

POGLUT2 AND POGLUT3: A TALE OF TWO NOVEL PROTEIN *O*-  
GLUCOSYLTRANSFERASES

by

DANIEL B. WILLIAMSON

(Under the Direction of Robert S. Haltiwanger)

ABSTRACT

*O*-glycosylation of Epidermal Growth Factor-like (EGF) repeats plays crucial roles in protein folding, trafficking and function. The Notch extracellular domain has been used as a model to study these mechanisms due to its many *O*-glycosylated EGF repeats. Three enzymes were previously known to *O*-glycosylate Notch EGF repeats: Protein *O*-Glucosyltransferase 1 (POGLUT1), Protein *O*-Fucosyltransferase 1 (POFUT1), and EGF Domain Specific *O*-Linked *N*-Acetylglucosamine Transferase (EOGT). All of these modifications affect Notch activity. Recently, we identified POGLUT2 and POGLUT3 as two novel *O*-glucosyltransferases that modify a few Notch EGF repeats at sites distinct from those modified by POGLUT1. Comparison of these modification sites revealed a putative consensus sequence which predicted modification of many extracellular matrix (ECM) proteins including fibrillins (FBNs) and Latent TGF- $\beta$ -binding proteins (LTBPs). These proteins are essential for formation of microfibrils and elastic fibers in the ECM, which provides structural support to respective tissues and regulates TGF- $\beta$  signaling. Glycoproteomic analysis revealed that approximately half of the 47 EGF repeats in FBN1 and FBN2, and half of the 18 EGF repeats in LTBP1, are modified by POGLUT2 and/or POGLUT3. Cellular assays showed that loss of modifications by POGLUT2

and/or POGLUT3 significantly reduces secretion of a FBN1 fragment, suggesting a possible mechanism for the effects of POGLUT2/3-mediated *O*-glucosylation. Alternatively, it is possible that these *O*-glucose modifications added by POGLUT2/3 affect protein-protein interactions, as has been demonstrated by research on POGLUT1 and POFUT1 modifications on Notch. Mouse knockouts of either *Poglut1* or *Pofut1* lead to embryonic lethality with Notch phenotypes. Deletion of both *Poglut2* and *Poglut3* leads to perinatal lethality in mice with developmental defects affecting mouse limbs and lung development similar to previously described *Fbn* and *Ltbp* mouse models. These results highlight the physiological importance of POGLUT2 and POGLUT3, and strongly suggests their role in development and proper function of the ECM. Here we discuss the identification and function of POGLUT2 and POGLUT3 and the ongoing research that continues to elucidate the biological significance of these novel enzymes.

INDEX WORDS: Glycosylation, *O*-glucose, EGF repeats, Notch, Fibrillin, LTBP, Extracellular matrix, Protein secretion, Protein folding, Mass spectrometry

POGLUT2 AND POGLUT3: A TALE OF TWO NOVEL PROTEIN *O*-  
GLUCOSYLTRANSFERASES

by

DANIEL B. WILLIAMSON

B.S., The University of North Carolina at Chapel Hill, 2012

A Dissertation Submitted to Graduate Faculty of The University of Georgia in Partial Fulfillment  
of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2022

Copyright © 2022

Daniel B. Williamson

All Rights Reserved



POGLUT2 AND POGLUT3: A TALE OF TWO NOVEL PROTEIN *O*-  
GLUCOSYLTRANSFERASES

by

DANIEL B. WILLIAMSON

Major Professor: Robert S. Haltiwanger

Committee: Nancy Manley  
Kelley Moremen  
Lance Wells

Electronic Version Approved:

Ron Walcott  
Vice Provost for Graduate Education and Dean of the Graduate School  
The University of Georgia  
August 2022

## DEDICATION

To my wife, Libby Villa, for her constant support and our pets, Dexter and Finn.

## ACKNOWLEDGEMENTS

I thank my mentor Bob Haltiwanger and all the members, current and past, of the Haltiwanger lab. Specifically, I'd like to thank Hideyuki Takeuchi, Megumi Takeuchi, and Atsuko Ito for providing me invaluable training at the beginning of my PhD career. I'd also like to thank Bernadette Holdener and the members of her lab whom we routinely collaborate with and whose experimental guidance and assistance was greatly appreciated.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS .....	v
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
 CHAPTER	
1 INTRODUCTION .....	1
2 LITERATURE REVIEW .....	4
2.1 <i>O</i> -glycosylation of EGF repeats .....	4
2.2 Identification of POGLUT2 and POGLUT3 .....	6
2.3 Multiple ECM proteins are <i>O</i> -glucosylated by POGLUT2 and POGLUT3 .....	8
2.4 Biological Significance of POGLUT2 and POGLUT3 .....	10
2.5 Figures .....	15
3 POGLUT2 AND POGLUT3 <i>O</i> -GLUCOSYLATE MULTIPLE EGF REPEATS IN FIBRILLIN-1, -2, AND LTBP1 AND PROMOTE SECRETION OF FIBRILLIN-1 .....	19
3.1 Abstract .....	20
3.2 Introduction .....	21
3.3 Results .....	23
3.4 Discussion .....	32
3.5 Experimental Procedures .....	35

3.6 Data Availability .....	40
3.7 Acknowledgements .....	40
3.8 Main Text Figures .....	41
3.9 Supplemental Tables and Figures.....	51
4 POGLUT2 AND POGLUT3 <i>O</i> -GLUCOSYLATION OF FBN1 AND FBN2 MAINTAINS THE STRUCTURE AND FUNCTION OF THE EXTRACELLAR MATRIX AND IS REQUIRED FOR PROPER TGF- $\beta$ SIGNALING AND DEVELOPMENT OF THE LUNG .....	58
4.1 Abstract .....	59
4.2 Introduction .....	60
4.3 Results .....	62
4.4 Discussion .....	69
4.5 Experimental Procedures.....	73
4.6 Data Availability .....	79
4.7 Main Text Tables and Figures .....	80
4.8 Supplemental Tables and Figures.....	91
5 CONCLUSIONS AND FUTURE DIRECTIONS .....	101
6 REFERENCES .....	105

## LIST OF TABLES

	Page
Table S3.1: FBN1, FBN2, and LTBP1 EGFs with variable residues at position 4 of the POGLUT2 and 3 consensus sequence .....	51
Table S3.2: POGLUT2 versus POGLUT3 preferred modification sites on FBN1, FBN2, and LTBP1 .....	52
Table S3.3: Primers used for site-directed mutagenesis of FBN1-N and FBN2-N .....	53
Table 4.1: Viability of heterozygous <i>Poglut2</i> and 3 null intercrosses .....	80
Table S4.1: Primers and PCR conditions used for genotyping <i>Poglut2</i> mice .....	91
Table S4.2: Primers and PCR conditions used for genotyping <i>Poglut3</i> mice .....	92
Table S4.3: <i>Poglut2</i> WT/KO intercross progeny viability .....	93
Table S4.4: <i>Poglut3</i> WT/KO intercross progeny viability .....	94
Table S4.5: <i>Poglut2</i> ; <i>Poglut3</i> double heterozygote intercross progeny viability .....	95

## LIST OF FIGURES

	Page
Figure 2.1: Schematic of an EGF repeat illustrating modification sites of identified <i>O</i> -glycosyltransferases .....	15
Figure 2.2: Ligand-binding domain EGF repeats from NOTCH1 and -3 are modified with novel <i>O</i> -Glc.....	16
Figure 2.3: POGLUT2 and POGLUT3 are responsible for <i>O</i> -glucosylation of mouse NOTCH1 EGF11 .....	17
Figure 2.4: Summary of FBN1, -2, and LTBP1 EGF repeats modified by POGLUT2 and 3.....	18
Figure 3.1: Predicted sites of modification by POGLUT2 and/or POGLUT3 on EGF repeats of human FBN1, FBN2, and LTBP1 based on original consensus sequence .....	41
Figure 3.2: Endogenous FBN1 is modified at high stoichiometry with hexose at more than the predicted number of EGF repeats .....	42
Figure 3.3: POGLUT2 and 3 extensively <i>O</i> -glucosylate EGF repeats on recombinant human FBN1 EGFs 1-26 (FBN1-N) expressed in HEK293T cells.....	43
Figure 3.4: POGLUT2 and 3 extensively <i>O</i> -glucosylate EGF repeats on recombinant human FBN2 EGFs 1-28 (FBN2-N) and LTBP1 in HEK293T cells .....	44
Figure 3.5: Summary POGLUT2 and 3 sites mapped on endogenous FBN1 from dermal fibroblasts and recombinant FBN1, FBN2, and LTBP1 expressed in HEK293T cells .....	45

Figure 3.6: POGLUT2 and 3 consensus sequence can be used to convert an unmodified EGF to a POGLUT2/3 modified EGF .....	46
Figure 3.7: POGLUT2 and POGLUT3 exhibit site-specificity on EGF repeats of recombinant FBN1-N, FBN2-N, and LTBP1 .....	47
Figure 3.8: POGLUT2 versus POGLUT3 consensus sequence .....	48
Figure 3.9: POGLUT2 and POGLUT3 site-specificity is not strictly defined by amino acid sequence between cysteines 3 and 4 of EGF repeats .....	49
Figure 3.10: Knockout of <i>POGLUT2</i> and/or <i>POGLUT3</i> leads to reduced secretion of recombinant human FBN1-N from HEK293T cells .....	50
Figure S3.1: Purification of recombinant FBN1-N, FBN1-C, FBN2-N, FBN2-C, and LTBP1 expressed in HEK293T cells.....	54
Figure S3.2: Clustal Omega protein sequence alignment of POGLUT2 versus POGLUT3 preferred EGF modification sites .....	55
Figure 4.1: <i>Poglut2/3 DKO</i> mice are smaller than littermates and have syndactyly .....	81
Figure 4.2: POGLUT2 and 3 modifications are lost in <i>Poglut2/3 DKO</i> mice.....	82
Figure 4.3: POGLUT2 and 3 are essential for blood vessel wall structure at E18.5 .....	83
Figure 4.4: POGLUT2 and 3 are essential for characteristic folded/ruffled structure of terminal bronchioles .....	84
Figure 4.5: POGLUT 2 and 3 are needed for normal terminal sacculle development at E18.5 .....	85
Figure 4.6: Loss of POGLUT2 and 3-mediated <i>O</i> -glucosylation on EGF repeats leads to reduced abundance and fragmented pattern of FBN1, -2, and elastin in lung blood vessels ....	86
Figure 4.7: Loss of POGLUT2 and 3 leads to reduced abundance and fragmented pattern of FBN1, -2, and elastin in the terminal bronchiole .....	87



Figure 4.8: Loss of POGLUT2 and 3 leads to reduced abundance and fragmented pattern of FBN1, -2, and elastin in the alveoli .....	88
Figure 4.9: TGF- $\beta$ signaling is significantly reduced in <i>Poglut2/3 DKO</i> lungs .....	89
Figure 4.10: Abundances of secreted FBN1 and other ECM proteins are affected by loss of POGLUT2 and 3 in dermal fibroblasts at E18.5.....	90
Figure S4.1: Targeting strategy and PCR design for <i>Poglut2</i> alleles .....	96
Figure S4.2: Targeting strategy and PCR design for <i>Poglut3</i> alleles .....	97
Figure S4.3: <i>Poglut2/3 DKO</i> mice display forelimb and hindlimb syndactyly .....	98
Figure S4.4: The alveolar space in <i>Poglut2/3 DKO</i> lungs is slightly increased compared to wild-type .....	99
Figure S4.5: Abundances of secreted ECM proteins are affected by loss of POGLUT2 and 3 in dermal fibroblasts at E18.5 .....	100
Figure 5.1: Working model of how loss of <i>O</i> -glucosylation by POGLUT2/3 could affect substrate interactions intracellularly and extracellularly that are required for ECM structure and function .....	104

## CHAPTER 1

### INTRODUCTION

#### **Purpose of the Study**

This study is aimed to evaluate the function and physiological relevance of POGLUT2 and POGLUT3 using *in vitro* and *in vivo* methods. We identified POGLUT2 and 3 following a surprising discovery of an *O*-glucose modification at a novel site (between Cys<sup>3</sup>-Cys<sup>4</sup>) of an EGF repeat within the extracellular domain (ECD) of NOTCH1, a transmembrane receptor protein (1, 2). An EGF repeat is a common protein domain found in many secreted proteins localized to the cell surface and extracellular space. EGF repeats consists of approximately 40 amino acids, including six conserved cysteines that form three disulfide bonds in a conserved pattern (Cys<sup>1</sup>-Cys<sup>3</sup>, Cys<sup>2</sup>-Cys<sup>4</sup>, Cys<sup>5</sup>-Cys<sup>6</sup>). The Notch ECD is known to be heavily *O*-glycosylated by multiple forms of *O*-glycosylation between these cysteines, such as POGLUT1 which adds an *O*-glucose between Cys<sup>1</sup>-Cys<sup>2</sup> (3-6); however, no enzymes were previously known to add an *O*-glucose to an EGF repeat between Cys<sup>3</sup>-Cys<sup>4</sup>. Using both *in vitro* enzyme assays and *POGLUT2* and 3 knockout HEK293T cells, we identified POGLUT2 and 3 as novel *O*-glucosyltransferases that modify a serine residue between Cys<sup>3</sup>-Cys<sup>4</sup> of an EGF repeat at the putative consensus sequence C<sup>3</sup>-X-N-T-X-G-S-F-C<sup>4</sup>. Both NOTCH1 and NOTCH3 are modified by POGLUT2 and 3 (1). Interestingly, the putative consensus sequence is a common protein motif found in many other secreted proteins (1). The fibrillin (FBN) and LTBP family of proteins have the highest occurrence of this motif, significantly more than Notch (1). This broadened the scope and impact of this study and suggested POGLUT2 and POGLUT3 play a more significant role in the

function of FBNs and LTBP<sub>s</sub> compared to Notch. Protein folding, stability, secretion, and extracellular interactions are significantly affected by *O*-glycosylation on EGF repeats (3, 7-10). Moreover, loss of *O*-glycosyltransferases like *Poglut1* or *Pofut1* leads to embryonic lethality with Notch-like phenotypes in mice (11, 12). Since FBNs and LTBP<sub>s</sub> are predicted to be extensively modified by POGLUT2/3, this raised the possibility and led to our hypothesis that POGLUT2/3-mediated *O*-glucosylation would have similar effects on these predicted substrates. To test his hypothesis, we designed experiments to address the following questions:

1. Are FBNs and LTBP<sub>s</sub> modified by POGLUT2/3 and to what extent?
  - a. What other proteins are similarly modified by POGLUT2/3?
2. Are POGLUT2/3 functionally redundant?
  - a. Is there EGF site-specificity or protein-specificity between these enzymes?
3. Does deletion of POGLUT2/3 affect secretion/function of substrates?
4. Are these findings recapitulated in a mouse model?
  - a. Do *Poglut2/3* knockout mice exhibit early lethality like *Poglut1* and *Pofut1* mice?
  - b. Are developmental phenotypes observed, and are they linked to impaired secretion/function of POGLUT2/3 substrates like FBNs and LTBP<sub>s</sub>?

In Chapter 2, a review of the identification, function, and biological relevance of POGLUT2 and POGLUT3 is presented. Portions of this chapter were published in Biochemical Society Transactions (BST) and The Proceedings of the National Academy of Science (PNAS) and are copied here with permission (1, 13). This chapter begins with a brief history of *O*-glycosylation of EGF repeats and how it affects protein function in the context of Notch. This is followed by a description of the initial identification of POGLUT2 and 3 as *O*-glucosyltransferases that modify Notch EGF repeats. The next section provides an overview of

the ECM environment, specifically focusing on the roles of FBN1, -2, and LTBP1, followed by the discovery that these proteins are extensively modified by POGLUT2 and 3. This chapter concludes with a discussion regarding the biological significance of POGLUT2 and 3 with respect to protein folding, secretion, and protein interactions concerning ECM-related disorders and cancer.

In Chapter 3, the discovery that POGLUT2 and 3 modify multiple ECM-related proteins and affect secretion of FBN1 is discussed. This work was published in *The Journal of Biological Chemistry*. Chapter 4 examines *Poglut2* and *Poglut3* single and double knockout mouse models. Data is presented confirming POGLUT2/3 mediated *O*-glucosylation of multiple ECM proteins is abolished in *Poglut2/3 DKO* animals. Phenotypes such as perinatal lethality, reduced size, syndactyly, and impaired lung development are also detailed. In Chapter 5, major results from the previous chapters are tied together, providing the foundation for future directions and experiments.

### **How this study is original**

POGLUT2 and 3, previously known as KDELC1 and 2, had unknown function and physiological relevance prior to this study. We were the first to identify KDELC1 and KDELC2 as Protein *O*-Glucosyltransferases, hence renaming them POGLUT2 and 3, respectively. After identifying the novel consensus sequence for modification by POGLUT2 and 3, we subsequently discovered additional proteins that are modified by POGLUT2 and 3 such as: FBN1, -2, LTBP1, fibulin-2 (FBLN2), nidogen-1(NID1), and fibulin-5 (FBLN5). We were also the first to generate and examine *Poglut2* and *Poglut3* single and double knockout mice, which are an indispensable resource we have used to begin to understand the physiological role and significance of *Poglut2* and *Poglut3*.

## CHAPTER 2

### LITERATURE REVIEW

#### **2.1 O-glycosylation of EGF repeats**

Addition of glucose or fucose to the hydroxyls of serine or threonine residues (*i.e.* *O*-linked) in Epidermal Growth Factor-like (EGF) repeats was discovered more than three decades ago. *O*-glucose was reported on EGF repeats of blood coagulation factors VII and IX (14), and soon after, *O*-fucose was discovered on an EGF repeat of urinary-type plasminogen activator (15). Since their initial discovery, many other proteins have been identified with EGF repeats modified by *O*-glucose or *O*-fucose, including Notch receptors and ligands (16-18), but these modifications are less common than the *O*-linked modifications found on proteoglycans and mucins. However, research over the past two decades has revealed significant biological roles for *O*-glycans on EGF repeats (16-23).

EGF repeats are small protein domains containing approximately 40 amino acids and are commonly found in many cell-surface and secreted proteins. Each EGF repeat contains six conserved cysteines that form three disulfide bonds in a specific pattern (Cys<sup>1</sup>-Cys<sup>3</sup>, Cys<sup>2</sup>-Cys<sup>4</sup>, Cys<sup>5</sup>-Cys<sup>6</sup>) that stabilize its structure (Figure 2.1). EGF repeats containing specific amino acid consensus sequences between these conserved cysteines can be modified by *O*-glucose, *O*-fucose, and/or *O*-GlcNAc (Figure 2.1) (3-6). *O*-glucose is added to a serine by Protein *O*-Glucosyltransferase 1 (POGLUT1) between Cys<sup>1</sup>-Cys<sup>2</sup> (3). This glucose can be elongated by two xylose residues to form a trisaccharide (24, 25). *O*-fucose is added to a serine or threonine by

Protein *O*-Fucosyltransferase 1 (POFUT1) between Cys<sup>2</sup>-Cys<sup>3</sup> (5, 8). The fucose can be modified by  $\beta$ 3-N-acetylglucosaminyltransferases of the Fringe family, and further elongated with galactose and sialic acid to form a tetrasaccharide (26-28). Finally, *O*-GlcNAc is added to a serine or threonine by EGF Domain Specific *O*-Linked *N*-Acetylglucosamine Transferase (EOGT) between Cys<sup>5</sup>-Cys<sup>6</sup> (29). This glycan can be elongated with galactose and sialic acid in mammals (28, 30). More recently, *O*-glucose was shown to be added to a distinct serine between Cys<sup>3</sup>-Cys<sup>4</sup> by POGLUT2 or POGLUT3, which is the focus of this review (1). This glucose has not been found to be elongated by xylose.

Notch was the first protein identified to be heavily modified at multiple sites with *O*-glucose and *O*-fucose (4, 27, 31). The Notch family of receptors are transmembrane proteins required for development in all metazoans; they have extracellular domains (ECD) consisting of 29-36 tandem EGF repeats. Due to the abundance of *O*-glucose and *O*-fucose modification sites on its ECD, the Notch receptors have been used as a model protein to elucidate the biological roles of *O*-glycosylation by POGLUT1 and POFUT1. In mice, elimination of either *Poglut1* or *Pofut1* results in embryonic lethality with Notch-like phenotypes (11, 12). In humans, numerous cancers and developmental disorders are associated with mutations in *POGLUT1* and *POFUT1* (20-23). Thus, modification of Notch receptors by POGLUT1 and POFUT1 is clearly necessary for their function.

Mechanistic studies have revealed two functions for *O*-glycans on Notch EGF repeats: direct interaction with ligands and folding of EGF repeats. Co-crystal structures of NOTCH1 with ligands revealed that the *O*-fucose residues on EGF repeats 8 and 12 are in direct contact with ligands, and mutation of these sites reduces ligand binding (2, 28, 32). In addition, mutation of the *O*-fucose site on EGF12 in *Notch1* in mice or *Notch* in *Drosophila* results in embryonic

lethality, demonstrating the biological importance of this modification (33, 34). In addition to being directly involved in ligand binding, *in vitro* studies have revealed specific roles these enzymes have in Notch folding and trafficking. POGLUT1 and POFUT1 are localized to the endoplasmic reticulum (ER) (3, 7), and both enzymes only modify properly folded EGF repeats (8, 9), suggesting that they serve as folding sensors for EGF repeats in the ER. The addition of *O*-glucose or *O*-fucose increases the stability of an EGF repeat compared to an unmodified EGF repeat (10). Additionally, elimination of *POGLUT1* and/or *POFUT1* significantly reduces cell surface expression of NOTCH1 in HEK293T cells (10), although this effect is not seen in all cells (12, 35). These results suggest that POGLUT1 and POFUT1 function as a quality control for folding of EGF repeats in the ER.

The recent identification of two novel POGLUTs, POGLUT2 and POGLUT3, adds to the list of *O*-glycosyltransferases that modify proteins with EGF repeats (1). POGLUT2 and POGLUT3 heavily *O*-glucosylate the extracellular matrix (ECM) proteins fibrillin-1 (FBN1), fibrillin-2 (FBN2), and latent transforming growth factor  $\beta$ -binding protein 1 (LTBP1) (36). These proteins are modified by POGLUT2 and 3 to a similar extent that Notch is modified by POGLUT1 and POFUT1 (27, 36). This suggests *O*-glucosylation by POGLUT2 and 3 may regulate folding, secretion, and function of these ECM proteins.

## **2.2 Identification of POGLUT2 and POGLUT3**

Two mammalian homologs of *POGLUT1*, *KDELC1* and *KDELC2*, of unknown function were identified approximately a decade ago (37). *KDELC1*, also known as ER protein 58 (EP58), had been identified previously as a novel ER-resident protein, but its function remained unclear (38). Both KDELC1 and KDELC2 contain a CAP10 domain common to glycosyltransferases

and an ER retention signal (37, 38). *POGLUT1*, the mammalian homolog to *Drosophila* Rumi, can rescue *rumi* mutant phenotypes in flies; however, *KDELC1* and *KDELC2*, which are not expressed in *Drosophila*, have no effect on *rumi* mutants (3, 37). Additionally, POGLUT1 assays using an EGF repeat known to be a POGLUT1 acceptor substrate is not a substrate for either KDELC1 or KDELC2 (37). Thus, while homologous to a known *O*-glucosyltransferase, these findings suggest *KDELC1* and *KDELC2* do not modify EGF repeats containing a POGLUT1 consensus sequence between Cys<sup>1</sup> and Cys<sup>2</sup> (Figure 2.1).

In 2013, an *O*-hexose modification at a site distinct from the POGLUT1 consensus on NOTCH1 was initially identified by mass spectral analysis (39). Subsequent co-crystallized structures of a portion of the NOTCH1 ECD bound to ligand confirmed an *O*-hexose at this novel site (2). The modification occurred on serine 435 (S435) between Cys<sup>3</sup> and Cys<sup>4</sup> of EGF11 (2), clearly different than the POGLUT1 predicted EGF modification sites (Figure 2.1). This suggested that either POGLUT1 could modify this novel site, or one or more unidentified enzyme(s) were responsible for this modification. *KDELC1* and *KDELC2* were logical candidates because of their homology to *POGLUT1*.

*In vitro* enzyme assays followed by mass spectral analysis confirmed *KDELC1* and *KDELC2*, rather than *POGLUT1*, are responsible for the *O*-hexose modification on NOTCH1 S435 (1). The *O*-hexose modification was also confirmed to be an *O*-glucose (1). As a result, *KDELC1* and *KDELC2* were renamed, *POGLUT2* and *POGLUT3* (1). Interestingly, the homologous EGF repeat from NOTCH2 (EGF11) does not have a serine in the same position, but the homologous EGF repeats from NOTCH3 (EGF10) and NOTCH4 (EGF11) do. NOTCH2 EGF11 is not modified with *O*-glucose, but NOTCH3 EGF10 is (Figure 2.2) (1).



Using *POGLUT2* and *POGLUT3* single and double knockout HEK293T cells generated with CRISPR, we also confirmed complete loss of *O*-glucosylation on NOTCH1 EGF11 (S435) only when both *POGLUT2* and *POGLUT3* were deleted (Figure 2.3). Single knockout cells revealed no change in *O*-glucosylation of NOTCH1 S435, suggesting partial redundancy in the function of *POGLUT2* and *POGLUT3* (Figure 2.3). Alignment of Notch sequences between Cys<sup>3</sup> and Cys<sup>4</sup> in these EGF repeats revealed a preliminary consensus sequence for *O*-glucosylation by *POGLUT2* and/or *POGLUT3*, C<sup>3</sup>-x-N-T-x-G-S-F-x-C<sup>4</sup> (Figure 2.2) (1). This sequence occurs in EGF repeats of more than 34 proteins, magnifying the possible biological impact of *POGLUT2* and 3. Notch ligands also contain EGF repeats. *POGLUT2* and 3 modification sites have not been confirmed on these ligands, but a few EGF repeats have a serine in the correct amino acid position between Cys<sup>3</sup> and Cys<sup>4</sup> to be modified.

### **2.3 Multiple ECM proteins are *O*-glucosylated by *POGLUT2* and *POGLUT3***

The extracellular matrix (ECM) contains a complex network of protein interactions required for development, tissue homeostasis, and tissue structural integrity (40-43). Many ECM proteins contain the preliminary *POGLUT2* and 3 consensus sequence, C<sup>3</sup>-x-N-T-x-G-S-F-x-C<sup>4</sup> (1, 36). The proteins with the highest occurrence of this sequence are from the fibrillin and latent transforming growth factor  $\beta$ -binding protein superfamily (36). Initial research efforts have focused specifically on *POGLUT2* and 3 mediated *O*-glucosylation of FBN1, FBN2, and LTBP1. These three proteins have clearly defined pathologies and well-studied biochemical interactions within the ECM environment (41, 42, 44-47). FBN1 is a major component of microfibrils in the ECM. These microfibrils act as a scaffold for elastin deposition, providing essential mechanical support and tissue integrity in elastic tissues including the heart, lungs, skin,

and musculoskeletal system (41, 42). In addition to the structural integrity of tissues, fibrillin microfibrils are critical for important cell signaling pathways in development, like TGF- $\beta$  (44, 48-52). Inactive TGF- $\beta$  is secreted from cells while bound to LTBP in a multiprotein complex (44, 53, 54). Fibrillin microfibrils bind this complex through protein-protein interactions with LTBP, serving as a reservoir of inactive TGF- $\beta$  and therefore regulating the availability of TGF- $\beta$  for cell signaling (44, 48, 55). *Fbn1* knockout mice die within two weeks of birth (46), and *Fbn1* and *Fbn2* double knockout mice do not survive past embryonic day 16 (46). *Ltbp1* knockout mice die perinatally (47). In all knockout conditions, cardiovascular abnormalities specifically associated with the aorta are present (46, 47). In humans, *FBN1* mutations are highly associated with Marfan syndrome (MFS), among other connective tissue disorders (45, 56, 57). Cardiovascular deformities and spontaneous aortic ruptures are hallmark manifestations of MFS (56).

Almost three decades ago, *O*-glycan modifications were identified on several FBN1 EGF repeats between C<sup>3</sup>-C<sup>4</sup> (58). At the time, these modifications were predicted to be *O*-xylose (58). Our recent mass spectral analysis reveals FBN1, FBN2, and LTBP1 are *O*-glucosylated by POGLUT2 and 3 (Figure 2.4A) (36) (See Chapter 3 for details). FBN1 and FBN2 each have 47 EGF repeats, and LTBP1 has 18 EGF repeats. Approximately half of the EGF repeats of each protein are *O*-glucosylated by POGLUT2 and 3 (Figure 2.4A) (36). Alignment of the sequences between Cys<sup>3</sup> and Cys<sup>4</sup> in these EGF repeats reveal tyrosine is commonly present at position 8 (50% of the time) in addition to the previously reported phenylalanine (Figure 2.4B) (36). These results necessitated broadening the POGLUT2 and 3 consensus sequence from C<sup>3</sup>-x-N-T-x-G-S-F-x-C<sup>4</sup> to C<sup>3</sup>-x-N-T-x-G-S-(FY)-x-C<sup>4</sup> (Figure 2.4B) (36). This sequence is not meant to encompass all modified sites since other amino acids can occur at multiple positions throughout

the putative consensus as indicated by the WebLogo (Figure 2.4B); however, it can be used as a tool to predict POGLUT2 and 3 modification sites. This new consensus sequence is found in 56 proteins, almost twice as many proteins as the previous consensus sequence (36), highlighting the physiological relevance of *O*-glucosylation by POGLUT2 and 3. Most of these proteins are associated with the extracellular matrix. This remains a putative consensus sequence, and additional changes could continue to expand the number of proteins *O*-glucosylated by POGLUT2 and 3.

POGLUT2 and 3 *O*-glucosylate EGF repeats in a site-specific manner. POGLUT2 demonstrates higher *in vitro* activity towards NOTCH3 EGF10 compared to NOTCH1 EGF11, whereas POGLUT3 shows a preference for NOTCH1 EGF11 (1). Site-specificity is also observed in multiple EGF repeats of FBN1, FBN2, and LTBP1 by comparing sites modified in *POGLUT2* or *POGLUT3* single knockout HEK293T cells (36). *O*-glucosylation is completely lost in *POGLUT2/POGLUT3*-double null cells, demonstrating that these are the only enzymes expressed in HEK 293T cells capable of modifying the site between Cys<sup>3</sup> and Cys<sup>4</sup>. This suggests that POGLUT2 and 3 act in tandem to efficiently modify protein substrates. Additional methods like tissue-specific gene expression analysis may elucidate differences in the biological roles between these two enzymes.

## **2.4 Biological Significance of POGLUT2 and POGLUT3**

POGLUT2 and 3 modify a large percentage of FBN1, FBN2, and LTBP1 EGF repeats, similar to the extensive amount of POGLUT1 and POFUT1 modifications found on Notch receptors (4, 27, 31, 36, 59). As mentioned above, *O*-fucosylation by POFUT1 is required for Notch to bind ligands (2, 32, 59), and *O*-glycosylation by both POGLUT1 and POFUT1

stabilizes EGF repeats and is required for proper Notch trafficking in certain contexts (10). Mutations in *POGLUT1* and *POFUT1* are causal to numerous human diseases (21), and elimination of either gene leads to embryonic lethality in mice (3, 11, 12). Recent analyses have demonstrated that *POGLUT2* and *POGLUT3* have similar roles with respect to *FBN1*, *FBN2*, and *LTBP1*.

As with *POGLUT1* and *POFUT1*, *POGLUT2* and 3 only modify properly folded EGF repeats (1, 8, 37, 60). *FBN1* secretion is significantly reduced in *POGLUT2* and 3 single and double knockout HEK293T cells similar to Notch trafficking defects in *POGLUT1* or *POFUT1* knockouts (10, 36) (See Chapter 3 for details). Both *Poglut2* and *Poglut3* have ER-retention signals (37, 38). Since secreted proteins fold in the ER, these findings strongly suggest that *POGLUT2* and 3 play a role as EGF repeat folding sensors and are essential for proper secretion of target proteins.

*O*-glycans also influence protein-protein interactions. As mentioned above, *NOTCH1* EGF12 is within the Notch ligand-binding domain (2). Mutation of the *O*-fucose site in *NOTCH1* EGF12 results in significantly reduced Notch signaling (28). S435 in *NOTCH1* EGF11, which is modified by *POGLUT2/3*, is also within the Notch ligand-binding domain (1, 2). Elimination of the *POGLUT2/3* site by mutation of S435 to alanine has little effect on *NOTCH1* activity; however, in combination with an EGF12 mutant lacking *O*-fucose, loss of *O*-glucose at S435 decreases *NOTCH1* activation further than the EGF12 mutant alone (1). The decrease in *NOTCH1* activation is not due to reduced cell surface expression (1). This shows that loss of an *O*-glucose added by *POGLUT2* and 3 negatively affects Notch activation which is dependent on Notch-ligand interactions. *POGLUT2* and 3 targets *FBN1*, *FBN2*, and *LTBP1* are secreted proteins involved in assembly and function of the ECM, a complex environment of

countless protein-protein interactions essential for proper signaling and development (40-42). Loss of POGLUT2 and 3-mediated *O*-glucose modifications on these proteins, in addition to other POGLUT2 and 3 targets yet to be confirmed, may lead to deleterious effects on protein function through impaired protein folding, trafficking, and/or protein-protein interactions resulting in disease.

MFS is one of the most well characterized ECM-related diseases. Mutations in *FBNI* resulting in MFS were first identified in the early 1990s (61-63). Since then, over 3000 *FBNI* genetic variants have been identified (64-66). Abnormal fibrillogenesis and impairment of elastogenesis can weaken tissue stability and lead to aortic rupture, which is commonly observed in MFS (45, 56).

Dysregulation of TGF- $\beta$  signaling is a proposed cause of MFS (44, 49, 56, 67). Other ECM-related disorders linked to aberrant TGF- $\beta$  levels include Loeys-Dietz Syndrome (LDS), Ehlers-Danlos Syndrome (EDS), Shprintzen-Goldberg Syndrome (SGS), Familial Thoracic Aortic Aneurysm and Dissection Syndrome (FTAAD), and Hereditary Hemorrhagic Telangiectasia (HHT) (57). Almost all of these disorders, like MFS, manifest in multiple organ systems (57). It is possible that *O*-glucosylation of fibrillins and LTBP $\beta$ s by POGLUT2 and 3 is critical to the function of these proteins, allowing for proper ECM structure, stability, and TGF $\beta$  signaling.

The fibrillin proteins contain more cysteine than any other extracellular protein (68). Many of these are found in 6-cysteine EGF repeats and 8-cysteine TGF- $\beta$ -binding protein-like (TB) domains within these proteins (63, 69-71). Most of the EGF repeats are of a subset that bind calcium and are predicted to be more rigid structurally. Many MFS mutations eliminate or add additional cysteines within EGF repeats of *FBNI*. Unpaired cysteines in wild-type EGF

repeat sequences are vanishingly rare, presumably due to the increased folding complexity from mixed disulfides. These mutations cause protein misfolding which reduces or abolishes FBN1 secretion and incorporation into microfibrils (72-74). Disulfide bonds between cysteines of an EGF repeat are critical for protein structure and stability, allowing for normal protein function. *O*-glycosylation, which occurs between cysteines of an EGF repeat, further stabilizes EGF repeats and promotes proper protein trafficking (10). As we have shown, FBN1, -2 and LTBP1 are heavily *O*-glucosylated by POGLUT2 and 3 (36). Elimination of *POGLUT2* and 3 in HEK293T cells reduces secretion of FBN1, resembling secretion defects of MFS variant FBN1 (36, 73, 74). Loss of *O*-glycans may also significantly disrupt protein-protein interactions in the extracellular space. In the absence of *O*-glucose, fibrillins may not oligomerize properly to form normal microfibrils. This could lead to dysregulated TGF- $\beta$  signaling and deleterious ECM remodeling in tissues like the heart and lungs. Loss of *O*-glucose may also prevent binding of the LTBP-TGF- $\beta$  complex to fibrillin microfibrils, causing aberrant TGF- $\beta$  signaling. Any of these mechanisms could lead to ECM disorders previously mentioned. Further research is needed to address these proposed mechanisms.

A unique set of MFS mutations within FBN1 EGF repeats affect calcium binding. The rigidity of FBN1 structure is largely due to calcium binding, which provides structural integrity and protection from proteolysis (75-78).  $\beta$ -hydroxylation of Asp/Asn is required for efficient calcium binding (79-81). Interestingly, the consensus sequence for  $\beta$ -hydroxylation overlaps with the POGLUT2 and 3 consensus sequence (36, 82). Most of the POGLUT2 and 3-modified EGF repeats of FBN1 are  $\beta$ -hydroxylated (36). This suggests *O*-glucosylation of calcium-binding EGF repeats by POGLUT2 and 3 may be linked to  $\beta$ -hydroxylation and calcium binding. It is possible that MFS mutations within the  $\beta$ -hydroxylation consensus sequence that

affect FBN1 structure and function could be due to loss of *O*-glucosylation instead of, or in addition to, effects on  $\beta$ -hydroxylation. FBN2 and LTBP1 are also composed of calcium-binding EGF repeats. More work is required to determine the effects of POGLUT2 and 3 *O*-glucosylation on calcium binding of FBN1, FBN2, and LTBP1.

A few papers have been published implicating POGLUT2 and 3 in cancers (83-88). Mutations in these genes make up approximately two percent of all simple somatic mutation cases reported in the NCI GDC (89). Interestingly, POGLUT2 and 3 mutations in reported cancer cases largely have opposite effects on genetic expression when compared to each other; POGLUT2 mutations more commonly lead to increased POGLUT2 genetic expression, whereas POGLUT3 mutations more commonly lead to decreased POGLUT3 genetic expression (89). The ECM has a role in cell fate, differentiation, proliferation, and migration (40, 43), and decades of research have focused on how perturbations of the ECM can lead to cancer. While we know that FBN1, FBN2, and LTBP1 are ECM proteins *O*-glucosylated by POGLUT2 and 3, many other ECM proteins including hemicentin-1 and fibulins contain the POGLUT2 and 3 consensus sequence for modification (36). Modulation of *POGLUT2* and/or *POGLUT3* expression could alter the ECM, therefore affecting cancer progression.

Aberrant Notch signaling is also associated with cancer development (21). POGLUT2 and 3-mediated *O*-glucosylation has already been shown to influence NOTCH1 signaling (1). *POGLUT3* knockdown in glioblastoma cells has been reported to downregulate expression of Notch receptors and high expression levels of *POGLUT3* correlate with poor prognosis in glioma patients (84). This implicates POGLUT2 and 3 as possible therapeutic targets to modulate Notch signaling.

## 2.5 Figures

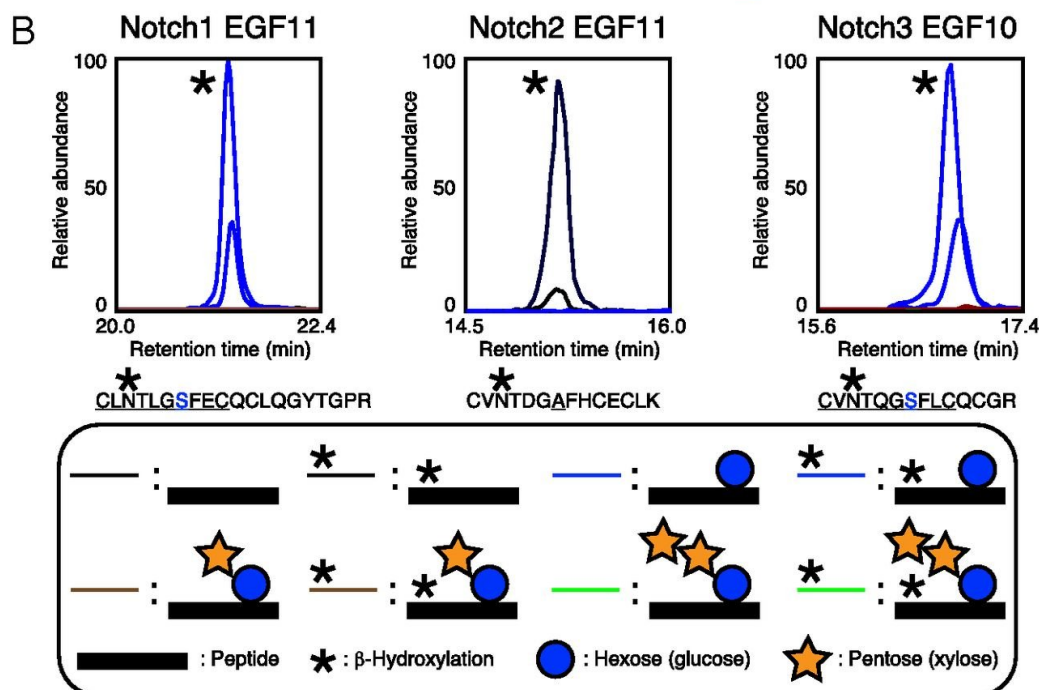


**Figure 2.1. Schematic of an EGF repeat illustrating modification sites of identified *O*-glycosyltransferases.** POGLUT2 and 3 add *O*-glucose to a site distinct from POGLUT1, POFUT1, and EOGT. *O*-fucose or *O*-GlcNAc are not extended past GlcNAc in *Drosophila*. Yellow (numbered), cysteines. \*, site of  $\beta$ -hydroxylation. Symbols for monosaccharides are shown below. Reproduced from (36) with permission.

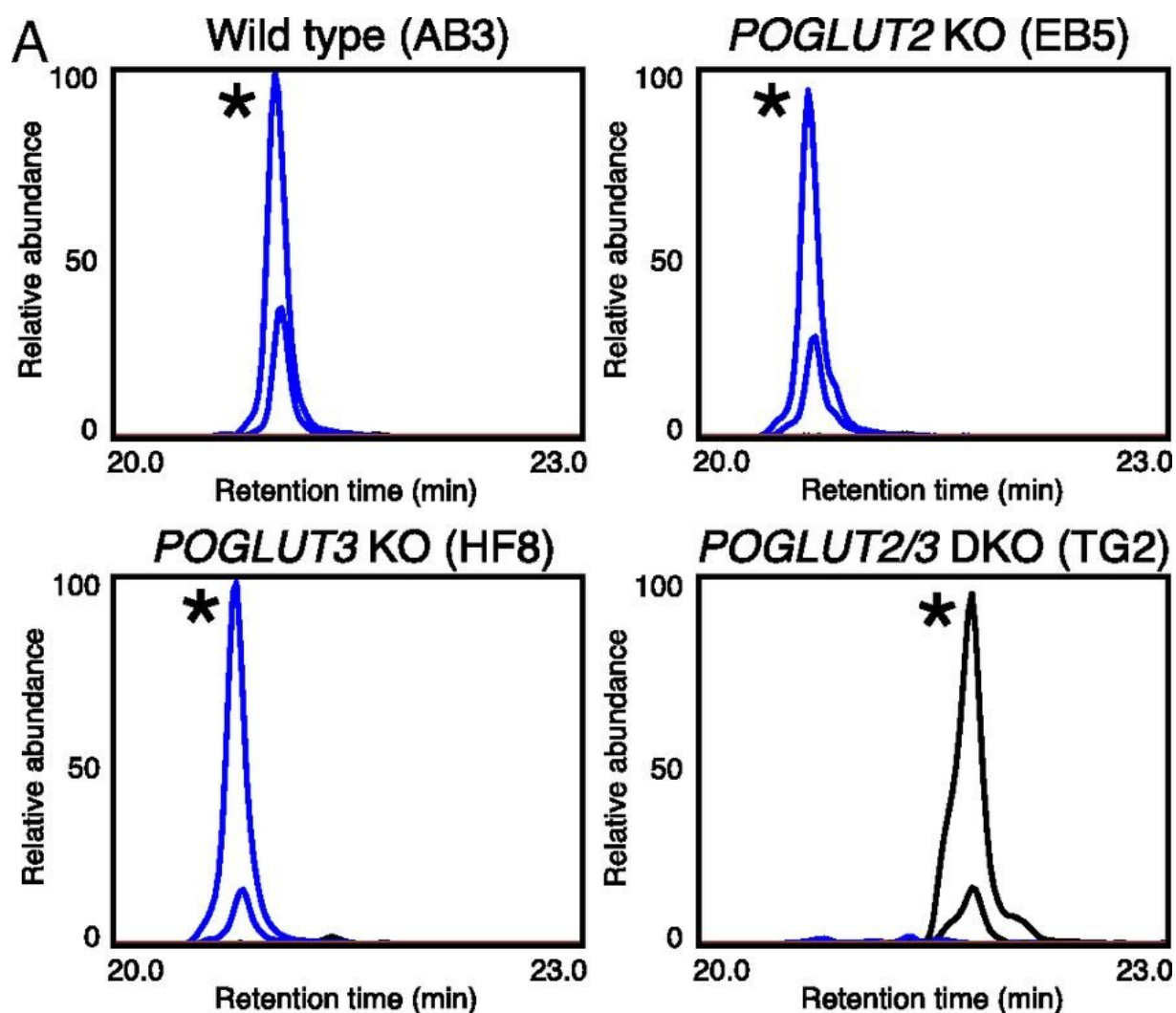


**A**

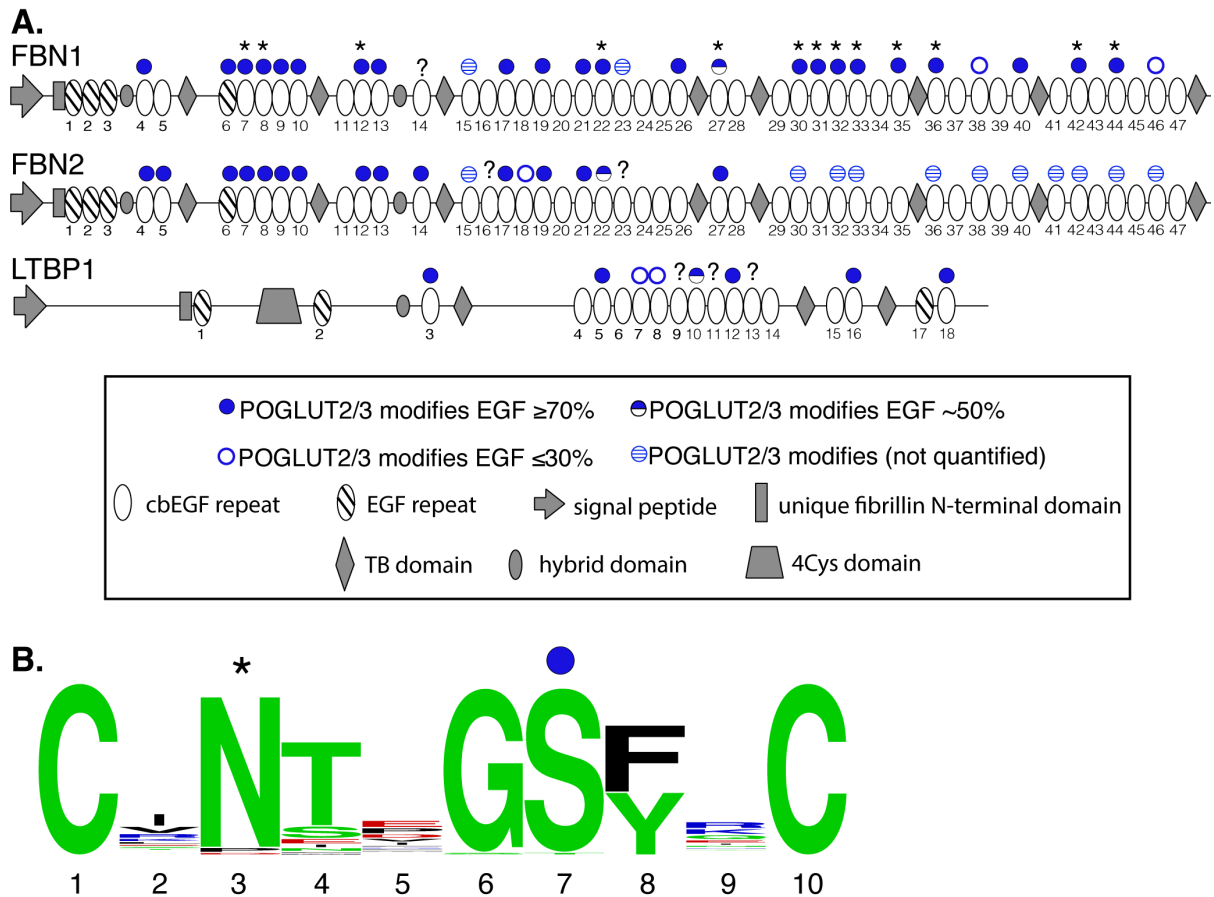
NOTCH1 EGF11: DVDE**C**DLGANP---**C**EHAGK**C**LNTLG**S**FECQ**C**LQGYTGPRCEI  
 NOTCH2 EGF11: DVDE**C**AMANSNP--**C**EHAGK**C**VNTDGA**F**H**C**E**C**LKGYAGPRCEM  
 NOTCH3 EGF10: DVDE**C**SIGANP---**C**EHLGR**C**VNTQ**G**S**F**LCQ**C**GRGYTGSRCEA  
 NOTCH4 EGF11: DLDE**C**MAQQGPSP**C**EHGG**S**CINTPG**S**FN**C**LCLPGYTGSRCEA  
 Putative Consensus: **CxNTxGSFxC**



**Figure 2.2. Ligand-binding domain EGF repeats from NOTCH1 and -3 are modified with novel *O*-Glc.** (A) Amino acid sequence alignment of EGF11 from mouse NOTCH1, -2, and -4 and of EGF10 from mouse NOTCH3. EGF repeats aligned based on conserved cysteines (boldface type). Modified serine is in blue. Putative consensus sequence is at the bottom. (B) EICs of glycoforms of the peptides from EGF11 of mouse NOTCH1 EGF1-36-Myc-His6, EGF11 of mouse NOTCH2 EGF1-36-Myc-His6, and EGF10 of mouse NOTCH3 EGF1-34-Myc-His6. Each was expressed in HEK293T cells, purified from the medium, reduced, alkylated, digested with trypsin, and analyzed by nano LC-MS/MS. Black line, unmodified peptide; blue line, peptide + *O*-Glc. A single asterisk indicates  $\beta$ -hydroxylation of asparagine in peptide. Reproduced from (1) with permission.



**Figure 2.3. POGLUT2 and POGLUT3 are responsible for *O*-glucosylation of mouse NOTCH1 EGF11.** (A) EICs of glycoforms of the EGF11 peptide from NOTCH1 EGF1-36-Myc-His6 produced in WT HEK293T cells (clone AB3), POGLUT2 KO cells (clone EB5), POGLUT3 KO cells (clone HF8), or POGLUT2 and POGLUT3 double-KO cells (clone TG2). Black line, unglycosylated peptide; blue line, *O*-glycosylated peptide. A single asterisk indicates  $\beta$ -hydroxylation of asparagine in peptide. Three independent transfection experiments were performed. Reproduced from (1) with permission.



**Figure 2.4. Summary of FBN1, FBN2, and LTBP1 EGF repeats modified by POGLUT2 and 3.**

**A.** WebLogo illustration of the current consensus sequence for modification by POGLUT2 and 3 based on all identified sites. *Amino acids* are colored according to their chemical properties: polar amino acids (G,S,T,Y,C,Q,N) are green, basic (K,R,H) are blue, acidic (D,E) are red, and hydrophobic (A,V,L,I,P,W,F,M) are black (90). *Numbers* below sequence represent amino acid position. \*, predicted site for  $\beta$ -hydroxylation. **B.** domain maps of FBN1, FBN2, and LTBP1 with predicted and identified POGLUT2 and 3 modification sites. \*, EGF repeats modified by POGLUT2 and 3 and  $\beta$ -hydroxylated greater than 50%. ?, unknown glycosylation status due to no protein sequence coverage with mass spectral analysis. Reproduced from (36) with permission.

## CHAPTER 3

### POGLUT2 AND POGLUT3 *O*-GLUCOSYLATE MULTIPLE EGF REPEATS IN FIBRILLIN-1, -2, AND LTBP1 AND PROMOTE SECRETION OF FIBRILLIN-1<sup>1</sup>

<sup>1</sup>Williamson DB, Sohn CJ, Ito A, Haltiwanger RS. POGLUT2 and POGLUT3 *O*-glucosylate multiple EGF repeats in fibrillin-1, -2, and LTBP1 and promote secretion of fibrillin-1. J Biol Chem. 2021 Sep;297(3):101055. Reprinted here with permission of the publisher.

### **3.1 Abstract**

Fibrillin-1 (FBN1) is the major component of extracellular matrix microfibrils, which are required for proper development of elastic tissues, including the heart and lungs. Through protein–protein interactions with latent transforming growth factor (TGF)  $\beta$ -binding protein 1 (LTBP1), microfibrils regulate TGF- $\beta$  signaling. Mutations within the 47 epidermal growth factor-like (EGF) repeats of FBN1 cause autosomal dominant disorders including Marfan Syndrome, which is characterized by disrupted TGF- $\beta$  signaling. We recently identified two novel protein *O*-glucosyltransferases, Protein *O*-glucosyltransferase 2 (POGLUT2) and 3 (POGLUT3), that modify a small fraction of EGF repeats on Notch. Here, using mass spectral analysis, we show that POGLUT2 and POGLUT3 also modify over half of the EGF repeats on FBN1, fibrillin-2 (FBN2), and LTBP1. While most sites are modified by both enzymes, some sites show a preference for either POGLUT2 or POGLUT3. POGLUT2 and POGLUT3 are homologs of POGLUT1, which stabilizes Notch proteins by addition of *O*-glucose to Notch EGF repeats. Like POGLUT1, POGLUT2 and 3 can discern a folded versus unfolded EGF repeat, suggesting POGLUT2 and 3 are involved in a protein folding pathway. *In vitro* secretion assays using the N-terminal portion of recombinant FBN1 revealed reduced FBN1 secretion in *POGLUT2* knockout, *POGLUT3* knockout, and *POGLUT2* and 3 double-knockout HEK293T cells compared with wild type. These results illustrate that POGLUT2 and 3 function together to *O*-glucosylate protein substrates and that these modifications play a role in the secretion of substrate proteins. It will be interesting to see how disease variants in these proteins affect their *O*-glucosylation.

### **3.2 Introduction**

Fibrillins are extracellular matrix (ECM) glycoproteins that form microfibrils, which act as a scaffold for elastin deposition (41). Depending on developmental stage, microfibrils consist of fibrillin-1 (FBN1) and/or fibrillin-2 (FBN2) and are commonly found in connective tissues such as the heart, lungs, and skin (41, 91). Through interactions with latent TGF $\beta$ -binding proteins (LTBPs) such as LTBP1, microfibrils also act as a reservoir of TGF $\beta$  (48, 50, 52). LTBP1 covalently binds the small latent complex (SLC) consisting of TGF $\beta$  bound to its latency-associated peptide (LAP). This interaction of LTBP1, LAP, and TGF $\beta$  comprises the large latent complex (LLC) and is necessary for secretion of TGF $\beta$  (53, 54). Previous studies have confirmed interactions between LTBP1 and FBN1 (48, 92, 93). FBN1 microfibrils bind the LLC, leading to a large reservoir of latent TGF $\beta$  in the ECM (94, 95). Once sequestered in the ECM, TGF $\beta$  can be activated through interactions between LTBP1 and integrins (96). Fibrillins and LTBPs comprise a protein superfamily, and mutations within this protein family cause various human developmental disorders including Marfan syndrome (MFS), which has been linked to dysregulated TGF $\beta$  signaling (49, 97). Fibrillins and LTBPs have similar domain structures consisting heavily of sequential calcium-binding epidermal growth factor repeats (cbEGFs). Each EGF repeat contains six conserved cysteine residues that form three disulfide bonds in a conserved pattern. A large percentage of EGF repeats in these proteins contain a consensus sequence for modification with an *O*-glucose sugar by two recently described enzymes: POGLUT2 and POGLUT3 (1).

Decades of research have been performed analyzing *O*-glycosylation of EGF repeats and the significance of the *O*-glycans. This research has been dominated by investigations of the Notch receptors (18, 21-23), which structurally resemble fibrillins and LTBPs in that they have

extracellular domains with as many as 36 tandem EGF repeats. Notch EGF repeats are heavily *O*-glycosylated by three enzymes: Protein *O*-Glucosyltransferase 1 (POGLUT1), Protein *O*-Fucosyltransferase 1 (POFUT1), and EGF Domain-Specific *O*-Linked N-Acetylglucosamine Transferase (EOGT) (Fig. 3.1A) (3-6, 27, 28, 31, 59). Elimination of either *Pofut1* or *Poglut1* in mice results in embryonic lethality with Notch-associated phenotypes (11, 12), illustrating the significance *O*-glycans have on Notch function.

POGLUT2 and 3 were identified as two *O*-glucosyltransferases that modify Notch EGF repeats between cysteines 3 and 4 of the EGFs with the consensus sequence C<sup>3</sup>-x-N-T-x-G-S-F-x-C<sup>4</sup>, in which an *O*-glucose is added to serine (1) (Fig. 3.1A). This site of modification is between cysteines 3 and 4 of an EGF repeat, distinct from POGLUT1 modifications that occur between cysteines 1 and 2. Only two sites on Notch proteins were modified by POGLUT 2 and 3 (1). FBN1 and -2 consist of 47 EGFs repeats, and LTBP1 has 18 EGF repeats. Using the original POGLUT2 and 3 consensus sequence, human FBN1, FBN2, and LTBP1 are predicted to have 12, 10, and 5 EGF repeats modified by POGLUT2 and 3, respectively (Fig. 3.1B). The small number of POGLUT2 and 3 sites on Notch receptors compared with the number of predicted sites on fibrillins and LTBP suggests POGLUT2 and 3 may have a more significant role in the function of these proteins compared with Notch. It has previously been shown that posttranslational modifications exist on several serine residues located between FBN1 EGFs29 to 35 (58), but neither the modifications nor the enzymes responsible were identified. Interestingly, these sites match the POGLUT2 and 3 consensus sequence.

The consensus sequence for POGLUT2 and 3 mediated *O*-glucosylation overlaps with the consensus sequence for  $\beta$ -hydroxylation, C<sup>3</sup>-x-(ND)-x-x-x-x-(FY)-x-C<sup>4</sup>, which has been linked to calcium binding that is essential for the structure and function of FBN1 (58, 79-82, 98).

Our work presents the first demonstration that FBN1, FBN2, and LTBP1 are extensively modified by *O*-glucose residues at high stoichiometry, added by POGLUT2 and/or POGLUT3. Semiquantitative mass spectral analysis was performed to determine sites of *O*-glucosylation and  $\beta$ -hydroxylation as well as the efficiency of modification at each site. Our results show that approximately half of the EGFs in these proteins are modified by POGLUT2 and/or 3, indicating that our original consensus sequence was too narrow. Additionally, we have identified site-specific preferences between POGLUT2 and 3 and demonstrated that POGLUT2 and 3 are possible regulators of protein folding and secretion.

### **3.3 Results**

#### **Endogenous human FBN1 is extensively O-glucosylated at high stoichiometry**

To confirm whether human FBN1 is modified with *O*-glucose, we first analyzed endogenous human FBN1 secreted from human dermal fibroblasts. Endogenous FBN1 was purified by immunoprecipitation of conditioned medium from 7-day post-confluent fibroblasts using a well-characterized polyclonal antibody that recognizes the unique N-terminal proline-rich domain of human FBN1 (74). A representative western blot illustrating FBN1 purification from conditioned media is shown in Figure 3.2A.

An established protocol for preparing purified proteins and digesting with trypsin for mass spectral analysis was used to identify *O*-glucosylation sites on purified human FBN1 (1). We obtained peptide matches covering greater than 70% of endogenous full-length FBN1. All peptides were identified using Byonic (Table S3.5 and Figs. S3.3–S3.29). For quantification of *O*-glycoforms, extracted ion chromatograms (EICs) were generated for each peptide modified with *O*-Hexose (Fig. 3.2B). Four glycoforms were searched based on mass to charge ratio ( $m/z$ )



for each EIC: unmodified peptide, unmodified peptide plus  $\beta$ -hydroxylation, peptide plus *O*-hexose, and peptide plus *O*-hexose and  $\beta$ -hydroxylation. The relative abundance of each glycoform of a specific peptide was calculated and plotted as a percentage of the total abundance (Fig. 3.2C).

Based on the original consensus sequence, 11 of 12 predicted sites were modified with *O*-hexose at high stoichiometry. EGF14 did not have sequence coverage, so it was not analyzed. Further analysis led to the identification of 13 additional EGFs modified with *O*-hexose, most of which were modified at high stoichiometry. Some EGFs had lower amounts of *O*-hexose modification (EGF38 and 46). We also identified EGFs that had a serine at position 7 of the POGLUT2/3 consensus sequence (Fig. 3.1A) but were unmodified (EGF5 and 29). The 13 additional modification sites included amino acids that were not identified in the original POGLUT2 and 3 consensus sequence, broadening the consensus sequence for POGLUT2 and 3 by allowing for a tyrosine at position 8 of the consensus sequence. In total, over half of the 47 EGF repeats of endogenous human FBN1 were modified by POGLUT2 and 3 (Fig. 3.2D). Many of the cbEGFs are also modified more than 50% by  $\beta$ -hydroxylation, with some notable exceptions (EGFs 4, 6, 9, 10, 13, 17, 21, and 40; Fig. 3.2C).

### ***POGLUT2 and/or 3 are responsible for addition of O-glucose to FBN1***

To confirm that POGLUT2 and 3 are indeed responsible for the modifications found on FBN1, and that these modifications are *O*-glucose, transient transfections with constructs encoding His-tagged N- and C-terminal FBN1 fragments (99) were performed in wild-type and POGLUT2/3 double-knockout HEK293T cells as previously described (1). The FBN1 N-terminal protein fragment (FBN1-N) is illustrated in Figure 3.3A with predicted and confirmed

sites of POGLUT2 and 3 modification identified. Recombinant FBN1-N was purified from conditioned media over a Ni-NTA column, digested, and prepared for mass spectral analysis (Fig. S3.1). Fifteen of the 26 EGFs were confirmed to be modified at similar stoichiometries to those seen in endogenous FBN1, confirming that HEK293T cells are a good system to monitor this form of modification. In addition to the modified EGFs identified on the N-terminal region of endogenous FBN1 (Fig. 3.2D), three additional modified EGFs were identified on recombinant FBN1-N (EGF19, 23, 26, Fig. 3.3A). These additional sites were identified due to a higher protein sequence coverage in our mass spectral analysis of overexpressed FBN1-N versus endogenous FBN1. POGLUT2 and 3 mediated modifications were lost when FBN1-N was expressed in POGLUT2 and 3 double-knockout HEK293T cells, confirming that these two enzymes are responsible for *O*-glucosylation of FBN1 and that the modifications are *O*-glucose (Fig. 3.3B). Interestingly, the level of  $\beta$ -hydroxylation is lower on recombinant FBN1-N compared with endogenous FBN1 (compare EGFs 7, 8, and 12 in Figs. 3.2C and 3.3B), suggesting that  $\beta$ -hydroxylation is less efficient in HEK293T cells than in dermal fibroblasts. All peptides were identified using Byonic (Tables S3.6–S3.10 and Figs. S3.3–S3.17).

The C-terminal half of FBN1 (FBN1-C) was also expressed in wild-type and POGLUT2/3 double-knockout HEK293T cells, purified from media (Fig. S3.1), and analyzed by mass spectrometry. The same POGLUT2 and 3 modified EGFs were identified between the C-terminal region of endogenous FBN1 and recombinant FBN1-C; however, the POGLUT2 and 3 modification efficiency deviated from what was observed with endogenous FBN1. It is possible there was a folding defect in FBN1-C protein, leading to lower stoichiometry of POGLUT2 and 3 modified peptides compared with endogenous FBN1. Due to this, we did not quantify the efficiency of modification for *O*-glucosylation by POGLUT2 and 3 on FBN1-C. All *O*-glucose

modifications on FBN1-C were lost when expressed in the POGLUT2/3 double-knockout cells. POGLUT2 and 3 modified peptides of recombinant FBN1-C can be found in Tables S3.11 and S3.12 (Figs. S3.18–S3.29).

***FBN2 and LTBP1 are O-glucosylated at multiple sites by POGLUT2 and 3 in HEK293T cells***

A FBN2 N-terminal fragment (FBN2-N) (100) and full-length LTBP1 (50) were expressed in wild-type and POGLUT2/3 double-knockout HEK293T cells (Fig. 3.4A). Each recombinant protein was purified using a Ni-NTA column from conditioned medium (Fig. S3.1), digested, and analyzed by mass spectrometry. As with FBN1, a greater number of EGFs than predicted with the original consensus sequence were modified by POGLUT2 and/or 3 for both proteins. For FBN2-N, 17 out of 28 total EGFs were *O*-glucosylated, and for LTBP1, eight out of 18 total EGFs were *O*-glucosylated (Fig. 3.4A). The majority of these POGLUT2 and 3 modifications were at high stoichiometry (Fig. 3.4, B and C). FBN2 EGF18 is an interesting site that has very low levels of POGLUT2 and/or 3 modification. This site almost matches the original POGLUT2 and 3 consensus sequence, but instead of a serine at position 7 of the consensus sequence, there is a threonine. This suggests that POGLUT2 and 3 can modify a threonine, but serine is highly preferred. All *O*-glucose modifications were lost when FBN2-N or LTBP1 were expressed in POGLUT2/3 double-knockout HEK293T cells, confirming that POGLUT2 and 3 are responsible for *O*-glucosylating FBN2 and LTBP1 (Fig. 3.4, B and C). Peptides identified using Byonic for FBN2-N can be found in Tables S3.13–S3.15 (Figs. S3.30–S3.46) and those for LTBP1 in Tables S3.16–S3.19 (Figs. S3.57–S3.64).

FBN2-C (100) was also expressed in wild-type and POGLUT2/3 double-knockout cells, and the resulting peptides identified for FBN2-C can be found in Tables S3.20 and S3.21 (Figs.

S3.47–S3.56). As with FBN1-C, most of the EGFs were modified 50% or less by POGLUT2 and 3. As previously mentioned, FBN1-C recombinant data did not correspond with endogenous FBN1 mass spectral data. Due to this, we did not quantify stoichiometry at each modified EGF, but instead only identified FBN2-C EGFs that were modified by POGLUT2 and/or 3. All *O*-glucose residues were lost when FBN2-C was expressed in POGLUT2/3 double-knockout cells (Table S3.21).

### **POGLUT2 and POGLUT3 modify EGF repeats with the consensus sequence**

#### **C<sup>3</sup>-x-N-T-x-G-S-(FY)-x-C<sup>4</sup>**

After compiling all confirmed POGLUT2 and 3 modification sites on NOTCH1, NOTCH3, FBN1, FBN2, and LTBP1, we generated a WebLogo summarizing sequences between cysteines 3 and 4 of all modified sites (Fig. 3.5A). The most notable difference compared with the original consensus sequence is the addition of a tyrosine at the eighth position in the sequence, C<sup>3</sup>-x-N-T-x-G-S-(FY)-x-C<sup>4</sup>. A summary of the predicted POGLUT2 and 3 *O*-glucosylation sites using the highly conserved residues of the new consensus sequence and confirmed modification sites on FBN1, -2, and LTBP1 is shown in Figure 3.5B. For all three proteins, approximately half of their EGF repeats are *O*-glucosylated by POGLUT2 and 3.

While all sites with the C<sup>3</sup>-x-N-T-x-G-S-(FY)-x-C<sup>4</sup> sequence were at least partially modified, other sites contain non-conserved residues at the third, fourth, and sixth position (Fig. 3.5A). For example, at the third position where asparagine is mostly conserved, FBN1 EGF6 has a proline and FBN1 EGF31 has an aspartic acid. Both of these EGFs were modified at high stoichiometry by POGLUT2 and/or 3. At the fourth position where threonine is mostly conserved, we also observed EGFs with asparagine, glutamate, serine, methionine, leucine,

isoleucine, and valine. *O*-glucosylation generally occurs at high stoichiometry with asparagine, glutamine, or serine at the fourth position. A significant decrease or no modification at all was observed when methionine, leucine, isoleucine, or valine was at the fourth position (Table S3.1). This suggests that hydrophobic amino acids at this position are not preferred by POGLUT2 and 3 for *O*-glucosylation of an EGF repeat. A serine in position 4 is usually highly modified (FBN1 EGF13; FBN2 EGF10, 14, 19), but LTBP1 EGF7 is poorly modified, suggesting that other factors in LTBP1 are reducing the efficiency of modification at EGF7. At the sixth position of the consensus sequence where glycine is mostly conserved, FBN1 EGF5 and FBN2 EGF6 had a serine. Interestingly, FBN1 EGF5 was modified at very low stoichiometry (<1%), but FBN2 EGF6 was modified at high stoichiometry, again indicating that additional unknown factors are contributing to efficiency of modification at individual EGFs.

To further analyze the POGLUT2 and 3 consensus sequence, we attempted to convert an unmodified EGF to a POGLUT2 and/or 3 modified EGF. We selected FBN1 EGF11, which has the sequence C<sup>3</sup>-x-N-L-x-G-T-Y-x-C<sup>4</sup>. Mass spectral analysis confirmed this site is not modified by POGLUT2 or 3 in WT cells (Fig. 3.6). This site may be unmodified due to the hydrophobic leucine residue at position 4 and/or because threonine is at position 7 instead of serine. Threonine and serine differ by a single methyl group. *O*-glycosyltransferases such as POFUT1 and EOGT have been shown to modify both a serine and threonine (4, 6), while POGLUT1 only modifies serine (4, 101). With site-directed mutagenesis we generated the following three mutants using FBN1-N plasmid: L744T (position 4), T747S (position 7, potential site of modification), and double-mutant L744T/T747S (Fig. 3.6). Each mutant was expressed in wild-type HEK293T cells and purified protein was analyzed by mass spectrometry. FBN1 EGF11 remained unmodified in the single mutants but was modified at high stoichiometry in the L744T/T747S double mutant

(Fig. 3.6). This further supports POGLUT2 and 3 do not prefer a hydrophobic residue at position 4 of the consensus sequence, and it also demonstrates that serine is preferred over threonine at position 7. The ability to convert an unmodified EGF to a POGLUT2 and/or 3 modified EGF further supports the importance of the highly conserved residues in the consensus sequence for POGLUT2 and 3. Peptides identified by Byonic for FBN1 EGF11 are in Tables S3.22–S3.25.

### **POGLUT2 and POGLUT3 display some site-specific differences**

Our previous studies have indicated that POGLUT2 and 3 show a unique preference for different EGFs of Notch proteins (1). To further investigate this site specificity, FBN1-N, FBN2-N, and LTBP1 were expressed in POGLUT2 single-knockout and POGLUT3 single-knockout cells and mass spectral analysis was performed (Tables S3.26–S3.33). These results were compared to wild-type data to identify sites where *O*-glucosylation was reduced by either knockout condition (Fig. 3.7). *O*-glucosylation needed to be reduced by 30% or more compared with wild type to be considered significant. This cutoff value was used to account for sample variation and mass spectrometer conditions that typically affect mass spectral quantification in the 10 to 20% range. Several EGFs were identified as POGLUT2 or POGLUT3 preferred sites on FBN1-N (Fig. 3.7A), FBN2-N (Fig. 3.7B), and LTBP1 (Fig. 3.7C). For example, FBN1 EGF9 is *O*-glucosylated at high stoichiometry in both wild-type and POGLUT3 knockout cells, but in POGLUT2 knockout cells the *O*-glucose modification is almost completely lost (Fig. 3.7A). Since *O*-glucosylation is significantly reduced only when POGLUT2 is knocked out, this suggests FBN1 EGF9 cannot be modified efficiently by POGLUT3 and is a POGLUT2 preferred EGF repeat. Conversely, LTBP1 EGF 10 is approximately 50% modified in wild-type and POGLUT2 knockout cells, but in POGLUT3 knockout cells the *O*-glucose modification is

almost completely lost (Fig. 3.7C). This suggests that LTBP1 EGF10 is a POGLUT3 preferred EGF repeat since *O*-glucosylation is significantly reduced only when POGLUT3 is knocked out. Some EGF repeats showed no preference by either enzyme. FBN2 EGF10 and 13 maintained *O*-glucosylation at high stoichiometry regardless of knockout condition (Fig. 3.7B). Ten POGLUT2 preferred EGFs were identified: FBN1 EGFs 4, 6, 9, 12, 21; FBN2 EGFs 9, 21, 22; and LTBP1 EGFs 5, and 18 (Table S3.2). Four POGLUT3 preferred sites were identified: FBN1 EGF8; FBN2 EGF 12 and 14; and LTBP1 EGF 10 (Table S3.2). Weblogo summaries of the sites modified by each enzyme are shown in Figure 3.8. There are no significant differences between the two consensus sequences, indicating that factors other than the amino acid sequence between cysteines 3 and 4 of an EGF repeat may contribute to site specificity. Overall, these data support our previous results (1) that while partially redundant, there is site specificity between POGLUT2 and 3 and that these enzymes work in a collaborative fashion to ensure *O*-glycosylation of substrate proteins.

Our site-mapping data showed the POGLUT2 and 3 consensus sequence can be used to accurately predict *O*-glucosylation sites, but we wanted to directly analyze whether changes in any position between cysteines 3 and 4 can drive site specificity between POGLUT2 and 3. For this study, FBN1 EGF12 and FBN2 EGF12 were selected. FBN1 EGF12 is POGLUT2 preferred, and FBN2 EGF 12 is POGLUT3 preferred (Fig. 3.7, A and B). These EGF repeats were selected because they only differ by two amino acids between cysteines 3 and 4, and they are in the same relative position of each protein (Fig. 3.9A, red). We found it interesting that two sites so similar exhibited variable enzyme specificity. F788Y/V789S double mutations were generated for FBN1-N, and Y832F/ S833V double mutations were generated for FBN2-N. Each mutant was expressed in wild-type, POGLUT2 knockout, and POGLUT3 knockout cells and mass spectral

analysis was performed on purified proteins (Tables S34–S39). Results were compared with wild-type FBN1-N and FBN2-N EGF12 data to determine changes in *O*-glucosylation by POGLUT2 and 3. No significant *O*-glucosylation differences were observed for either mutant compared with wild type, and only slight changes were seen in the POGLUT2 or POGLUT3 knockouts (Fig. 3.9B). These data strongly suggest that the site specificity of POGLUT2 and 3 is dictated by sequences outside of those between cysteines 3 and 4, and the context of an entire EGF repeat needs to be considered. We used Clustal Omega to align whole EGF sequences of POGLUT2 and POGLUT3 preferred EGFs (Fig. S3.2). No significant differences were identified when comparing whole EGF sequences. This suggests the primary sequence alone cannot be used to predict site specificity, but instead, other factors such as the position of an EGF in the context of protein structure need to be considered as well.

#### **Knockout of POGLUT2 and 3 reduces secretion of FBN1-N in HEK293T cells**

We have previously shown that *O*-glucosylation by enzymes such as POFUT1 and POGLUT1 stabilizes EGF repeats and assists secretion of substrate proteins such as Notch (10). An established cell-based secretion assay protocol (102) was used to analyze whether *O*-glucosylation by POGLUT2 and 3 is necessary for efficient secretion of FBN1-N. The same FBN1-N construct used for mass spectral analysis was used for secretion assays. We attempted secretion assays using full-length protein, but it was difficult to consistently express for accurate quantification. Figure 3.10A is a representative western blot illustrating secretion of FBN1-N and co-expressed human IgG. Compared with wild-type HEK293T cells, secretion of FBN1-N was reduced approximately twofold in POGLUT2 knockout or POGLUT3 knockout cells, while secretion of FBN1-N was reduced approximately fourfold in POGLUT2/3 double-knockout cells



(Fig. 3.10B). Secretion of human IgG, a control protein without EGF repeats, was unchanged across all conditions. These data suggest that *O*-glucose addition by POGLUT2 and/or POGLUT3 is necessary for efficient secretion of FBN1-N.

### **3.4 Discussion**

Our work is the first to analyze *O*-glucosylation of FBN1, FBN2, and LTBP1. Through mass spectral analysis, we confirmed FBN1, FBN2, and LTBP1 are heavily modified with *O*-glucose by POGLUT2 and/or POGLUT3. We also identified site specificity between POGLUT2 and POGLUT3, which provides a potential rationale for why mammals possess two enzymes with functional redundancy. Finally, our data show that addition of *O*-glucose is necessary for proper secretion of FBN1-N, suggesting a biologically significant role for these enzymes. In total, we mapped 62 POGLUT2 and 3 mediated *O*-glucosylation sites out of 112 EGF repeats of FBN1, -2, and LTBP1. The original POGLUT2 and 3 consensus sequence, C<sup>3</sup>-x-N-T-x-G-S-F-x-C<sup>4</sup>, predicted less than half these many *O*-glucose modification sites (compare Fig. 3.1 to Fig. 3.5) (1). Based on our analysis, most of the residues in the original consensus sequence were highly conserved, except that tyrosine was found in place of phenylalanine at nearly half of the modified sites. All sites containing the broadened consensus sequence, C<sup>3</sup>-x-N-T-x-G-S-(FY)-x-C<sup>4</sup>, were at least partially modified, suggesting that this sequence accurately predicts modification of an EGF repeat. Database searches with this sequence reveal that many other proteins also had increases in the predicted number of modification sites including hemicentin-1, fibulin-1, fibulin-2, and LTBP-4 (Table S3.4). The total number of human proteins predicted to be modified by POGLUT2 and 3 increased from 34 to 56, and the total number of predicted sites increased from 86 to 167 (Table S3.4). Nonetheless, our data show that residues other than N

(position 3), T (position 4), or G (position 6) occur between cysteines 3 and 4 of modified sites, and there are likely additional factors outside of this sequence that play a role, indicating that more work needs to be done to define exactly what POGLUT2 and 3 recognize in an EGF repeat. To avoid missing additional proteins or sites modified by POGLUT2 and 3, a more general search string of C<sup>3</sup>-x-x-x-x-x-S-x-x-C<sup>4</sup> can be used, although this sequence will also identify sites that are not modified.

Sequence logo graphs have previously been generated using over 1000 human EGF-like domain sequences found in the PROSITE database (60). Asparagine at position 3, glycine at position 6, serine or glycine at position 7, and tyrosine or phenylalanine at position 8 are the most commonly observed amino acids between cysteines 3 and 4 (60). Interestingly, this is very similar to the POGLUT2 and 3 consensus sequence. This suggests that all amino acids within the POGLUT2 and 3 consensus sequence may not be required for *O*-glucosylation, but instead, some of these amino acids exist within the consensus sequence simply because they are common in most EGFs. This is supported by the exceptions we found to the POGLUT2 and 3 consensus sequence.

Our analyses using POGLUT2 and POGLUT3 single-knockout cells provided insight into site-specific preferences between these two enzymes. We identified ten EGFs between FBN1, FBN2, and LTBP1 that are preferred by POGLUT2 and four preferred by POGLUT3. This suggests that POGLUT2 could have a more significant role in modifying FBN1, -2, and LTBP1, whereas POGLUT3 may be targeted toward other proteins yet to be identified as *O*-glucosylated. Thus, POGLUT2 and POGLUT3 may not be strictly redundant. Comparison of sequences between cysteines 3 and 4 modified by POGLUT2 or POGLUT3 did not reveal any significant differences, indicating that sequences outside of the consensus sequence may be

influencing specificity (Fig. 3.8). This was further supported by our attempt to change site specificity of EGF12 of FBN1 and -2 (Fig. 3.9). Transcriptional analysis of POGLUT2 and POGLUT3 in embryonic stem cells and epiblast-like cells further highlights these enzymes are not completely redundant: POGLUT3 mRNA transcript levels were greatly increased in epiblast-like cells compared with POGLUT2 (103).

The POGLUT2 and 3 consensus sequence, C<sup>3</sup>-x-N-T-x-G-S-(FY)-x-C<sup>4</sup>, overlaps the consensus sequence for  $\beta$ -hydroxylation, C<sup>3</sup>-x-(ND)-x-x-x-x-(FY)-x-C<sup>4</sup> (82). The majority of POGLUT2 and/or 3 modified EGFs on endogenous FBN1 were also  $\beta$ -hydroxylated. Interestingly,  $\beta$ -hydroxylation was reduced in recombinant FBN1 expressed in HEK293T cells compared with endogenous FBN1 secreted from human dermal fibroblasts. Previous studies show that  $\beta$ -hydroxylation is necessary for efficient calcium binding to EGF repeats (79-81). Calcium-binding affinities are highly variable across FBN1 (75). Binding of calcium provides rigidity to the structure of FBN1 and affords protection from proteolysis (75). A subset of MFS patient mutations in FBN1 affect residues involved in calcium binding, and these mutations have been linked to increased susceptibility to proteolysis (76, 77). FBN2 and LTBP1 have similar domain organization to FBN1 and are primarily composed of calcium-binding EGF repeats. It is possible that POGLUT2 and 3 *O*-glucosylation of EGF repeats is connected to  $\beta$ -hydroxylation and calcium binding of FBN1, FBN2, and LTBP1. Some of the defects in FBN1 structure or function related to MFS mutations could be due to loss of *O*-glucosylation rather than changes in  $\beta$ -hydroxylation. Further studies will need to be performed to investigate this.

Lastly, we have shown that secretion of FBN1-N is reduced in POGLUT2, POGLUT3, and POGLUT2 and 3 knockout cells. This coincides with previous work demonstrating that *O*-glycosylation by POGLUT1 and POFUT1 stabilizes Notch EGF repeats and promotes cell-

surface expression of the receptor (10). Like POGLUT1 and POFUT1, POGLUT2 and 3 have an endoplasmic reticulum (ER) retention sequence, localizing them to the ER where protein folding for the secretory pathway occurs (104). POGLUT2 and POGLUT3 also only modify folded EGF repeats, indicating that they can differentiate a folded from an unfolded structure (1). Combined with our secretion assay results, this indicates that *O*-glucosylation by POGLUT2 and 3 assists folding of substrate proteins and is necessary for efficient secretion. It is also possible that these *O*-glucose modifications are involved in the many binding interactions that occur in the ECM between FBN1, -2, and LTBP1 and other proteins required for proper development. *O*-glycans such as *O*-fucose modifications found on Notch receptor are directly involved in Notch-ligand binding, affecting Notch signaling (2, 28, 32, 35, 59). This provides a precedent for the involvement *O*-glycans on EGF repeats in protein–protein interactions. Microfibrils are a complex network of many protein–protein interactions that *O*-glucosylation by POGLUT2 and 3 may affect (41, 105). Further work is needed to understand the role of such *O*-glycans in the context of the ECM.

### **3.5 Experimental Procedure**

#### **Plasmids and mutagenesis**

pDNBP-rf16 (FBN1-N), pcDNA-rf6h (FBN1-C), pcDNA-rFBN2-2 (FBN2-N), pDNBP-rFBN2-1 (FBN2-C), and pCEP-Pu hLTBP1 FL-His have been described previously (58, 79, 80). The FBN1 and 2 plasmids were kindly gifted by Dr. Dieter Reinhardt (McGill University), and the LTBP1 plasmid was kindly gifted by Dr. Clair Baldock (The University of Manchester). All FBN1-N and FBN2-N mutants were generated using site-directed mutagenesis with CloneAmp HiFi Premix (Takara Bio, USA). Primers used to generate each mutant are listed in Table S3.3.

PCR products were digested with DpnI for 1 h at 37 °C and transformed into DH5 $\alpha$ -competent cells (Invitrogen). All mutated plasmids were confirmed by sequencing. pKR5-IgG plasmid encoding human IgG was used as a transfection control for secretion assays as previously described (10, 102).

### **Cell culture**

All cells were maintained at 37 °C and 5% CO<sub>2</sub> in DMEM high glucose media supplemented with 10% bovine calf serum and 1% penicillin/streptomycin. HEK293T cells were obtained from ATCC. *POGLUT2* and *POGLUT3* single and double knockout HEK293T cells have been previously described (1). Primary human dermal fibroblasts derived from adult foreskin were purchased from ATCC.

### **Antibodies**

Rabbit derived IgG against the proline-rich N-terminal region of human FBN1 (106) was generously provided by Dr. Penny Handford (Oxford University).

### **Immunoprecipitation of endogenous FBN1 from primary fibroblasts**

Fibroblasts were cultured to confluency in a 10 cm dish without changing media. Approximately 7 d post confluency, culture media was collected and cleared. For immunoprecipitation, 50  $\mu$ l of magnetic beads (Protein G Dynabeads, Invitrogen) were incubated with 4  $\mu$ l anti-FBN1 antibody (diluted in 200  $\mu$ l 1X PBS, 0.02% Tween) for 20 min rotating at room temp. One wash was performed with 200  $\mu$ l 1X PBS, 0.02% Tween. Beads were resuspended in 1 mL conditioned culture media and incubated for 1 h rotating at room temp. Three washes were performed with

1X PBS, 0.02% Tween. FBN1 was eluted from beads in 30  $\mu$ l of 8 M urea in 400 mM ammonium bicarbonate at 37 C for 15 min. Samples were stored at -20 C.

### **Protein expression and purification**

For each replicate, five 10 cm dishes of approximately 90% confluent HEK293T cells were transiently transfected. This was done for each plasmid and cell condition (wild type, single knockout, and double knockout cells). Prior to transfection, complete media was removed, cells were washed once with 1X PBS, and 6 mL Opti-MEM (Invitrogen) reduced serum media was added. Cells were transfected using PEI and 5  $\mu$ g plasmid per 10 cm dish (DNA:PEI 1:6) in 500  $\mu$ l OPTI-MEM. Media was harvested after approximately 3 d. Media was cleared by centrifugation followed by a 0.45  $\mu$ m syringe filter. 5 M NaCl and 1 M imidazole were added to a final concentration of 500 mM and 10 mM, respectively. For purification, a 350-400  $\mu$ l Ni-NTA bead volume (700-800  $\mu$ l 50% slurry) was used (Qiagen). Wash buffer consisted of 500 mM NaCl and 10 mM imidazole in 1X TBS. Proteins were eluted using 250 mM imidazole in 1X TBS. Purified proteins were analyzed by SDS-gel electrophoresis (Fig. S3.1).

### **Mass spectral analysis**

Purified protein (15-20  $\mu$ l) was transferred to Protein LoBind tubes. Tubes were placed in speed vacuum for 10-15 min to concentrate samples. Proteins were denatured and reduced using 10  $\mu$ l of reducing buffer containing 8 M Urea, 400 mM ammonium bicarbonate, and 10 mM TCEP at 50 °C for 5 min. Alkylation was performed at room temp in the dark with 100 mM iodoacetamide in 50 mM TrisHCl for 30 min to 1 h. Mass spectral grade water (45  $\mu$ l) was added to each sample. Trypsin (cleaves C-terminal to lysine and arginine, Thermo Scientific Pierce

90057) and/or V8 (cleaves C-terminal glutamic and aspartic acid, Millipore Sigma P6181) or chymotrypsin (cleaves C-terminal to tyrosine, phenylalanine, tryptophan, leucine, or isoleucine, Sigma 90056) protease was added without exceeding 500 ng of total enzyme per sample. Samples were incubated in 37 °C water bath for 4-6 h. Formic acid (7 µl of 5%) was added and samples were sonicated for 20 min. Samples were desalted with Millipore C18 Zip Tip Pipette Tips. After elution in 50% acetonitrile, 0.1% acetic acid, samples were diluted to an approximate concentration of 10 ng/µl, 17% acetonitrile, and 0.1% formic acid. Approximately 10 ng of each sample was injected on a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Fisher) with an Easy nano-LC HPLC system with a C18 EasySpray PepMap RSLC C18 column (50 µm × 15 cm, Thermo Fisher Scientific). A 30 min binary gradient solvent system (Solvent A: 0.1% formic acid in water and Solvent B: 90% acetonitrile, 0.1% formic acid in water) with a constant flow of 300 nL/min was used. Positive polarity mode was used with a m/z range of 350-2,000 at a resolution of 35,000 and automatic gain control set to  $1 \times 10^6$ . Higher energy collisional dissociation-tandem mass spectrometry (HCD-MS/MS) was used on the top 10 precursor ions in each full scan (collision energy set to 27%,  $2 \times 10^5$  gain control, isolation window m/z 3.0, dynamic exclusion enabled, and 17,500 fragment resolution. PMi-Byonic (v.2.10.5) and Proteome Discoverer (v2.1) were used to identify peptides. Fixed modifications: Carbamidomethyl +57.021464 at C. Variable modifications: Oxidation +15.994915 at M,H,N,D, Deamidated +0.984016 at N, and Ammonia-loss -17.026549 at N-Term C. Precursor mass tolerance was set to 20 ppm and fragment mass tolerance was set to 30 ppm. Two missed cleavages were allowed. Protein and peptide false discovery rates were set to a threshold of 1% and calculated in Byonic software version 2.10.5 (Protein Metrics) using the 2-dimensional target decoy strategy as described (107). All data was searched against either a human FBN1

database (Uniprot accession number P35555 version 4 updated April 10, 2019, 1 entry), FBN2 (Uniprot accession number P35556 version 3 updated May 26, 2009, 1 entry), or LTBP1 (Uniprot accession number Q14766 version 4 updated March 2, 2010, 1 entry). Xcalibur Qual Browser (v2.0.3) was used to generate EICs for all identified peptides. For each peptide, area under the curve was calculated for each peak corresponding to searched glycoforms. Relative abundance was calculated by comparing area under the curve for a single glycoform to the total area under curve for all searched glycoforms of a specific peptide. Glycoforms searched: unmodified peptide, unmodified peptide plus  $\beta$ -hydroxylation, modified peptide with *O*-hexose, and modified peptide with *O*-hexose plus  $\beta$ -hydroxylation.

#### **Cell-based secretion assays**

Wild type, single knockout, and double knockout HEK293T cells were transiently co-transfected with the plasmid encoding human FBN1-N and human IgG as described (10, 102).

Approximately  $8\text{--}9 \times 10^5$  cells in DMEM High Glucose media supplemented with 10% BCS were seeded into 6 well plates approximately 24 h prior to transfection. Media was changed to 800  $\mu$ l of Opti-MEM reduced serum media before adding transfection mix. Each transfection mix consisted of 1.5  $\mu$ g FBN-N plasmid, 0.1  $\mu$ g of IgG plasmid, 9  $\mu$ l PEI, and 150  $\mu$ l Opti-MEM. Cells were incubated for approximately 48 h after transfection. Media was harvested and cleared by centrifugation.

#### **Western blotting**

Cleared conditioned medium (36  $\mu$ l) was mixed with 12  $\mu$ l 4X reducing sample buffer containing 2-mercaptoethanol. Samples were incubated on heat block at 100 °C for 10 min.



Gradient gels (10-lane, 4-20%, Bio-Rad) were used to resolve proteins. Semi-dry transfer was performed using Invitrogen Power Blotter XL, nitrocellulose power blotter select transfer stacks, and 1X power blotter transfer buffer. Membranes were blocked at 4 °C overnight in 5% milk in 1X TBS 0.1% Tween 20. Membranes were incubated with mouse anti-His antibody (Bio-Rad catalog number MCA1396) at 1:1,000 dilution in blocking solution for 1 h at room temp. Secondary antibodies, IRDye®680RD goat (polyclonal) anti-mouse IgG (LI-COR catalog number 925-68070) and IRDye®800CW goat (polyclonal) anti-human IgG (H + L) (LI-COR catalog number 925-32232), were incubated with membrane at 1:10,000 dilution for 1 h at room temp. Blots were imaged and quantified using a LI-COR Odyssey CLx and Odyssey Imager software.

### **Experimental design and statistical rationale**

See Table S3.40. Available by request or with original manuscript.

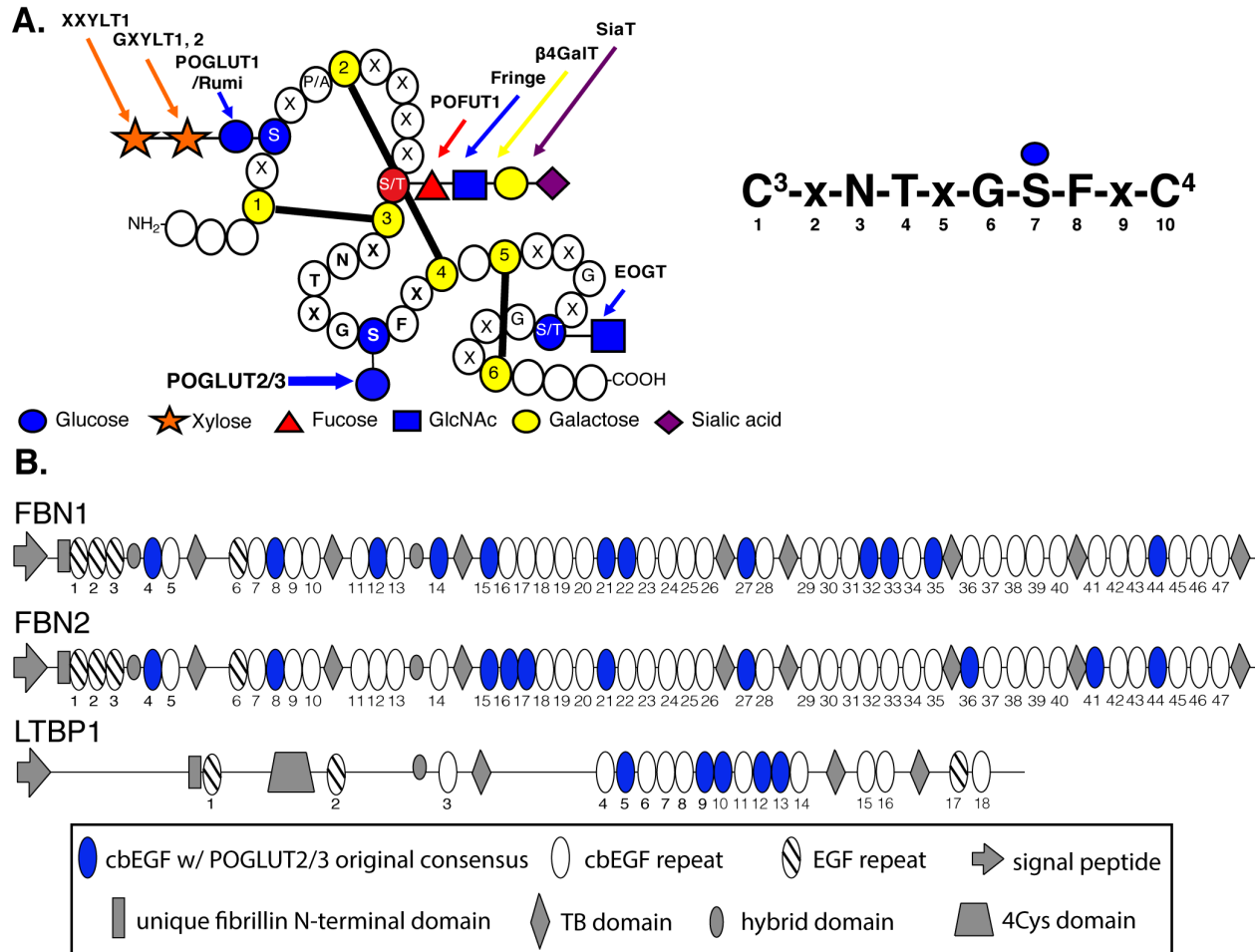
### **3.6 Data Availability**

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium *via* the PRIDE (108) partner repository with the data set identifier PXD025725.

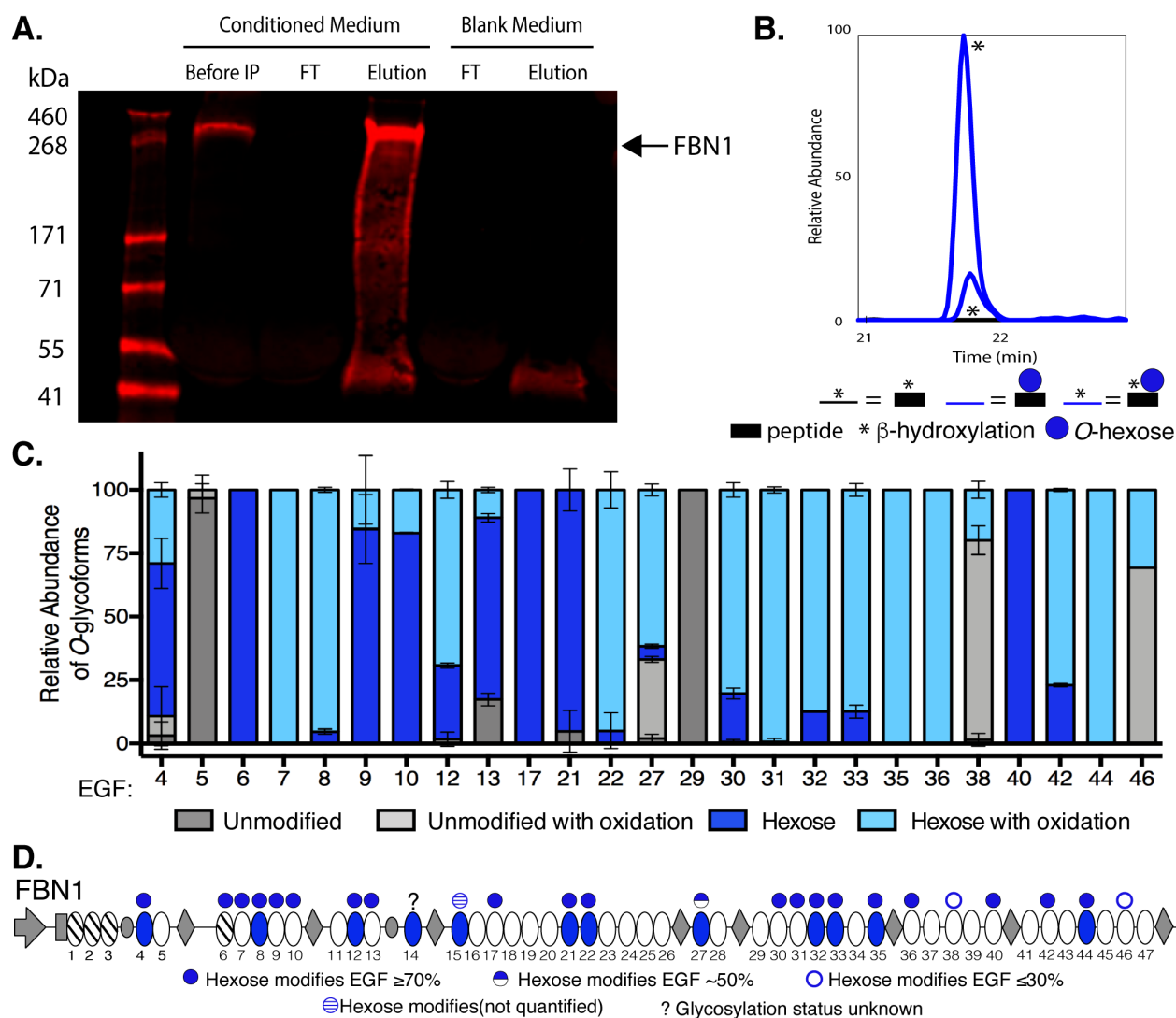
### **3.7 Acknowledgements**

Thank you to the labs of Drs. Penny Handford, Dieter Reinhardt, and Clair Baldock for providing plasmids and antibodies. We also thank members of the Holdener and Haltiwanger labs for their discussions and input.

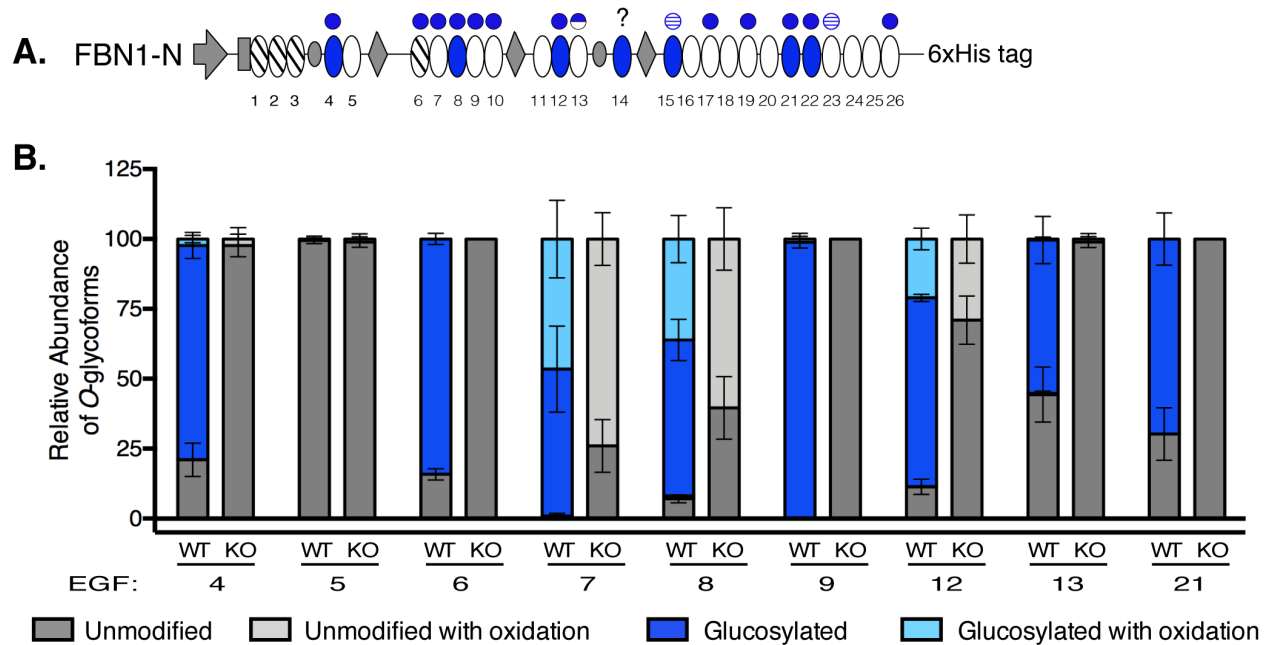
### 3.8 Main Text Figures



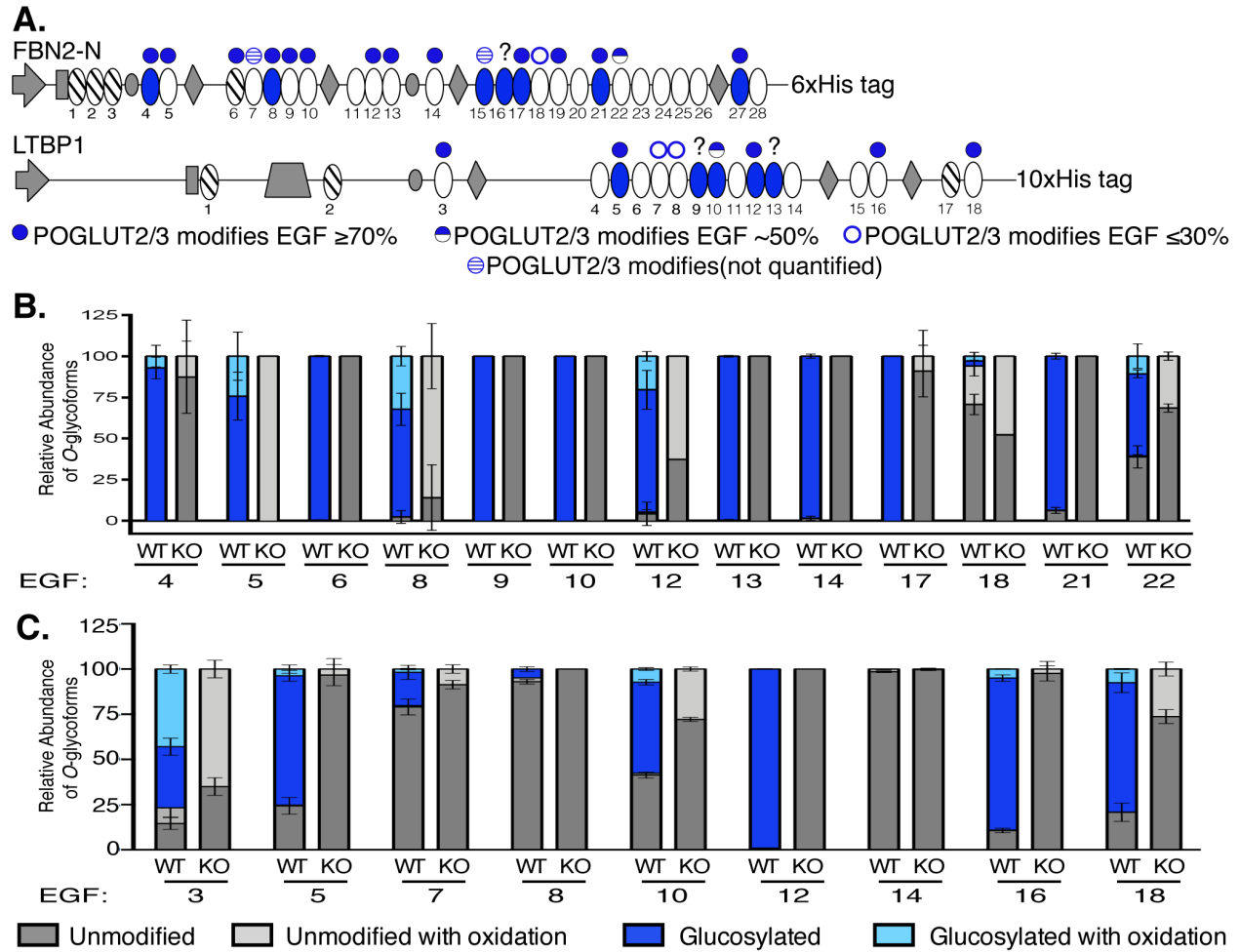
**Figure 3.1. Predicted sites of modification by POGLUT2 and/or POGLUT3 on EGF repeats of human FBN1, FBN2, and LTBP1 based on original consensus sequence.** *A* (left) A cartoon of an EGF repeat showing modification sites of several *O*-glycosyltransferases. POGLUT2 and 3 modify EGF repeats at a site distinct from POGLUT1, POFUT1, and EOGT. *Yellow*, cysteines. Adapted with permission from (1). *A* (right) The original consensus sequence with positions between cysteines 3 and 4 numbered. *B*, Domain maps of human FBN1, FBN2, and LTBP1. *Blue ovals*, EGF repeat with original POGLUT2 and 3 consensus sequence, C<sup>3</sup>-X-N-T-X-G-S-F-X-C<sup>4</sup>. *White ovals*, EGF repeat without original POGLUT2 and 3 consensus sequence.



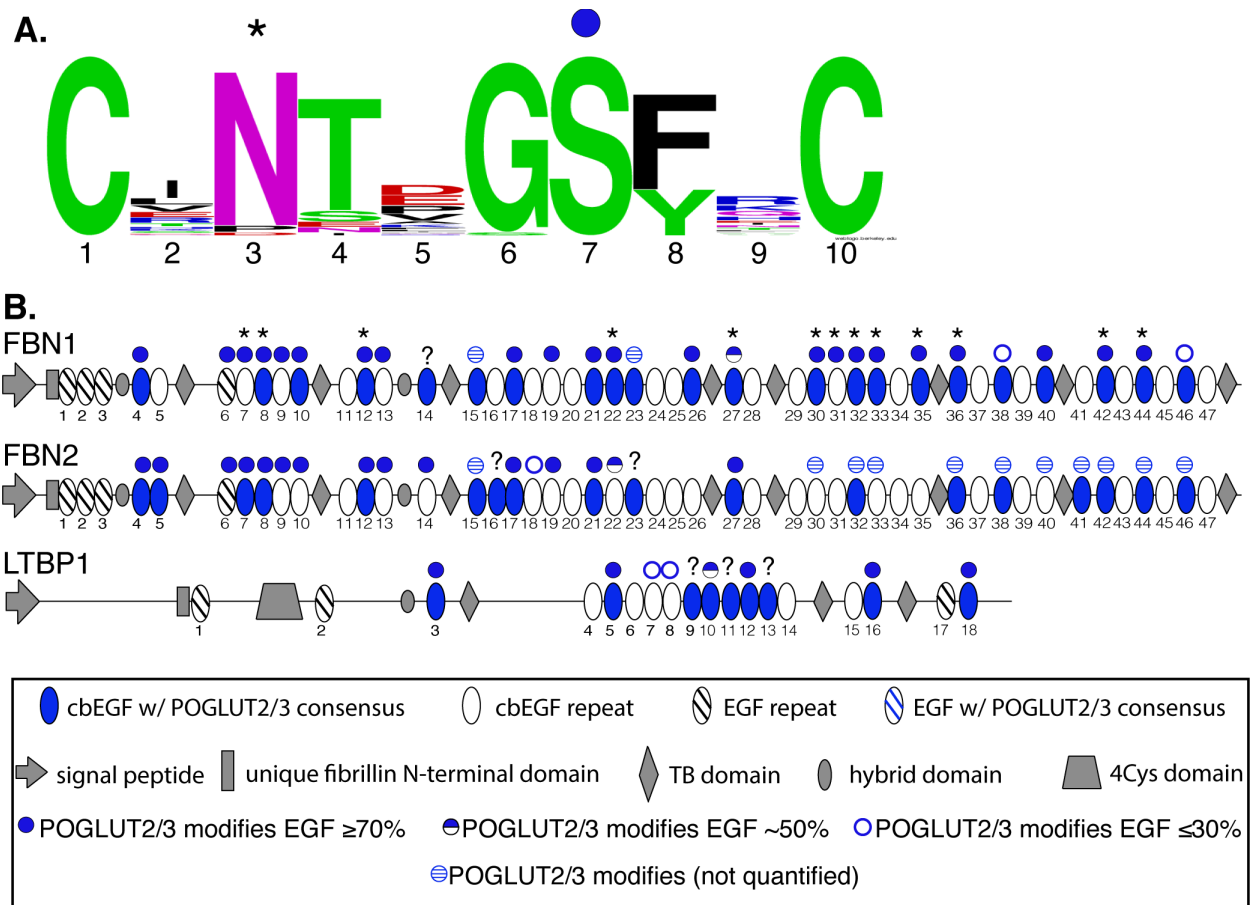
**Figure 3.2. Endogenous FBN1 is modified at high stoichiometry with hexose at more than the predicted number of EGF repeats.** *A*, Western blot illustrating successful immunoprecipitation of endogenous FBN1 from conditioned medium of human dermal fibroblasts. *Before IP*, fibroblast conditioned medium before immunoprecipitation. *FT*, flow-through after antibody application. *Elution*, human FBN1 eluted off magnetic beads. *Red channel*, anti-proline rich N-terminal region of human FBN1. Arrow marks migration position of FBN1. *B*, representative EIC of the peptide from FBN1 EGF33 modified by a hexose at high stoichiometry. EICs were generated by searching for m/z values of glycoforms of the tryptic peptide from EGF33, R.CVNTYGSYECK. *C*, Mass spectral data in Table S5. *C*, Relative quantification of hexose modification and  $\beta$ -hydroxylation on peptides derived from the indicated FBN1 EGFs averaged from three biological replicates. Error bars show  $\pm$  SD. Mass spectral data in Table S5. *D*, domain map of FBN1 illustrating predicted (based on original consensus sequence C<sup>3</sup>-X-N-T-X-G-S-F-X-C<sup>4</sup>) and confirmed sites of modification by hexose. *Blue oval*, EGF with original POGLUT2 and 3 consensus sequences. *Blue circle*, EGF confirmed modified with hexose. *?*, glycosylation status unknown due to lack of sequence coverage from mass spectral data. EGF15 is modified, but quantification was not possible because the unmodified peptide m/z is too small for the mass spectrometer to detect.



**Figure 3.3. POGLUT2 and 3 extensively O-glucosylate EGF repeats on recombinant human FBN1 EGFs 1-26 (FBN1-N) expressed in HEK293T cells.** A, schematic of recombinant protein expressed in HEK293T cells as in Figures 1B and 2D. B, Relative quantification of major O-glycoforms and  $\beta$ -hydroxylation of peptides derived from EGFs of FBN1-N overexpressed in wild-type or *POGLUT2/3* double knockout HEK293T cells. WT, wild type. KO, *POGLUT2/3* double knockout cells. Only EGFs with at least 2 replicate values are plotted in the bar graph. Error bars show +/- SD. Mass spectral data in Tables S6-S10.

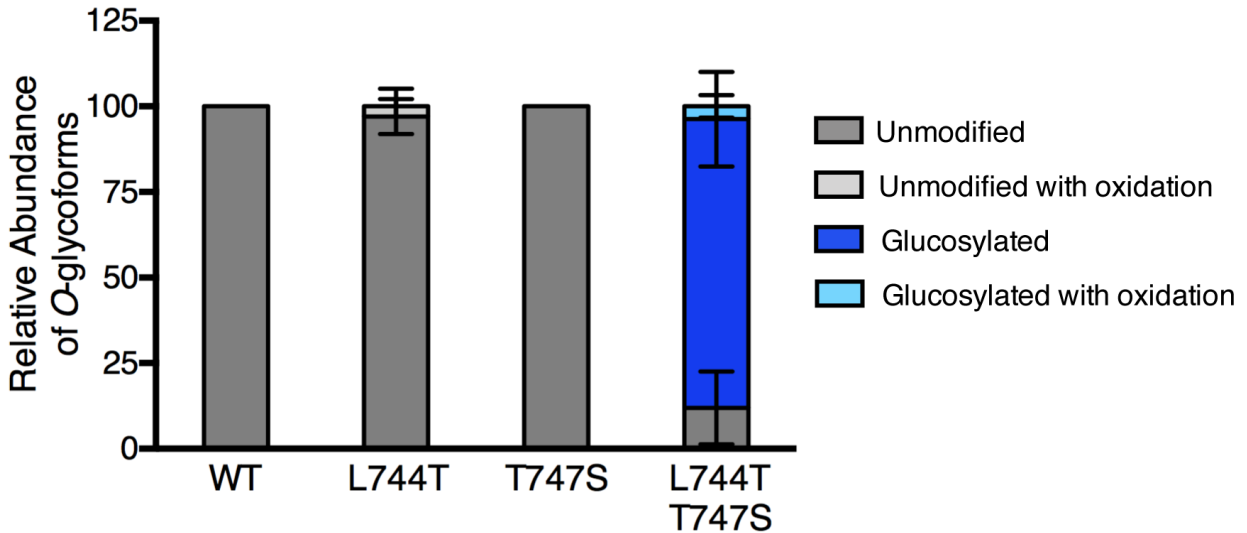


**Figure 3.4. POGLUT2 and 3 extensively O-glucosylate EGF repeats on recombinant human FBN2 EGFs 1-28 (FBN2-N) and LTBP1 in HEK293T cells.** A, schematic of recombinant FBN2-N and LTBP1 expressed HEK293T cells as in Figures 1B and 2D. Relative quantification of major O-glycoforms and  $\beta$ -hydroxylation of peptides derived from EGFs of FBN2-N (B) and LTBP1 (C) overexpressed in wild-type and *POGLUT2/3* double knockout HEK 293T cells. WT, wild type. KO, *POGLUT2/3* double knockout cells. Only EGFs with at least 2 replicate values are plotted in the bar graph. Error bars show  $\pm$  SD. Mass spectral data for FBN2-N are in Tables S13-S15, and for LTBP1 are in Tables S16-S19.

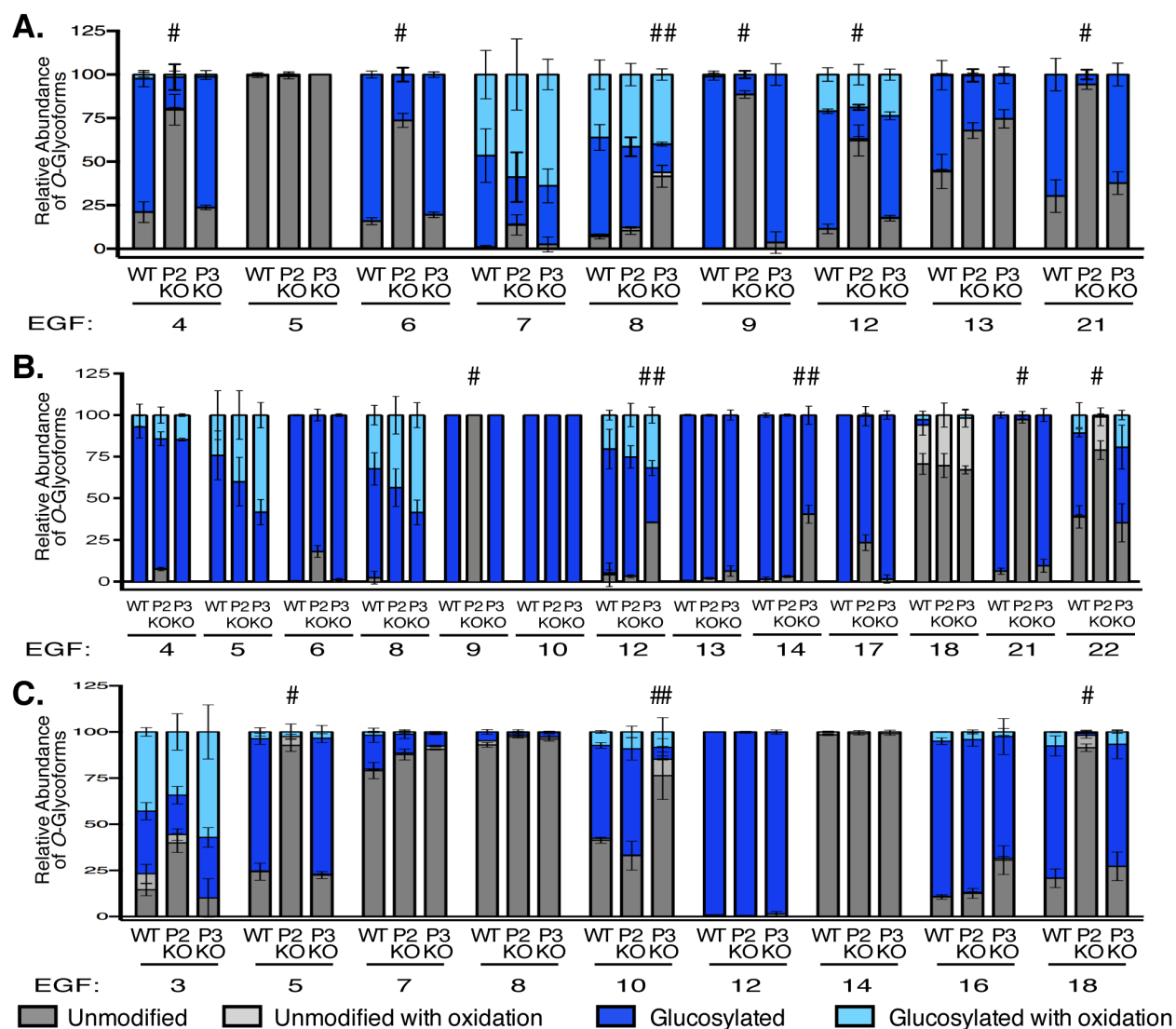


**Figure 3.5. Summary POGLUT2 and 3 sites mapped on endogenous FBN1 from dermal fibroblasts and recombinant FBN1, FBN2, and LTBP1 expressed in HEK293T cells.** *A*, WebLogo summary of all identified POGLUT2 and 3 modification sites. Numbers below amino acid indicate position within consensus sequence between cysteines three and four of an EGF repeat. \*, *POGLUT2* and/or 3 modified EGFs of endogenous FBN1 that are also  $\beta$ -hydroxylated  $>50\%$ . *Blue circle*, *O*-glucose modified site. *B*, domain maps of FBN1, FBN2, and LTBP1 illustrating predicted and confirmed sites of POGLUT2 and 3 modification. *Blue ovals*, EGF with revised consensus sequence, C<sup>3</sup>-x-N-T-x-G-S-(FY)-x-C<sup>4</sup>. \*, FBN1 EGFs that are *O*-glucosylated by POGLUT2/3 and  $\beta$ -hydroxylated more than 50%. ?, glycosylation status unknown due to lack of sequence coverage from mass spectral data.

Wild type FBN1 EGF11: C<sup>3</sup>-E-N-L-R-G-T-Y-K-C<sup>4</sup>  
 L744T FBN1 EGF11: C<sup>3</sup>-E-N-T-R-G-T-Y-K-C<sup>4</sup>  
 T747S FBN1 EGF11: C<sup>3</sup>-E-N-L-R-G-S-Y-K-C<sup>4</sup>  
 L744T+T747S FBN1 EGF11: C<sup>3</sup>-E-N-T-R-G-S-Y-K-C<sup>4</sup>

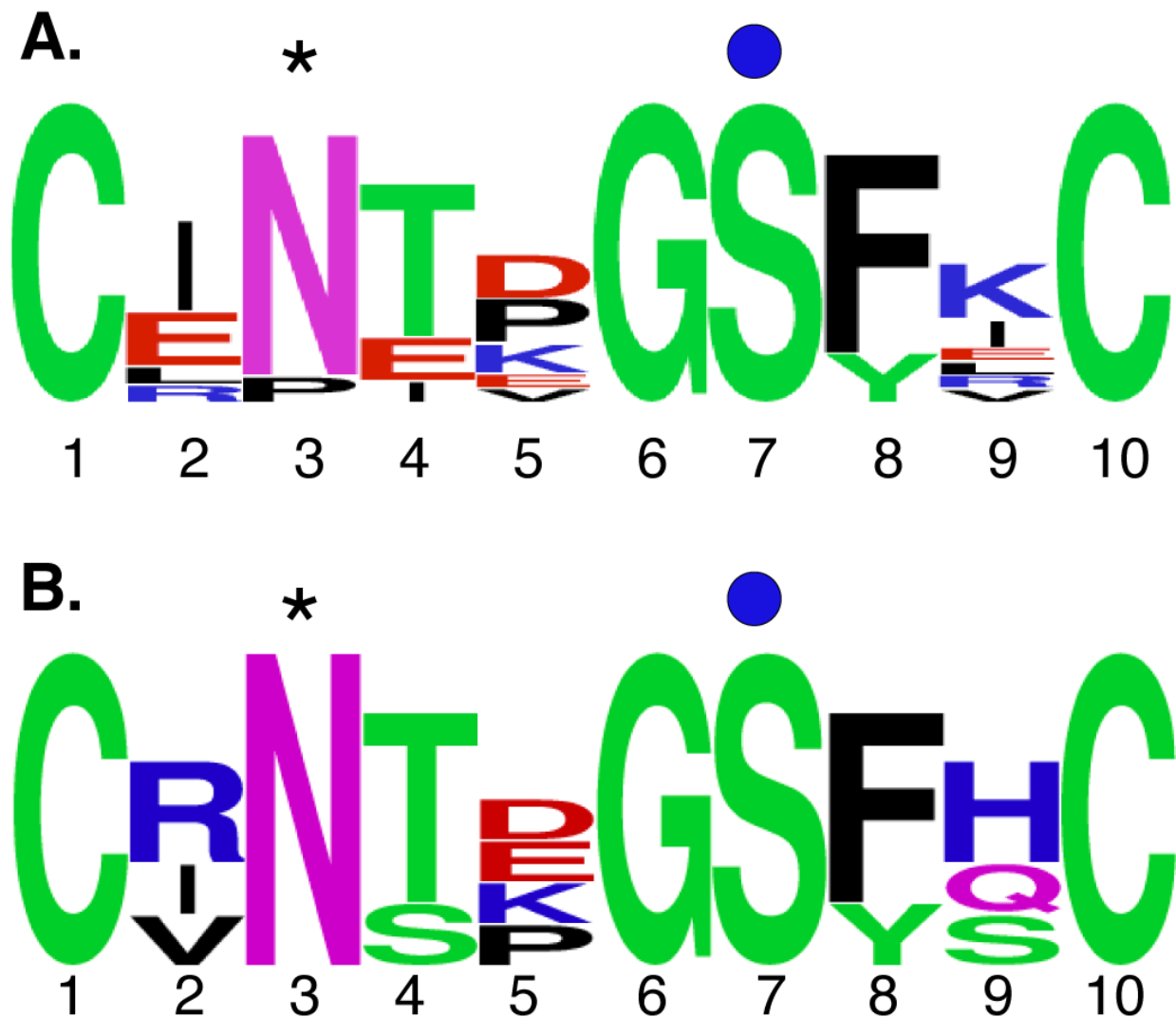


**Figure 3.6. POGLUT2 and 3 consensus sequence can be used to convert an unmodified EGF to a POGLUT2/3 modified EGF.** Wild type and mutant amino acid sequence between cysteines 3 and 4 of FBN1 EGF11. *Red underlined*, mutated sites. Relative quantification of O-glycoforms and  $\beta$ -hydroxylation of peptides from wild type and mutant FBN1 EGF11. Error bars show +/- SD. Mass spectral data are in Tables S22-S25.



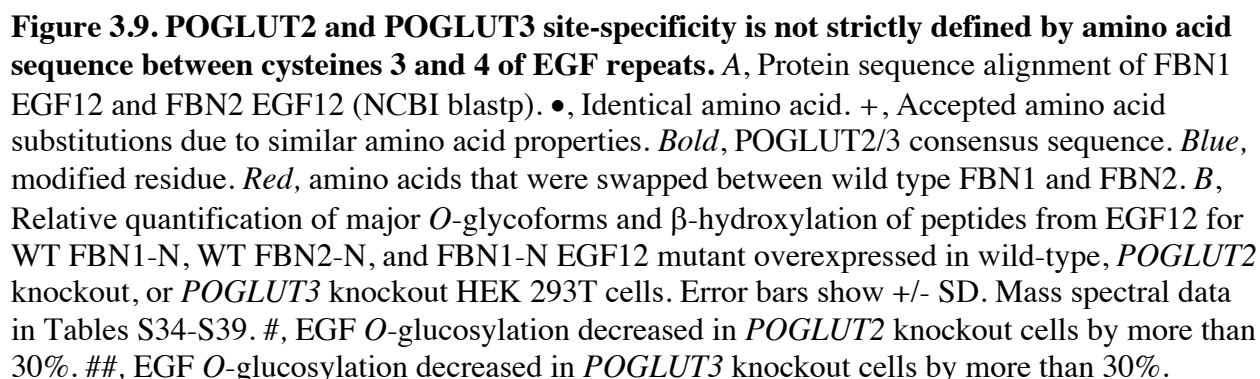
**Figure 3.7. POGLUT2 and POGLUT3 exhibit site-specificity on EGF repeats of recombinant FBN1-N, FBN2-N, and LTBP1.** Relative quantification of major O-glycoforms and  $\beta$ -hydroxylation of peptides from EGF repeats. A, FBN1-N. Mass spectral data in Tables S26-S27. B, FBN2-N. Mass spectral data in Tables S28-29. C, LTBP1. Mass spectral data in Tables S30-S33. Proteins were overexpressed in wild-type (WT), *POGLUT2* knockout (P2 KO), or *POGLUT3* knockout (P3 KO) HEK293T cells. #, O-glucosylation decreased in *POGLUT2* knockout cells by more than 30%. ##, O-glucosylation decreased in *POGLUT3* knockout cells by more than 30%. Only EGFs with at least 2 replicate values are plotted in the bar graph. Error bars show  $\pm$  SD.

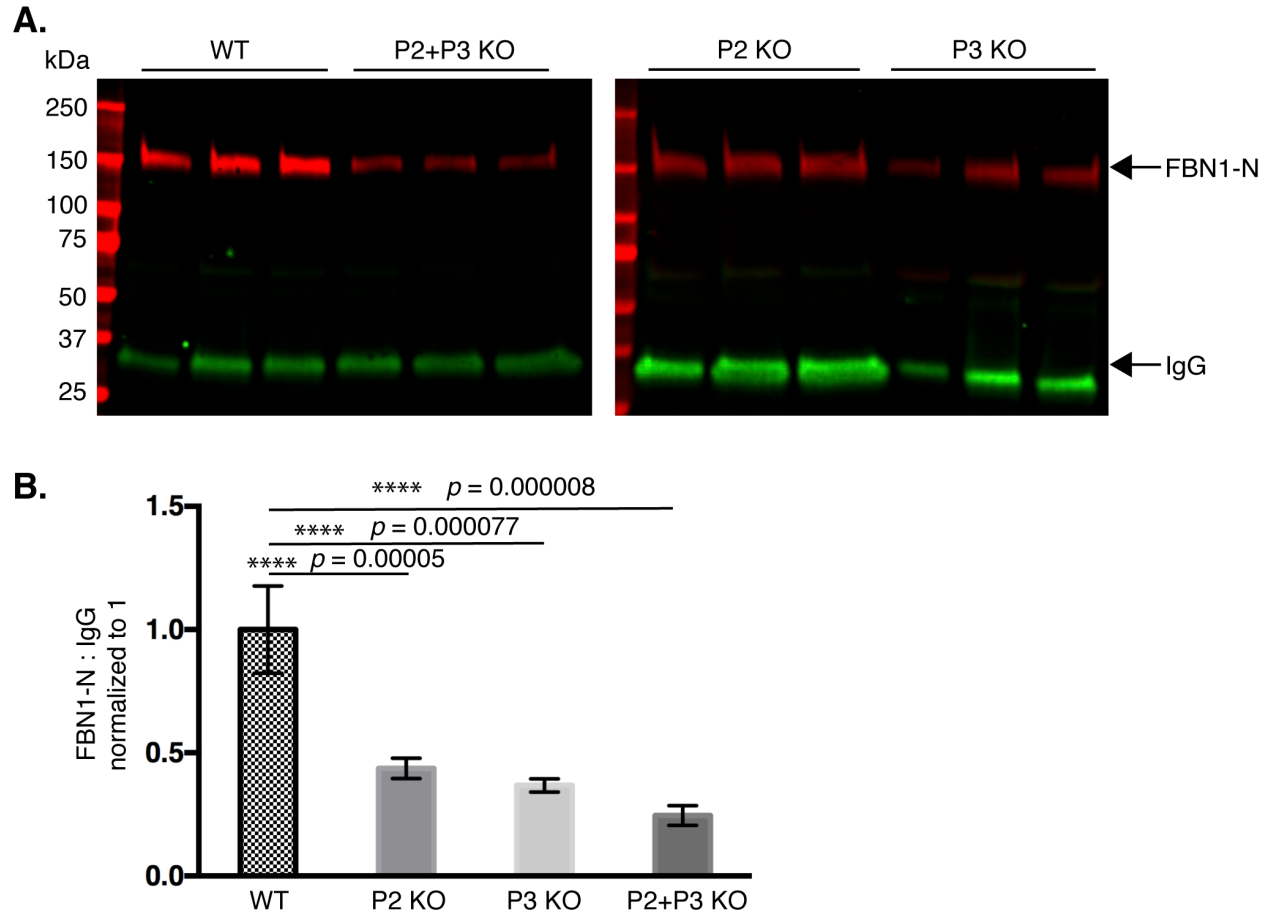




**Figure 3.8. POGLUT2 versus POGLUT3 consensus sequence.** *A*, WebLogo summarizing sequences between cysteines 3 and 4 of the EGF repeat preferred by POGLUT2 based on mass spectral analysis of FBN1-N, FBN2-N, and LTBP1. *B*, WebLogo summarizing sequences preferred by POGLUT3 based on mass spectral analysis of FBN1-N, FBN2-N, and LTBP1. \*, predicted site for  $\beta$ -hydroxylation. *Blue circle*, *O*-glucose modification site.

FBN1 EGF12: DINECVLNSLLCDNGQ**CRNTPGSFV**CTCPKGFYKPD LKTCE  
 ••+••++• ••••• ••••••+•••• ••••••+••••••+••••••+••••••  
 FBN2 EGF12: DIDECLVNRLLCDNGL**CRNTPGSYS**CTCPPGYVFR TETETCE






**Figure 3.10. Knockout of *POGLUT2* and/or *POGLUT3* leads to reduced secretion of recombinant human FBN1-N from HEK293T cells.** A, Secretion assay with hFBN1-N expressed in wild type, *POGLUT2* knockout, *POGLUT3* knockout, or *POGLUT2/3* double knockout HEK293T cells. IgG was co-transfected as a transfection and secretion control. Red channel, anti-His antibody. Green channel, anti-IgG antibody. B, Quantification of Western blot replicates. FBN1-N signal was normalized to IgG signal. One-way ANOVA in Excel was used to calculate significance. \*\*\*\*,  $p < 0.0001$ . Three biological replicates were performed, and three technical triplicates were performed on wild type and double knockout samples. Two biological replicates and three technical triplicates were performed on single knockout samples. Error bars show  $\pm$  SD.

### 3.9 Supplemental Tables and Figures

\*Tables S3.4-3.40 and Figures S3.3-3.64 can be found in Supporting Information of the original manuscript (36).

**Table S3.1: FBN1, FBN2, and LTBP1 EGFs with variable residues at position 4 of the POGLUT2 and 3 consensus sequence.** Sequence at the bottom is a WebLogo summary for EGFs in the table with low or no modification by POGLUT2 and 3.

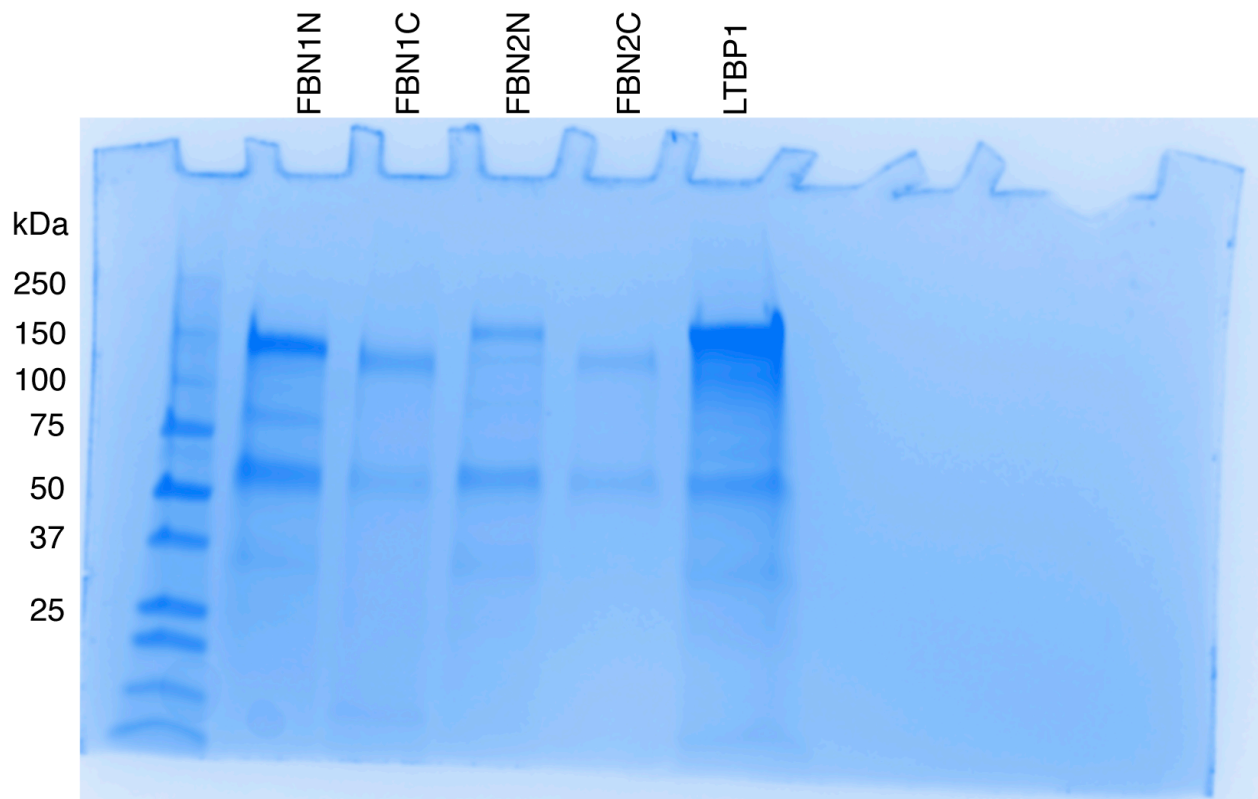
Protein	EGF repeat	Sequence C <sup>3</sup> -C <sup>4</sup>	O-glucosylation Efficiency
<b>FBN-1</b>	7	CIN <b>N</b> QGSYTC	High
	9	CIN <b>E</b> DGSFKC	High
	13	CKN <b>S</b> PGSFIC	High
	29	CIN <b>M</b> VGSFRC	unmodified
	34	CQN <b>L</b> DGSYRC	unmodified
<b>FBN-2</b>	9	CIN <b>E</b> DGSFKC	High
	10	CIN <b>S</b> EGSFRC	High
	13	CRN <b>N</b> LGSFNC	High
	14	CVN <b>S</b> KGSFHC	High
	19	CTN <b>S</b> EGSYEC	High
	22	CLN <b>I</b> PGSFKC	medium
<b>LTBP1</b>	7	CVN <b>S</b> PGSYQC	low
	8	CSN <b>L</b> EGSYMC	low
	14	CEN <b>V</b> EGSFLC	unmodified
Consensus Sequence:			

**Table S3.2. POGLUT2 versus POGLUT3 preferred modification sites on FBN1, FBN2, and LTBP1.**

<b>POGLUT2-preferred sites</b>	<b>Protein</b>	<b>EGF</b>	<b>Sequence (C<sup>3</sup>-C<sup>4</sup>)</b>
	FBN1	4	CINTVGSFEC
		6	CIPTPGSYRC
		9	CINEDGSFKC
		12	CRNTPGSFVC
		21	CENTKGSFIC
	FBN2	9	CINEDGSFKC
		21	CENTKGSFIC
		22	CLNIPGSFKC
	LTBP1	5	CENTEGSFLC
		18	CINTDGSYKC
<b>POGLUT3-preferred sites</b>	FBN1	8	CINTDGSFHC
	FBN2	12	CRNTPGSYSC
		14	CVNSKGSFHC
	LTBP1	10	CRNTEGSFQC

**Table S3.3. Primers used for site-directed mutagenesis of FBN1-N and FBN2-N.**

Template	Mutation	Direction	Sequence	T <sub>m</sub>
rhFBN1-N	F788Y V789S	FWD	5'-TGGAAGTTATAGCTGTACCTGCCCCAAGGGATTATC-3'	74 C
		REV	5'-AGGTACAGCTATAACTTCCAGGAGTATTTCTACATTGTCCA-3'	73.7 C
rhFBN2-N	Y832F S833V	FWD	5'-AGGAAGTTTCGTCTGTACGTGCCACCAGGGTAT-3'	74.5 C
		REV	5'-ACGTACAGACGAACTTCCTGGCGTGTTCGGCA-3'	74.5 C
rhFBN1-N	L744T	FWD	5'-GTGAAAACACTCGTGGGACCTATAAATGTATATGCAATTCAGGATATGAAG-3'	78 C
		REV	5'-GGTCCCACGAGTGTTTTTCACAGATTCATTTGGGCAATATC-3'	77.3 C
rhFBN1-N	T747S	FWD	5'-CTTCGTGGGTCCTATAAATGTATATGCAATTCAGGATATGAAGTG-3'	79.8 C
		REV	5'-CATTATAGGACCCACGAAGGTTTTTCACAGATTCATTTGG-3'	79.6 C
rhFBN1-N L744T	L744T + T747S	FWD	5'-ACTCGTGGGTCCTATAAATGTATATGCAATTCAGGATATGAAGTG-3'	79.8 C
		REV	5'-CATTATAGGACCCACGAGTGTTTTTCACAGATTCATTTGG-3'	79.6 C

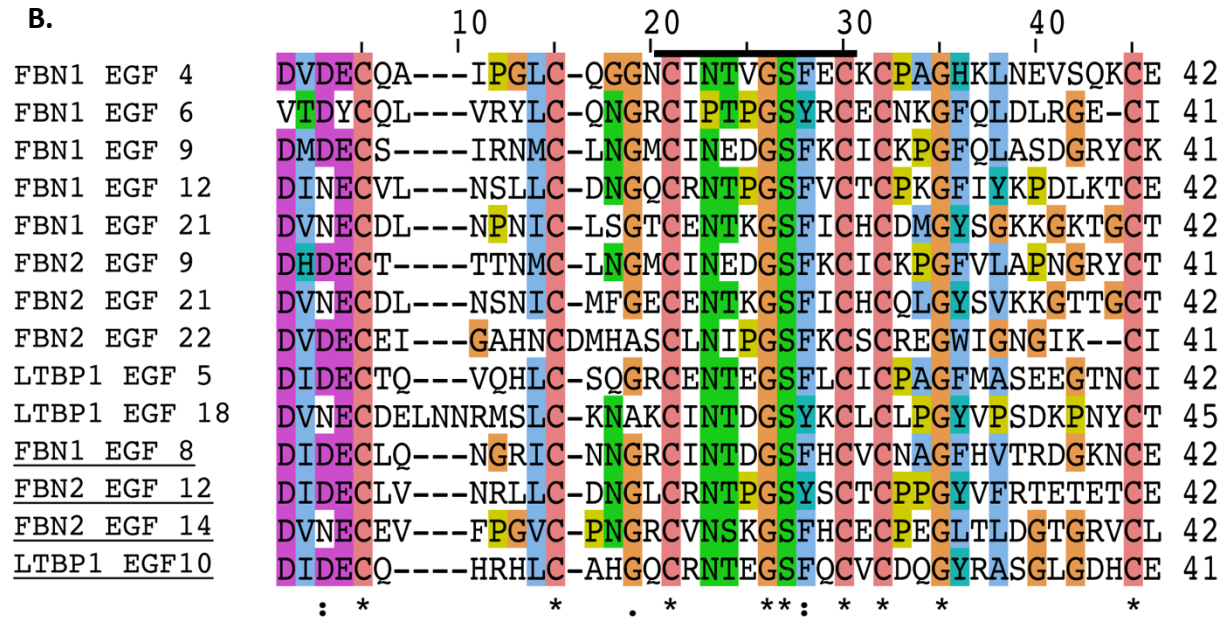


**Figure S3.1. Purification of recombinant FBN1-N, FBN1-C, FBN2-N, FBN2-C, and LTBP1 expressed in HEK293T cells.** Purified proteins were separated on 4-20% gradient gel under reducing conditions followed by Gel Code Blue staining (Invitrogen). Approximate, predicted molecular weights of each protein: FBN1-N (165 kDa), FBN1-C (140 kDa), FBN2-N (186 kDa), FBN2-C (137 kDa), and LTBP1 (187 kDa). Band between 50 and 75 kDa is bovine serum albumin leftover from culture medium supplemented with bovine calf serum.

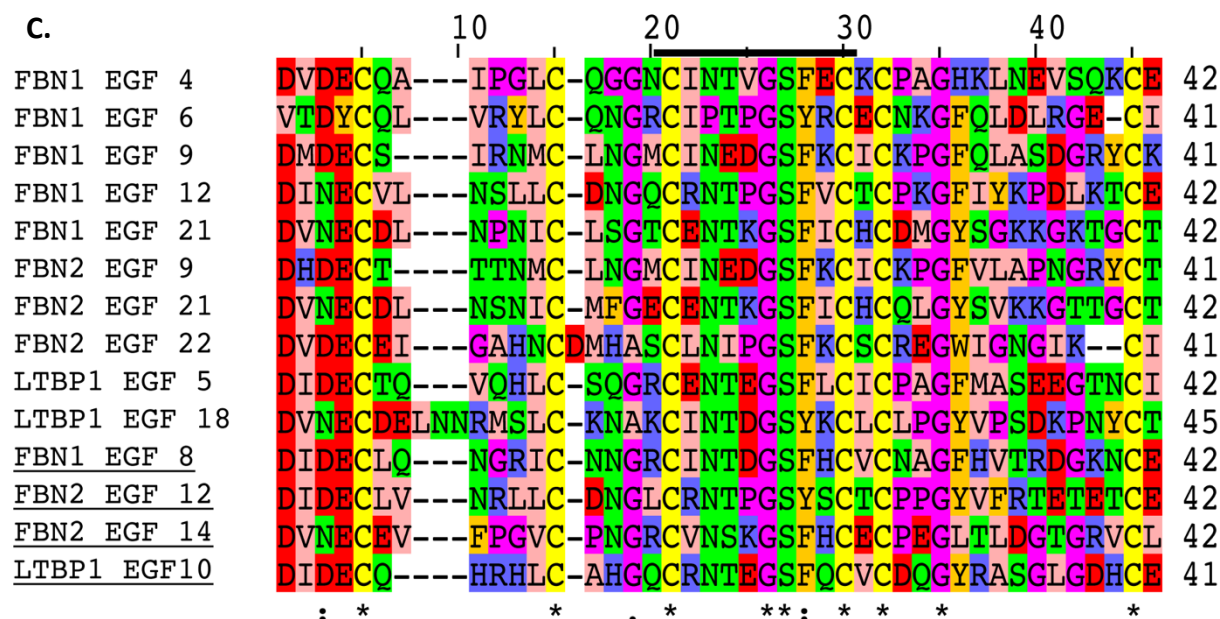
**A.**

			10	20	30	40	
FBN1	EGF	4	DVDECQA---	IPGLC-QGGNCINTVGSFECKCPAGHKLNEVSQKCE	42		
FBN1	EGF	6	VTDYCQL---	VRYLK-QNGRCIPTPGSYRCECNKGFQLDLRGE-CI	41		
FBN1	EGF	9	DMDECS----	IRNMC-LNGMCINEDGSFKCICKPGFQLASDGRYCK	41		
FBN1	EGF	12	DINECVL---	NSLLC-DNGQCRNTPGSFVCTCPKGFIIYKPDCLKTCE	42		
FBN1	EGF	21	DVNECDL---	NPNIC-LSGTCENTKGSFICHCDMGYSGKKGKTGCT	42		
FBN2	EGF	9	DHDECT----	TTNMC-LNGMCINEDGSFKCICKPGFVLAPNGRYCT	41		
FBN2	EGF	21	DVNECDL---	NSNIC-MFGECENTKGSFICHCQLGYSVKKGTTGCT	42		
FBN2	EGF	22	DVDECEI---	GAHNCDMHASCLNIPGSFKCSCREGWIGNGIK--CI	41		
LTBP1	EGF	5	DIDECTQ---	VQHLC-SQGRCENTEGSFLCICPAGFMASEEGTNCI	42		
LTBP1	EGF	18	DVNECDELNNRMSLC-	KNAKCINTDGSYKCLCLPGYVPSDKPNYCT	45		
<u>FBN1</u>	<u>EGF</u>	<u>8</u>	DIDECLO---	NGRIC-NNGRCINTDGSFHCVCNAGFHVTRDGKNCE	42		
<u>FBN2</u>	<u>EGF</u>	<u>12</u>	DIDECLO---	NRLLC-DNGLCRNTPGSYSCTCPPGYVFRTEETETCE	42		
<u>FBN2</u>	<u>EGF</u>	<u>14</u>	DVNECEV---	FPGVC-PNGRCVNSKGSFHCECPEGLTLDGTGRVCL	42		
<u>LTBP1</u>	<u>EGF</u>	<u>10</u>	DIDECQ----	HRHLC-AHGQCRNTEGSFQCVCDQGYRASGLGDHCE	41		
			:	*	*	.	*
					**	:	*
					*	*	*
					*	*	*





Clustal X Default Colouring			
Category	Colour	Residue at position	{ Threshold, Residue group }
Hydrophobic	BLUE	A,I,L,M,F,W,V	{>60%, WLVI MAFCHP}
		C	{>60%, WLVI MAFCHP}
Positive charge	RED	K,R	{>60%,KR},{>80%, K,R,Q}
Negative charge	MAGENTA	E	{>60%,KR},{>50%,QE},{>85%,E,Q,D}
		D	{>60%,KR},{>85%, K,R,Q},{>50%,ED}
Polar	GREEN	N	{>50%, N},{>85%, N,Y}
		Q	{>60%,KR},{>50%,QE},{>85%,Q,E,K,R}
		S,T	{>60%, WLVI MAFCHP},{>50%, TS},{>85%,S,T}
Cysteines	PINK	C	{>85%, C}
Glycines	ORANGE	G	{>0%, G}
Prolines	YELLOW	P	{>0%, P}
Aromatic	CYAN	H,Y	{>60%, WLVI MAFCHP},{>85%, W,Y,A,C,P,Q,F,H,I,L,M,V}
Unconserved	WHITE	any / gap	If none of the above criteria are met



Aliphatic/hydrophobic	ILVAM
Aromatic	FWY
Positive	KRH
Negative	DE
Hydrophilic	STNQ
conformationally special	PG
Cysteine	C

**Figure S3.2. Clustal Omega protein sequence alignment of POGLUT2 versus POGLUT3 preferred EGF modification sites.** EGFs with POGLUT3 preferred sites are underlined. Bold line above section of alignment indicates position of POGLUT2 and 3 consensus sequence. \*, single, fully conserved residue. : (colon), conserved amino acids of strongly similar properties. . (period), conserved amino acids of weakly similar properties. *A*, raw alignment. *B*, alignment with Clustal X default colors. *C*, alignment with Zappo colors. Jalview was used for color schemes.

## CHAPTER 4

POGLUT2 AND POGLUT3 *O*-GLUCOSYLATION OF FBN1 AND FBN2 MAINTAINS THE STRUCTURE AND FUNCTION OF THE EXTRACEULLAR MATRIX AND IS REQUIRED FOR PROPER TGF- $\beta$  SIGNALING AND DEVELOPMENT OF THE LUNG<sup>1</sup>

<sup>1</sup>Williamson DB, Neupane S, Grady RC, Holdener BC, Haltiwanger RS. Manuscript in preparation. To be submitted to Developmental Cell.

#### **4.1 Abstract**

*Poglut2* and *Poglut3* encode two Protein *O*-glucosyltransferases (POGLUT) that modify epidermal growth factor (EGF)-like repeats containing the putative consensus sequence C<sup>3</sup>-N-T-x-G-S-F/Y-C<sup>4</sup>. Based on this putative consensus sequence, POGLUT2 and 3 are predicted to modify up to 56 proteins, many of which are involved in the extracellular matrix (ECM). To date, we have identified fibrillin-1 (FBN1), fibrillin-2 (FBN2), and latent transforming growth factor  $\beta$  binding protein-1 (LTBP1) as the proteins most heavily modified by POGLUT2 and 3. Deletion of *POGLUT2* and *POGLUT3* in HEK293T cells reduces secretion of human FBN1, but how loss of POGLUT2 and 3 affects mammalian development is unknown. Here we show that *Poglut2* and *Poglut3* double knockout mice died perinatally and had defects in lung, vascular, and limb development. Blocking POGLUT2/3-mediated *O*-glucosylation significantly reduced FBN1, FBN2, and elastin levels in developing lung with concomitant reduction of TGF- $\beta$  signaling. Glycoproteomic analyses confirmed that EGFs of predicted POGLUT2/3 substrates FBN1, fibulin-2 (FBLN2), fibulin-5 (FBLN5), and nidogen-1 (NID1) were modified *in vivo*. Moreover, double knockout of *Poglut2* and *Poglut3* completely eliminated the *O*-glucose modification. Proteomic analyses revealed loss of POGLUT2 and 3 reduced the abundance of FBN1 found in fibroblast media and affected the abundance of other ECM-related proteins. These data highlight *O*-glucosylation by POGLUT2 and 3 impact the structure and function of the ECM through its effects on POGLUT2 and 3 substrates FBN1 and FBN2, leading to deleterious developmental defects through dysregulation of TGF- $\beta$  signaling. Future work will elaborate on this mechanism to understand the specific role POGLUT2 and 3-mediated *O*-glucosylation has in the ECM environment.

## **4.2 Introduction**

Protein *O*-glycosylation of EGF repeats is required for proper mammalian development (21, 109). Deletion of *O*-glycosyltransferases that modify EGF repeats leads to embryonic lethality in mice (11, 12), and mutations in these glycosyltransferases result in human diseases (21, 109). These findings focus on Protein *O*-Fucosyltransferase 1 (POFUT1) and Protein *O*-Glucosyltransferase 1 (POGLUT1), which both modify Notch EGF repeats. Loss of these enzymes leads to Notch phenotypes associated with dysregulation of Notch signaling, setting a precedent for the physiological impact of *O*-glycosylation of EGF repeats on protein function. POGLUT2 and 3 are two recently identified glycosyltransferases that add *O*-glucose to a serine residue on an EGF repeat between cysteines 3 and 4 at the putative consensus sequence C<sup>3</sup>-N-T-x-G-S-F/Y-C<sup>4</sup> (36) (see Figure 2.1). POGLUT2 and 3 seem functionally redundant based on cell culture experiments, though each enzyme exhibits site-specificity for certain EGF repeats (36). This suggests these enzymes may deviate from one another physiologically. Fifty-six proteins have POGLUT2 and 3 modification sites based on the consensus sequence, and of these, at least 41 proteins function in the extracellular matrix (ECM) (36). The fibrillin (FBN) and latent transforming growth factor  $\beta$ -binding proteins (LTBPs) are predicted to contain the greatest number of POGLUT2/3-modified EGFs, and we confirmed that approximately half of the EGF repeats of FBN1, -2, and LTBP1 are modified by POGLUT2 and 3 (36).

FBN1 and FBN2 are major structural components of microfibrils in the ECM (100). These microfibrils act as a scaffold for elastogenesis, which is essential for proper development of tissues like the lungs, heart, and skin (110, 111). FBN microfibrils also serve as a reservoir of inactive TGF- $\beta$  through interactions with LTBPs (42). FBN and LTBP function are required for efficient secretion, activation, and regulation of TGF- $\beta$  signaling (112). In humans, Marfan

syndrome (MFS), caused by *FBNI* mutations, is one of the most extensively studied connective tissue disorders. The pathological features of MFS, which commonly manifest in the heart and lungs, are attributed to dysregulation of TGF- $\beta$  signaling (49, 56, 57). Similarly, *LTBP* genetic variants are also associated with dysregulation of TGF- $\beta$  signaling (44, 112). In humans, *LTBP* mutations lead to bone growth defects and developmental defects in the aorta, lungs, eyes, and muscles, and these phenotypes have been recapitulated in mice (44, 112). Additionally, loss of *Fbn1*, *Ltbp1*, -3, or -4 leads to premature lethality in mice due to aortic dissection (46, 47, 113, 114). These studies highlight the biological significance of FBNs and LTBP and their interactions with one another, and it brings to the forefront the need to understand the physiological role of POGLUT2 and 3 *O*-glucosylation on these proteins.

*O*-glycosylation stabilizes EGF repeats and promotes secretion of proteins in which they occur (10). Notably, *O*-glycans on EGF repeats of NOTCH1 are also involved in ligand-receptor interactions that are required for efficient binding and signaling (28). Based on these observations, we hypothesized that loss of POGLUT2 and 3-mediated *O*-glucosylation could impact both the efficiency of substrate secretion as well as protein interactions in the extracellular space. To test this hypothesis, we generated *Poglut2* and *Poglut3* mouse knockouts (Supplemental Figures 4.1 and 4.2) and evaluated the impact of the mutations in double knockout (*Poglut2/3 DKO*) animals. Loss of POGLUT2/3-mediated *O*-glucosylation caused neonatal death likely resulting from defects in cardiovascular and lung development. In developing lung, we detected abnormal FBN distribution with abnormalities in elastogenesis and TGF- $\beta$  signaling. Mass spectral analyses identified more global alterations in ECM composition including proteins not predicted to be POGLUT2 and 3 substrates. The results from these studies suggest that POGLUT2/3-mediated *O*-glucose modifications were important not only for

optimum secretion of substrates, but also raised the possibility that these glycans were required for function of POGLUT2/3 substrates in the extracellular space and maintenance of a normal ECM environment.

### **4.3 Results**

#### **Loss of both *Poglut2* and 3 caused perinatal lethality, reduced size, and syndactyly in mice**

To assess the physiological function of POGLUT2 and 3, we generated *Poglut2* and 3 mouse models (Figures S4.1 and S4.2, Tables S4.1 and S4.2). The *Poglut2 tm1d* allele was the result of deletion of exons 3-4, and the *Poglut3 tm1d* allele resulted from the deletion of exons 2-3. The *tm1d* allele for both *Poglut2* and 3 was used to generate single knockout (*SKO: Poglut2 KO/KO; Poglut3 WT/WT* and *Poglut2 WT/WT; Poglut3 KO/KO*) and double knockout (*DKO: Poglut2 KO/KO; Poglut3 KO/KO*) animals. *Poglut2* and *Poglut3 SKOs* were viable and fertile with no obvious phenotypes and present at expected Mendelian ratios (Table S4.3-4.4). *Poglut2/3 DKO* mice were statistically underrepresented at weaning and significantly smaller than littermates (Figure 4.1A-B, Tables 4.1 and S4.5). *Poglut2/3 DKOs* had fully penetrant syndactyly between digits three and four of one or more hind- and/or forelimbs (Figures 4.1C, S4.3). Other genotypes also significantly deviated from the Mendelian ratio, such as *Poglut2 WT/KO; Poglut3 WT/KO* and *Poglut2 WT/WT; Poglut3 KO/KO* (Table S4.5), but more work needs to be done to explain these observations. In contrast, at embryonic day (E) 18.5 *Poglut2/3 DKOs* were present at expected Mendelian ratio (Table 4.1), suggesting *Poglut2/3 DKO* causes lethality soon after birth (within 3 days of birth). It was difficult to pinpoint an exact day when lethality occurred. In some cases, deceased *Poglut2/3 DKO* pups were found in cages (ages P1-P3 based off lack of fur), but in most instances they were not observed. *Poglut2/3 DKO* mice

were also significantly smaller than littermates at E18.5 (Figure 4.1D). These observations are consistent with previous *Fbn* and *Ltbp* mouse models. *Fbn1* knockout mice die around two weeks after birth, and *Ltbp1* and -4 single knockout mice exhibit perinatal or neonatal death, respectively (46, 115). Mice with *Fbn2* genetic variants have varying degrees of syndactyly (116), and *Ltbp3* and -4 single knockout mice are decreased in size compared to controls (114, 115, 117). The overlapping of these phenotypes between *Poglut2/3* double knockouts and *Fbn* and *Ltbp* mice suggests loss of POGLUT2/3-mediated *O*-glucosylation inhibits the normal function of FBNs and LTBP.

***O*-glucosylation of EGF repeats by POGLUT2 and 3 was completely lost in *Poglut2/3* DKO mice**

We performed mass spectral glycoproteomic analyses to confirm that deletion of *Poglut2/3* in DKO mice resulted in loss of *O*-glucosylation of POGLUT2 and 3 substrates. Here, we analyzed POGLUT2 and 3 *O*-glucosylation of several secreted EGF domain-containing proteins from wild type and *Poglut2/3* DKO mouse lung primary fibroblast cultures. Medium was collected after three days and analyzed by mass spectrometry as described in Methods. A subset of the most abundant proteins predicted to be modified by POGLUT2 and 3 were analyzed for glycosylation. All glucosylated peptides were identified using Byonic. Extracted ion chromatograms (EICs) were generated for glucosylated, glucosylated plus  $\beta$ -hydroxylated, unmodified, and unmodified plus  $\beta$ -hydroxylated forms of each peptide. EGFs of FBN1, FBLN2, FBLN5, and NID1 had expected POGLUT2/3 mediated *O*-glucose modifications based on the presence of the consensus sequence, and most of the EGF repeats detected were modified at high stoichiometry (Figure 4.2 A-D). Loss of POGLUT2 and 3 (-/-) completely abolished *O*-



glucosylation of these EGF repeats between C<sup>3</sup>-C<sup>4</sup> (Figure 4.2 A-D). FBLN2, FBLN5, and NID1 were previously not known to be *O*-glucosylated by POGLUT2 and 3. Other secreted proteins such as FBN2 and LTBPs are predicted to be modified by POGLUT2 and 3 based on the putative consensus sequence (36), but they were detected at a lower abundance compared to FBN1, FBLN2, FBLN5, and NID1. Human FBN2 and LTBP1 are heavily *O*-glucosylated by POGLUT2 and 3 (36), but the limited detection of these proteins from mouse fibroblast media prevented quantification of peptides containing the POGLUT2 and 3 putative consensus sequence. Two FBN1 EGF repeats not previously identified were modified by Protein *O*-Fucosyltransferase 1 (POFUT1) (Figure 4.2E). The relative amount of *O*-fucosylation did not change between wild type and *Poglut2/3 DKO* animals (Figure 4.2E). Combined, these observations suggested that no other enzymes can modify the POGLUT2/3 consensus and that loss of *Poglut2* and 3 in mice did not affect other forms of *O*-glycosylation.

#### ***Poglut2/3 DKO causes abnormal lung blood vessel, airway, and alveoli development***

To determine whether defects in lung development contributed to perinatal death in *Poglut2/3 DKO*s, we compared histological sections of E18.5 lung sections from wild type and *Poglut2/3 DKO* lungs (Figures. 4.3-4.5). At E18.5, developing lungs are at the saccular stage (118). In control lungs, the blood vessels were comprised of well-defined inner intima (endothelium), media, and external adventitia layers that were separated by internal and external elastic lamina (Figure 4.3A) (119). Terminal bronchioles had very distinct mucosal (epithelium and lamina propria), smooth muscle and adventitia layers, where the arrangement of epithelium and lamina propria create the characteristic epithelial folds of terminal bronchiole (Figure 4.4A, B) (118, 120). The saccules at this time are the most distal aspects of the lung airway and were

comprised of an airway epithelial layer consisting of alveolar epithelial type I (AEC1) and type II (AEC2) cells supported by outer fibroblast cells, where some AEC2 cells have direct contact with AEC1 cells and others remain isolated in the alveolar chord (Figure 4.5A) (121). In contrast to controls, *Poglut2/3 DKO* blood vessels layers showed poorly arranged endothelium, loose media and aberrant internal and external lamina (Figure 4.3C). The terminal bronchioles of *Poglut2/3 DKO* lungs had flatter epithelial folds and reduced lamina propria (Fig. 4.4C), and the saccules appeared enlarged compared to controls (Figure 4.5C). These abnormalities likely resulted from aberrant trafficking or function of POGLUT2/3 substrates.

***POGLUT2/3 mediated O-glucose modification of EGF repeats was essential for maintaining the level and distribution FBN1, FBN2, and elastin in developing lung***

We used immunohistochemistry to evaluate the distribution of FBN1 and FBN2 and their colocalization with elastin in developing lung of wild type and *Poglut2/3 DKO*s. In E18.5 control embryos, FBN1 and FBN2 were characteristically localized in blood vessels, terminal bronchiole, and developing alveoli saccules (Figures 4.6-4.8). In the lung blood vessels, FBN1, FBN2, and Alexa-fluor 633 (considered a stain for elastin) colocalized to the distinct layers of the internal and external elastic lamina (Fig. 4.6 A-C), and diffuse FBN staining which did not colocalize with elastin was detected in media and adventitia layers of the blood vessels (Fig. 4.6 A-C). In *Poglut2/3 DKO* developing lung blood vessels FBN1, FBN2, and elastin levels were significantly reduced (Figure 4.6D-F, G-I). In addition, FBN1 and FBN2 were nearly absent from the internal elastic lamina and staining in the external lamina was not contiguous. Consistent with the role of FBNs networks in laying the foundation for elastic fiber formation

(41, 46, 122), Alexa-fluor 633 was significantly reduced in *Poglut2/3 DKO*s in the region of the internal elastic lamina (Figure 4.6F and I).

In wild-type terminal bronchiole, strong FBN1 and FBN2 staining was detected in the basement membrane underlying the characteristically folded epithelial layer and smooth muscle layers with reduced staining in the lamina propria (Figure 4.7A-B). Strong Alexa-fluor 633 staining colocalized to the epithelial layer basement membrane with moderate staining underlying the smooth muscle layer (Figure 4.7C). In contrast, in the flattened *Poglut2/3 DKO* terminal bronchioles, FBN1, FBN2, and Alexa-fluor 633 levels were significantly reduced and staining appeared punctate and disorganized compared to controls (Figure 4.7D-I).

In the developing saccules of the alveolar region of the airway, FBN1, FBN2, and Alexa-fluor 633 staining localized to the matrix underlying the airway epithelial cells (Figure 4.8 A-C). In the developing saccules of *Poglut2/3 DKO* lungs, FBN1, FBN2, and elastin levels were significantly reduced (Figure 4.8 D-I), and FBN2 networks appeared discontinuous. The reduced complexity of Alexa-fluor 633 staining in developing saccules of *Poglut2/3 DKO*s suggested that subdivision of the saccules into alveoli was impaired and is consistent with the enlarged airway space observed in Figure 4.5C. In the saccular stage, fibroblasts deposit elastin in the alveolar wall (123) possibly at the base of the AEC I cells (Figure 4.5A). At the site of elastin deposition in the alveolar wall, the secondary septa arise during the alveolar maturation (121, 124). Taken together, these findings suggest a critical role for POGLUT2/3-mediated *O*-glucosylation of EGF repeats in lung development and assembly of FBN/Elastin networks in the developing lung.

**TGF- $\beta$  signaling was reduced in blood vessels, terminal bronchioles and alveoli of the *Poglut2/3 DKO* lung at E18.5**

Since POGLUT2/3 substrates like FBN1 and LTBP1 also influence TGF- $\beta$  signaling through interacting with each other to generate a reservoir of latent TGF- $\beta$  (53, 54, 112, 125, 126), we sought to determine whether the loss of POGLUT2/3-mediated *O*-glucosylation of EGFs impacted TGF- $\beta$  signaling. As an output for TGF- $\beta$  signaling levels we used immunohistochemistry to compare the levels and distribution of pSMAD2, a downstream effector of signaling (127, 128), in lung sections of wild-type and *Poglut2/3 DKO* embryos (Figure 4.9). In both the wild-type and the *Poglut2/3 DKO* lungs, pSMAD2 was mostly localized in the nucleus in the endothelial layers of the blood vessel, the smooth muscle layers of the terminal bronchiole, and epithelial cells of the alveoli (Figure 4.9 A-F, H). In *Poglut2/3 DKO*, pSMAD2 immunolocalization was visually observed in fewer cells in different regions of lung (Figure 4.9D-F) when compared to similar regions of wild type lungs (Figure 4.9A-C). Overall, there was a reduced level of pSMAD2 in the *Poglut2/3 DKO* lung (Figure 4.9G). Dysregulation of TGF- $\beta$  signaling affects lung alveologenesis (114, 115, 129, 130), and our data suggests that loss of POGLUT2 and 3, and hence *O*-glucosylation of EGF substrates like FBN1 and FBN2, ultimately affects lungs alveolar maturation through reduced TGF- $\beta$  signaling. Again, this is further evidenced by the trend of an increase in size of the alveolar spaces in *Poglut2/3 DKO* lungs compared to wildtype (Figure 4.5C and S4.4), though this data was not statistically significant, suggesting an alveolar septation defect.

**Abundance of FBN1 and other ECM proteins are differentially altered from Poglut2/3 DKO mouse fibroblasts compared to wild type**

We recently showed that loss of POLGUT2 and/or 3 in HEK293T cells reduces secretion of an overexpressed fragment of FBN1 (36). Here, we utilized a mass spectral-based proteomics approach to analyze whether the abundance of FBN1 and other secreted proteins were affected in E18.5 *Poglut2/3 DKO* mouse lung and dermal fibroblasts. A standard label-free quantitative analysis method within Proteome Discoverer (v2.5) was used to identify peptides and quantify relative protein abundances (131). Notably, FBN1 abundance is reduced by approximately thirty percent compared to wild type (Figure 4.10), which was similar but less severe compared to the previously reported data (36). Several other proteins crucial to the assembly and function of the ECM were also decreased in abundance such as collagen  $\alpha 1(V)$  (CO5A1), sushi, von Willebrand factor type A (SVEP1), collagen  $\alpha 1(VI)$  (CO6A1), and laminin subunit gamma-1 (LAMC1) (Figure 4.10, Table S4.6). Several molecular chaperones had reduced abundance as well, including Serpin H1 (SERPH), protein disulfide isomerase (PDIA1), endoplasmic reticulum chaperone BiP (BiP), heat shock protein HSP 90-beta (HSP90B), and endoplasmin (ENPL) (Figure 4.10, Table S4.6). Molecular chaperones are not generally thought of as secreted proteins; however, they have been found in the extracellular space (132-134). A few ECM-related proteins were increased in abundance in the medium of *Poglut2/3 DKO* fibroblasts including fibulin-4 (FBLN4, 58% increase), stromelysin-1 (MMP3, 75% increase), and gelsolin (GELS, 46% increase) (Figure 4.10). A complete list of all identified proteins from dermal fibroblasts is found in Table S4.6.

These analyses were also performed using E18.5 lung fibroblasts (Figure S4.5, Table S4.6). Surprisingly, FBN1 was only reduced by about 15% and was not statistically significant.

This differs from the observation in dermal fibroblasts and previously reported data using HEK293T cells (36), suggesting loss of POGLUT2/3 may affect substrate secretion in tissues differentially. Basement membrane-specific heparan sulfate proteoglycan core protein (HSPG2 or perlecan) and decorin (PGS2) abundance were reduced by about 35-40%. Perlecan and decorin both impact assembly of fibrillin microfibrils (135, 136). Perlecan and FBN1 colocalize and interact with one another, and perlecan-deficient mice (embryonic lethal) display a decreased amount of FBN1 microfibrils (135). This raises the possibility that POGLUT2/3-mediated *O*-glucosylation may have a role in the interactions of FBNs with proteoglycans in the ECM. Interestingly, cathepsins S, B, and Z were all increased in abundance in *Poglut2/3 DKO* lung fibroblasts, though this mechanism is unclear.

#### **4.4 Discussion**

Previous studies suggested a role for POGLUT2/3-mediated *O*-glucosylation of EGF repeats for efficient folding and secretion of POGLUT2/3 substrates like FBN1, FBN2, and LTBP1 (10, 36). However, these experiments were performed using protein fragments overexpressed in cultured HEK293T cells and raised the question of whether POGLUT2/3-mediated *O*-glucosylation of EGF substrates was essential for protein secretion *in vivo*, where full length proteins (rather than fragments) would be expressed at endogenous levels in cells adapted for their secretion (fibroblasts versus HEK293T cells). Additionally, it is still unknown whether *O*-glucosylation by POGLUT2/3 is required for proper function of protein substrates in the extracellular environment. To address these questions, we took advantage of KOMP *Poglut2* and *Poglut3* knockout mouse models. The absence of EGF *O*-glucosylation between Cys3-Cys4 on several proteins in mouse *Poglut2/3 DKO* lung fibroblasts demonstrated unequivocally that

only POGLUT2/3 can modify the consensus sequence. Moreover, analyses of the effects of *Poglut2/3* mutations on mouse embryo development support a critical role for POGLUT2/3-mediated *O*-glucosylation of EGFs in the assembly of fibrillin microfibrils critical for elastogenesis and TGF- $\beta$  signaling. In addition, mass spectral analyses of conditioned medium from primary lung fibroblasts demonstrated that predicted POGLUT2/3 substrates FBN1, FBLN2, FBLN5, and NID1 were *O*-glucosylated, raising the possibility that lethality, cardiovascular, lung, limb, and size phenotypes observed in *Poglut2/3* *DKOs* stem from the combined effects of the loss of POGLUT2/3-mediated *O*-glucosylation on multiple substrates.

Since FBN1 and FBN2 have the most POGLUT2/3 modified sites (36), it is likely they are the major substrates impacted in *Poglut2/3* *DKO* mice. *Poglut2/3* *DKO* significantly reduced the levels of FBN1 and -2 in developing lung blood vessels, airway, and alveoli. This observation was consistent with our previous observation where secretion of FBN1 constructs containing N-terminal EGFs 1-26 was reduced approximately 75% in *POGLUT2/3* *DKO* HEK293T cells (36) and our observation that FBN1 was reduced by approximately 30% percent in *Poglut2/3* *DKO* dermal fibroblasts. *Poglut2/3* *DKO* mice have overlapping phenotypes affecting the lung, limbs, and blood vessels with *Fbn1* *mgR* and *Fbn2* mice. *Fbn1* *mgR* is a Marfan syndrome (MFS) mouse model where *Fbn1* expression is roughly 20% of wild-type (137). At postnatal day 7 and 56, *Fbn1* *mgR* airspaces are significantly increased (138). The airway spaces of *Poglut2/3* *DKO* lungs at E18.5 were slightly increased, but not statistically significant (Figure S4.4); however, since this is a progressive phenotype, it is possible we did not observe a severe effect in *Poglut2/3* *DKO* mice since the mutations are perinatal lethal. FBN1 and 2 are required for elastogenesis (41, 46, 122, 139, 140), and elastin fragmentation is significantly increased in the ascending aorta of *Fbn1* *mgR*, which is consistent with our

observations of reduced and punctate elastin IHC signal in the blood vessels of *Poglut2/3 DKO* lung. Moreover, *Fbn2* mutant mice exhibit varying levels of syndactyly (116, 141, 142), which is consistent with syndactyly we observed in *Poglut2/3 DKO* limbs. Together, these observations strongly suggest loss of *O*-glucosylation by POGLUT2/3 compromises FBN1 and FBN2 function.

Our observations that FBN2 levels were not reduced in medium from *Poglut2/3 DKO* dermal or lung fibroblasts and FBN1 levels were unaltered in medium from *Poglut2/3 DKO* lung fibroblasts raises the possibility that reduced function of secreted, unmodified POGLUT2/3 substrates also contributes to developmental abnormalities. Possible mechanisms by which the *O*-linked glycan could impact FBN function include coordination of the calcium ion-rigidity, fibrillin microfibril assembly, increased susceptibility to degradation, and interactions with binding partners. Future work will address these mechanisms.

Changes in the fibrillin microfibril network result in numerous connective tissue disorders that are associated with dysregulated cell signaling pathways such as TGF- $\beta$  (44, 57). For disorders such as MFS, hyperactive TGF- $\beta$  signaling is observed due to *FBN1* mutations (49, 56). Similar elevated levels of TGF- $\beta$  signaling were observed in aorta isolated from *Fbn1 mgR* mice at 60 days (137). However, closer examination at day 16 aorta demonstrated that TGF- $\beta$  signaling is initially reduced (137), suggesting that elevated TGF- $\beta$  signaling in MFS patients is caused by disease progression. For this reason, the reduced TGF- $\beta$  signaling that we observed in *Poglut2/3 DKO* E18.5 lung is consistent with the impact of *Fbn1* mutations on TGF- $\beta$  signaling.

Similar to FBN1 and FBN2, loss of *O*-glucosylation by POGLUT2/3 likely affects the function of other substrates such as LTBP and fibulins (FBLNs), which could contribute to the phenotypes we observed in *Poglut2/3 DKO*s. LTBP are necessary for efficient secretion of



TGF- $\beta$  through interaction with the latency associated peptide (LAP)-TGF- $\beta$  binary complex called the small latent complex (SLC) (112). This forms a ternary complex, named the large latent complex (LLC), which is secreted into the extracellular space (112). Through interactions of the LLC with FBN microfibrils, LTBP $s$  facilitate sequestration and activation of TGF- $\beta$  (53, 54, 112, 126). TGF- $\beta$  is mainly activated through integrins on the cell surface, which requires mechanical tension to free TGF- $\beta$  from the LLC (112). Loss of POGLUT2/3 *O*-glucosylation on LTBP $s$  could impair TGF- $\beta$  signaling through multiple mechanisms. In *Poglut2/3 DKO* $s$ , formation of the LLC could be impaired, limiting secretion of latent TGF- $\beta$ . Additionally, once in the extracellular space, unglucosylated LTBP may be more susceptible to protease degradation which would prevent necessary LLC incorporation into the ECM for integrin-mediated activation of TGF- $\beta$ . Moreover, loss of POLGUT2/3 *O*-glycans on LTBP $s$  and FBN $s$  could reduce the overall stiffness of the ECM and LLC, which may reduce the efficiency of TGF- $\beta$  activation due to a possible increase in the amount of integrin-mediated tension required to free TGF- $\beta$  from the LLC. *Fbln4* and *Fbln5* are required for elastogenesis (143, 144). Deletion of *Fbln4* causes perinatal lethality in mice, whereas *Fbln5* mice age normally; however, in all cases, developmental abnormalities are present in the lungs and blood vessels. Loss of *O*-glycans on these proteins in *Poglut2/3 DKO* $s$  may reduce secretion and/or inhibit their interaction with tropoelastin. *Fbln4* and *Fbln5* are also linked to TGF- $\beta$  signaling (143), which suggests *Poglut2/3 DKO* effects on their function could contribute to dysregulation of TGF- $\beta$  signaling we observed.

Dysregulation of TGF- $\beta$  signaling could subsequently lead to transcriptional changes inside the cell, altering the abundance of secreted ECM proteins. Proteomic analysis of the lung

and dermal fibroblast secretome revealed numerous changes in ECM-related proteins that are not predicted to be modified by POGLUT2 and 3 such as collagens, laminin, lysyl oxidase-like proteins, matrix metalloproteinases, cathepsins, and lysozymes. This also highlights a possible mechanism to explain why we did not observe a more severe secretion defect of POGLUT2/3 substrates like FBNs. Transcription of these genes may have been increased due to dysregulation of TGF- $\beta$  signaling, which could correct for an initial reduction in FBN secretion from *Poglut2/3 DKO* fibroblasts. Transcriptional analyses of these cells and/or tissue will need to be performed to confirm this.

In conclusion, this work provides mechanistic insight into the effects of POGLUT2/3 mediated *O*-glucosylation of FBNs and other POGLUT2/3 substrates. Due to the complex, dynamic nature of the ECM, it is likely *O*-glucosylation by POGLUT2/3 will affect substrate function differentially depending on context. Regardless, we provide evidence suggesting FBN1 and FBN2 function were primarily affected in *Poglut2/3 DKO*s. Future work will attempt to uncover mechanisms to explain these observations in the context of FBNs as well as other POGLUT2/3 substrates.

## **4.5 Experimental Procedures**

### **Mice and genotyping**

*Poglut2* heterozygous ES cells were purchased from EuMMCR (*Poglut2*<sup>tm2a(EUCOMM)Hmgu</sup> (MGI:1919300)). *Poglut3* targeting vector was purchased from EuMMCR. *Poglut3* heterozygous ES cells were generated at Texas A&M Institute for Genomic Medicine (*Poglut3*<sup>tm380258(L1L2\_Bact\_P)</sup> (MGI:1923765)). *Poglut2* and 3 alleles were generated at the Mouse Transgenic and Gene Targeting Core at Emory University. Heterozygous targeted ES cells were injected into

C57BL/6N blastocysts and resultant chimeras were mated to females of the same strain of mice to generate animals heterozygous for the *Knockout First* allele (Fig. S4.1 and 4.2). Maps of alleles and primers used for genotyping are in Figures S4.1 and 4.2. Detailed primer information and PCR conditions are found in Tables S4.1 and 4.2. Only knockout allele *tm1d* from *Poglut2* and 3 single and double knockout mice was analyzed for this paper. All alleles for *Poglut2* and 3 were maintained at the University of Georgia through backcrossing to C57BL/6J.

### **Primary fibroblast cultures**

Lung and tail tissues were collected at E18.5. Tissues were minced with scissors into approximately 1 mm pieces. Tissue pieces were transferred to 1 mL vials containing 2.5 mg/mL Collagenase D (product info) in DMEM high glucose media supplemented with 10% fetal calf serum and 1% penicillin/streptomycin (complete medium). Samples were incubated 45-60 minutes, shaking 200 rpm, at 37°C. Seventy-micron cell strainers attached to 50 mL conicals were pre-wet with 10 mL of complete media. Tissues were passed through cell strainers using a double-sided pestle. Cells were pelleted at 600 x g for 6 minutes at room temperature. Supernatant was removed and cells were washed once with 10 mL of complete media. Cells were pelleted again at 600 x g for 6 minutes at room temperature. Cells were resuspended in 1 mL of complete medium. Cell suspension was added to 1 mL of complete medium in a single well of a 6-well plate or to 2 mL of complete medium in a 6 cm dish. After 24-48 hours, medium was removed, cells were washed with PBS, and fresh complete medium was added. Medium was changed every 2-3 days until confluency. All cells were maintained at 37 °C and 5% CO<sub>2</sub> in DMEM high glucose media supplemented with 10% fetal calf serum and 1% penicillin/streptomycin. For glycoproteomic and proteomic experiments, fibroblast cultures after

one passage were grown to confluency in 6 cm dishes. At confluency, medium was switched to DMEM high glucose media with no serum or antibiotics. After 3 days, medium was collected, cleared, and stored at -20 °C for mass spectral analysis.

### **Mass spectral analysis**

BCA assays were used to determine protein concentration of each media sample. Media volume equivalent to 10 µg was acetone precipitated in ice-cold acetone (1:4 media:acetone) overnight at -20 °C. Samples were spun down at max speed, 10 minutes, 4 °C. Supernatant was removed and pellet was processed for mass spectral analysis. Proteins were denatured and reduced using 15 µL of reducing buffer containing 8 M Urea, 400 mM ammonium bicarbonate, and 10 mM TCEP at 50 °C for 5 min. Alkylation was performed at room temp in the dark with 7.5 µL of 100 mM iodoacetamide in 50 mM TrisHCl for 30 min to 1 h. Mass spectral grade water (60 µl) was added to each sample. 500 ng of trypsin (cleaves C-terminal to lysine and arginine, Thermo Scientific Pierce 90057) protease was added per sample. Samples were incubated in 37 °C water bath for 3-4 h. Formic acid (10.5 µl of 5%) was added and samples were sonicated for 15 min. Samples were desalted with Millipore C18 Zip Tip Pipette Tips. After elution in 50% acetonitrile, 0.1% acetic acid, samples were diluted to an approximate concentration of 1 µg/µL, 15% acetonitrile, and 0.1% formic acid. Approximately 2-3 µg of each sample was injected on a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Fisher) with an Easy nano-LC HPLC system with a C18 EasySpray PepMap RSLC C18 column (50 µm × 15 cm, Thermo Fisher Scientific). A 90 min binary gradient solvent system (Solvent A: 0.1% formic acid in water and Solvent B: 90% acetonitrile, 0.1% formic acid in water) with a constant flow of 300 nL/min was used. Positive polarity mode was used with a m/z range of 350-2,000 at a resolution of 35,000 and automatic

gain control set to  $1 \times 10^6$ . Higher energy collisional dissociation-tandem mass spectrometry (HCD-MS/MS) was used on the top 10 precursor ions in each full scan (collision energy set to 27%,  $2 \times 10^5$  gain control, isolation window  $m/z$  3.0, dynamic exclusion enabled, and 17,500 fragment resolution. For glycoproteomic analysis, PMi-Byonic (v4.1.10) was used to identify peptides. Fixed modifications: Carbamidomethyl +57.021464 at C. Variable modifications: Oxidation +15.994915 at M,N,D, Deamidated +0.984016 at N,Q, and Ammonia-loss -17.026549 at N-Term C. Precursor mass tolerance was set to 20 ppm and fragment mass tolerance was set to 10 ppm. Two missed cleavages were allowed. Protein and peptide false discovery rates were set to a threshold of 1% and calculated using the 2-dimensional target decoy strategy as described (ref.). All data was searched against a mouse FBN1 database (Uniprot accession number Q61554 version 174 updated July 6, 2016, 1 entry), FBLN2 (Uniprot accession number P37889 version 185 updated October 3, 2012, 1 entry), NID1 (Uniprot accession number P10493 version 217 updated July 27, 2011, 1 entry), HMCN1 (Uniprot accession number D3YXG0 version 97 updated April 20, 2010, 1 entry), FBLN5 (Uniprot accession number QPWVH9 version 181 updated November 1, 1999, 1 entry), and SVEP1 (Uniprot accession number A2AVA0 version 108 updated February 5, 2008, 1 entry). Xcalibur Qual Browser (v2.0.3.2) was used to generate EICs for all identified peptides. For each peptide, area under the curve was calculated for each peak corresponding to searched glycoforms. Relative abundance was calculated by comparing area under the curve for a single glycoform to the total area under curve for all searched glycoforms of a specific peptide. Glycoforms searched: unmodified peptide, unmodified peptide plus  $\beta$ -hydroxylation, modified peptide with *O*-hexose, and modified peptide with *O*-hexose plus  $\beta$ -hydroxylation. For proteomics analysis Proteome Discoverer (v2.5) was used to identify and quantify peptides. The default Comprehensive Enhanced Annotation LFQ and Precursor

Quantitation workflow was used. Precursor abundance was based on peak intensity. Protein abundance calculations were performed using the TopN approach where the average of the five most abundant distinct peptide groups were calculated for each protein. Minimum peptide length was set to 6, and false discovery rate was set to 1%. Precursor mass tolerance was 10 ppm and fragment mass tolerance 0.02 Da. Two missed cleavages were allowed for trypsin digestion. Static modifications: Carbamidomethyl +57.021464 at C. Dynamic modifications: Oxidation / +15.995 Da (D, M, N, P), Deamidated / +0.984 Da (N, Q), Hex / +162.053 Da (S, T), Acetyl / +42.011 Da (N-Terminus), Gln->pyro-Glu / -17.027 Da (Q), Met-loss / -131.040 Da (M), Met-loss+Acetyl / -89.030 Da (M). All data was searched against the reviewed *Mus musculus* proteome from UniProt (Reviewed 17,1114 proteins Swiss Prot). Data was moved into Excel for statistical analysis. A PSM cutoff of 170 (dermal fibroblasts) and 120 (lung fibroblasts) was used. Total protein abundance for each sample was calculated by summing all protein abundances. All samples were normalized to the sample with the highest protein abundance. Average abundance was calculated for each protein. Fold change was determined by dividing the knockout average by the wild-type average for each protein. A one direction t-test was used to calculate statistical significance.

### **Histology and immunohistochemistry**

Histology and immunostaining was performed on 5 mm thin paraffin sections prepared from E18.5 left lungs that were either fixed in 4% paraformaldehyde or 5% acetic acid in ethanol (volume/volume) as previously described (145). Briefly, paraformaldehyde fixed lungs were dehydrated through graded series of ethanol :30%, 50%, 70%, 80%, 90%, 95% and 100% and then xylene whereas acetic acid fixed tissues were directly transferred to 100% ethanol and then

xylene. Acetic acid-fixed tissues were then processed through xylene and paraffin (1:1) mix and infiltrated with and embedded in paraffin and processed for sectioning using a microtome.

Sections were routinely stained with hematoxylin and eosin (H&E) staining as previously described (146). Briefly, sections were deparaffinized in xylenes and rehydrated using ethanol (100%, 95%, 80% and 70%), washed with reverse osmosis water for 10 minutes and stained in Mayer's hematoxylin (Sigma, Cat. no. MHS32-1 L) for 10 minutes and washed in running tap water for 20 minutes. The sections were again passed through 70% ethanol for 2 minutes and counterstained with eosin (Fisher Scientific Cat. no. E511-25) for a minute. Finally, sections were dehydrated for 2 minutes each in 70%, 95%, 100%, 100%, and 100% ethanol, and cleared in xylenes before mounting. Sections were mounted with Secure Mount™ (Fisher Scientific, Cat. no. 23-022208) and coverslipped.

For immunohistochemistry, the sections were deparaffinized using xylenes and rehydrated using ethanol (100%, 95%, 80% and 70%), rinsed with water and finally with 1X PBS. The immunohistochemistry of FBN1 and FBN2 was performed on 5% acetic acid-ethanol fixed tissues and pSMAD2 immunohistochemistry was performed on 4% paraformaldehyde fixed tissues as previously described (145). Briefly, tissue sections were permeabilized with 0.1% Triton-X-100 in PBS for 30 minutes for FBN1 and FBN2 and for pSMAD2 were boiled under pressure for 10 minutes in tris-EDTA buffer (pH 9.0; Tris 1.21 g, EDTA 0.37 g, distilled water 1 L, adjust pH 9.0, add 0.5 ml of tween 20 and mix well). The sections (cooled and tap water washed) were blocked with 2.5% goat serum (Vector laboratories, Cat no. s1012) for 30 min and incubated overnight with anti-FBN1 [1:100, Rabbit polyclonal (pAb 9543) (Courtesy of Dr.

Lynn Sakai, Oregon Health and Science University, Portland, OR)], anti-FBN2 [1:500, mFib2-Gly (Gly-Rich domain) polyclonal antiserum to mouse Fibrillin 2 (Courtesy of Dr. Robert Mecham, Washington University School of Medicine, St. Louis, MO)], and anti-pSMAD2 (1:300, Phospho-SMAD2, 3108S, Cell Signaling Technology). The FBN1 and FBN2 immunoreactivity was detected using Alexa Fluor® 488 tagged secondary antibody (1:400, Goat Anti-Rabbit IgG H&L, Abcam, Cat. no. ab150077). The pSMAD2 immunoreactivity was detected using Alexa Fluor® 568 tagged secondary antibody (1:400, Goat Anti-Rabbit IgG H&L, Abcam, Cat. no. ab175471). Elastin was detected using Alexa Fluor™ 633 Hydrazide (0.2 µM/L, A30634, Thermo scientific). Then the slides were mounted with DAPI Fluormount-G® (SouthernBiotech, Cat. no. 0100–20) and coverslipped.

The H&E-stained histological sections were photographed using a Nikon Optiphot microscope, AxioCam MRc camera and AxioVisionLE program (Zeiss). Fluorescent images were taken at 63× using Leica TCS SP8 X scanning confocal microscope (Leica, Germany). ImageJ (<http://imagej.net/>) was used to measure the level of fluorescence of the images. Fluorescence was measured in sections obtained from three animals ( $n = 3$ ) for each group with a total of six to nine sections per genotype. Data were evaluated for significance using unpaired, two-tailed *t*-test.

#### **4.6 Data Availability**

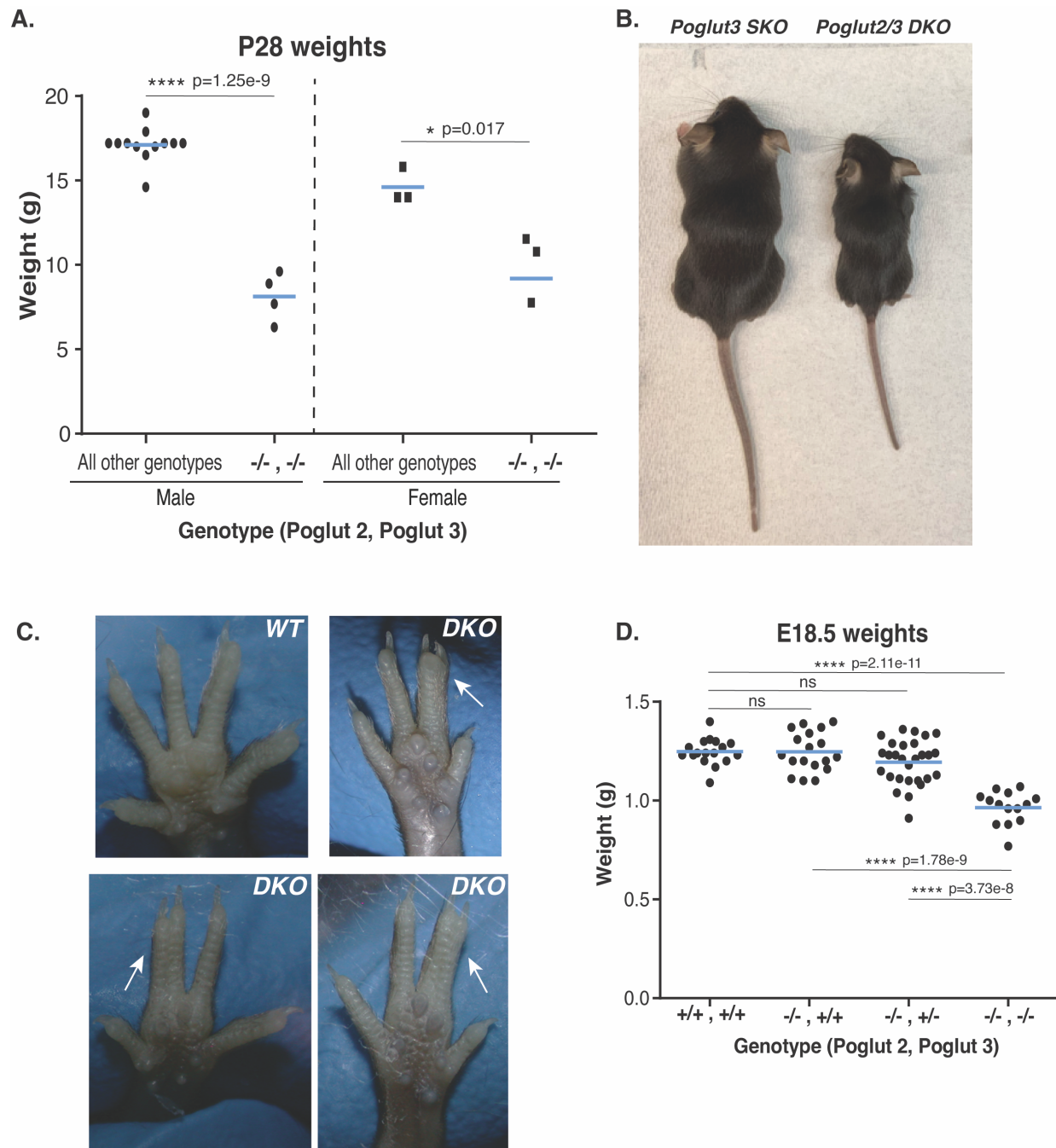
Glycoproteomics data and proteomic secretome analysis (Table S4.6) available by request.



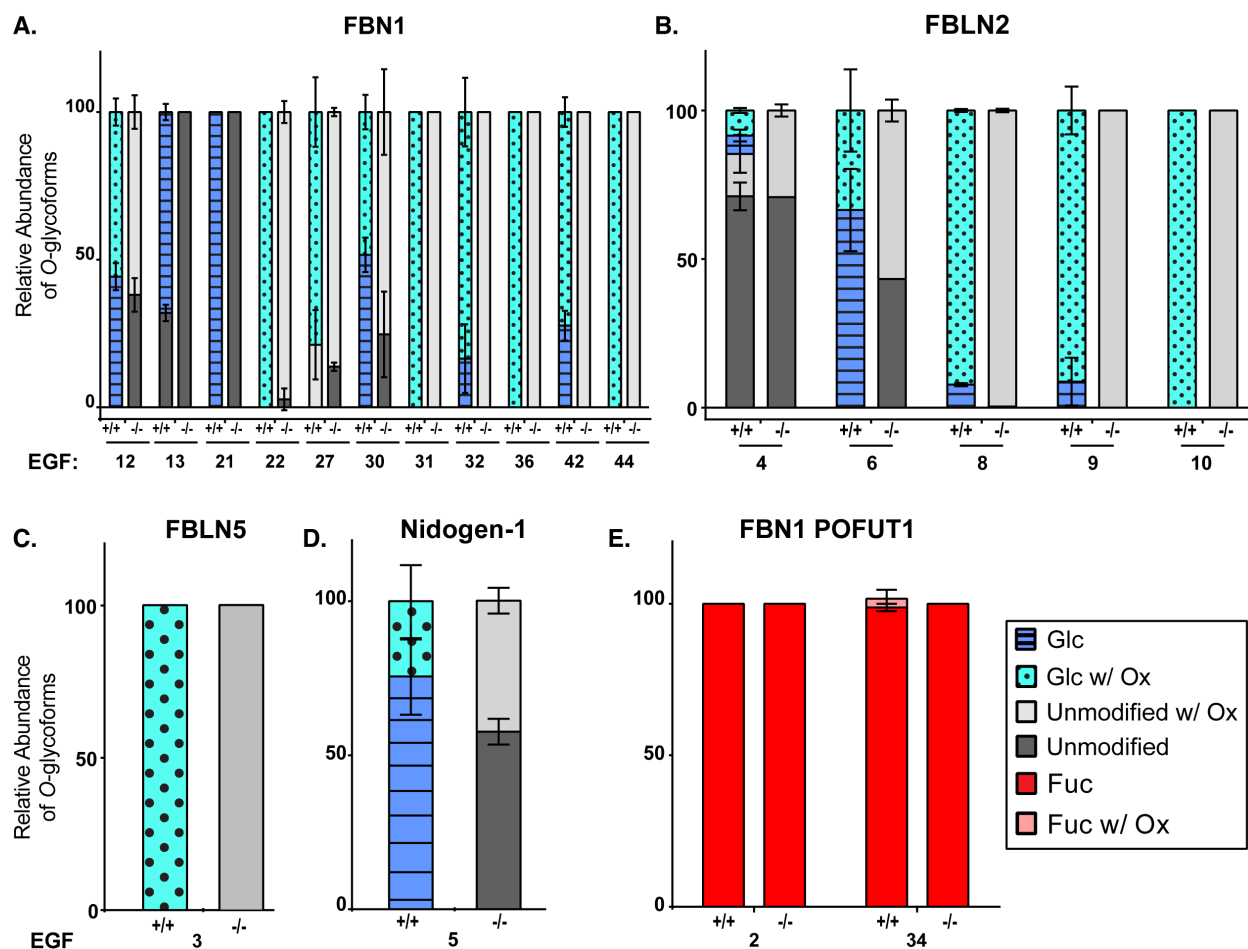
#### 4.7 Main Text Tables and Figures

**Table 4.1. Viability of heterozygous *Poglut2* and 3 null intercrosses.** <sup>1</sup>Animals were generated from intercrosses using C57BL/6J backcross generation N0 through N2. <sup>2</sup>Chi squared with 2 degrees of freedom.

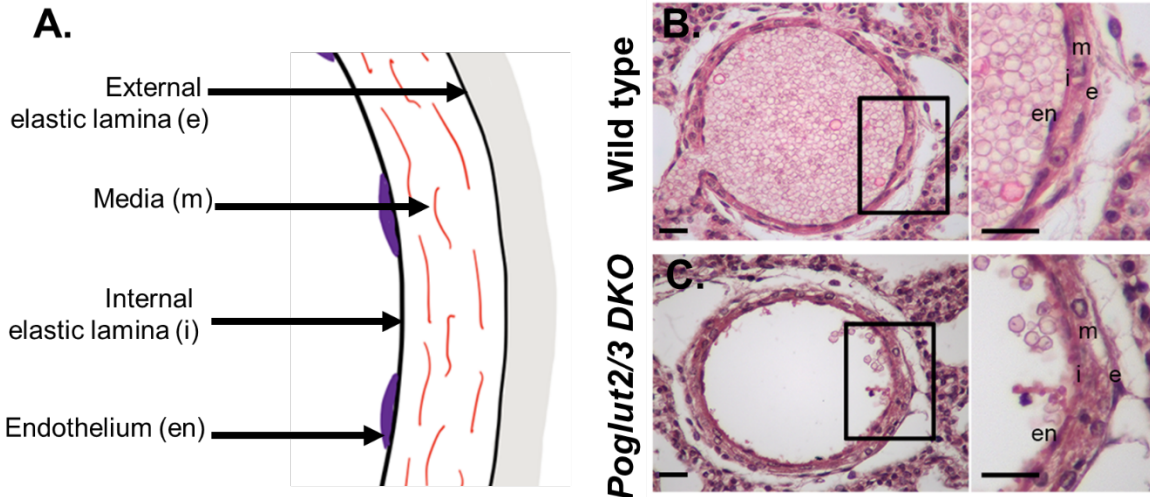
Intercross <sup>1</sup> : <i>Poglut2</i> <sup>-/-</sup> ; <i>Poglut3</i> <sup>+/-</sup>					Intercross <sup>1</sup> : <i>Poglut2</i> <sup>+/-</sup> ; <i>Poglut3</i> <sup>-/-</sup>				
Age	Genotype ( <i>Poglut2</i> , <i>Poglut3</i> )	# Weaning	# Expected	Chi square <sup>2</sup> (p)	Age	Genotype ( <i>Poglut2</i> , <i>Poglut3</i> )	# Weaning	# Expected	Chi square <sup>2</sup> (p)
P21	-/- , +/+	39	24.25	19.907(5e-5)		+/+ , -/-	31	25.25	6.921(0.03)
	-/- , +/-	50	48.5		P21	+/- , -/-	56	50.5	
	-/- , -/-	8	24.25			-/- , -/-	14	25.25	
E18.5	-/- , +/+	22	18.25	1.027(0.60)		+/+ , -/-	10	11	0.364(0.83)
	-/- , +/-	34	36.5		E18.5	+/- , -/-	24	22	
	-/- , -/-	17	18.25			-/- , -/-	10	11	



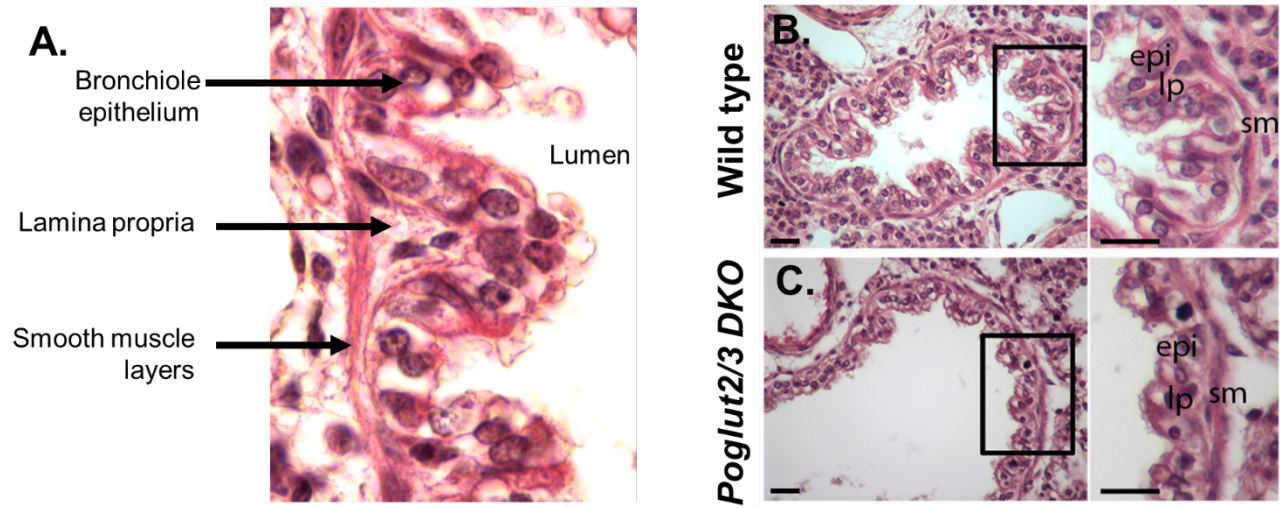
**Figure 4.1. *Poglut2/3* DKO mice are smaller than littermates and have syndactyly.** (A) Body weight of *Poglut2/3* DKO animals compared to all other possible genotypes from *Poglut2* and *Poglut3* intercrosses at P28. Twenty animals across four litters were weighed. Mean weight of each genotype group is indicated by a blue horizontal line. (B) Comparison of a *Poglut3* KO male SKO (left) and a *Poglut2/3* DKO male (right) at P28. (C) Hindlimb syndactyly in *Poglut2/3* DKO mice compared to wild-type hindlimb. White arrow, site of syndactyly. (D) Comparison of body weight across multiple genotypes at E18.5. Eight litters (60 embryos) were used from *Poglut2* KO/KO;*Poglut3* WT/KO intercrosses, and two litters (17 embryos) were used from wild-type intercrosses. (B) \*  $p \leq 0.01$ , \*\*\*\*  $p \leq 0.0001$ , ns not significant. Pictures in panel (C) were taken by Richard Grady, Stony Brook University.



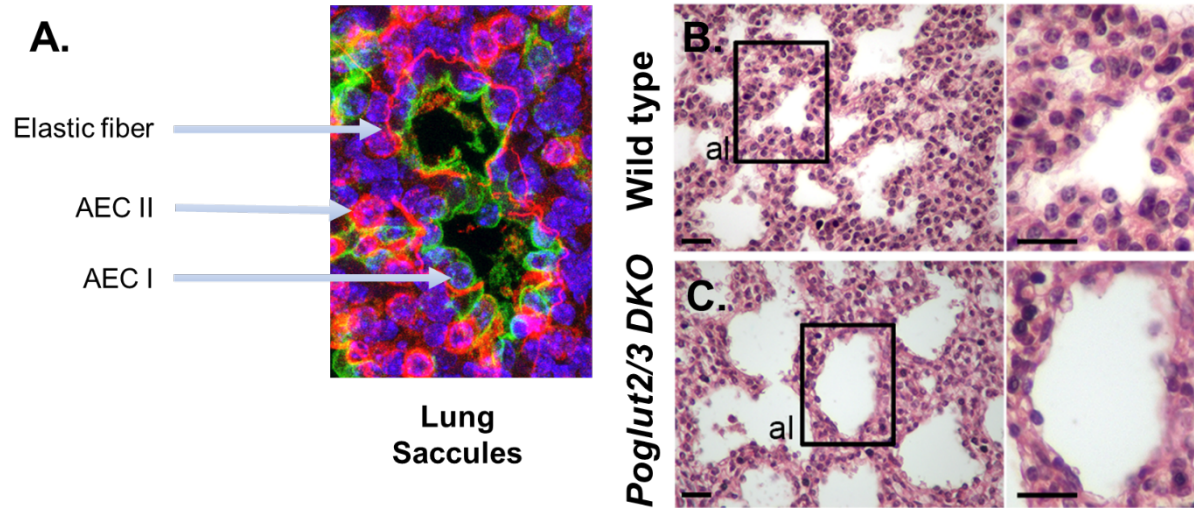
**Figure 4.2. POGLUT2 and 3 modifications are lost in *Poglut2/3* DKO mice.** (A-D) Relative quantification of *O*-hexose modification and  $\beta$ -hydroxylation on peptides from EGF repeats of endogenous (A) FBN1, (B) FBLN2, (C) FBLN5, and (D) Nidogen-1 secreted from lung fibroblasts established at E18.5. (E) Relative quantification of *O*-fucose modification and  $\beta$ -hydroxylation on peptides from EGF repeats of endogenous FBN1 secreted from lung fibroblasts at E18.5. The relative abundance of glycoforms was calculated based on area under the curve from extracted ion chromatograms. These were plotted as a percentage of the total abundance for each peptide. Averages were taken from three biological replicates. Error bars show  $\pm$  SD. Abbreviations: +/+, wild type; -/-, *Poglut2/3* DKO; Glc, glucose; Fuc, fucose; Ox, oxidation ( $\beta$ -hydroxylation).



**Figure 4.3. POGLUT2 and 3 are essential for blood vessel wall structure at E18.5.** (A) Drawing of blood vessel layers. (B-C) Hematoxylin and eosin staining of blood vessels from E18.5 wild type (B) and *Poglut2/3 DKO* (C) lungs. Rectangles indicate digitally magnified regions shown on the right of each image. Abbreviations: en; endothelium, m; media, i; internal elastic lamina, e; external elastic lamina. Scale bars 20  $\mu$ m. Data was generated by Sanjiv Neupane, Stony Brook University.

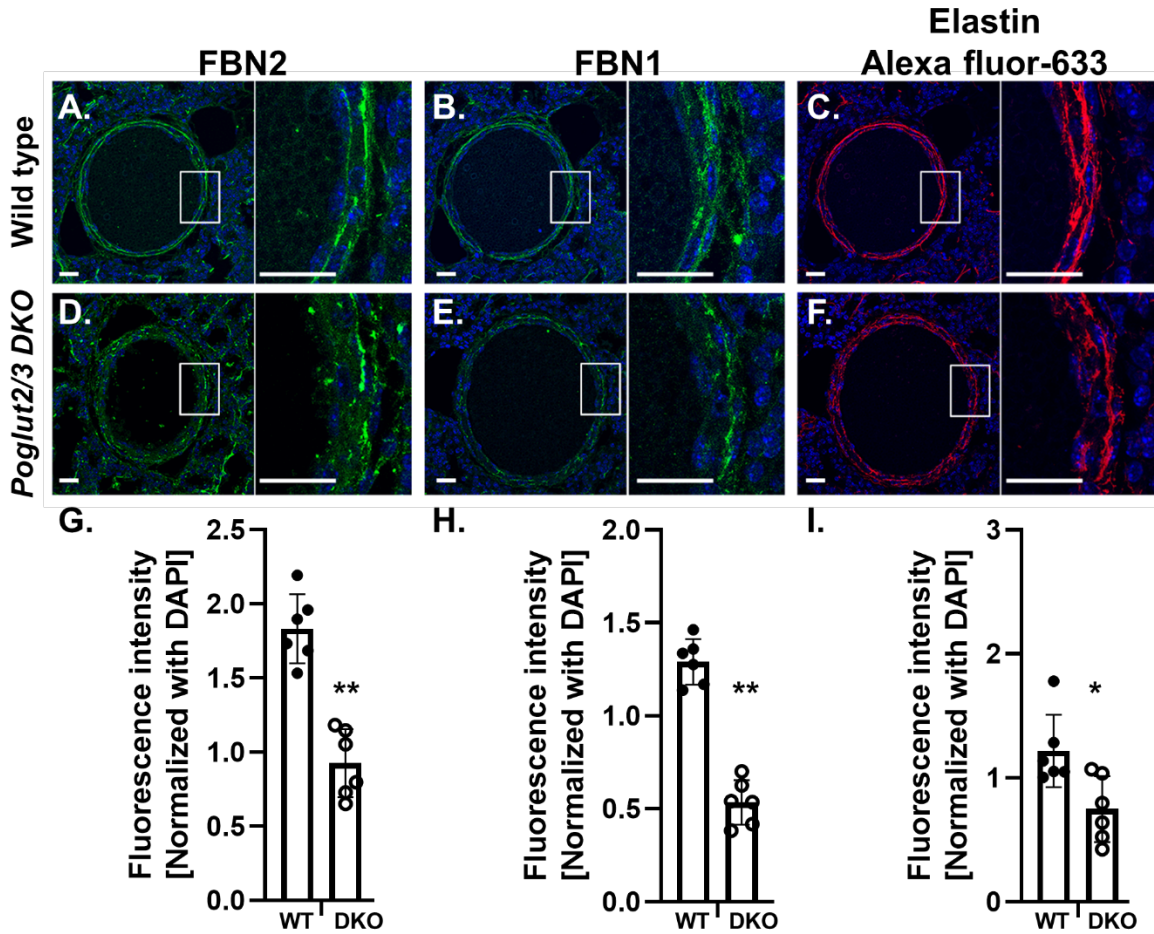


**Figure 4.4. POGLUT2 and 3 are essential for characteristic folded/ruffled structure of terminal bronchioles.** (A) Photograph showing different layers of terminal bronchiole. (B-C) Hematoxylin and eosin staining of terminal bronchioles from E18.5 wild type (B) and *Poglut2/3 DKO* (C) lungs. Rectangles indicate digitally magnified regions shown on the right of each image. Abbreviations: epi; epithelium, lp; lamina propria, i; sm; smooth muscles. Scale bars 20  $\mu$ m. Data was generated by Sanjiv Neupane, Stony Brook University.

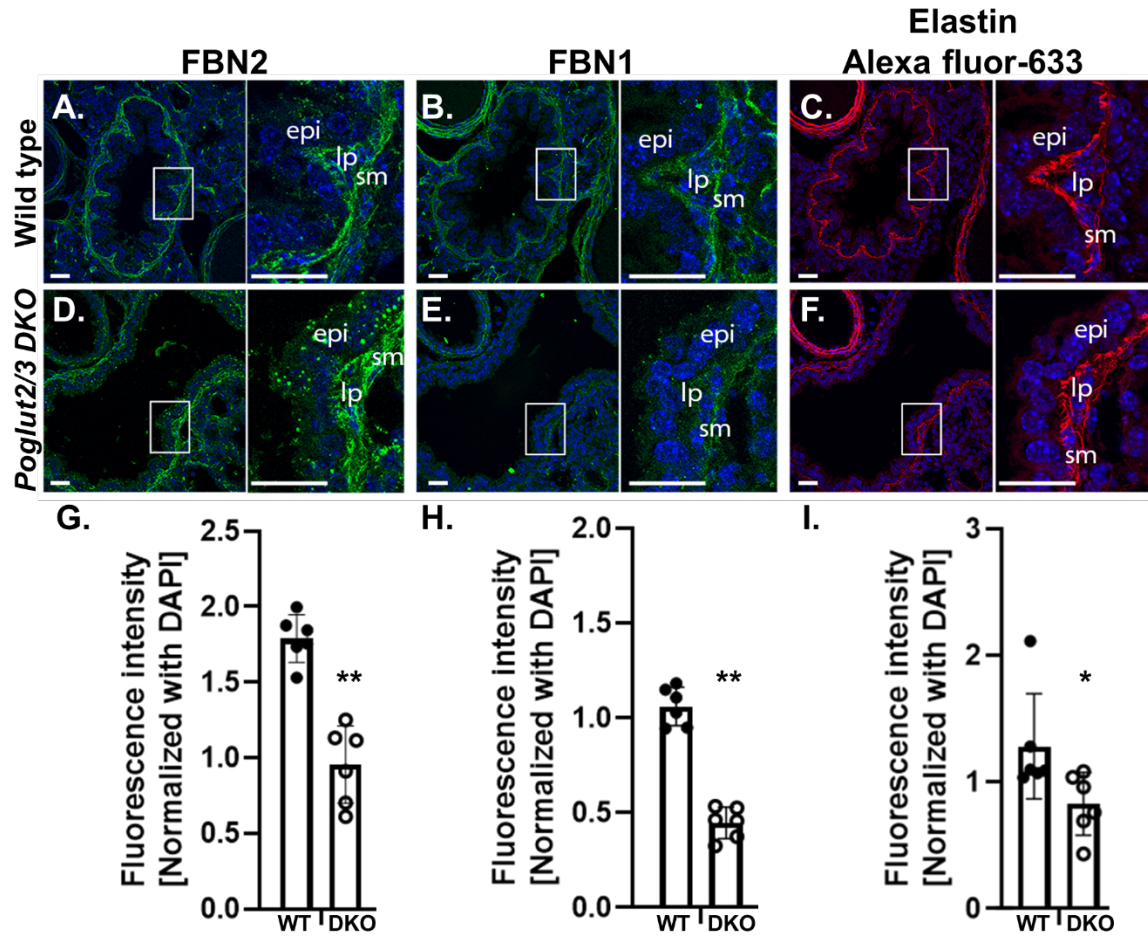


**Figure 4.5. POGLUT 2 and 3 are needed for normal terminal saccule development at E18.5.** (A) Lung saccule at E18.5 showing elastic fiber, alveolar epithelial cell type I (AEC I) and alveolar epithelial cell type II (AEC II). Hematoxylin and eosin staining of alveolar region from E18.5 wild type (B) and *Poglut2/3 DKO* (C) lungs. Rectangles indicate digitally magnified regions shown on the right of each image. Abbreviations: al; alveolus. Scale bars 20  $\mu$ m. Data was generated by Sanjiv Neupane, Stony Brook University.



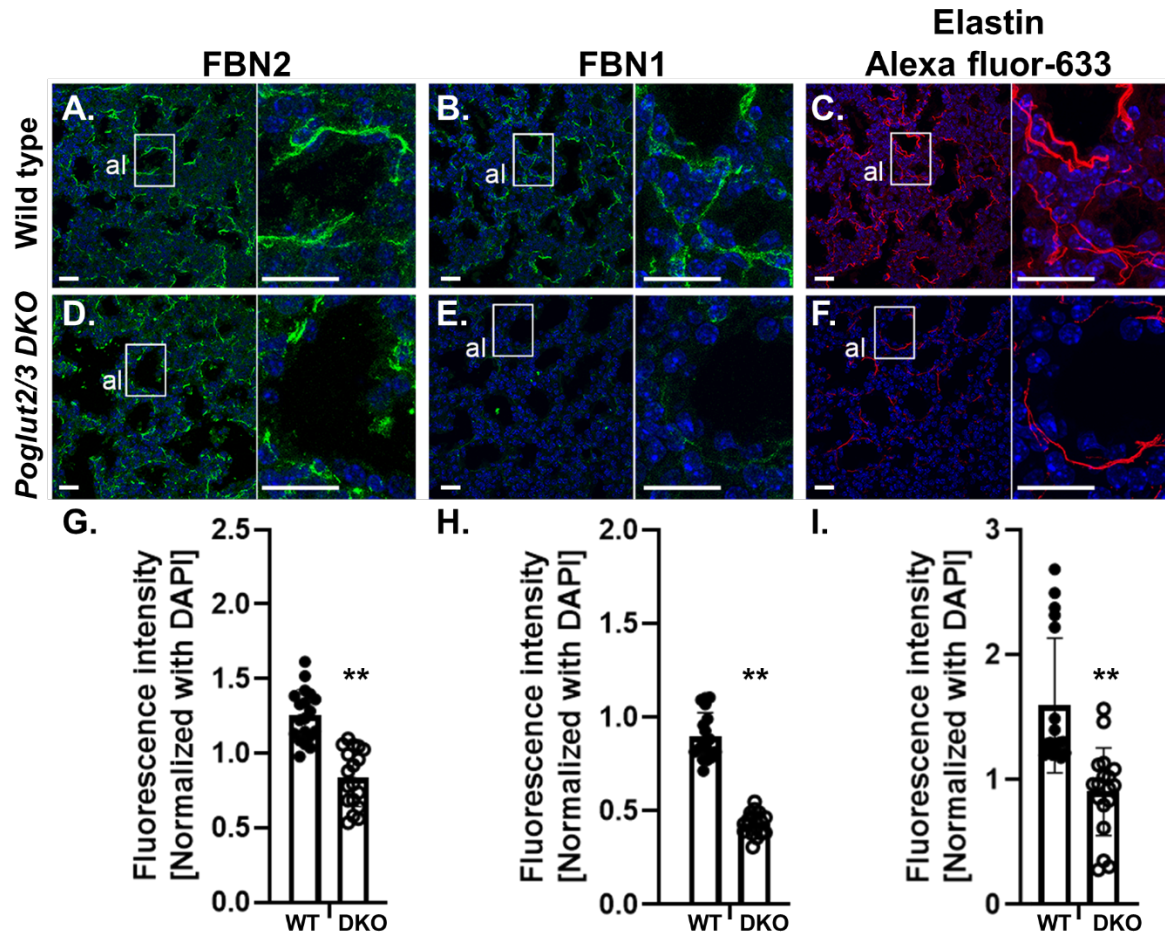


**Figure 4.6. Loss of POGLUT2 and 3-mediated *O*-glucosylation on EGF repeats leads to reduced abundance and fragmented pattern of FBN1, -2, and elastin in lung blood vessels.** (A-F) Representative maximum projection images of fibrillin-2 (FBN2, green) (A and D) and fibrillin-1 (FBN1, green) (B and E) immunolocalization, and elastin (red) fluorescence (C and F) and counterstained with DAPI (blue) in blood vessels from E18.5 wild type (WT) (A-C) and *Poglut2/3* DKO (D-F) lungs. Rectangles indicate digitally magnified regions shown on the right of each image. Quantification of FBN2 (G), and FBN1 (H) immunofluorescence, and elastin (I) fluorescence signals (G-I). Analyses were performed from three embryos per genotype (n=3) with 2 sections from each lung. Data from control (solid black circle) and *Poglut2/3* DKO (open circles) were evaluated for statistical significance using unpaired, two-tailed *t*-test: \*p ≤ 0.05 and \*\*p ≤ 0.01. Scale bars 20 μm. Data was generated by Sanjiv Neupane, Stony Brook University.

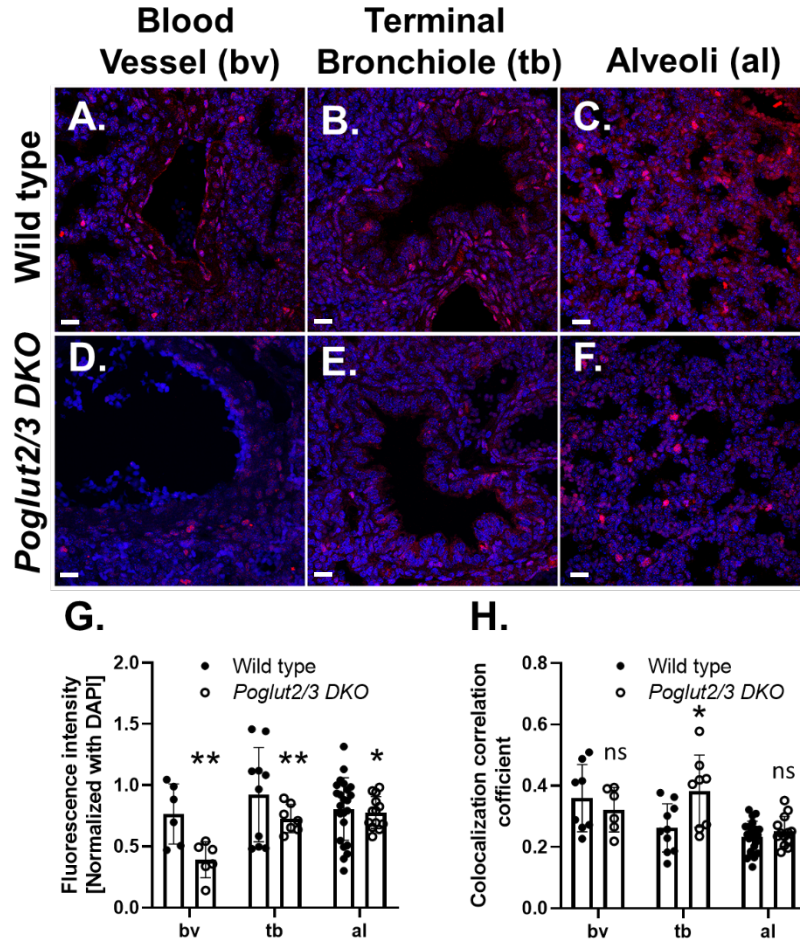


**Figure 4.7. Loss of POGLUT2 and 3 leads to reduced abundance and fragmented pattern of FBN1, -2, and elastin in the terminal bronchiole.** (A-F) Representative maximum projection images of fibrillin-2 (FBN2, green) (A and D) and fibrillin-1 (FBN1, green) (B and E) immunolocalization, and elastin (red) fluorescence (C and F) and counterstained with DAPI (blue) in blood vessels from E18.5 wild type (WT) (A-C) and *Poglut2/3* DKO (D-F) lungs. Rectangles indicate digitally magnified regions shown on the right of each image. Quantification of FBN2 (G), and FBN1 (H) immunofluorescence, and elastin (I) fluorescence signals (G-I). Analyses were performed from three embryos per genotype (n=3) with 2 sections from each lung. Data from control (solid black circle) and *Poglut2/3* DKO (open circles) were evaluated for statistical significance using unpaired, two-tailed *t*-test: \* $p \leq 0.05$  and \*\* $p \leq 0.01$ . Abbreviations: en; endothelium, sm; smooth muscle, lp; lamina propria. Scale bars 20 μm. Data was generated by Sanjiv Neupane, Stony Brook University.

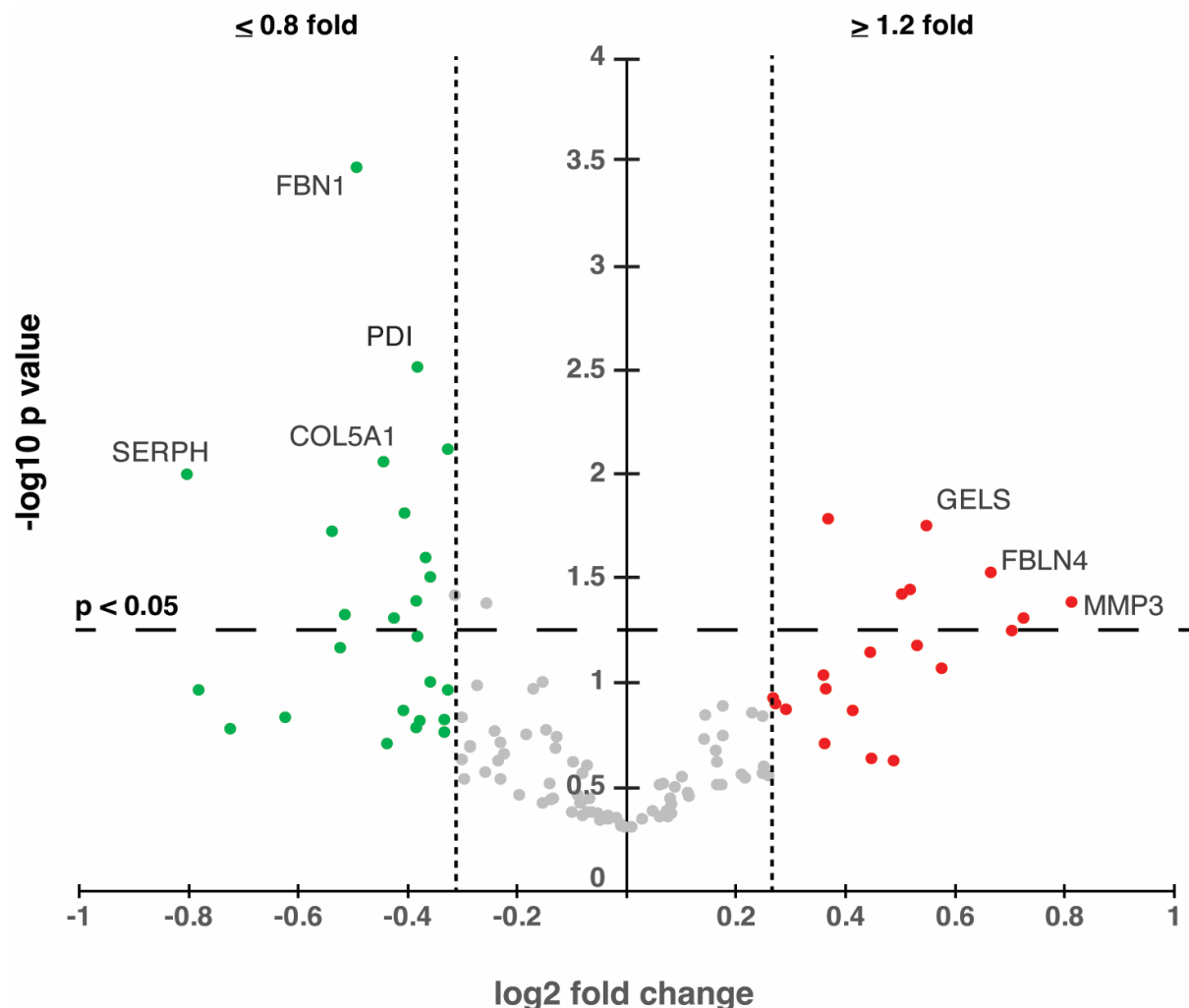




**Figure 4.8. Loss of POGLUT2 and 3 leads to reduced abundance and fragmented pattern of FBN1, -2, and elastin in the alveoli.** (A-F) Representative maximum projection images of fibrillin-2 (FBN2, green) (A and D) and fibrillin-1 (FBN1, green) (B and E) immunolocalization, and elastin (red) fluorescence (C and F) and counterstained with DAPI (blue) in blood vessels from E18.5 wild type (WT) (A-C) and *Poglut2/3* DKO (D-F) lungs. Rectangles indicate digitally magnified regions shown on the right of each image. Quantification of FBN2 (G), and FBN1 (H) immunofluorescence, and elastin (I) fluorescence signals (G-I). Analyses were performed from three embryos per genotype (n=3) with 2 sections (3 images/section) from each lung. Data from control (solid black circle) and *Poglut2/3* DKO (open circles) were evaluated for statistical significance using unpaired, two-tailed *t*-test: \*\* $p \leq 0.01$ . Abbreviation: al; alveolus. Scale bars 20  $\mu$ m. Data was generated by Sanjiv Neupane, Stony Brook University.



**Figure 4.9. TGF- $\beta$  signaling is significantly reduced in *Poglut2/3* DKO lungs. (A-F)** Representative maximum projection images of pSMAD2 (red) immunolocalization and DAPI (blue) at E18.5 surrounding blood vessels (**A and D**), terminal bronchioles (**B and E**), and alveoli (**C and F**) of the lung from wild type (**A-C**) and *Poglut2/3* DKO (**D-F**). (**G-H**) Quantification of pSMAD2 immunofluorescence signal relative to DAPI (**G**) and pSMAD2-DAPI nuclear colocalization by estimation of colocalization correlation coefficient (**H**). Analyses were performed from three embryos per genotype (n=3) with 2-4 sections per group. Data from control (solid black circle) and *Poglut2/3* DKO (open circles) were evaluated for statistical significance using unpaired, two-tailed *t*-test: \*,  $p \leq 0.05$ , \*\*,  $p \leq 0.01$ , ns; not significant. Scale bars 20  $\mu$ m. Data was generated by Sanjiv Neupane, Stony Brook University.



**Figure 4.10. Abundances of secreted FBN1 and other ECM proteins are affected by loss of *POGLUT2* and *3* in dermal fibroblasts at E18.5.** A TopN data dependent acquisition strategy was utilized to quantify each protein abundance. The peak intensities of the top 5 precursors for each protein were used. All samples were normalized to the sample with highest total protein abundance. Fold change (FC) for each protein abundance was calculated by dividing *Poglut2/3* *DKOs* by wild-type. A one-tailed t-test was used to calculate statistical significance. A cut-off value of  $\pm 20\%$  was used to separate proteins that were increased or decreased in abundance compared to wild-type. *Green*, proteins with reduced abundance; *grey*, proteins with no change in abundance; *red*, proteins with increased abundance. *WT* n=8 total replicates, 4 biological replicates run in duplicate. *Poglut2/3* *DKO* n=9, 5 biological replicates.

#### 4.8 Supplemental Tables and Figures

**Table S4.1. Primers and PCR conditions used for genotyping *Poglut2* mice.**

<i>Poglut2</i> allele	Primer Name & [concentration]	Primer Sequence 5'→3'	PCR Product size (bp)	PCR Conditions
<i>Wild type</i>	K1 5' F (a) [0.25 μm]	gcctttagtagctgagacatctctc	496	95°C—3 min  95°C—30 sec 60°C—30 sec 70°C—30 sec (35 cycles)  70°C—5 min  4°C—hold
	K1 3' R (b) [0.25 μm]	cacagcgggtgaacatctcacttaacacc		
<i>Knockout-first (tm1a)</i>	K1 5' F (a) [0.25 μm]	gcctttagtagctgagacatctctc	279	
	LAR3 (c) [0.25 μm]	caacggggttcttctgttagtcc		
<i>lacZ-Δ3-4 (tm1b)</i>	CSD Lac F (d) [0.25 μm]	gctaccattaccagttggtctggtgtca	386	
	K1 3 <sup>rd</sup> LoxP R (e) [0.25 μm]	caccacaatgaactgatggcgag		
<i>Floxed3 (tm1c)</i>	K1 5' F (a) [0.25 μm]	gcctttagtagctgagacatctctc	630	
	K1 3' R (b) [0.25 μm]	cacagcgggtgaacatctcacttaacacc		
<i>Δ3-4 (tm1d)</i>	K1 5' F (a) [0.25 μm]	gcctttagtagctgagacatctctc	259	
	K1 3 <sup>rd</sup> LoxP R (e) [0.25 μm]	caccacaatgaactgatggcgag		

**Table S4.2. Primers and PCR conditions used for genotyping *Poglut3* mice.**

<i>Poglut3</i> allele	Primer Name & [concentration]	Primer Sequence 5'→3'	PCR Product size (bp)	PCR Conditions
<i>Wild type</i>	K2 5' F (f) [0.25 μm]	ggttactagacacttcaatggc	353	95°C—3 min  95°C—30 sec 59°C—30 sec 72°C—30 sec (35 cycles)  72°C—3 min  4°C—hold
	K2 3' R (g) [0.25 μm]	aacgttgccatttcctgatcatt		
<i>Knockout-first (tm1a)</i>	K2 5' F (f) [0.25 μm]	ggttactagacacttcaatggc	241	
	KOMP R (h) [0.25 μm]	ggtggtgtgggaaagggttc		
<i>lacZ-Δ3-4 (tm1b)</i>	CSD Lac F (i) [0.25 μm]	gctaccattaccagttggtctggtgtca	497	
	K2 b_d R (j) [0.25 μm]	actcgtgtgagatccttgcc		
<i>Floxed3 (tm1c)</i>	K2 5' F (f) [0.25 μm]	ggttactagacacttcaatggc	561	
	K2 3' R (g) [0.25 μm]	aacgttgccatttcctgatcatt		
<i>Δ3-4 (tm1d)</i>	K1 5' F (f) [0.25 μm]	ggttactagacacttcaatggc	524	
	K2 new b_d R (k) [0.25 μm]	tcaccagctgctcatgcaata		

**Table S4.3. *Poglut2* WT/KO intercross progeny viability.**

<b>Intercross<sup>1</sup>: <i>Poglut2</i> WT/KO</b>			
<b>Genotype</b>	<b># Weaned</b>	<b># Expected</b>	<b>Chi square<sup>2</sup> (p)</b>
<i>WT/WT</i>	18	18.75	
<i>WT/KO</i>	34	37.5	
<i>KO/KO</i>	23	18.75	1.32 (0.5169)

<sup>1</sup>Animals were generated from intercrosses using C57BL/6J backcross generation N1 through N6

<sup>2</sup>Chi squared with 2 degrees of freedom.

**Table S4.4. *Poglut3* WT/KO intercross progeny viability.**

<b>Intercross<sup>1</sup>: <i>Poglut3</i> WT/KO</b>			
<b>Genotype</b>	<b># Weaned</b>	<b># Expected</b>	<b>Chi square<sup>2</sup> (p)</b>
<i>WT/WT</i>	18	16.5	
<i>WT/KO</i>	34	33	
<i>KO/KO</i>	14	16.5	0.545 (0.7613)

<sup>1</sup>Animals were generated from intercrosses using C57BL/6J backcross generation N0 through N3

<sup>2</sup>Chi squared with 2 degrees of freedom.

**Table S4.5. *Poglut2* WT/KO;*Poglut3* WT/KO double heterozygote intercross progeny viability.**

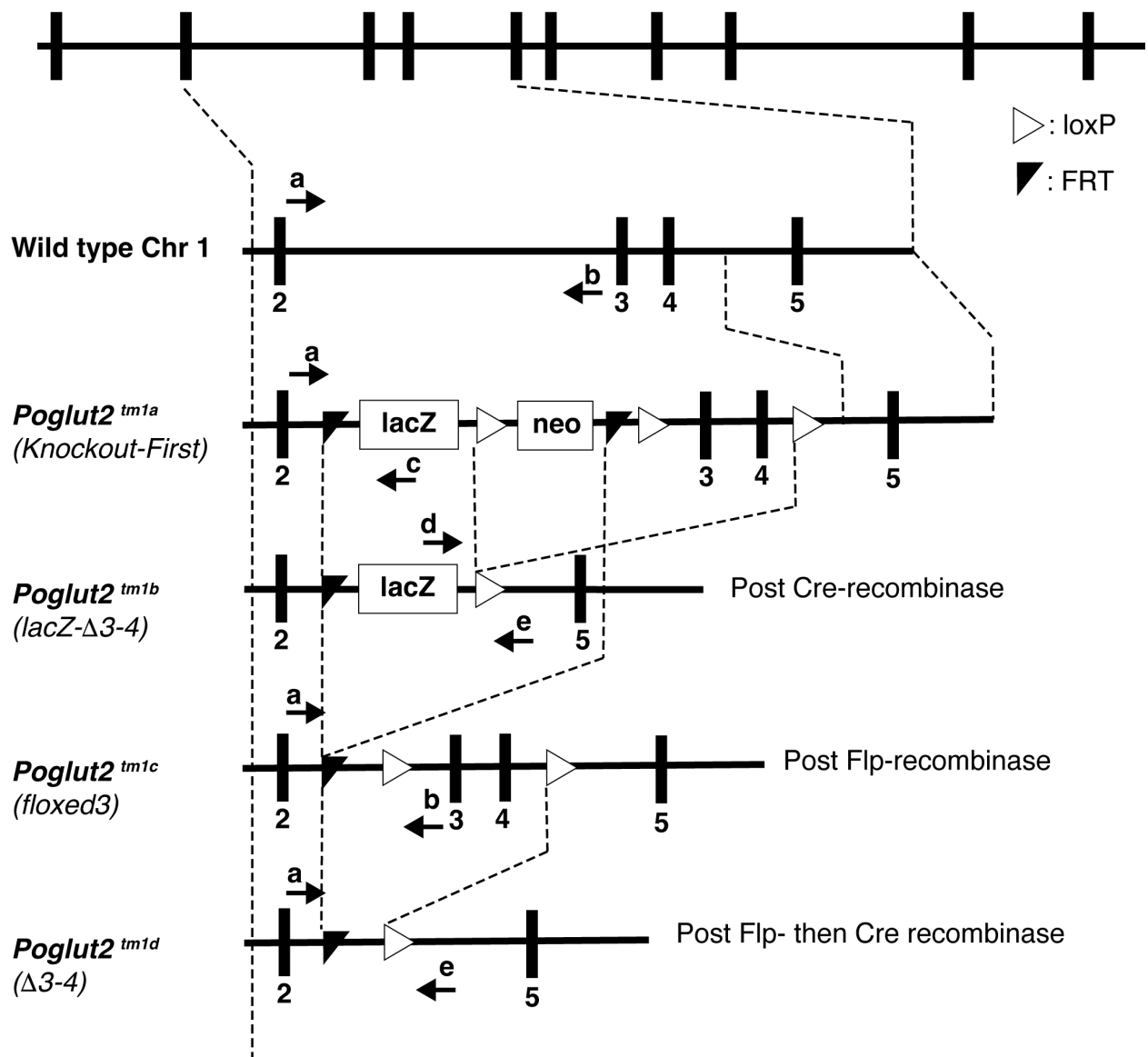
<b>Intercross<sup>1</sup>: <i>Poglut2</i> WT/KO;<i>Poglut3</i> WT/KO</b>					
<b>Genotype</b>	<b># Weaning</b>	<b># Expected</b>	<b>Frequency (%)</b>		<b>Chi square<sup>2</sup> (p)</b>
<b><i>Poglut2</i>;<i>Poglut3</i></b>			<b># Weaning</b>	<b># Expected</b>	
<i>WT/KO</i> ; <i>WT/KO</i>	18	28	16.07	25	
<i>WT/WT</i> ; <i>WT/KO</i>	10	14	8.92	12.5	
<i>WT/KO</i> ; <i>WT/WT</i>	19	14	16.96	12.5	
<i>WT/KO</i> ; <i>KO/KO</i>	17	14	15.18	12.5	
<i>KO/KO</i> ; <i>WT/KO</i>	16	14	14.29	12.5	
<i>WT/WT</i> ; <i>WT/WT</i>	8	7	7.14	6.25	
<i>WT/WT</i> ; <i>KO/KO</i>	15	7	13.39	6.25	
<i>KO/KO</i> ; <i>WT/WT</i>	7	7	6.25	6.25	
<i>KO/KO</i> ; <i>KO/KO</i>	2	7	1.79	6.25	20.29 (0.0093)

<sup>1</sup>Animals were generated from intercrosses using C57BL/6J backcross generation N0 through N2

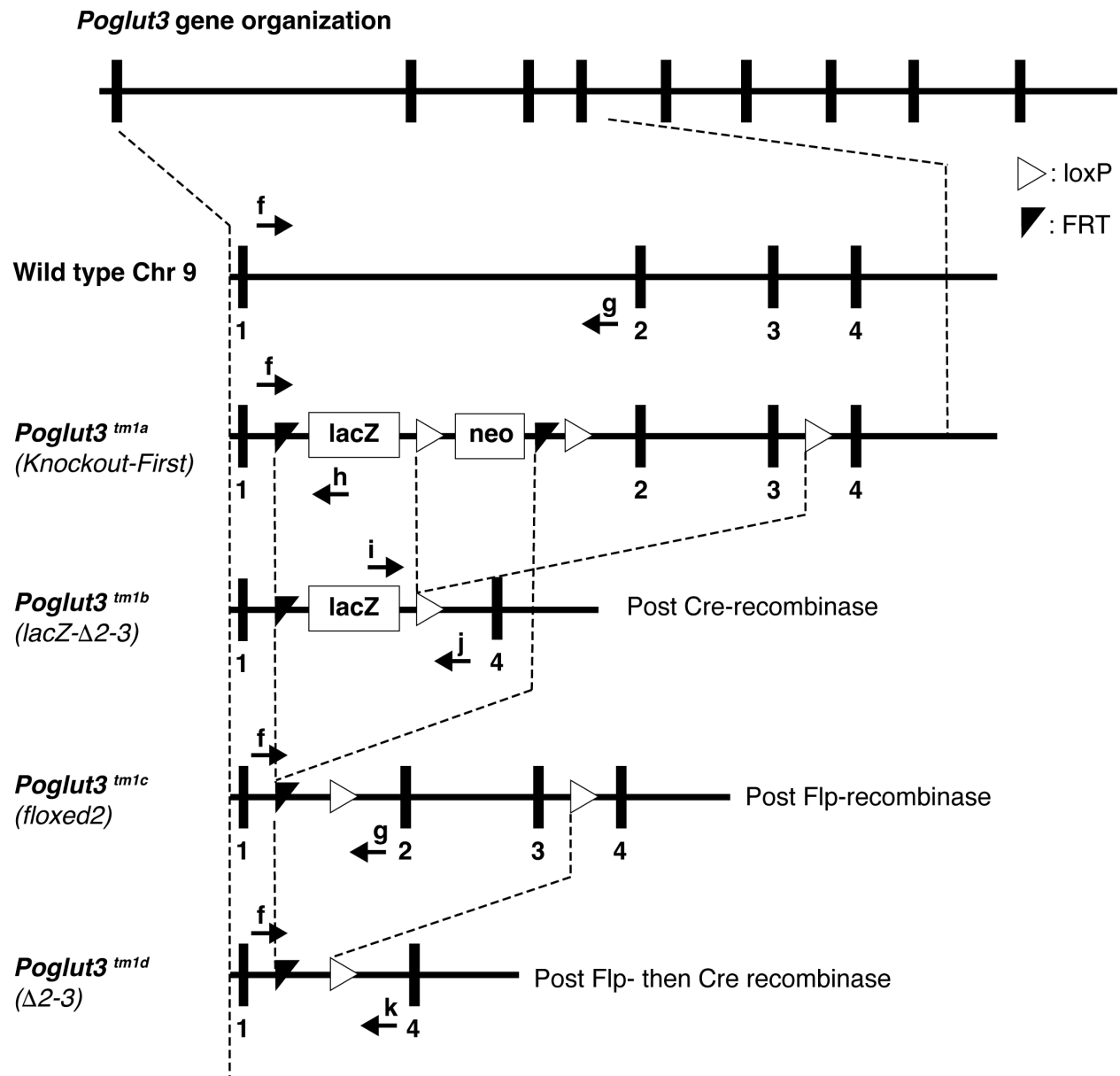
<sup>2</sup>Chi squared with 8 degrees of freedom.



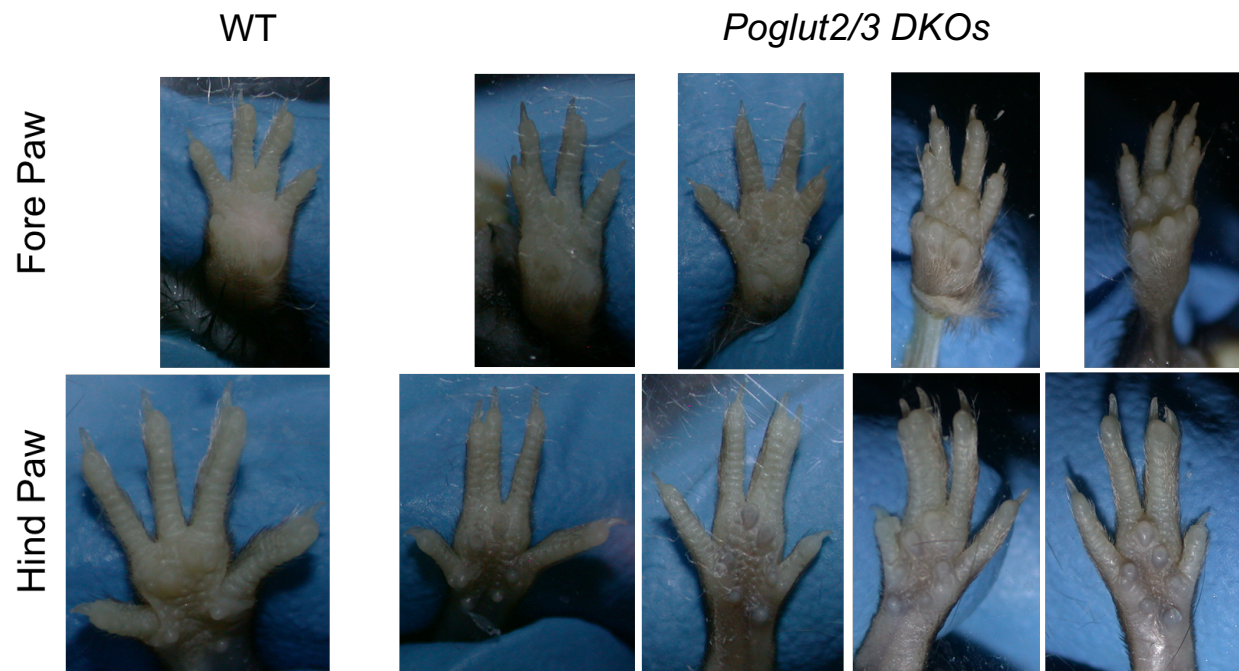
# ***Poglut2* gene organization**



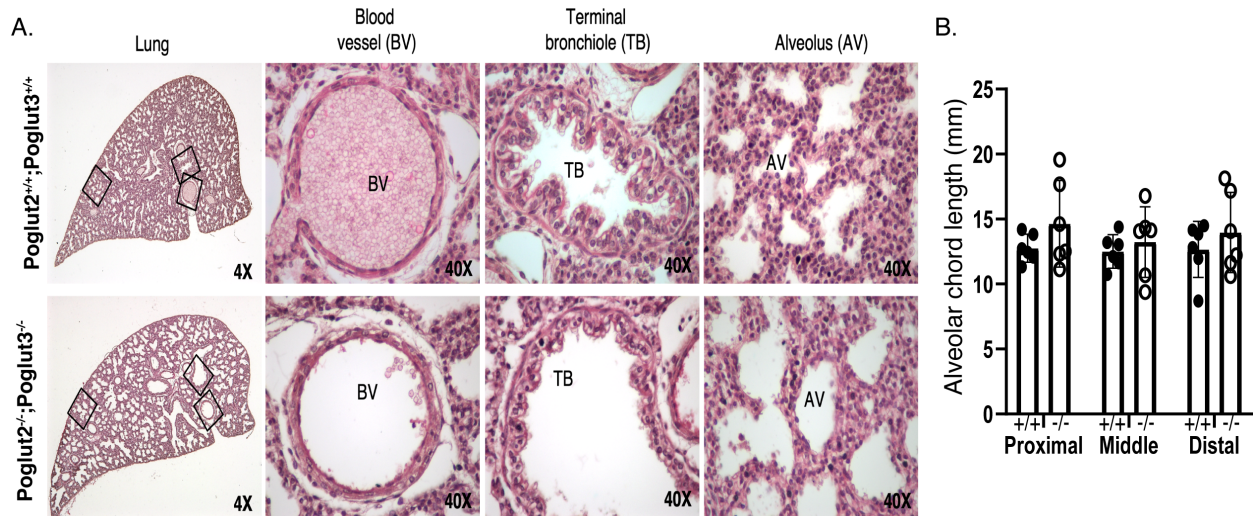
**Figure S4.1. Targeting strategy and PCR design for *Poglut2* alleles.**



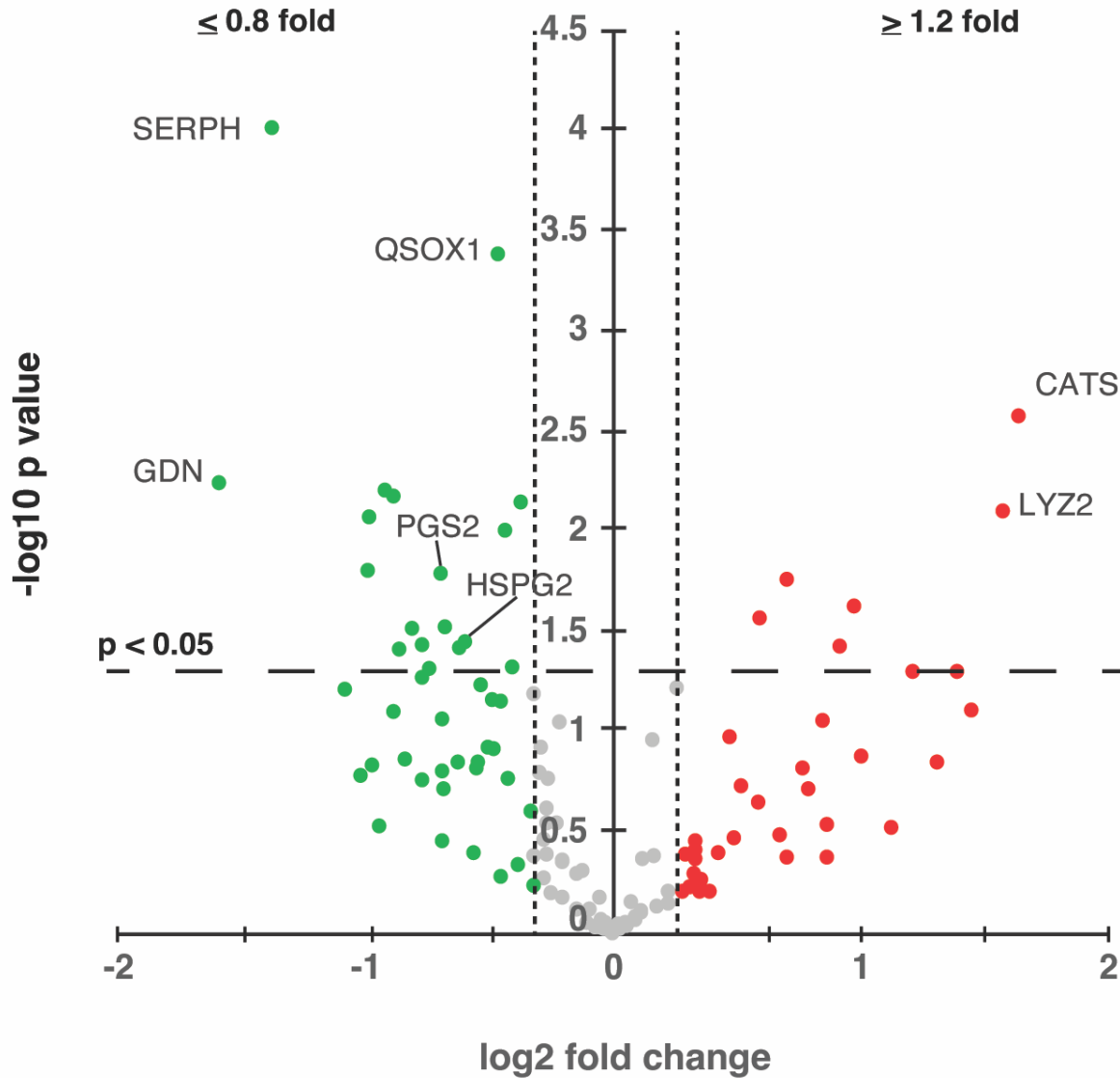
**Figure S4.2. Targeting strategy and PCR design for *Poglut3* alleles.**



**Figure S4.3. *Poglut2/3* DKO mice display forelimb and hindlimb syndactyly.** Pictures were taken by Richard Grady, Stony Brook University.



**Figure S4.4. The alveolar space in *Poglut2/3* DKO lungs is slightly increased compared to wild-type. (A)** Representative image of H&E stain of lung sections at low (4X) and high (40X) magnification. **(B)** Comparison of alveolar chord length between wild-type and *Poglut2/3* DKO lungs. Three biological replicates, two sections per replicate, were quantified for each genotype. Data were generated by Sanjiv Neupane, Stony Brook University.



**Figure S4.5 Abundances of secreted ECM proteins are affected by loss of POGLUT2 and 3 in dermal fibroblasts at E18.5.** A TopN data dependent acquisition strategy was utilized to quantify each protein abundance. The peak intensities of the top 5 precursors for each protein were used. All samples were normalized to the sample with highest total protein abundance. Fold change (FC) for each protein abundance was calculated by dividing *Poglut2/3 DKO* values by wild-type. A one-tailed t-test was used to calculate statistical significance. A cut-off value of  $\pm 20\%$  was used to separate proteins that were increased or decreased in abundance compared to wild-type. *Green*, proteins with reduced abundance; *grey*, proteins with no change in abundance; *red*, proteins with increased abundance. *WT* n=6 total replicates, 3 biological replicates run in duplicate. *Poglut2/3 DKO* n=5, 3 biological replicates.

## CHAPTER 5

### CONCLUSIONS AND FUTURE DIRECTIONS

POGLUT2 and POGLUT3 (previously known as KDELC1 and KDELC2, respectively) *O*-glucosylate EGF repeats between C<sup>3</sup>-C<sup>4</sup> at the putative consensus sequence C<sup>3</sup>-x-N-T-x-G-S-F/Y-x-C<sup>4</sup> (1, 36). We identified the first POGLUT2 and 3 modification site on NOTCH1 EGF11 and subsequently identified additional sites on NOTCH3, leading to a preliminary consensus sequence of C<sup>3</sup>-x-N-T-x-G-S-F-x-C<sup>4</sup> for modification by POGLUT2 and 3 (1). A motif search using this preliminary consensus sequence revealed the fibrillins and LTBP1s contained the largest number of predicted POGLUT2 and 3 modification sites, and many other ECM proteins were predicted to be modified (1). We confirmed that over half of FBN1, -2, and LTBP1 EGF repeats are modified by POGLUT2 and 3 at high stoichiometry (36). These data broadened the POGLUT2 and 3 consensus sequence to C<sup>3</sup>-x-N-T-x-G-S-F/Y-x-C<sup>4</sup> (36). Future biochemical studies will continue to elaborate on the consensus sequence for modification by POGLUT2 and 3, which is a useful tool for identifying additional POGLUT2 and 3 substrates. Additionally, a subset of reported MFS patient mutations is within the POGLUT2 and 3 putative consensus sequence. Generating these variants *in vitro* and performing biochemical analysis could provide evidence implicating POGLUT2 and 3 *O*-glucosylation in the pathology of MFS.

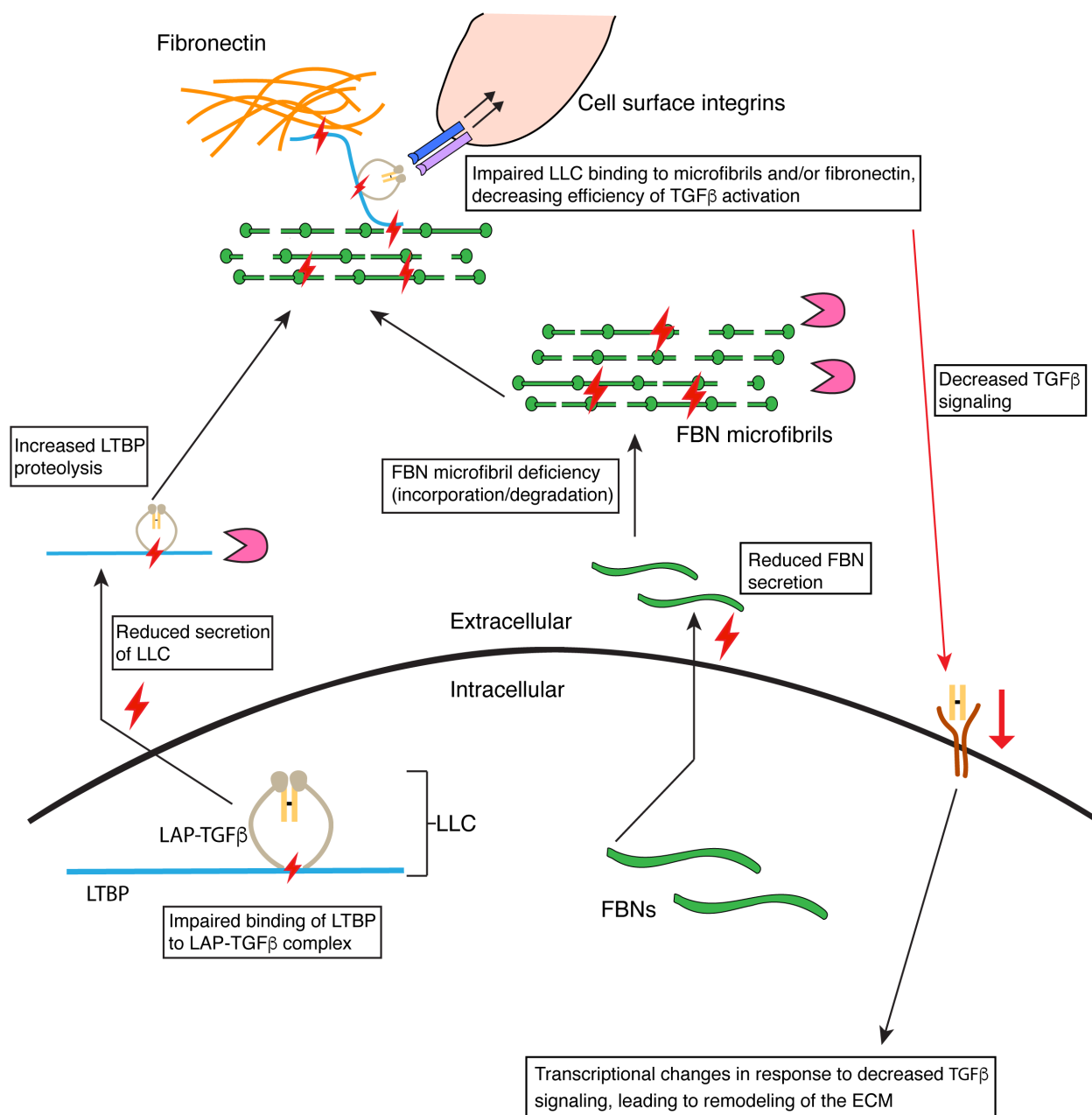
To understand the physiological importance of *Poglut2* and 3, we generated the first *Poglut2* and 3 mouse models. Based on the amount of POGLUT2 and 3 modifications on FBN1, -2, and LTBP1 and the number of other ECM proteins predicted to be modified POGLUT2 and 3, we anticipated phenotypes in elastic tissues such as the lungs, heart, and skin where the

structure and function of ECM is crucial for proper development. Our previous analysis revealed secretion of a FBN1 fragment from *Poglut2/3* HEK293T cells was significantly reduced (36), which led us to expect similar observations *in vivo*. *Poglut2* single knockout and *Poglut3* single knockout animals displayed no obvious phenotypes, which is possibly due to functional redundancy of POGLUT2 and 3 even though we have several examples of site-specificity between these two enzymes (36). Deletion of both *Poglut2* and 3 led to perinatal lethality, reduced size, syndactyly, ECM remodeling in the lungs, and reduced TGF- $\beta$  signaling. Our results had overlapping phenotypes with other mouse models with mutations in *Fbn1*, *Fbn2*, *Ltbp1*, *Ltbp3*, *Ltbp4*, *Fbln4*, and *Fbln5* (112, 114, 116, 137, 147-150). This suggested that loss of POGLUT2/3 *O*-glucosylation on these proteins impacted their function. We clearly demonstrated the pattern and level of FBNs and elastin were disturbed in *Poglut2/3 DKO* lungs, which likely explains the reduction in TGF- $\beta$  signaling we observed since FBN microfibrils are essential for TGF- $\beta$  activation (Figure 5.1). Interestingly, we did not measure a significant reduction in FBN secretion from *Poglut2/3 DKO* lung fibroblasts even though we saw a significant decrease in FBN1 secretion from *Poglut2/3 DKO* dermal fibroblasts. Future research will focus on exploring how loss of POGLUT2/3-mediated *O*-glucosylation lead to dysregulation of TGF- $\beta$  signaling in the context of FBNs and other POGLUT2/3 substrates like FBLNs and LTBP. Several possible mechanisms are illustrated in Figure 5.1. Future experiments will include *in vitro* microfibril incorporation assays, proteolysis assays, additional IHC staining, transcriptome analysis, and proteomics on whole lung tissue to address these possibilities.

*In vitro* microfibril incorporation assays will provide a qualitative analysis of FBN microfibrils in *Poglut2/3 DKO*s. These assays can be performed over a time course to evaluate a rate of microfibril formation. *In vitro* proteolysis assays will assess whether loss of *O*-

glucosylation by POGLUT2/3 enhances substrate susceptibility to enzymatic degradation, which could also explain the decrease in FBN levels we observed with IHC staining of mouse lungs. Enhanced proteolysis of either LTBP1 or FBNs could lead to an overall reduction of latent TGF- $\beta$  incorporation into the ECM, which is required for efficient TGF- $\beta$  activation. Fluorescent IHC staining of LTBP1, -3, and -4 could corroborate this mechanism. If reduced levels of LTBP1 are found in the ECM, then it would suggest *O*-glucosylation by POGLUT2 and 3 permits efficient incorporation of the LTBP- TGF- $\beta$  complex in the ECM. As mentioned previously, the reduction in TGF- $\beta$  signaling in *Poglut2/3 DKO*s likely leads to transcriptional changes in the cell. Transcriptome analysis (RNA-Seq) could be used to verify this. Mass spectral proteomics of whole lung tissue could measure the effect of these transcriptional changes on a protein level, which would provide additional insight into the cardiovascular and lung pathology of *Poglut2/3 DKO*s.





**Figure 5.1. Working model of how loss of *O*-glycosylation by POGLUT2/3 could affect substrate interactions intracellularly and extracellularly that are required for ECM structure and function.** The efficiency of substrate transport through the secretory pathway may be reduced due to the loss of *O*-glucose on POGLUT2/3 substrates, leading to secretion defects of proteins like FBNs and LTBPs. Loss of *O*-glycosylation by POGLUT2/3 could also reduce secretion of the LLC due to reduced LTBP interactions intracellularly. Alternatively, *O*-glucose modifications added by POGLUT2/3 may play a role in microfibril assembly as well as the countless protein interactions that take place in the extracellular space that are required for processes like TGF-β signaling. Model is depicted for LTBPs and FBNs but could apply to other POGLUT2/3 substrates as well. *Red lightning bolt*, defects due to loss of POGLUT2/3-mediated *O*-glycosylation.

## References

1. Takeuchi H, Schneider M, Williamson DB, Ito A, Takeuchi M, Handford PA, et al. Two novel protein O-glucosyltransferases that modify sites distinct from POGLUT1 and affect Notch trafficking and signaling. *Proceedings of the National Academy of Sciences of the United States of America*. 2018;115(36):E8395-E402.
2. Luca VC, Jude KM, Pierce NW, Nachury MV, Fischer S, Garcia KC. Structural biology. Structural basis for Notch1 engagement of Delta-like 4. *Science (New York, NY)*. 2015;347(6224):847-53.
3. Acar M, Jafar-Nejad H, Takeuchi H, Rajan A, Ibrani D, Rana NA, et al. Rumi is a CAP10 domain glycosyltransferase that modifies Notch and is required for Notch signaling. *Cell*. 2008;132(2):247-58.
4. Moloney DJ, Shair LH, Lu FM, Xia J, Locke R, Matta KL, et al. Mammalian Notch1 is modified with two unusual forms of O-linked glycosylation found on epidermal growth factor-like modules. *The Journal of biological chemistry*. 2000;275(13):9604-11.
5. Wang Y, Shao L, Shi S, Harris RJ, Spellman MW, Stanley P, et al. Modification of epidermal growth factor-like repeats with O-fucose. Molecular cloning and expression of a novel GDP-fucose protein O-fucosyltransferase. *The Journal of biological chemistry*. 2001;276(43):40338-45.
6. Matsuura A, Ito M, Sakaidani Y, Kondo T, Murakami K, Furukawa K, et al. O-linked N-acetylglucosamine is present on the extracellular domain of notch receptors. *The Journal of biological chemistry*. 2008;283(51):35486-95.
7. Luo Y, Haltiwanger RS. O-fucosylation of notch occurs in the endoplasmic reticulum. *The Journal of biological chemistry*. 2005;280(12):11289-94.
8. Wang Y, Spellman MW. Purification and characterization of a GDP-fucose:polypeptide fucosyltransferase from Chinese hamster ovary cells. *The Journal of biological chemistry*. 1998;273(14):8112-8.
9. Takeuchi H, Kantharia J, Sethi MK, Bakker H, Haltiwanger RS. 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. *The Journal of biological chemistry*. 2012;287(41):33934-44.
10. Takeuchi H, Yu H, Hao H, Takeuchi M, Ito A, Li H, et al. O-Glycosylation modulates the stability of epidermal growth factor-like repeats and thereby regulates Notch trafficking. *The Journal of biological chemistry*. 2017;292(38):15964-73.
  11. Shi S, Stanley P. Protein O-fucosyltransferase 1 is an essential component of Notch signaling pathways. *Proceedings of the National Academy of Sciences of the United States of America*. 2003;100(9):5234-9.
  12. Fernandez-Valdivia R, Takeuchi H, Samarghandi A, Lopez M, Leonardi J, Haltiwanger RS, et al. Regulation of mammalian Notch signaling and embryonic development by the protein O-glucosyltransferase Rumi. *Development (Cambridge, England)*. 2011;138(10):1925-34.
  13. Williamson DB, Haltiwanger RS. Identification, function, and biological relevance of POGLUT2 and POGLUT3. *Biochem Soc Trans*. 2022;50(2):1003-12.
  14. Hase S, Kawabata S, Nishimura H, Takeya H, Sueyoshi T, Miyata T, et al. A new trisaccharide sugar chain linked to a serine residue in bovine blood coagulation factors VII and IX. *J Biochem*. 1988;104(6):867-8.
  15. Kentzer EJ, Buko A, Menon G, Sarin VK. Carbohydrate composition and presence of a fucose-protein linkage in recombinant human pro-urokinase. *Biochemical and biophysical research communications*. 1990;171(1):401-6.
  16. Holdener BC, Haltiwanger RS. Protein O-fucosylation: structure and function. *Current opinion in structural biology*. 2019;56:78-86.
  17. Yu H, Takeuchi H. Protein O-glucosylation: another essential role of glucose in biology. *Current opinion in structural biology*. 2019;56:64-71.
  18. Saiki W, Ma C, Okajima T, Takeuchi H. Current Views on the Roles of O-Glycosylation in Controlling Notch-Ligand Interactions. *Biomolecules*. 2021;11(2):309.
  19. Haltiwanger RS, Wells L, Freeze HH, Stanley P. Other Classes of Eukaryotic Glycans. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, et al., editors. *Essentials of Glycobiology*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press

20. Harvey BM, Haltiwanger RS. Regulation of Notch Function by O-Glycosylation. *Advances in experimental medicine and biology*. 2018;1066:59-78.
21. Matsumoto K, Luther KB, Haltiwanger RS. Diseases related to Notch glycosylation. *Mol Aspects Med*. 2020;100938.
22. Stanley P. Roles of Notch Glycosylation in Signaling. *The FASEB Journal*. 2020;34(S1):1-.
23. Pandey A, Niknejad N, Jafar-Nejad H. Multifaceted regulation of Notch signaling by glycosylation. *Glycobiology*. 2020;31(1):8-28.
24. Sethi MK, Buettner FF, Krylov VB, Takeuchi H, Nifantiev NE, Haltiwanger RS, et al. Identification of glycosyltransferase 8 family members as xylosyltransferases acting on O-glucosylated notch epidermal growth factor repeats. *The Journal of biological chemistry*. 2010;285(3):1582-6.
25. Sethi MK, Buettner FF, Ashikov A, Krylov VB, Takeuchi H, Nifantiev NE, et al. Molecular cloning of a xylosyltransferase that transfers the second xylose to O-glucosylated epidermal growth factor repeats of notch. *The Journal of biological chemistry*. 2012;287(4):2739-48.
26. Moloney DJ, Panin VM, Johnston SH, Chen J, Shao L, Wilson R, et al. Fringe is a glycosyltransferase that modifies Notch. *Nature*. 2000;406(6794):369-75.
27. Harvey BM, Rana NA, Moss H, Leonardi J, Jafar-Nejad H, Haltiwanger RS. Mapping Sites of O-Glycosylation and Fringe Elongation on Drosophila Notch. *The Journal of biological chemistry*. 2016;291(31):16348-60.
28. Kakuda S, Haltiwanger RS. Deciphering the Fringe-Mediated Notch Code: Identification of Activating and Inhibiting Sites Allowing Discrimination between Ligands. *Developmental cell*. 2017;40(2):193-201.
29. Sakaidani Y, Nomura T, Matsuura A, Ito M, Suzuki E, Murakami K, et al. O-Linked-N-acetylglucosamine on extracellular protein domains mediates epithelial cell-matrix interactions. *Nature Communications*. 2011;2(1):583.

30. Ogawa M, Senoo Y, Ikeda K, Takeuchi H, Okajima T. Structural Divergence in O-GlcNAc Glycans Displayed on Epidermal Growth Factor-like Repeats of Mammalian Notch1. *Molecules*. 2018;23(7).
31. Rana NA, Nita-Lazar A, Takeuchi H, Kakuda S, Luther KB, Haltiwanger RS. O-glucose trisaccharide is present at high but variable stoichiometry at multiple sites on mouse Notch1. *The Journal of biological chemistry*. 2011;286(36):31623-37.
32. Luca VC, Kim BC, Ge C, Kakuda S, Wu D, Roenigk-Peikar M, et al. Notch-Jagged complex structure implicates a catch bond in tuning ligand sensitivity. *Science (New York, NY)*. 2017;355(6331):1320-4.
33. Pandey A, Harvey BM, Lopez MF, Ito A, Haltiwanger RS, Jafar-Nejad H. Glycosylation of Specific Notch EGF Repeats by O-Fut1 and Fringe Regulates Notch Signaling in *Drosophila*. *Cell Rep*. 2019;29(7):2054-66.e6.
34. Varshney S, Wei HX, Batista F, Nauman M, Sundaram S, Siminovitch K, et al. A modifier in the 129S2/SvPasCrl genome is responsible for the viability of Notch1<sup>[12f/12f]</sup> mice. *BMC Dev Biol*. 2019;19(1):19.
35. Stahl M, Uemura K, Ge C, Shi S, Tashima Y, Stanley P. Roles of Pofut1 and O-fucose in mammalian Notch signaling. *The Journal of biological chemistry*. 2008;283(20):13638-51.
36. Williamson DB, Sohn CJ, Ito A, Haltiwanger RS. POGLUT2 and POGLUT3 O-glucosylate multiple EGF repeats in fibrillin-1, -2, and LTBP1 and promote secretion of fibrillin-1. *The Journal of biological chemistry*. 2021;297(3):101055.
37. Takeuchi H, Fernandez-Valdivia RC, Caswell DS, Nita-Lazar A, Rana NA, Garner TP, et al. Rumi functions as both a protein O-glucosyltransferase and a protein O-xylosyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(40):16600-5.
38. Kimata Y, Ooboki K, Nomura-Furuwatari C, Hosoda A, Tsuru A, Kohno K. Identification of a novel mammalian endoplasmic reticulum-resident KDEL protein using an EST database motif search. *Gene*. 2000;261(2):321-7.
39. Andrawes MB, Xu X, Liu H, Ficarro SB, Marto JA, Aster JC, et al. Intrinsic selectivity of Notch 1 for Delta-like 4 over Delta-like 1. *The Journal of biological chemistry*. 2013;288(35):25477-89.

40. Walker C, Mojares E, Del Río Hernández A. Role of Extracellular Matrix in Development and Cancer Progression. *Int J Mol Sci.* 2018;19(10).
41. Godwin ARF, Singh M, Lockhart-Cairns MP, Alanazi YF, Cain SA, Baldock C. The role of fibrillin and microfibril binding proteins in elastin and elastic fibre assembly. *Matrix biology : journal of the International Society for Matrix Biology.* 2019;84:17-30.
42. Thomson J, Singh M, Eckersley A, Cain SA, Sherratt MJ, Baldock C. Fibrillin microfibrils and elastic fibre proteins: Functional interactions and extracellular regulation of growth factors. *Seminars in cell & developmental biology.* 2019;89:109-17.
43. Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nature Communications.* 2020;11(1):5120.
44. Rifkin DB, Rifkin WJ, Zilberberg L. LTBP in biology and medicine: LTBP diseases. *Matrix biology : journal of the International Society for Matrix Biology.* 2018;71-72:90-9.
45. Sakai LY, Keene DR, Renard M, De Backer J. FBN1: The disease-causing gene for Marfan syndrome and other genetic disorders. *Gene.* 2016;591(1):279-91.
46. Carta L, Pereira L, Arteaga-Solis E, Lee-Arteaga SY, Lenart B, Starcher B, et al. Fibrillins 1 and 2 perform partially overlapping functions during aortic development. *The Journal of biological chemistry.* 2006;281(12):8016-23.
47. Todorovic V, Frendewey D, Gutstein DE, Chen Y, Freyer L, Finnegan E, et al. Long form of latent TGF-beta binding protein 1 (Ltbp1L) is essential for cardiac outflow tract septation and remodeling. *Development (Cambridge, England).* 2007;134(20):3723-32.
48. Isogai Z, Ono RN, Ushiro S, Keene DR, Chen Y, Mazzieri R, et al. Latent Transforming Growth Factor  $\beta$ -binding Protein 1 Interacts with Fibrillin and Is a Microfibril-associated Protein\*. *Journal of Biological Chemistry.* 2003;278(4):2750-7.
49. Neptune ER, Frischmeyer PA, Arking DE, Myers L, Bunton TE, Gayraud B, et al. Dysregulation of TGF-beta activation contributes to pathogenesis in Marfan syndrome. *Nature genetics.* 2003;33(3):407-11.

50. Massam-Wu T, Chiu M, Choudhury R, Chaudhry SS, Baldwin AK, McGovern A, et al. Assembly of fibrillin microfibrils governs extracellular deposition of latent TGF $\beta$ . *Journal of Cell Science*. 2010;123(17):3006-18.
51. Creamer TJ, Bramel EE, MacFarlane EG. Insights on the Pathogenesis of Aneurysm through the Study of Hereditary Aortopathies. *Genes (Basel)*. 2021;12(2).
52. Ono RN, Sengle G, Charbonneau NL, Carlberg V, Bächinger HP, Sasaki T, et al. Latent transforming growth factor beta-binding proteins and fibulins compete for fibrillin-1 and exhibit exquisite specificities in binding sites. *The Journal of biological chemistry*. 2009;284(25):16872-81.
53. Brunner AM, Marquardt H, Malacko AR, Lioubin MN, Purchio AF. Site-directed mutagenesis of cysteine residues in the pro region of the transforming growth factor beta 1 precursor. Expression and characterization of mutant proteins. *The Journal of biological chemistry*. 1989;264(23):13660-4.
54. Miyazono K, Olofsson A, Colosetti P, Heldin CH. A role of the latent TGF-beta 1-binding protein in the assembly and secretion of TGF-beta 1. *Embo j*. 1991;10(5):1091-101.
55. Dallas SL, Sivakumar P, Jones CJ, Chen Q, Peters DM, Mosher DF, et al. Fibronectin regulates latent transforming growth factor-beta (TGF beta) by controlling matrix assembly of latent TGF beta-binding protein-1. *The Journal of biological chemistry*. 2005;280(19):18871-80.
56. Zeigler SM, Sloan B, Jones JA. Pathophysiology and Pathogenesis of Marfan Syndrome. In: Halper J, editor. *Progress in Heritable Soft Connective Tissue Diseases*. Cham: Springer International Publishing; 2021. p. 185-206.
57. Wheeler JB, Ikonomidis JS, Jones JA. Connective Tissue Disorders and Cardiovascular Complications: The Indomitable Role of Transforming Growth Factor- $\beta$  Signaling. In: Halper J, editor. *Progress in Heritable Soft Connective Tissue Diseases*. Cham: Springer International Publishing; 2021. p. 161-84.
58. Glanville RW, Qian RQ, McClure DW, Maslen CL. Calcium binding, hydroxylation, and glycosylation of the precursor epidermal growth factor-like domains of fibrillin-1, the Marfan gene protein. *The Journal of biological chemistry*. 1994;269(43):26630-4.

59. Kakuda S, LoPilato RK, Ito A, Haltiwanger RS. Canonical Notch ligands and Fringes have distinct effects on NOTCH1 and NOTCH2. *The Journal of biological chemistry*. 2020;295(43):14710-22.
60. Li Z, Han K, Pak JE, Satkunarajah M, Zhou D, Rini JM. Recognition of EGF-like domains by the Notch-modifying O-fucosyltransferase POFUT1. *Nature chemical biology*. 2017;13(7):757-63.
61. Dietz HC, Cutting GR, Pyeritz RE, Maslen CL, Sakai LY, Corson GM, et al. Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. *Nature*. 1991;352(6333):337-9.
62. Dietz HC, Pyeritz RE, Hall BD, Cadle RG, Hamosh A, Schwartz J, et al. The Marfan syndrome locus: confirmation of assignment to chromosome 15 and identification of tightly linked markers at 15q15-q21.3. *Genomics*. 1991;9(2):355-61.
63. Maslen CL, Corson GM, Maddox BK, Glanville RW, Sakai LY. Partial sequence of a candidate gene for the Marfan syndrome. *Nature*. 1991;352(6333):334-7.
64. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, et al. Human Gene Mutation Database (HGMD): 2003 update. *Hum Mutat*. 2003;21(6):577-81.
65. Stengl R, Bors A, Ágg B, Pólos M, Matyas G, Molnár MJ, et al. Optimising the mutation screening strategy in Marfan syndrome and identifying genotypes with more severe aortic involvement. *Orphanet J Rare Dis*. 2020;15(1):290.
66. Collod-Bérout G, Le Bourdelles S, Ades L, Ala-Kokko L, Booms P, Boxer M, et al. Update of the UMD-FBN1 mutation database and creation of an FBN1 polymorphism database. *Hum Mutat*. 2003;22(3):199-208.
67. Dawson A, Li Y, Li Y, Ren P, Vasquez HG, Zhang C, et al. Single-Cell Analysis of Aneurysmal Aortic Tissue in Patients with Marfan Syndrome Reveals Dysfunctional TGF- $\beta$  Signaling. *Genes*. 2022;13(1):95.
68. Halper J. Basic Components of Connective Tissues and Extracellular Matrix: Fibronectin, Fibrinogen, Laminin, Elastin, Fibrillins, Fibulins, Matrilins, Tenascins and Thrombospondins. *Advances in experimental medicine and biology*. 2021;1348:105-26.



69. Corson GM, Chalberg SC, Dietz HC, Charbonneau NL, Sakai LY. Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure and alternatively spliced exons at the 5' end. *Genomics*. 1993;17(2):476-84.
70. Downing AK, Knott V, Werner JM, Cardy CM, Campbell ID, Handford PA. Solution structure of a pair of calcium-binding epidermal growth factor-like domains: implications for the Marfan syndrome and other genetic disorders. *Cell*. 1996;85(4):597-605.
71. Yuan X, Downing AK, Knott V, Handford PA. Solution structure of the transforming growth factor beta-binding protein-like module, a domain associated with matrix fibrils. *Embo j*. 1997;16(22):6659-66.
72. Suk JY, Jensen S, McGettrick A, Willis AC, Whiteman P, Redfield C, et al. Structural consequences of cysteine substitutions C1977Y and C1977R in calcium-binding epidermal growth factor-like domain 30 of human fibrillin-1. *The Journal of biological chemistry*. 2004;279(49):51258-65.
73. Jensen SA, Atwa O, Handford PA. Assembly assay identifies a critical region of human fibrillin-1 required for 10-12 nm diameter microfibril biogenesis. *PLoS One*. 2021;16(3):e0248532.
74. Jensen SA, Iqbal S, Bulsiewicz A, Handford PA. A microfibril assembly assay identifies different mechanisms of dominance underlying Marfan syndrome, stiff skin syndrome and acromelic dysplasias. *Human molecular genetics*. 2015;24(15):4454-63.
75. Handford PA. Fibrillin-1, a calcium binding protein of extracellular matrix. *Biochimica et biophysica acta*. 2000;1498(2-3):84-90.
76. Reinhardt DP, Ono RN, Notbohm H, Müller PK, Bächinger HP, Sakai LY. Mutations in Calcium-binding Epidermal Growth Factor Modules Render Fibrillin-1 Susceptible to Proteolysis: A POTENTIAL DISEASE-CAUSING MECHANISM IN MARFAN SYNDROME \*. *Journal of Biological Chemistry*. 2000;275(16):12339-45.
77. Kirschner R, Hubmacher D, Iyengar G, Kaur J, Fagotto-Kaufmann C, Brömme D, et al. Classical and neonatal Marfan syndrome mutations in fibrillin-1 cause differential protease susceptibilities and protein function. *The Journal of biological chemistry*. 2011;286(37):32810-23.

78. Haller SJ, Roitberg AE, Dudley AT. Steered molecular dynamic simulations reveal Marfan syndrome mutations disrupt fibrillin-1 cbEGF domain mechanosensitive calcium binding. *Scientific reports*. 2020;10(1):16844.
79. Rees DJ, Jones IM, Handford PA, Walter SJ, Esnouf MP, Smith KJ, et al. The role of beta-hydroxyaspartate and adjacent carboxylate residues in the first EGF domain of human factor IX. *Embo j*. 1988;7(7):2053-61.
80. Handford PA, Mayhew M, Baron M, Winship PR, Campbell ID, Brownlee GG. Key residues involved in calcium-binding motifs in EGF-like domains. *Nature*. 1991;351(6322):164-7.
81. Handford PA, Mayhew M, Brownlee GG. Calcium binding to fibrillin? *Nature*. 1991;353(6343):395-.
82. Stenflo J, Ohlin AK, Owen WG, Schneider WJ. beta-Hydroxyaspartic acid or beta-hydroxyasparagine in bovine low density lipoprotein receptor and in bovine thrombomodulin. *The Journal of biological chemistry*. 1988;263(1):21-4.
83. Shang HS, Lu HF, Lee CH, Chiang HS, Chu YL, Chen A, et al. Quercetin induced cell apoptosis and altered gene expression in AGS human gastric cancer cells. *Environmental toxicology*. 2018;33(11):1168-81.
84. Tsai YL, Chang HH, Chen YC, Chang YC, Chen Y, Tsai WC. Molecular Mechanisms of KDELC2 on Glioblastoma Tumorigenesis and Temozolomide Resistance. *Biomedicines*. 2020;8(9).
85. Li C, Yao Y, Long D, Lin X. KDELC1 and TRMT1 Serve as Prognosis-Related SARS-CoV-2 Proteins Binding Human mRNAs and Promising Biomarkers in Clear Cell Renal Cell Carcinoma. *Int J Gen Med*. 2021;14:2475-90.
86. Wang J, Ma X, Ma J. Identification of Four Enhancer-Associated Genes as Risk Signature for Diffuse Glioma Patients. *J Mol Neurosci*. 2021.
87. Yan K, Wang Y, Lu Y, Yan Z. Coexpressed Genes That Promote the Infiltration of M2 Macrophages in Melanoma Can Evaluate the Prognosis and Immunotherapy Outcome. *J Immunol Res*. 2021;2021:6664791.

88. Carreras J, Nakamura N, Hamoudi R. Artificial Intelligence Analysis of Gene Expression Predicted the Overall Survival of Mantle Cell Lymphoma and a Large Pan-Cancer Series. *Healthcare (Basel)*. 2022;10(1).
89. Grossman RL, Heath AP, Ferretti V, Varmus HE, Lowy DR, Kibbe WA, et al. Toward a Shared Vision for Cancer Genomic Data. *N Engl J Med*. 2016;375(12):1109-12.
90. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res*. 2004;14(6):1188-90.
91. Zhang H, Hu W, Ramirez F. Developmental expression of fibrillin genes suggests heterogeneity of extracellular microfibrils. *Journal of Cell Biology*. 1995;129(4):1165-76.
92. Unsöld C, Hyytiäinen M, Bruckner-Tuderman L, Keski-Oja J. Latent TGF-beta binding protein LTBP-1 contains three potential extracellular matrix interacting domains. *J Cell Sci*. 2001;114(Pt 1):187-97.
93. Sengle G, Tsutsui K, Keene DR, Tufa SF, Carlson EJ, Charbonneau NL, et al. Microenvironmental regulation by fibrillin-1. *PLoS Genet*. 2012;8(1):e1002425.
94. Sengle G, Sakai LY. The fibrillin microfibril scaffold: A niche for growth factors and mechanosensation? *Matrix Biology*. 2015;47:3-12.
95. Chaudhry SS, Cain SA, Morgan A, Dallas SL, Shuttleworth CA, Kielty CM. Fibrillin-1 regulates the bioavailability of TGFβ1. *Journal of Cell Biology*. 2007;176(3):355-67.
96. Wipff P-J, Hinz B. Integrins and the activation of latent transforming growth factor β1 – An intimate relationship. *European Journal of Cell Biology*. 2008;87(8):601-15.
97. Doyle JJ, Gerber EE, Dietz HC. Matrix-dependent perturbation of TGFβ signaling and disease. *FEBS Letters*. 2012;586(14):2003-15.
98. Cardy CM, Handford PA. Metal ion dependency of microfibrils supports a rod-like conformation for fibrillin-1 calcium-binding epidermal growth factor-like domains. *Journal of molecular biology*. 1998;276(5):855-60.

99. Reinhardt DP, Keene DR, Corson GM, Pöschl E, Bächinger HP, Gammie JE, et al. Fibrillin-1: organization in microfibrils and structural properties. *Journal of molecular biology*. 1996;258(1):104-16.
100. Lin G, Tiedemann K, Vollbrandt T, Peters H, Bätge B, Brinckmann J, et al. Homo- and Heterotypic Fibrillin-1 and -2 Interactions Constitute the Basis for the Assembly of Microfibrils\*. *Journal of Biological Chemistry*. 2002;277(52):50795-804.
101. Shao L, Luo Y, Moloney DJ, Haltiwanger R. O-glycosylation of EGF repeats: identification and initial characterization of a UDP-glucose: protein O-glucosyltransferase. *Glycobiology*. 2002;12(11):763-70.
102. Zhang A, Berardinelli SJ, Leonhard-Melief C, Vasudevan D, Liu TW, Taibi A, et al. O-Fucosylation of ADAMTSL2 is required for secretion and is impacted by geleophysic dysplasia-causing mutations. *The Journal of biological chemistry*. 2020;295(46):15742-53.
103. Pecori F, Yokota I, Hanamatsu H, Miura T, Ogura C, Ota H, et al. A defined glycosylation regulatory network modulates total glycome dynamics during pluripotency state transition. *Scientific reports*. 2021;11(1):1276.
104. Braakman I, Hebert DN. Protein folding in the endoplasmic reticulum. *Cold Spring Harb Perspect Biol*. 2013;5(5):a013201.
105. Jensen SA, Handford PA. New insights into the structure, assembly and biological roles of 10-12 nm connective tissue microfibrils from fibrillin-1 studies. *The Biochemical journal*. 2016;473(7):827-38.
106. Kettle S, Yuan X, Grundy G, Knott V, Downing AK, Handford PA. Defective calcium binding to fibrillin-1: consequence of an N2144S change for fibrillin-1 structure and function. *Journal of molecular biology*. 1999;285(3):1277-87.
107. Bern MW, Kil YJ. Two-Dimensional Target Decoy Strategy for Shotgun Proteomics. *Journal of Proteome Research*. 2011;10(12):5296-301.
108. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Research*. 2018;47(D1):D442-D50.

109. Urata Y, Takeuchi H. Effects of Notch glycosylation on health and diseases. *Dev Growth Differ*. 2020;62(1):35-48.
110. Asano K, Cantalupo A, Sedes L, Ramirez F. The Multiple Functions of Fibrillin-1 Microfibrils in Organismal Physiology. *Int J Mol Sci*. 2022;23(3).
111. Peeters S, De Kinderen P, Meester JAN, Verstraeten A, Loeys BL. The fibrillinopathies: New insights with focus on the paradigm of opposing phenotypes for both FBN1 and FBN2. *Hum Mutat*. 2022.
112. Rifkin D, Sachan N, Singh K, Sauber E, Tellides G, Ramirez F. The role of LTBP3 in TGFβ signaling. *Dev Dyn*. 2022;251(1):95-104.
113. Horiguchi M, Todorovic V, Hadjiolova K, Weiskirchen R, Rifkin DB. Abrogation of both short and long forms of latent transforming growth factor-β binding protein-1 causes defective cardiovascular development and is perinatally lethal. *Matrix biology : journal of the International Society for Matrix Biology*. 2015;43:61-70.
114. Colarossi C, Chen Y, Obata H, Jurukovski V, Fontana L, Dabovic B, et al. Lung alveolar septation defects in *Ltbp-3*-null mice. *Am J Pathol*. 2005;167(2):419-28.
115. Dabovic B, Robertson IB, Zilberberg L, Vassallo M, Davis EC, Rifkin DB. Function of latent TGFβ binding protein 4 and fibulin 5 in elastogenesis and lung development. *J Cell Physiol*. 2015;230(1):226-36.
116. Chaudhry SS, Gazzard J, Baldock C, Dixon J, Rock MJ, Skinner GC, et al. Mutation of the gene encoding fibrillin-2 results in syndactyly in mice. *Human molecular genetics*. 2001;10(8):835-43.
117. Bult CJ, Blake JA, Smith CL, Kadin JA, Richardson JE. Mouse Genome Database (MGD) 2019. *Nucleic Acids Res*. 2019;47(D1):D801-d6.
118. Hussain M, Xu C, Lu M, Wu X, Tang L, Wu X. Wnt/β-catenin signaling links embryonic lung development and asthmatic airway remodeling. *Biochim Biophys Acta Mol Basis Dis*. 2017;1863(12):3226-42.
119. Kloc M, Ghobrial RM. Chronic allograft rejection: A significant hurdle to transplant success. *Burns Trauma*. 2014;2(1):3-10.

120. Mescher AL. The Respiratory System. Junqueira's Basic Histology: Text and Atlas, 15e. New York, NY: McGraw-Hill Education; 2018.
121. Morrissey EE, Hogan BL. Preparing for the first breath: genetic and cellular mechanisms in lung development. *Developmental cell*. 2010;18(1):8-23.
122. Trask TM, Trask BC, Ritty TM, Abrams WR, Rosenbloom J, Mecham RP. Interaction of tropoelastin with the amino-terminal domains of fibrillin-1 and fibrillin-2 suggests a role for the fibrillins in elastic fiber assembly. *The Journal of biological chemistry*. 2000;275(32):24400-6.
123. Mižíková I, Morty RE. The Extracellular Matrix in Bronchopulmonary Dysplasia: Target and Source. *Front Med (Lausanne)*. 2015;2:91.
124. Tschanz SA, Salm LA, Roth-Kleiner M, Barré SF, Burri PH, Schittny JC. Rat lungs show a biphasic formation of new alveoli during postnatal development. *J Appl Physiol* (1985). 2014;117(1):89-95.
125. Robertson IB, Dias HF, Osuch IH, Lowe ED, Jensen SA, Redfield C, et al. The N-Terminal Region of Fibrillin-1 Mediates a Bipartite Interaction with LTBP1. *Structure (London, England : 1993)*. 2017;25(8):1208-21 e5.
126. Yoshinaga K, Obata H, Jurukovski V, Mazziere R, Chen Y, Zilberberg L, et al. Perturbation of transforming growth factor (TGF)-beta1 association with latent TGF-beta binding protein yields inflammation and tumors. *Proceedings of the National Academy of Sciences of the United States of America*. 2008;105(48):18758-63.
127. Tzavlaki K, Moustakas A. TGF- $\beta$  Signaling. *Biomolecules*. 2020;10(3):487.
128. Nakao A, Imamura T, Souchelnytskyi S, Kawabata M, Ishisaki A, Oeda E, et al. TGF-beta receptor-mediated signalling through Smad2, Smad3 and Smad4. *Embo j*. 1997;16(17):5353-62.
129. Chen H, Zhuang F, Liu YH, Xu B, Del Moral P, Deng W, et al. TGF-beta receptor II in epithelia versus mesenchyme plays distinct roles in the developing lung. *Eur Respir J*. 2008;32(2):285-95.

130. Miao Q, Chen H, Luo Y, Chiu J, Chu L, Thornton ME, et al. Abrogation of mesenchyme-specific TGF- $\beta$  signaling results in lung malformation with prenatal pulmonary cysts in mice. *Am J Physiol Lung Cell Mol Physiol*. 2021;320(6):L1158-168.
131. Orsburn BC. Proteome Discoverer-A Community Enhanced Data Processing Suite for Protein Informatics. *Proteomes*. 2021;9(1).
132. Eckersley A, Mellody KT, Pilkington S, Griffiths CEM, Watson REB, O'Cualain R, et al. Structural and compositional diversity of fibrillin microfibrils in human tissues. *The Journal of biological chemistry*. 2018;293(14):5117-33.
133. Ashworth JL, Kelly V, Wilson R, Shuttleworth CA, Kielty CM. Fibrillin assembly: dimer formation mediated by amino-terminal sequences. *J Cell Sci*. 1999;112 ( Pt 20):3549-58.
134. Ponamarczuk H, Popielarski M, Stasiak M, Bednarek R, Studzian M, Pulaski L, et al. Contribution of activated beta3 integrin in the PDI release from endothelial cells. *Front Biosci (Landmark Ed)*. 2018;23(9):1612-27.
135. Tiedemann K, Sasaki T, Gustafsson E, Göhring W, Bätge B, Notbohm H, et al. Microfibrils at basement membrane zones interact with perlecan via fibrillin-1. *The Journal of biological chemistry*. 2005;280(12):11404-12.
136. Kielty CM, Whittaker SP, Shuttleworth CA. Fibrillin: evidence that chondroitin sulphate proteoglycans are components of microfibrils and associate with newly synthesised monomers. *FEBS Lett*. 1996;386(2-3):169-73.
137. Pereira L, Lee SY, Gayraud B, Andrikopoulos K, Shapiro SD, Bunton T, et al. Pathogenetic sequence for aneurysm revealed in mice underexpressing fibrillin-1. *Proceedings of the National Academy of Sciences of the United States of America*. 1999;96(7):3819-23.
138. Jespersen K, Liu Z, Li C, Harding P, Sestak K, Batra R, et al. Enhanced Notch3 signaling contributes to pulmonary emphysema in a Murine Model of Marfan syndrome. *Scientific reports*. 2020;10(1):10949.
139. Papke CL, Yanagisawa H. Fibulin-4 and fibulin-5 in elastogenesis and beyond: Insights from mouse and human studies. *Matrix biology : journal of the International Society for Matrix Biology*. 2014;37:142-9.

140. Wang R, de Kort BJ, Smits AIPM, Weiss AS. Elastin in Vascular Grafts. In: Walpoth B, Bergmeister H, Bowlin G, Kong D, Rotmans J, Zilla P, editors. *Tissue-Engineered Vascular Grafts*. Cham: Springer International Publishing; 2019. p. 1-32.
141. Miller G, Neilan M, Chia R, Gheryani N, Holt N, Charbit A, et al. ENU mutagenesis reveals a novel phenotype of reduced limb strength in mice lacking fibrillin 2. *PLoS One*. 2010;5(2):e9137.
142. Yin W, Kim HT, Wang S, Gunawan F, Li R, Buettner C, et al. Fibrillin-2 is a key mediator of smooth muscle extracellular matrix homeostasis during mouse tracheal tubulogenesis. *Eur Respir J*. 2019;53(3).
143. Zhang X, Alanazi YF, Jowitt TA, Roseman AM, Baldock C. Elastic Fibre Proteins in Elastogenesis and Wound Healing. *Int J Mol Sci*. 2022;23(8).
144. Choudhury R, McGovern A, Ridley C, Cain SA, Baldwin A, Wang MC, et al. Differential regulation of elastic fiber formation by fibulin-4 and -5. *The Journal of biological chemistry*. 2009;284(36):24553-67.
145. Neupane S, Berardinelli SJ, Cameron DC, Grady RC, Komatsu DE, Percival CJ, et al. O-fucosylation of thrombospondin type 1 repeats is essential for ECM remodeling and signaling during bone development. *Matrix biology : journal of the International Society for Matrix Biology*. 2022;107:77-96.
146. Neupane S, Goto J, Berardinelli SJ, Ito A, Haltiwanger RS, Holdener BC. Hydrocephalus in mouse B3glct mutants is likely caused by defects in multiple B3GLCT substrates in ependymal cells and subcommissural organ. *Glycobiology*. 2021;31(8):988-1004.
147. Yanagisawa H, Davis EC, Starcher BC, Ouchi T, Yanagisawa M, Richardson JA, et al. Fibulin-5 is an elastin-binding protein essential for elastic fibre development in vivo. *Nature*. 2002;415(6868):168-71.
148. McLaughlin PJ, Chen Q, Horiguchi M, Starcher BC, Stanton JB, Broekelmann TJ, et al. Targeted disruption of fibulin-4 abolishes elastogenesis and causes perinatal lethality in mice. *Mol Cell Biol*. 2006;26(5):1700-9.
149. Bultmann-Mellin I, Conradi A, Maul AC, Dinger K, Wempe F, Wohl AP, et al. Modeling autosomal recessive cutis laxa type 1C in mice reveals distinct functions for Ltbp-4 isoforms. *Disease models & mechanisms*. 2015;8(4):403-15.



150. Sterner-Kock A, Thorey IS, Koli K, Wempe F, Otte J, Bangsow T, et al. Disruption of the gene encoding the latent transforming growth factor-beta binding protein 4 (LTBP-4) causes abnormal lung development, cardiomyopathy, and colorectal cancer. *Genes Dev.* 2002;16(17):2264-73.