

Supplemental Figure 1

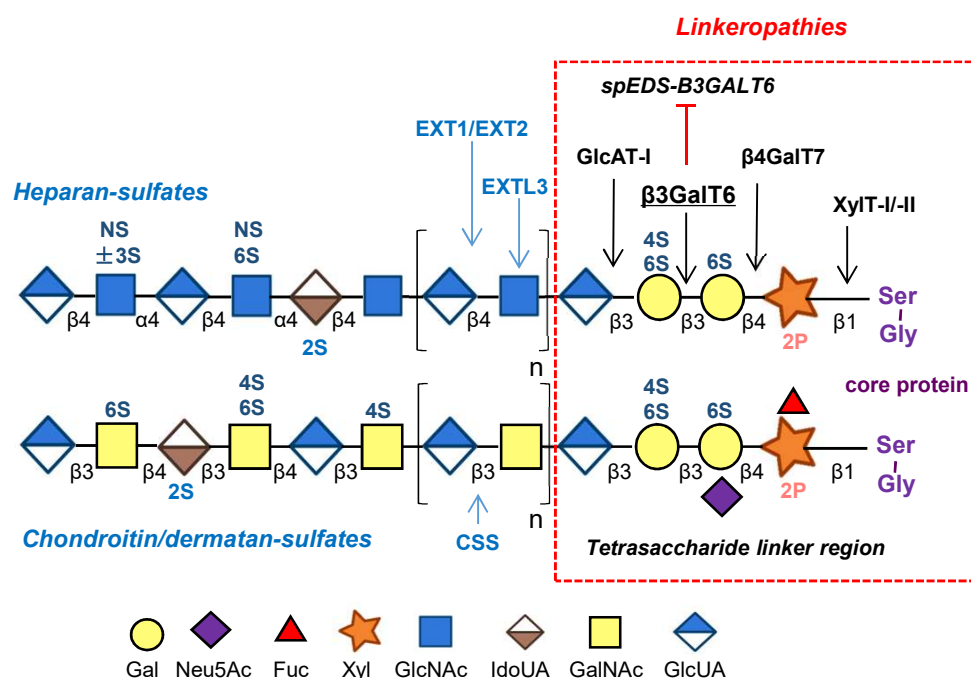


Figure 1. Schematic representation of GAG synthesis and associated “linkeropathies”. From right to left, each saccharide residue is sequentially added onto the PG core-protein by a reaction catalyzed by a different GT. The GTs involved in the formation of the tetrasaccharide linker region and associated genetic diseases called “linkeropathies” are framed in a red dotted box. The linkeropathies associated with *B3GALT6* mutations include spEDS, EDS-like, spondyloepimetaphyseal dysplasia with joint laxity type 1 (SEMD-JL1) and Al-Gazali syndrome. XylT-I/-II, xylosyltransferase I/II; $\beta 4$ GalT7, $\beta 1,4$ -galactosyltransferase 7; $\beta 3$ GalT6, $\beta 1,3$ -galactosyltransferase 6; GlcAT-I, $\beta 1,3$ -glucuronosyltransferase 3; EXTL3, exotosin-like glycosyltransferase 3; EXT1, exotosin glycosyltransferase 1; EXT2, exotosin glycosyltransferase 2; CSS, chondroitin sulfate synthase; NS, N-deacetylated-N-sulfate; S, sulfate and the digit before S corresponds to the position of the sulfate group on the sugar ring; 2P, 2-O-phosphate; Xyl, xylose; Gal, galactose; GlcUA, glucuronic acid; GalNAc, N-acetylgalactosamine; IdoUA, iduronic acid; GlcNAc, N-acetylglucosamine; Fuc, fucose; Neu5Ac, N-acetylneuraminic acid.

Supplemental Figure 2

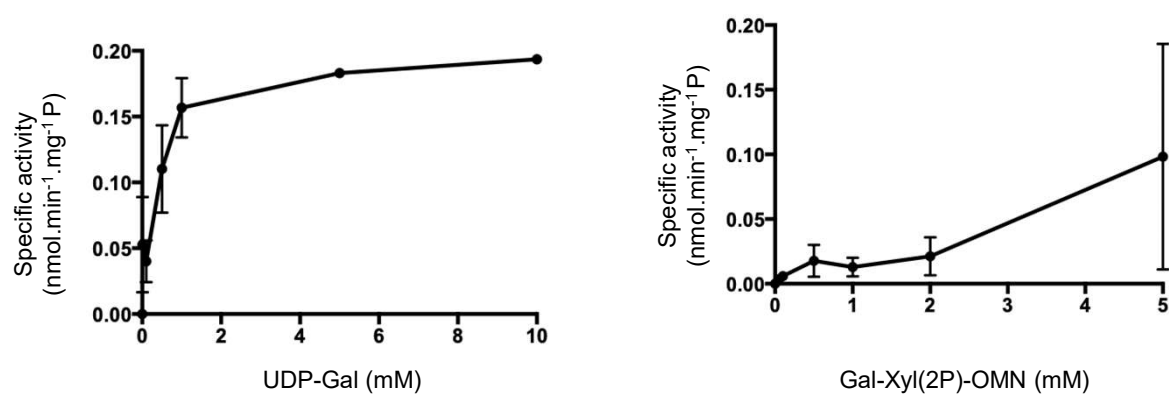


Figure 2. Kinetics of MBP- β 3GalT6 Δ N_{ter},50-WT towards the donor substrate UDP-Gal (left panel) and the acceptor substrate Gal-Xyl(2P)-OMN (right panel). Data are expressed as mean $\pm$ SD (n=3 experiments).

A

H1 B1 H2 B2 H3
MKLLRRRAWRRRAALGLGTLALCGAALLYLARCAAEPGDPRAMSGRSPPPPPAPARAAAFVLVLVASAPRAAERRSVIRSTWLARRGAPGDVWARFAVGTAGLGAERRALE

B3 H4 B4 B5 H5 H6 B6 H B7
REQARHGDL LLLPALRDAYENLTAKVLAMLAWLDEHVAFEFVLKADDDSFARLDALLAELRAREPARRRRRLYWGFFSGRGRVKPGGRWREAAWQLCDYYLPYALGGGYVL

H7 H8 B8 B9 H9
SADLVHYLR LSRDYLRRAWHSEDVSLGAWLAPVDVQREHDPFRDTEYRSRGCSNQYLVT HKQSL EDMLEKHATLAREGR LCKREVQLRLSYVYDWSAPPSQCCQRREGIP

B

Diagram of the protein structure showing helices (H1-H9) and beta-strands (B1-B9). The N-terminus (N) and C-terminus (C) are indicated.

C

Sequence alignment of hB3GALT6 and hB3GNT2. The alignment shows the protein structure with helices (H1-H9) and beta-strands (B1-B9). The N-terminus (N) and C-terminus (C) are indicated.

hB3GALT6 DEHV-AFEFVLKADDDSFARLDALLAELRAREPARRRRRLYWGFFSGRGRVKP--GGRWRE 200
hB3GNT2 STSCPDTEFVFVKGDDDFVFNTHILNLSLTKAKDLFIGDVIHNAGPHRDKCLKLYYI 291

hB3GALT6 AAWQLCDYYLPYALGGGYVLSADLVHYLR LSRDYLRRAWHSEDVSLGAWLAPVDVQREHDP 260
hB3GNT2 PEVVYSGLYPPYAGGGGFLYSGHLALRLYHITDQVHLYPIDDVYTGMC LQKLGLVPEKHK 351

Key residues and motifs are highlighted: D156, D158, Y182C, P194, G188, C206, P211, G216, B7, alpha7, alpha8, K193, D207H, G217S, D242.

(A) Primary sequence and secondary structure of $\beta 3\text{GalT6}$. Blue arrows and cylinders represent β strands (B) and α helices (H), respectively. Pathogenic variants p.(Y182C), p.(D207H) and p.(G217S) are indicated in red. **(B)** Topology diagram of $\beta 3\text{GalT6}$. Arrows and cylinders represent β strands (B) and α helices (H), respectively. **(C)** Sequence alignment of the catalytic site of the human $\beta 3\text{GalT6}$ (hB3GALT6, UniProtKB: Q96L58) and human $\beta 1,3$ -N-acetylglucosaminyltransferase 2 (hB3GNT2, UniProtKB: Q9NY97) using Clustal Omega software (29). Residues are coloured by the software according to their biochemical properties. Asterisks (*) indicate a fully conserved residue, colons (:) indicate strong similarity and periods (.) indicate weak similarity. β strands and α helices are indicated above the $\beta 3\text{GalT6}$ sequence by black dotted lines. Cyan dotted lines indicate two important bonds of $\beta 3\text{GalT6}$, respectively a hydrogen bond between D156 and G216 and a salt bridge between K193 and D242. Black full line indicates a long loop between G188 and P211 of $\beta 3\text{GalT6}$, with an orange plain line indicating a high fluctuation domain within this loop. The DxD motif is represented by a black full and bold line. Red and bold annotations indicate the three point mutations of $\beta 3\text{GalT6}$.

Supplemental Figure 4

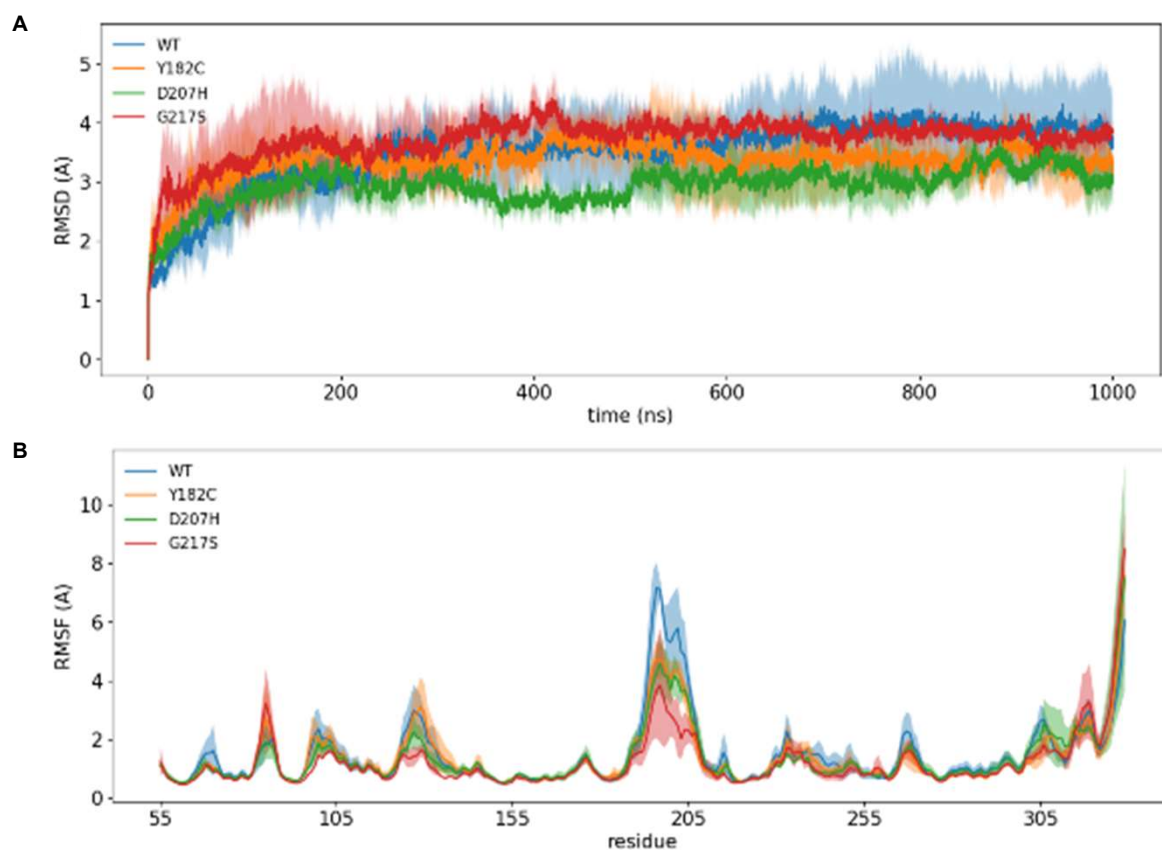


Figure 4. The average RMSD and RMSF of the studied systems from MD simulations. (A) The RMSD from the equilibrated structure was computed on the backbone (C, Ca, N, O) atoms and averaged over all the three replicates of WT and mutants. The average values are shown as lines and the shades correspond to the standard deviations. (B) The RMSF was measured on the backbone (C, Ca, N, O) atoms with respect to the average conformation and averaged by residue, considering the last 900 ns of the MD simulations. The values are averaged over the three replicates of each system. The average values are shown as lines and the shades correspond to the standard deviations.

Supplemental Figure 5

A

HeLa	GCCGCCGCCTTCCTGGCAGT	GCTGGTGGCCAGCGCGCCCCGCGCCGAGCGCCGAGC
H7	GCCGCCGCCTTCCTGGCAG	-----
HeLa	GTGATCCGCAGCACGTGGCTT	GCGCGGCGCGGGGCCCCGGGCGACGTGTGGGCGCGCTTT
H7	-----CACGTGGCTT	GCGCGGCGCGGGGCCCCGGGCGACGTGTGGGCGCGCTTT
HeLa	GCCGTGGGCACGGCCGGCCT	TGGGCGCCGAGGAGCGGCGGCCCTGGAGCGGGAGCAGGCG
H7	GCCGTGGGCACGGCCGGCCT	TGGGCGCCGAGGAGCGGCGGCCCTGGAGCGGGAGCAGGCG
HeLa	CGGCACGGGGACCTGCTGCT	GCTGCCGCGCTGCGCGACGCCTACGAAAACCTCACGGCC
H7	CGGCACGGGGACCTGCTGCT	GCTGCCGCGCTGCGCGACGCCTACGAAAACCTCACGGCC
HeLa	AAGGTGCT	
H7	AAGGTGCT	

B

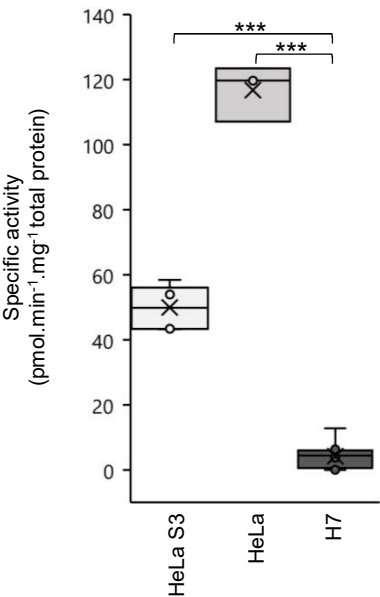


Figure 5. Invalidation of *B3GALT6* gene in *B3GALT6*-KO H7 cells and in vitro enzymatic activity of $\beta 3\text{GalT6}$. (A) Sequence alignment of *B3GALT6*-KO H7 clone with HeLa cells (WT). Genomic DNA was extracted from *B3GALT6*-KO H7 and HeLa cells (WT) and the CRISPR-Cas9 target region was amplified by PCR. The PCR products were subcloned into a cloning vector and sent for sequencing. The alignment reveals a 52pb deletion in the *B3GALT6* sequence. The cDNA region including the sequence targeted by the single guide RNA was amplified by PCR and sequenced by Sanger method. Chromatograms of the PCR products from control and *B3galT6*-KO ATDC5 showed a heterogeneous profile near the PAM sequence, indicating the presence of indel events. (B) Analysis of in vitro $\beta 3\text{GalT6}$ activity in total cell lysates from HeLa cells and *B3GALT6*-KO H7 cells. Enzymatic assays were performed using 1 mM acceptor substrate (Gal-Xyl(2P)-OMN), 5 mM donor substrate (UDP-Gal) and 10 μg total cell lysates. HeLa cells (2 cell lines) are considered as positive control. Data are presented as box plot with the mean symbolised by an X (n=2 experiments for HeLa S3, n=1 experiment for HeLa and n=3 experiments for H7, one triplicate per cell lysate); *** $p < 0.001$ when compared to H7 clone (Student's t test).

Supplemental Figure 6

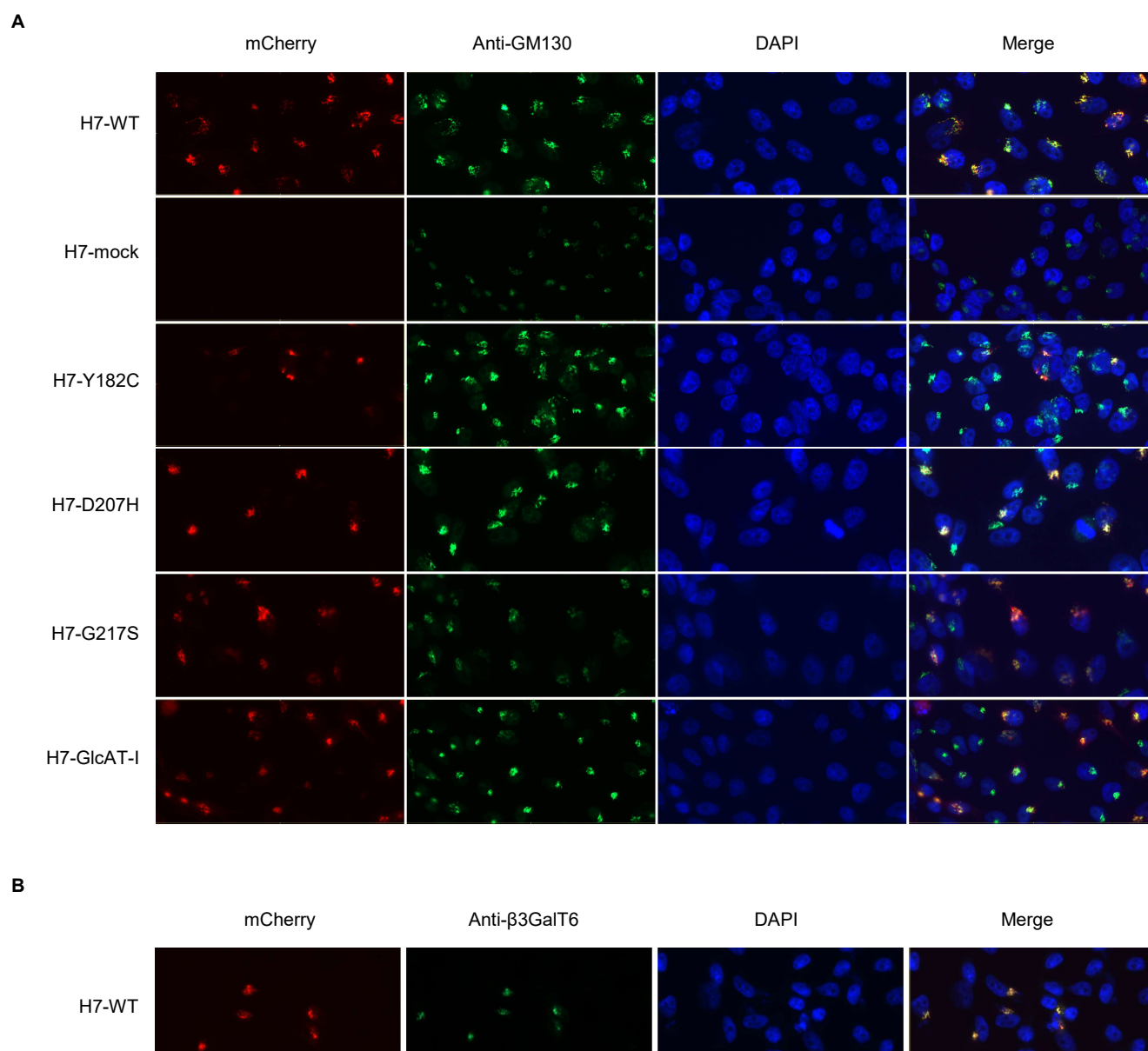


Figure 6. Representative image of the expression and localisation of WT and β 3GalT6 mutants in *B3GALT6*-KO H7 cells. (A) *B3GALT6*-KO H7 cells were transiently transfected with plasmids allowing the expression of mCherry fusion proteins corresponding to WT, β 3GalT6 mutants, GlcAT-I (Sartorius Stedim, Aubagne, France) or an empty vector (negative control). The GM130 cis-Golgi marker was used for protein co-localisation (B) Analysis of β 3GalT6 expression and localisation using a commercial anti- β 3GalT6 antibody. Nuclei were counterstained with DAPI (n=3 independent experiments).

Supplemental Figure 7

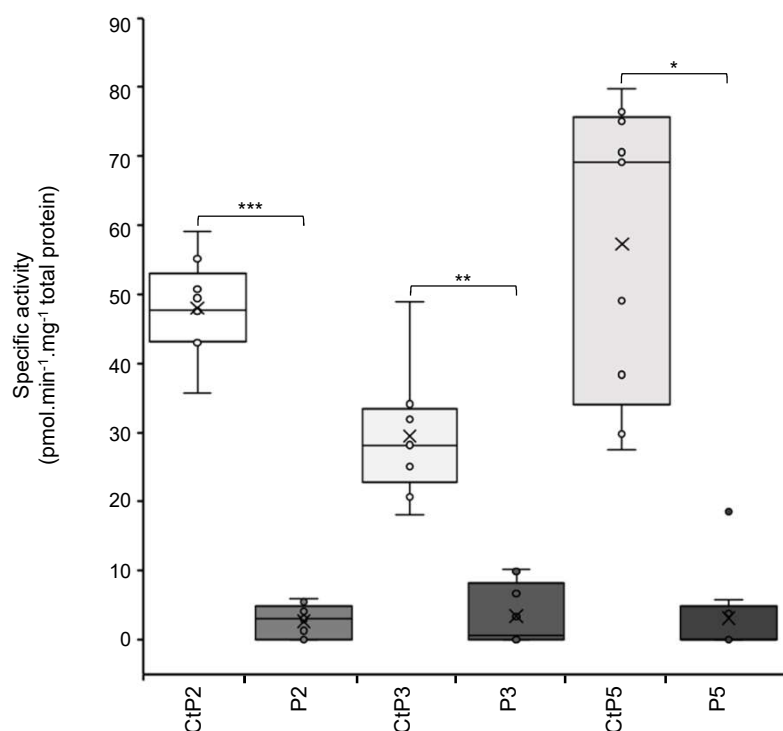


Figure 7. In vitro β 3GalT6 activity in spEDS patient fibroblasts. β 3GalT6 activity was measured using total cell lysates (10 μ g) prepared from control and patient fibroblasts. Gal-Xyl(2P)-OMN (2 mM) was used as acceptor substrate and UDP-Gal (1 mM) as donor substrate. Data are presented as box plot with the mean symbolised by an X (n=3 experiments; one triplicate per cell lysate); * p <0.05; ** p <0.01; *** p <0.001 when P2, P3 and P5 were compared to CtP2, CtP3 and CtP5, respectively (Student's t test).

Supplemental Figure 8

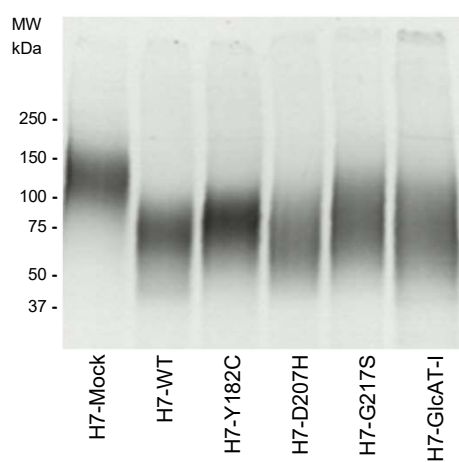


Figure 8. Representative profile of radiolabelled decorin in co-transfected *B3GALT6*-KO H7 cells. Radiolabelled decorin was secreted in conditioned media from *B3GALT6*-KO H7 cells transiently co-transfected with plasmids encoding WT, mutant β 3GalT6, GlcAT-I or empty vector and decorin (n=3 experiments in triplicate). In order to visualize the profile of GAG residual synthesis, the same amount of radioactivity (8,000 dpm) was loaded in each lane.

Supplemental Figure 9

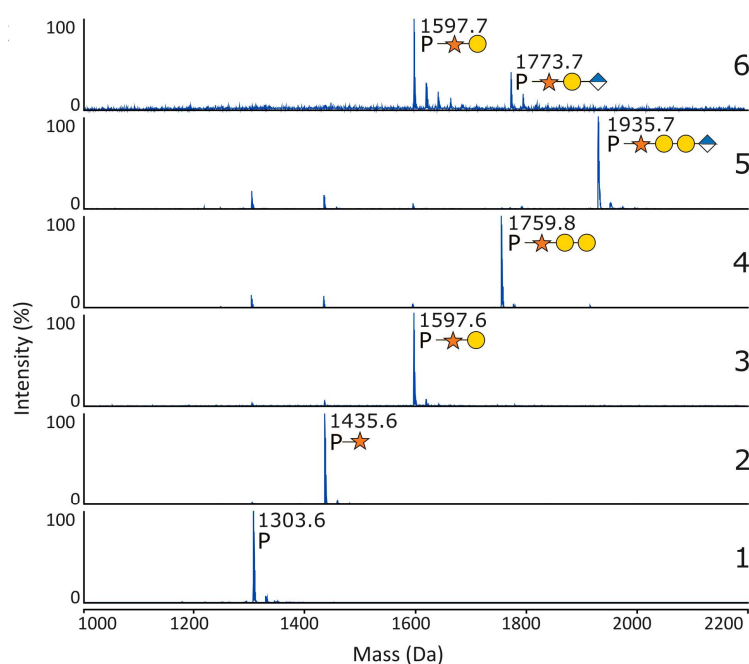


Figure 9. Analysis of reaction products by MALDI-MS. The composition and mass of the main peaks are indicated alongside. MALDI-MS determines the m/z (mass-to-charge) values of the products. For instance, the peptide 306EEASGEAS (referred to as P) has a mass of 1303.6 Da, and the corresponding peak of its singly charged ion has been detected at 1304.6 m/z . The addition of one xylose unit increases the mass by 132 Da (resulting in 1435.6 Da). Each galactose unit adds 162 Da (1597.6 and 1759.8 Da). One glucuronic acid unit further increases the mass by 176 Da (to 1935.7 Da). The major peak in reaction mixture 6 is a [Gal β 1-4Xyl]-peptide (1597.7 Da), and the second-highest peak corresponds to a [GlcA-3Gal β 1-4Xyl]-peptide (1773.7 Da).

Supplemental Figure 10

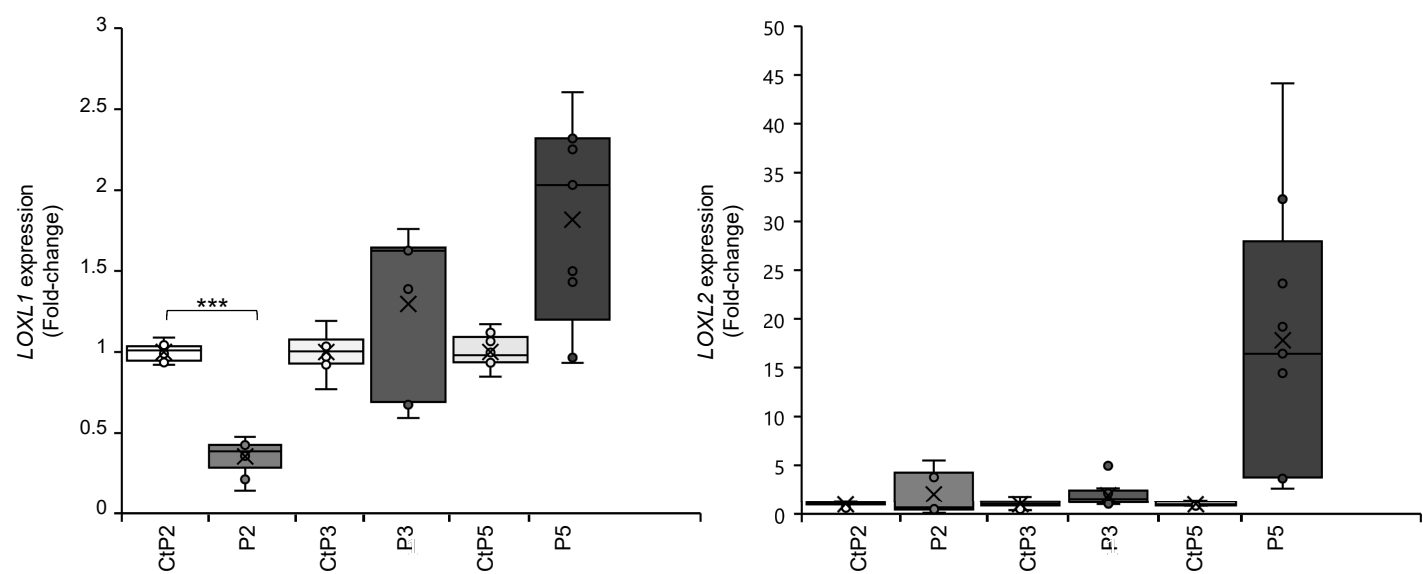


Figure 10. Analysis of gene expression in spEDS fibroblasts versus age and sex-matched controls. Results are presented as box plot with the mean symbolised by an X. Fold-changes were calculated for each patient using their corresponding control (n=3 experiments in triplicate); ***p<0.001 when comparing P2 to CtP2 (Student's t test).

Supplemental Figure 11

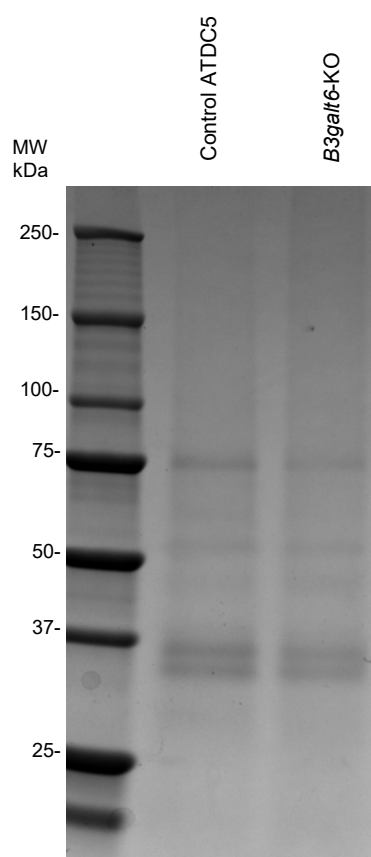


Figure 11. Comparison of collagen CNBr fragments obtained from control and *B3gal/t6*-KO ATDC5 pseudo-tissues. The collagen fraction was obtained by cell fractionation, digested with CNBr and analyzed by SDS-PAGE coupled to Coomassie staining. The image shown is representative of an analysis performed on 3 pseudotissues.

Supplemental Table 1

Table 1. Summary of patient characteristics and biochemical data from previous studies

Patient	DNA Change	Protein mutation	Syndrome subtype or disease	Residual in vitro β 3GalT6 activity ¹	Residual GAG synthesis from 4-MUX ²	HS expression ³	Decorin glycanation ⁴
P2	c.619G>C c.619G>C	p.(D207H) p.(D207H)	EDS-like	ND	60%	↘	↘↘↘
P3	c.619G>C c.323_344del	p.(D207H) p.(A108Gfs*163)	EDS-like	ND	16%	↘↘↘	↘↘↘
P5	c.649G>A c.649G>A	p.(G217S) p.(G217S)	EDS-like	ND	33%	↘	↘
PVII.1	c.545A>G c.545A>G	p.(Y182C) p.(Y182C)	SEMD-JL1	2%	19%	50%	↘↘↘

Data from (14) for patients P1, P2 and P3 and (16) for PVI.1

¹ Evaluated in vitro in patient fibroblast lysates using Gal-Xyl(2P)-OMN as acceptor substrate

² Measured in cellulo after Na₂[³⁵SO₄] radiolabeling in 4-MUX-primed GAG chains (10 μ M)

³ Visualised by immunofluorescence

⁴ Evaluated by immunoblot from conditioned media from patient fibroblast cultures

ND stands for not determined and HS for heparan-sulfates

Supplemental Table 2

Table 2. List of primers used for cloning β 3GalT6- $\Delta N_{ter}29$ or β 3GalT6- $\Delta N_{ter}50$ in the pETM41 expression vector and GlcAT-I- $\Delta N_{ter}76$ in the pETM30 expression vector		
MBP- β 3GalT6- $\Delta N_{ter}29$ -WT	Sense	5' -CGCCATGGCACGCTGCGCGGCCGAGCCC-3'
	Antisense	5' -TGCTCGAGTCAGGGGATGCCCTCCCTTCTC-3'
MBP- β 3GalT6- $\Delta N_{ter}50$ -WT	Sense	5' -CGCCATGGCAGCGCCCGCGCGCGCC-3'
	Antisense	5' -TGCTCGAGTCAGGGGATGCCCTCCCTTCTCTGGCA-3'
GlcAT-I- $\Delta N_{ter}76$ -WT	Sense	5' -TGCATAGCCATGGGAACG-3'
	Antisense	5' -TGCATGCGGCCGCTCA-3'
MBP- $\Delta N_{ter}29$ -Y182C	Sense	5' -CCGCCTCT G CTGGGGC-3'
	Antisense	5' -GCCCCAG C AGAGGCGG-3'
MBP- $\Delta N_{ter}29$ -D207H	Sense	5' -GGCAACTCGCC C ACTACTACCT-3'
	Antisense	5' -AGGTAGTAGT G GCAGAGTTGCC-3'
MBP- $\Delta N_{ter}29$ -G217S	Sens	5' -GCGCTGGGCGGC A GCTACGTGCTCTC-3'
	Antisense	5' -GAGAGCACGTAGC T GCCGCCCAGCGC-3'
The three last sets of primers were used for site-directed mutagenesis using MBP- β 3GalT6 $\Delta N_{ter}29$ construct. The mutated position is indicated in bold.		