Cadigan, K.M., and Liu, Y.I. (2006). Wnt signaling: complexity at the surface. *J. Cell Sci.* *119*, 395–402.
- Chen, J., Moloney, D.J., and Stanley, P. (2001). Fringe modulation of Jagged1-induced Notch signaling requires the action of 4galactosyltransferase-1. *Proc. Natl. Acad. Sci.* *98*, 13716–13721.
- Chen, S., Lee, B.H., and Bae, Y. (2014). Notch Signaling in Skeletal Stem Cells. *Calcif. Tissue Int.* *94*, 68–77.
- Conboy, I.M., and Rando, T.A. (2002). The Regulation of Notch Signaling Controls Satellite Cell Activation and Cell Fate Determination in Postnatal Myogenesis. *Dev. Cell* *3*, 397–409.
- Conlon, R.A., Reaume, A.G., and Rossant, J. (1995). Notch1 is required for the coordinate segmentation of somites. *Development* *121*, 1533–1545.
- Coutinho, P.M., Deleury, E., Davies, G.J., and Henrissat, B. (2003). An Evolving Hierarchical Family Classification for Glycosyltransferases. *J. Mol. Biol.* *328*, 307–317.
- Dell, A., Galadari, A., Sastre, F., and Hitchen, P. (2010). Similarities and Differences in the Glycosylation Mechanisms in Prokaryotes and Eukaryotes. *Int. J. Microbiol.* *2010*, 1–14.
- Der Vartanian, A., Audfray, A., Al Jaam, B., Janot, M., Legardinier, S., Maftah, A., and Germot, A. (2015). Protein O-Fucosyltransferase 1 Expression Impacts Myogenic C2C12 Cell Commitment via the Notch Signaling Pathway. *Mol. Cell. Biol.* *35*, 391–405.
- Descamps, S., Arzouk, H., Bacou, F., Bernardi, H., Fedon, Y., Gay, S., Reyne, Y., Rossano, B., and Levin, J. (2008). Inhibition of myoblast differentiation by Sfrp1 and Sfrp2. *Cell Tissue Res.* *332*, 299–306.
- Dou, G.-R., Wang, L., Wang, Y.-S., and Han, H. (2011). Notch Signaling in Ocular Vasculature Development and Diseases. *Mol. Med.* *18*, 47–55.



- Gebauer, J.M., Müller, S., Hanisch, F.-G., Paulsson, M., and Wagener, R. (2008). *O* - Glucosylation and *O* -Fucosylation Occur Together in Close Proximity on the First Epidermal Growth Factor Repeat of AMACO (VWA2 Protein). *J. Biol. Chem.* 283, 17846–17854.
- Hallgren, P., Lundblad, A., and Svensson, S. (1975). A new type of carbohydrate-protein linkage in a glycopeptide from normal human urine. *J. Biol. Chem.* 250, 5312–5314.
- Haltiwanger, R.S., Wells, L., Freeze, H.H., and Stanley, P. (2015). Other Classes of Eukaryotic Glycans. In *Essentials of Glycobiology*, A. Varki, R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, N.H. Packer, J.H. Prestegard, et al., eds. (Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press), p.
- Harris, R.J., Ling, V.T., and Spellman, M.W. (1992). O-linked fucose is present in the first epidermal growth factor domain of factor XII but not protein C. *J. Biol. Chem.* 267, 5102–5107.
- Hart, G.W. (1997). Dynamic O-Linked Glycosylation of Nuclear and Cytoskeletal Proteins. *Annu. Rev. Biochem.* 66, 315–335.
- He, B., Reguart, N., You, L., Mazieres, J., Xu, Z., Lee, A.Y., Mikami, I., McCormick, F., and Jablons, D.M. (2005). Blockade of Wnt-1 signaling induces apoptosis in human colorectal cancer cells containing downstream mutations. *Oncogene* 24, 3054–3058.
- Heu, K. (2013). Caractérisation de la lignée murine hypomorphe Pofut1cax : un modèle d'étude in vivo du rôle de la O-fucosylation de Notch au cours de la myogenèse. thesis. Limoges.
- Hofsteenge, J., Müller, D.R., de Beer, T., Löffler, A., Richter, W.J., and Vliegthart, J.F. (1994). New type of linkage between a carbohydrate and a protein: C-glycosylation of a specific tryptophan residue in human RNase Us. *Biochemistry* 33, 13524–13530.
- Hofsteenge, J., Blommers, M., Hess, D., Furmanek, A., and Miroshnichenko, O. (1999). The Four Terminal Components of the Complement System Are C- Mannosylated on Multiple Tryptophan Residues. *J. Biol. Chem.* 274, 32786–32794.
- Hofsteenge, J., Huwiler, K.G., Macek, B., Hess, D., Lawler, J., Mosher, D.F., and Peter-Katalinic, J. (2001). C -Mannosylation and O -Fucosylation of the Thrombospondin Type 1 Module. *J. Biol. Chem.* 276, 6485–6498.
- Hori, K., Sen, A., and Artavanis-Tsakonas, S. (2013). Notch signaling at a glance. *J. Cell Sci.* 126, 2135–2140.
- Hsieh, J.-C., Kodjabachian, L., Rebbert, M.L., Rattner, A., Smallwood, P.M., Samos, C.H., Nusse, R., Dawid, I.B., and Nathans, J. (1999). A new secreted protein that binds to Wnt proteins and inhibits their activities. *Nature* 398, 431–436.
- Hu, J., Dong, A., Fernandez-Ruiz, V., Shan, J., Kawa, M., Martínez-Ansó, E., Prieto, J., and Qian, C. (2009). Blockade of Wnt Signaling Inhibits Angiogenesis and Tumor Growth in Hepatocellular Carcinoma. *Cancer Res.* 69, 6951–6959.



- Hunter, D.D., Zhang, M., Ferguson, J.W., Koch, M., and Brunken, W.J. (2004). The extracellular matrix component WIF-1 is expressed during, and can modulate, retinal development. *Mol. Cell. Neurosci.* 27, 477–488.
- Inoue, T., Oz, H.S., Wiland, D., Gharib, S., Deshpande, R., Hill, R.J., Katz, W.S., and Sternberg, P.W. (2004). *C. elegans* LIN-18 Is a Ryk Ortholog and Functions in Parallel to LIN-17/Frizzled in Wnt Signaling. *Cell* 118, 795–806.
- Janghorban, M., Xin, L., Rosen, J.M., and Zhang, X.H.-F. (2018). Notch Signaling as a Regulator of the Tumor Immune Response: To Target or Not To Target? *Front. Immunol.* 9.
- Kakuda, S., and Haltiwanger, R.S. (2017). Deciphering the Fringe-Mediated Notch Code: Identification of Activating and Inhibiting Sites Allowing Discrimination between Ligands. *Dev. Cell* 40, 193–201.
- Kawano, Y., and Kypta, R. (2003). Secreted antagonists of the Wnt signalling pathway. *J. Cell Sci.* 116, 2627–2634.
- Kentzer, E.J., Buko, A., Menon, G., and Sarin, V.K. (1990). Carbohydrate composition and presence of a fucose-protein linkage in recombinant human pro-urokinase. *Biochem. Biophys. Res. Commun.* 171, 401–406.
- Kerekes, K., Bányai, L., and Patthy, L. (2015). Wnts grasp the WIF domain of Wnt Inhibitory Factor 1 at two distinct binding sites. *FEBS Lett.* 589, 3044–3051.
- Kim, M.-L., Chandrasekharan, K., Glass, M., Shi, S., Stahl, M.C., Kaspar, B., Stanley, P., and Martin, P.T. (2008). O-fucosylation of muscle agrin determines its ability to cluster acetylcholine receptors. *Mol. Cell. Neurosci.* 39, 452–464.
- Kopan, R., and Ilagan, M.X.G. (2009). The Canonical Notch Signaling Pathway: Unfolding the Activation Mechanism. *Cell* 137, 216–233.
- Leonhard-Melief, C., and Haltiwanger, R.S. (2010). O-Fucosylation of Thrombospondin Type 1 Repeats. In *Methods in Enzymology*, (Elsevier), pp. 401–416.
- Li, Z., Han, K., Pak, J.E., Satkunarajah, M., Zhou, D., and Rini, J.M. (2017). Recognition of EGF-like domains by the Notch-modifying O-fucosyltransferase POFUT1. *Nat. Chem. Biol.* 13, 757–763.
- Liepinsh, E., Bányai, L., Patthy, L., and Otting, G. (2006). NMR Structure of the WIF Domain of the Human Wnt-Inhibitory Factor-1. *J. Mol. Biol.* 357, 942–950.
- Lira-Navarrete, E., and Hurtado-Guerrero, R. (2018). A perspective on structural and mechanistic aspects of protein O-fucosylation. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* 74, 443–450.
- Lira-Navarrete, E., Valero-González, J., Villanueva, R., Martínez-Júlvez, M., Tejero, T., Merino, P., Panjikar, S., and Hurtado-Guerrero, R. (2011). Structural Insights into the Mechanism of Protein O-Fucosylation. *PLoS ONE* 6, e25365.



- Loriol, C., Dupuy, F., Rampal, R., Dlugosz, M.A., Haltiwanger, R.S., Maftah, A., and Germot, A. (2006). Molecular evolution of protein O-fucosyltransferase genes and splice variants. *Glycobiology* *16*, 736–747.
- Loriol, C., Audfray, A., Dupuy, F., Germot, A., and Maftah, A. (2007). The two N-glycans present on bovine Pofut1 are differently involved in its solubility and activity: Role of *Bos taurus* Pofut1 N-glycans. *FEBS J.* *274*, 1202–1211.
- Lu, D., Dong, W., Zhang, X., Quan, X., Bao, D., Lu, Y., and Zhang, L. (2013). WIF1 causes dysfunction of heart in transgenic mice. *Transgenic Res.* *22*, 1179–1189.
- Luo, Y., and Haltiwanger, R.S. (2005). O-Fucosylation of Notch Occurs in the Endoplasmic Reticulum. *J. Biol. Chem.* *280*, 11289–11294.
- Malinauskas, T. (2008). Docking of Fatty Acids into the WIF Domain of the Human Wnt Inhibitory Factor-1. *Lipids* *43*, 227–230.
- Malinauskas, T., and Jones, E.Y. (2014). Extracellular modulators of Wnt signalling. *Curr. Opin. Struct. Biol.* *29*, 77–84.
- Malinauskas, T., Aricescu, A.R., Lu, W., Siebold, C., and Jones, E.Y. (2011). Modular mechanism of Wnt signaling inhibition by Wnt inhibitory factor 1. *Nat. Struct. Mol. Biol.* *18*, 886–893.
- Martínek, V., Sklenář, J., Dračinský, M., Šulc, M., Hofbauerová, K., Bezouška, K., Frei, E., and Stiborová, M. (2010). Glycosylation Protects Proteins against Free Radicals Generated from Toxic Xenobiotics. *Toxicol. Sci.* *117*, 359–374.
- Mazieres, J., He, B., You, L., Xu, Z., Lee, A.Y., Mikami, I., Reguart, N., Rosell, R., McCormick, F., and Jablons, D.M. (2004). Wnt Inhibitory Factor-1 Is Silenced by Promoter Hypermethylation in Human Lung Cancer. *Cancer Res.* *64*, 4717–4720.
- Meloty-Kapella, L., Shergill, B., Kuon, J., Botvinick, E., and Weinmaster, G. (2012). Notch Ligand Endocytosis Generates Mechanical Pulling Force Dependent on Dynamin, Epsins, and Actin. *Dev. Cell* *22*, 1299–1312.
- Mohr, O.L. (1919). Character Changes Caused by Mutation of an Entire Region of a Chromosome in *Drosophila*. *Genetics* *4*, 275–282.
- Moloney, D.J., Panin, V.M., Johnston, S.H., Chen, J., Shao, L., Wilson, R., Wang, Y., Stanley, P., Irvine, K.D., Haltiwanger, R.S., et al. (2000a). Fringe is a glycosyltransferase that modifies Notch. *Nature* *406*, 369–375.
- Moloney, D.J., Shair, L.H., Lu, F.M., Xia, J., Locke, R., Matta, K.L., and Haltiwanger, R.S. (2000b). Mammalian Notch1 Is Modified with Two Unusual Forms of O-Linked Glycosylation Found on Epidermal Growth Factor-like Modules. *J. Biol. Chem.* *275*, 9604–9611.
- Moremen, K.W., Tiemeyer, M., and Nairn, A.V. (2012). Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* *13*, 448–462.



- Müller, J., Rana, N.A., Serth, K., Kakuda, S., Haltiwanger, R.S., and Gossler, A. (2014). O-fucosylation of the Notch Ligand mDLL1 by POFUT1 Is Dispensable for Ligand Function. *PLoS ONE* 9, e88571.
- Ng, R.C.-L., Matsumaru, D., Ho, A.S.-H., Garcia-Barceló, M.-M., Yuan, Z.-W., Smith, D., Kodjabachian, L., Tam, P.K.-H., Yamada, G., and Lui, V.C.-H. (2014). Dysregulation of Wnt inhibitory factor 1 (Wif1) expression resulted in aberrant Wnt- $\beta$ -catenin signaling and cell death of the cloaca endoderm and anorectal malformations. *Cell Death Differ.* 21, 978–989.
- Nishimura, H., Takao, T., Hase, S., Shimonishi, Y., and Iwanaga, S. (1992). Human factor IX has a tetrasaccharide O-glycosidically linked to serine 61 through the fucose residue. *J. Biol. Chem.* 267, 17520–17525.
- Oka, O.B.V., and Bulleid, N.J. (2013). Forming disulfides in the endoplasmic reticulum. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1833, 2425–2429.
- Okajima, T., Xu, A., Lei, L., and Irvine, K.D. (2005). Chaperone Activity of Protein O-Fucosyltransferase 1 Promotes Notch Receptor Folding. *Science* 307, 1599–1603.
- Panin, V.M., Shao, L., Lei, L., Moloney, D.J., Irvine, K.D., and Haltiwanger, R.S. (2002). Notch Ligands Are Substrates for Protein O-Fucosyltransferase-1 and Fringe. *J. Biol. Chem.* 277, 29945–29952.
- Patthy, L. (2000). The WIF module. *Trends Biochem. Sci.* 25, 12–13.
- Pennarubia, F., Al Jaam, B., Pinault, E., Brun, C.E., Maftah, A., Germot, A., and Legardinier, S. (2018a). Mouse WIF1 is only O-fucosylated on EGF-III despite two evolutionarily conserved consensus sites. *Sci. Rep. Manuscr. Submitted Publ.*
- Pennarubia, F., Pinault, E., Maftah, A., and Legardinier, S. (2018b). In vitro acellular method to reveal O-fucosylation on 1 EGF-like domains. *Glycobiology.*
- Piccolo, S., Agius, E., Leyns, L., Bhattacharyya, S., Grunz, H., Bouwmeester, T., and Robertis, E.M.D. (1999). The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. *Nature* 397, 707–710.
- Qin, C., Wang, Z., Shang, J., Bekkari, K., Liu, R., Pacchione, S., McNulty, K.A., Ng, A., Barnum, J.E., and Storer, R.D. (2010). Intracisternal A particle genes: Distribution in the mouse genome, active subtypes, and potential roles as species-specific mediators of susceptibility to cancer. *Mol. Carcinog.* 49, 54–67.
- Rabinovich, G.A., van Kooyk, Y., and Cobb, B.A. (2012). Glycobiology of immune responses: Glycobiology of immune responses. *Ann. N. Y. Acad. Sci.* 1253, 1–15.
- Rampal, R., Arboleda-Velasquez, J.F., Nita-Lazar, A., Kosik, K.S., and Haltiwanger, R.S. (2005). Highly Conserved O-Fucose Sites Have Distinct Effects on Notch1 Function. *J. Biol. Chem.* 280, 32133–32140.
- Rattner, A., Hsieh, J.-C., Smallwood, P.M., Gilbert, D.J., Copeland, N.G., Jenkins, N.A., and Nathans, J. (1997). A family of secreted proteins contains homology to the cysteine-rich ligand-binding domain of frizzled receptors. *Proc. Natl. Acad. Sci.* 94, 2859–2863.



- Russell, D., Oldham, N.J., and Davis, B.G. (2009). Site-selective chemical protein glycosylation protects from autolysis and proteolytic degradation. *Carbohydr. Res.* *344*, 1508–1514.
- Sakaidani, Y., Nomura, T., Matsuura, A., Ito, M., Suzuki, E., Murakami, K., Nadano, D., Matsuda, T., Furukawa, K., and Okajima, T. (2011). O-Linked-*N*-acetylglucosamine on extracellular protein domains mediates epithelial cell–matrix interactions. *Nat. Commun.* *2*, 583.
- Sasamura, T., Ishikawa, H.O., Sasaki, N., Higashi, S., Kanai, M., Nakao, S., Ayukawa, T., Aigaki, T., Noda, K., Miyoshi, E., et al. (2007). The O-fucosyltransferase O-fut1 is an extracellular component that is essential for the constitutive endocytic trafficking of Notch in *Drosophila*. *Development* *134*, 1347–1356.
- Sawa, M., Hsu, T.-L., Itoh, T., Sugiyama, M., Hanson, S.R., Vogt, P.K., and Wong, C.-H. (2006). Glycoproteomic probes for fluorescent imaging of fucosylated glycans in vivo. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 12371–12376.
- Schiffer, S.G., Foley, S., Kaffashan, A., Hronowski, X., Zichittella, A., Yeo, C.-Y., Miatkowski, K., Adkins, H.B., Damon, B., Whitman, M., et al. (2001). Fucosylation of Cripto is required for its ability to facilitate Nodal signaling. *J. Biol. Chem.*
- Schneider, M., Al-Shareffi, E., and Haltiwanger, R.S. (2017). Biological functions of fucose in mammals. *Glycobiology* *27*, 601–618.
- Schuster-Gossler, K., Harris, B., Johnson, K.R., Serth, J., and Gossler, A. (2009). Notch signalling in the paraxial mesoderm is most sensitive to reduced Pofut1 levels during early mouse development. *BMC Dev. Biol.* *9*, 6.
- Shcherbakova, A., Tiemann, B., Buettner, F.F.R., and Bakker, H. (2017). Distinct C-mannosylation of netrin receptor thrombospondin type 1 repeats by mammalian DPY19L1 and DPY19L3. *Proc. Natl. Acad. Sci.* *114*, 2574–2579.
- Shi, S., and Stanley, P. (2003). Protein O-fucosyltransferase 1 is an essential component of Notch signaling pathways. *Proc. Natl. Acad. Sci. U. S. A.* *100*, 5234–5239.
- Shi, S., Ge, C., Luo, Y., Hou, X., Haltiwanger, R.S., and Stanley, P. (2007). The Threonine That Carries Fucose, but Not Fucose, Is Required for Cripto to Facilitate Nodal Signaling. *J. Biol. Chem.* *282*, 20133–20141.
- Song, E., and Mechref, Y. (2013). LC–MS/MS Identification of the O-Glycosylation and Hydroxylation of Amino Acid Residues of Collagen  $\alpha$ -1 (II) chain from Bovine Cartilage. *J. Proteome Res.* *12*, 3599–3609.
- Stahl, M., Uemura, K., Ge, C., Shi, S., Tashima, Y., and Stanley, P. (2008). Roles of Pofut1 and O-Fucose in Mammalian Notch Signaling. *J. Biol. Chem.* *283*, 13638–13651.
- Stenflo, J. (1991). Structure-function relationships of epidermal growth factor modules in vitamin K-dependent clotting factors. *Blood* *78*, 1637–1651.



- Surmann-Schmitt, C., Widmann, N., Dietz, U., Saeger, B., Eitzinger, N., Nakamura, Y., Rattel, M., Latham, R., Hartmann, C., von der Mark, H., et al. (2009). Wif-1 is expressed at cartilage-mesenchyme interfaces and impedes Wnt3a-mediated inhibition of chondrogenesis. *J. Cell Sci.* *122*, 3627–3637.
- Takeuchi, H., Kantharia, J., Sethi, M.K., Bakker, H., and Haltiwanger, R.S. (2012). Site-specific O-Glucosylation of the Epidermal Growth Factor-like (EGF) Repeats of Notch. *J. Biol. Chem.* *287*, 33934–33944.
- Takeuchi, H., Yu, H., Hao, H., Takeuchi, M., Ito, A., Li, H., and Haltiwanger, R.S. (2017). O-Glycosylation modulates the stability of epidermal growth factor-like repeats and thereby regulates Notch trafficking. *J. Biol. Chem.* *292*, 15964–15973.
- Takeuchi, H., Schneider, M., Williamson, D.B., Ito, A., Takeuchi, M., Handford, P.A., and Haltiwanger, R.S. (2018). Two novel protein O -glucosyltransferases that modify sites distinct from POGlut1 and affect Notch trafficking and signaling. *Proc. Natl. Acad. Sci.* *115*, E8395–E8402.
- Tian, E., and Ten Hagen, K.G. (2009). Recent insights into the biological roles of mucin-type O-glycosylation. *Glycoconj. J.* *26*, 325–334.
- Urakami, S. (2006). Epigenetic Inactivation of Wnt Inhibitory Factor-1 Plays an Important Role in Bladder Cancer through Aberrant Canonical Wnt/ -Catenin Signaling Pathway. *Clin. Cancer Res.* *12*, 383–391.
- Urban, S. (2002). Intramembrane proteolysis controls diverse signalling pathways throughout evolution. *Curr. Opin. Genet. Dev.* *12*, 512–518.
- Üren, A., Reichsman, F., Anest, V., Taylor, W.G., Muraiso, K., Bottaro, D.P., Cumberland, S., and Rubin, J.S. (2000). Secreted Frizzled-related Protein-1 Binds Directly to Wingless and Is a Biphasic Modulator of Wnt Signaling. *J. Biol. Chem.* *275*, 4374–4382.
- Vaes, B., Dechering, K., Vansomeren, E., Hendriks, J., Vandeven, C., Feijen, A., Mummery, C., Reinders, M., Olijve, W., and Vanzoelen, E. (2005). Microarray analysis reveals expression regulation of Wnt antagonists in differentiating osteoblasts. *Bone* *36*, 803–811.
- Valliere-Douglass, J.F., Brady, L.J., Farnsworth, C., Pace, D., Balland, A., Wallace, A., Wang, W., Treuheit, M.J., and Yan, B. (2008). O-Fucosylation of an antibody light chain: Characterization of a modification occurring on an IgG1 molecule. *Glycobiology* *19*, 144–152.
- Varki, A. (2017). Biological roles of glycans. *Glycobiology* *27*, 3–49.
- Vasudevan, D., and Haltiwanger, R.S. (2014). Novel roles for O-linked glycans in protein folding. *Glycoconj. J.* *31*, 417–426.
- Wang, Y., and Spellman, M.W. (1998). Purification and Characterization of a GDP-fucose:Polypeptide Fucosyltransferase from Chinese Hamster Ovary Cells. *J. Biol. Chem.* *273*, 8112–8118.



- Wang, H., Zang, C., Liu, X.S., and Aster, J.C. (2015). The Role of Notch Receptors in Transcriptional Regulation: NOTCH AND TRANSCRIPTION. *J. Cell. Physiol.* *230*, 982–988.
- Wang, Y., Lee, G.F., Kelley, R.F., and Spellman, M.W. (1996). Identification of a GDP-L-fucose:polypeptide fucosyltransferase and enzymatic addition of *O*-linked fucose to EGF domains. *Glycobiology* *6*, 837–842.
- Wouters, M.A., Rigoutsos, I., Chu, C.K., Feng, L.L., Sparrow, D.B., and Dunwoodie, S.L. (2005). Evolution of distinct EGF domains with specific functions. *Protein Sci.* *14*, 1091–1103.
- Wu, J., Fang, J., Yang, Z., Chen, F., Liu, J., and Wang, Y. (2012). Wnt inhibitory factor-1 regulates glioblastoma cell cycle and proliferation. *J. Clin. Neurosci.* *19*, 1428–1432.
- Yang, Z., Wang, Y., Fang, J., Chen, F., Liu, J., Wu, J., and Wang, Y. (2010). Expression and aberrant promoter methylation of Wnt inhibitory factor-1 in human astrocytomas. *J. Exp. Clin. Cancer Res. CR* *29*, 26.
- Zhang, L., and Ten Hagen, K.G. (2011). The cellular microenvironment and cell adhesion: a role for O-glycosylation. *Biochem. Soc. Trans.* *39*, 378–382.
- Zhang, R., Engler, A., and Taylor, V. (2018). Notch: an interactive player in neurogenesis and disease. *Cell Tissue Res.* *371*, 73–89.
- Zhu, J., Ren, J., and Tang, L. (2018). Genistein inhibits invasion and migration of colon cancer cells by recovering WIF1 expression. *Mol. Med. Rep.* *17*, 7265–7273.
- Zielinska, D.F., Gnad, F., Wiśniewski, J.R., and Mann, M. (2010). Precision Mapping of an In Vivo N-Glycoproteome Reveals Rigid Topological and Sequence Constraints. *Cell* *141*, 897–907.



## ANNEXE

Catégorie du test	Test	Tendance ♀ Cax	Tendance ♂ Cax	Tendance Cax	p-value ♀	p-value ♂	p-value
Comportement (locomotion)	Nombre d'explorations verticales	↑	↑	↑	X	X	<b>&lt; 0.001</b>
Comportement (Anxiété)	% de temps passé au centre	↑	↑	↑	X	X	<b>0.001</b>
Comportement (Habituation)	Amplitude de sursaut acoustique	↓	↓	↓	X	X	<b>0.001</b>
	Inhibition de la pré- impulsion à 81 dB	↓	↓	↓	X	X	<b>0.026</b>
Neurology	Activité locomotrice	↑	↑	↑	X	X	<b>0.05</b>
	Élévation de la queue	↑	↑	↑	X	X	<b>0.035</b>
	Urination durant la période d'observation	↑	↑	↑	X	X	<b>0.009</b>
	Latence de la chute	↓	→	→	<b>&lt; 0.05</b>	NS	NS
	Résistance à l'anesthésie	↑	↑	↑	X	X	X
Métabolisme énergétique	Prise alimentaire	↑	↑	↑	X	X	<b>0.002</b>
	Consommation minimum O <sub>2</sub>	↑	↑	↑	X	X	<b>0.005</b>
	Ratio d'échange respiratoire VCO <sub>2</sub> /VO <sub>2</sub>	↑	↑	↑	X	X	<b>&lt; 0.001</b>
Œil	Épaisseur de la rétine	↑	↑	↑	<b>0.001</b>	<b>&lt; 0.001</b>	X
	Épaisseur de la lentille	↓	↓	↓	<b>0.04</b>	<b>0.002</b>	X
	Longueur axiale de l'œil	↓	↓	↓	<b>0.008</b>	<b>0.008</b>	X
Hématologie et chimie clinique	AUC [Glucose] 0-30	↓	↓	↓	X	X	<b>0.014</b>
	AUC [Glucose] 30-120	↓	↓	↓	X	X	<b>0.001</b>
	Concentration d'insuline plasmatique	↓	↑	↓	<b>0.023</b>	NS	NS
	Glycémie à jeun	↑	↑	↑	X	X	<b>&lt; 0.001</b>
	Concentration ASAT / GOT	↑	↑	↑	X	X	<b>0.01</b>



Hématologie et chimie clinique	Glycémie	↓	↓	↓			< 0.001
	Concentration LDH	↑	↑	↑			< 0.001
	Concentration en Fer	↓	↓	↓			0.002
	Capacité de liaison du fer insaturé	↑	↑	↑			0.004
	Saturation en transferrine	↓	↓	↓			0.001
	Nombre d'hématies (Test 1)	↓	↓	↓			0.01
	% Hémoglobine (Test 1)	↓	↓	↓			0.038
	Moyenne corpusculaire (Test 1)	↑	↑	↑			0.02
	Nombre de leucocytes (Test 1)	↑	↑	↑			0.014
	Volume plaquettaire moyen (Test 1)	↑	↑	↑			< 0.001
	Nombre d'hématies (Test 2)	→	→	→			NS
	% Hémoglobine (Test 2)	→	→	→			NS
	Moyenne corpusculaire (Test 2)	→	→	→			NS
	Nombre de leucocytes (Test 2)	→	→	→			NS
	Volume plaquettaire moyen (Test 2)	→	→	→			NS
Autres	Perte d'eau transépidermique	↓	↑	→	0.006	0.001	NS
	Température moyenne à la surface du corps	↑	→	→	0.031	NS	NS
	Masse corporelle	↓	↓	↓	NS	0.016	NS
	Masse du cœur	↓	↓	↓	NS	0.008	NS
	Masse du foie	↓	↓	↓	NS	0.008	NS

**Annexe 1** : Données significativement différentes pour les souris Pofut1<sup>cax/cax</sup> révélées par le crible primaire réalisé à la clinique de la souris en Allemagne (www.mouseclinic.de/)





## Résumé

---

La *O*-fucosylation, catalysée par Pofut1, est une glycosylation rare qui consiste en l'ajout d'un fucose *O*-lié sur la sérine ou la thréonine d'une séquence consensus ( $C^2X_4(\underline{S}/\underline{T})C^3$ ), portée par un domaine EGF-like (ELD) d'une glycoprotéine membranaire ou sécrétée. Notre analyse de la lignée murine Pofut1<sup>cax/cax</sup>, hypomorphe pour le gène *Pofut1*, a révélé une hypertrophie musculaire post-natale associée à une diminution du pool de cellules satellites. Ce phénotype est en partie associé à un défaut d'interaction entre les récepteurs NOTCH hypo-*O*-fucosylés des myoblastes dérivés de cellules satellites (MDCS) et leurs ligands DSL, ce qui aboutit à une plus faible activation de la signalisation Notch. D'autres protéines potentiellement impliquées dans la myogenèse peuvent également être la cible de POFUT1. C'est notamment le cas de la protéine Wnt inhibitory factor 1 (WIF1), qui dispose de cinq ELDs, dont deux sont potentiellement aptes à recevoir un *O*-fucose (ELDs III et V). Par une approche phylogénétique, nous avons montré la conservation de ces deux sites de *O*-fucosylation et de deux sites de *N*-glycosylation chez la plupart des bilatériens. Nos expériences démontrent l'occupation de tous ces sites, excepté le site de *O*-fucosylation de l'ELD V, chez la protéine WIF1 murine. La capacité de l'ELD III, produit de manière isolée, à recevoir un fucose *O*-lié a été démontrée après *O*-fucosylation *in vitro*, par l'association de cycloaddition azide-alcyne assistée au cuivre (CuAAC) et de spectrométrie de masse en mode MRM. Cette nouvelle approche expérimentale a par la suite été standardisée et sa sensibilité évaluée en comparant deux autres ELDs (ELDs 12 et 26 de NOTCH1) connus pour être *O*-fucosylés mais présentant des affinités différentes pour POFUT1. De façon surprenante, l'ELD V de WIF1 ne peut être *O*-fucosylé, probablement en raison d'un clash stérique entre cet ELD et POFUT1, prévenant ainsi leur interaction. L'analyse de la protéine WIF1 entière a confirmé les résultats obtenus sur les ELDs isolés et démontre l'occupation des deux sites de *N*-glycosylation. Enfin, nos résultats montrent également l'importance de ces deux *N*-glycanes, mais également celle du *O*-fucose de l'ELD III, pour une sécrétion optimale de la protéine WIF1 murine.

---

**Mots clés:** POFUT1, *O*-fucosylation, EGF-like, WIF1, CuAAC

## Abstract

The, Pofut1-catalyzed *O*-fucosylation, is a rare glycosylation which consists of the addition of an *O*-linked fucose to the serine or threonine of a consensus sequence ( $C^2X_4(\underline{S}/\underline{T})C^3$ ), carried by an EGF-like domain (ELD) of a membrane or secreted glycoprotein. Our analysis of the murine line Pofut1<sup>cax/cax</sup>, hypomorphic for the *Pofut1* gene, revealed post-natal muscle hypertrophy associated with a decrease in the satellite cell pool. This phenotype was partly associated with a lack of interaction between hypo-*O*-fucosylated NOTCH receptors of satellite cell-derived myoblasts (SCDM) and their DSL ligands, which resulted in a lower activation of Notch signaling. Other proteins potentially involved in myogenesis may also be the target of POFUT1. This is indeed the case for the protein Wnt inhibitory factor 1 (WIF1), which has five ELDs, whose only two are potentially able to receive an *O*-fucose (ELDs III and V). Using a phylogenetic approach, we showed in most bilaterians that these two *O*-fucosylation sites and two *N*-glycosylation sites were conserved. Our experiments showed the occupation of all these sites, except for the *O*-fucosylation site of murine WIF1 protein ELD V. The ability of the ELD III, produced as an isolated protein, to receive *O*-linked fucose was demonstrated after an *in vitro* *O*-fucosylation by combination of copper-catalysed azide-alkyne cycloaddition (CuAAC) and MRM-mass spectrometry. This new experimental approach was then standardized and its sensitivity was evaluated by comparing two other ELDs (NOTCH1 ELDs 12 and 26) known to be *O*-fucosylated but with different affinities for POFUT1. Surprisingly, WIF1's ELD V could not be *O*-fucosylated, probably due to a steric clash between this ELD and POFUT1, thus preventing their interaction. The analysis of the full-length WIF1 protein confirmed our results obtained with isolated ELDs and demonstrated the occupation of the two *N*-glycosylation sites. Finally, our results also showed the importance of these two *N*-glycans, but also the importance of ELD III's *O*-fucose, for optimal secretion of the murine WIF1 protein.

---

**Keywords:** POFUT1, *O*-fucosylation, EGF-like, WIF1, CuAAC

