

# Supplementary Information for

## **Substrate binding mode and catalytic mechanism of human heparan sulfate D-glucuronyl C5 epimerase**

Claire Debarnot, Yoan R Monneau, Véronique Roig-Zamboni, Vincent Delauzun, Christine Le Narvor, Emeline Richard, Jérôme Hénault, Adeline Goulet, Firas Fadel, Romain R Vivès, Bernard Priem, David Bonnaffé, Hugues Lortat-Jacob, Yves Bourne

Yves Bourne

Email: [yves.bourne@afmb.univ-mrs.fr](mailto:yves.bourne@afmb.univ-mrs.fr)

### **This PDF file includes:**

- Supplementary text
- Supplementary Materials and Methods
- Figs. S1 to S5
- Tables S1 to S2
- References for SI reference citations

## Supplementary Information Text

### Structural comparisons

Searches for structural homologues using the Dali server revealed that structures most closely related to each of the two large Glce domains are those of carbohydrate-binding modules (CBM) from family 29 (Z-score~10, 17% sequence identity, accession code 1oh3, (1)), family 22 (2wys, (2)) and the  $\beta$ -sandwich domain of glycoside hydrolase (GH) family 50 exo- $\beta$ -agarase (1bg4, (3)), as well as the  $(\alpha/\alpha)_6$  barrel of cellobiose 2-epimerase (Z-score ~13, 3vw5, (4)) or various GH8 glucanases (5gy3, (5)) or glucuronyl hydrolase (3wiw, (6)) for the  $(\alpha/\alpha)_4$  barrel domain. CBM29 has the particularity to display a “dimeric” binding mode with the oligosaccharide chain sandwiched at the concave face sites (CFS) of two CBM molecules (1). Given the relative high structural homology, the corresponding CFS in the  $\beta$ -sandwich domain of Glce is fully accessible and could represent potential recognition sites for a longer polysaccharide substrate or a protein partner (*SI Appendix*, Fig. S1A). However, the  $\beta$ -sandwich domain of Glce lacks the aromatic cradle that typically sits within the CFS of CBMs to interact with polysaccharides (7). Remarkably, the location of the bivalent ion near connecting loops is well conserved between Glce and CBM22. Similarly, the location of the substrate binding clefts is conserved between Glce and GH8 enzymes but its topology varies considerably (*SI Appendix*, Fig. S1B).

## Supplementary Materials and Methods

### Protein expression and purification

A codon-optimized synthetic DNA (Genscript) encoding a soluble form of hGlce flanked with a N-terminal FLAG epitope (DYKDDDDK) was cloned into the pYD7

expression vector (8). HEK-EBNA cells (NRCC, Biotechnology Research Institute) grown in complete medium were transfected with the pYD7-hGlce plasmid using linear polyethyleneimine (Polysciences). Clones were selected using 30  $\mu\text{g}/\text{mL}$  blasticidin (Euromedex) and analyzed by western blot using a monoclonal anti-FLAG M2-peroxidase antibody (Sigma-Aldrich). Selected cells were then amplified in DMEM/GlutaMAX medium supplemented with 10% fetal bovine serum, 25  $\mu\text{g}/\text{mL}$  G418 (PAA), 5  $\mu\text{g}/\text{mL}$  blasticidin in T175 culture flasks (37°C, 5%  $\text{CO}_2$ ). Once confluence was reached, the medium was replaced by CHO-S-SFMII medium (Gibco) supplemented with 25  $\mu\text{g}/\text{mL}$  G418, 5  $\mu\text{g}/\text{mL}$  blasticidin, and the flasks were transferred to 32°C to prevent protein aggregation. The hGlce-containing medium was collected after 5 days, centrifuged (5000 g, 10 min, 4°C) and dialyzed against 20 mM HEPES pH 7.0, 10 mM NaCl (buffer A) (overnight, 4°C) prior to purification. Similar protocols were used to generate HEK-EBNA cell lines stably producing the hGlce Tyr578Phe and Tyr500Phe mutants. The Glu499Gln mutant was transiently expressed from Expi293F cells cultured in the Expi293 expression medium (37°C, 8%  $\text{CO}_2$ ).

hGlce and the Tyr578Phe and Tyr500Phe mutants were purified by cation-exchange chromatography on HiTrap SP (HP column, 5ml, GE Healthcare) equilibrated with buffer A and eluted with the same buffer complemented to 1M NaCl, followed by gel-filtration on Superdex-200 (16/60 column, GE Healthcare) in buffer A complemented to 150 mM NaCl. The Glu499Gln mutant was purified by cation-exchange chromatography as above described, followed by affinity chromatography on a monoclonal anti-FLAG M1 antibody resin (Sigma-Aldrich) equilibrated in 50 mM Tris-HCl, pH 7.6, 150 mM NaCl, 2mM  $\text{CaCl}_2$  and eluted with the same buffer containing 2 mM EDTA instead of  $\text{CaCl}_2$ . The purified proteins were concentrated to ~6 mg/ml (hGlce and Tyr mutants) and ~1.6 mg/ml (Glu mutant) by ultrafiltration.

## Preparation of oligosaccharides

Heparosan, a bacterial polysaccharide composed of regular repeats of 4- $\beta$ -glucuronyl-(1 $\rightarrow$ 4)- $\alpha$ -N-acetylglucosaminyl-1 was recombinantly produced in *Escherichia coli* K-12 as previously described (9). To favor intracellular degradation of the polysaccharide, the *E. coli* K-12 was also modified to express a K5-lyase, resulting in the accumulation of disaccharide (dp2) to decasaccharide (dp10) composed of the repeated motif [GlcA $\beta$ 1-4GlcNAc $\alpha$ 1,4] with an unsaturated GlcA at the non-reducing end. After a first anion exchange chromatography, a mixture of dp6, dp8 and dp10 heparosan samples were loaded into two serial Superdex peptide 10/300GL columns (GE Healthcare) equilibrated with H<sub>2</sub>O and 250 mM NaCl as the mobile phase. Fraction of dp6, dp8 and dp10 were separately pooled (*SI Appendix*, Fig S2A), dialyzed against H<sub>2</sub>O and freeze-dried. The fractions corresponding to dp10 were pooled and a <sup>1</sup>H NMR spectrum was recorded at 700 MHz and 37 °C to assess its chemical composition (*SI Appendix*, Fig. S2B). Relative NMR signals of anomeric CH were used to assess the chain length. The materials were next re-suspended at 0.5 mg/mL into 2M NaOH, de-*N*-acetylated overnight at 56°C, dialyzed against H<sub>2</sub>O and freeze-dried. Glucosamine *N*-sulfation was performed by incubating the oligosaccharides at 1 mg/mL into sodium carbonate and sulfur trioxide-pyridine complex, both at 30 mg/mL, for 16 hours at 56 °C. To ensure complete *N*-sulfation the samples were treated three times. After extensive dialysis against H<sub>2</sub>O, the samples were freeze-dried overnight. The purity and structure of the resulting product (GlcA-GlcNS)<sub>5</sub> at 0.25 mg/mL was then assessed by recording a <sup>1</sup>H NMR spectrum on a 0.25 mg/mL sample at 700 MHz and 37°C (*SI Appendix*, Fig. S2B). The material was also characterized and quantified using disaccharide analysis. Briefly, the oligosaccharides were exhaustively digested into disaccharides using

heparinase. The resulting products were analyzed by RPIP-HPLC coupled to online 2-cyanoacetamide derivatization, as previously described (10). Commercial disaccharide standards were used for the qualitative and quantitative calibration.

The (IdoA-GlcNS)<sub>4</sub> octasaccharide was prepared in five steps from a protected octasaccharide, as previously described (*SI Appendix*, Fig. S3A) (11). Briefly, after deacetylation using Zemplén conditions and reduction of azido groups using 1,3-propanedithiol, the glucosamine moieties were *N*-sulfated using triethylamine (40 equiv.) and sulfur trioxide-pyridine complex (8 equiv.) in MeOH (0.4 M) at room temperature. Sulfur trioxide-pyridine complex was regularly added until complete consumption of the starting material. The resulting *N*-sulfated compound was obtained as a sodium salt after purification using RP-C18 flash chromatography followed by ion exchange on a Na<sup>+</sup> resin, and saponification of the methyl esters was performed using aqueous KOH in the presence of H<sub>2</sub>O<sub>2</sub>. After RP-C18 chromatography and ion exchange on Na<sup>+</sup> resin, the benzylated *N*-sulfated octasaccharide was afforded as a sodium salt. Hydrogenolysis using Pd(OH)<sub>2</sub> (20 % on charcoal) in a phosphate buffer / TBA mixture yielded the targeted *N*-sulfated octasaccharide (2.7 mg) as a sodium salt after filtration, desalting on a G-25 column, ion exchange on Na<sup>+</sup> resin and lyophilization. The three oligosaccharide species were characterized by <sup>1</sup>H, <sup>13</sup>C NMR spectroscopy and high-resolution mass spectrometry (*SI Appendix*, Fig. S3B). The (IdoA-GlcNS)<sub>4</sub> octasaccharide was quantified by NMR spectroscopy (*SI Appendix*, Fig. S3C).

### **Enzyme activity**

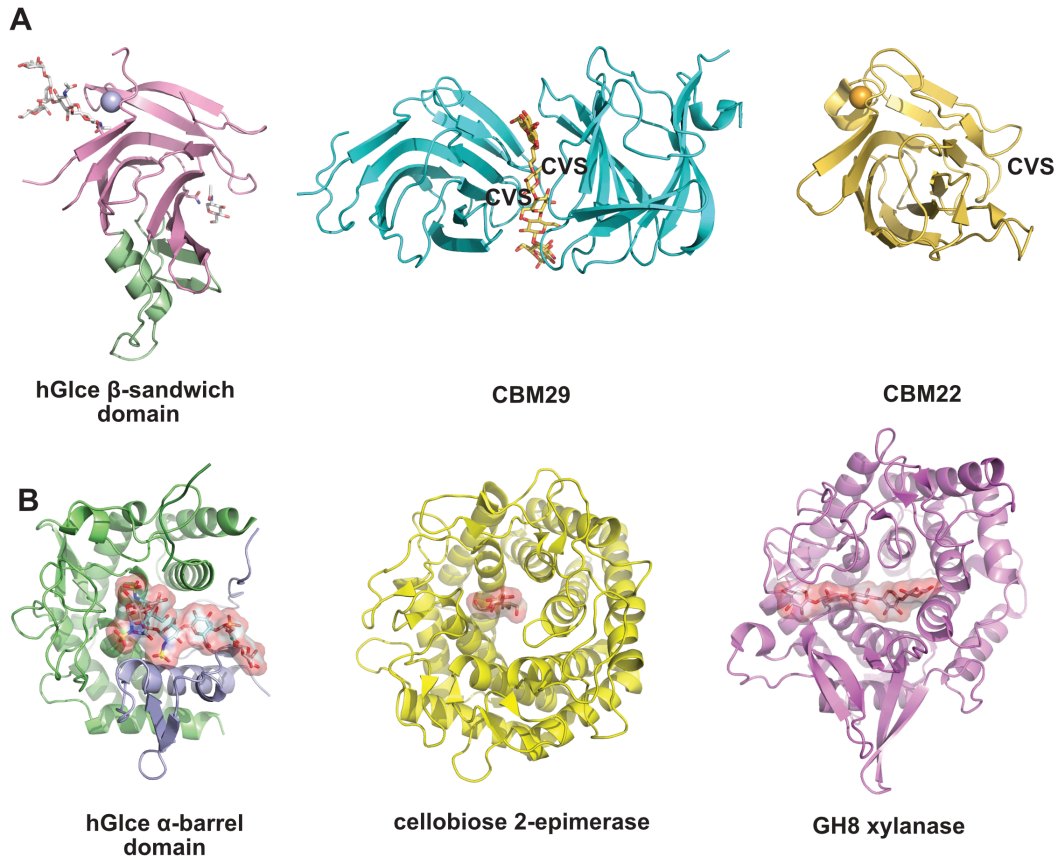
hGlcE (40 nM), the Tyr578Phe or Glu499Gln mutant (40 nM and 400 nM) or the Tyr500Phe mutant (50 nM) in 25 mM MES, pH 7.0, 5 mM CaCl<sub>2</sub> was incubated at

37°C with 1 mg/mL of  $^{13}\text{C}$ -labelled *N*-sulfated heparosan, prepared from *E. coli* K5 cells grown in M9 minimal medium containing  $^{13}\text{C}$  glucose as carbon source (12). The GlcA to IdoA conversion was followed by recording successive  $^{13}\text{C}$ - $^1\text{H}$  correlation experiments and integrating over time the corresponding peak volumes using sparky integration tools and a homemade script. The proportion of IdoA ( $x_{IdoA}^\infty$ ) achieved at equilibrium was determined by fitting the molar fractions of IdoA over time ( $t$ ), using the least squares method (Levenberg-Marquardt algorithm provided by Igor Pro version 6.03), against the following equation: (1)  $x_{IdoA} = x_{IdoA}^0 + (x_{IdoA}^\infty - x_{IdoA}^0)(1 - e^{-kt})$  with  $x_{IdoA}^0$ , molar fraction of IdoA at  $t=0$  and  $k$ , pseudo-first order rate constant of the reaction. The molar fractions of IdoA were determined at each time by calculating the ratio of the IdoA-characteristic NMR signal to the sum of the corresponding IdoA and GlcA signals. The initial rates,  $V_0$ , were determined using the first derivative of formal equations at  $t=0$ :  $V_0 = k (x_{IdoA}^\infty - x_{IdoA}^0)$ .

### **Crystallization and Structure Determination**

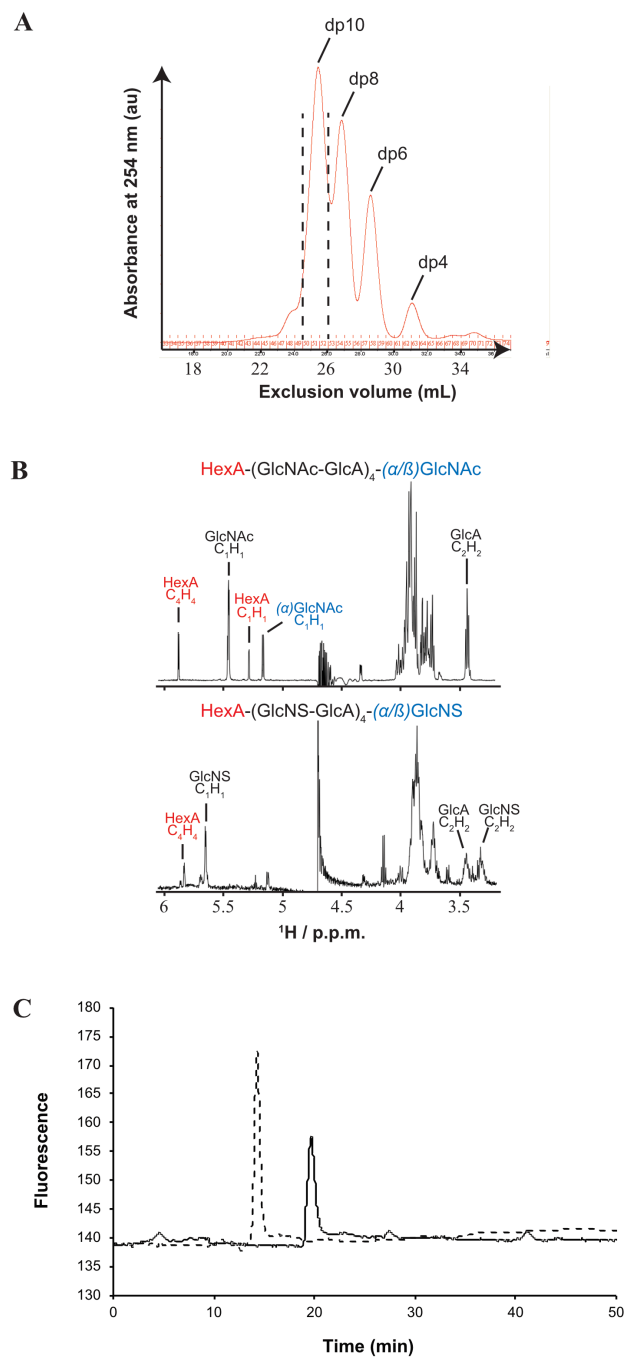
A thermal shift assay led to maintain hGlcE in its gel-filtration buffer for crystallization. Initial crystallization assays were performed at 20°C using commercial screens and a Mosquito robot (TTP Labtech). Initial needle-like crystals grew from a PEG2K condition. After several steps of optimization and rescreening, larger 3D crystals were obtained from a mix of 21.5% P2K MME, 0.1 M Na cacodylate pH 6.0 with 1.25 M lithium acetate, 0.1M MES pH 6.5, in a 2:1 (v/v) ratio. A similar protocol was used for crystallization of the purified Tyr578Phe mutant bound with the dp10 substrate or the dp8 product in 5-fold molar ratios. Suitable crystals were obtained from a mix of 24% MPEG 5K, 0.1M Na acetate pH 5.5 with 1.25 M lithium acetate, 0.1 M MES pH 6.5, in a 2:1 (v/v) ratio. Crystals of hGlcE were directly flash-cooled in liquid nitrogen, while

those of the Tyr578Phe complexes were briefly soaked within the reservoir solution supplemented with 25% glycerol prior to flash cooling. Diffraction data were processed with XDS (13) and scaled and merged using the CCP4 (14) suite of programs Pointless, Aimless and Truncate. Initial phases for hGlce were obtained by molecular replacement with Phenix (15) and the structure of zGlce (accession code 4PW2, 80% sequence identity with hGlce (16)) as a poly-Ala search model. The model was improved by manual adjustments with Coot (17) and refined with Phenix and Refmac (18) including NCS restraints and TLS refinement, with each subunit defining a TLS group. Structures of the Tyr578Phe mutant bound with the dp10 substrate or dp8 product were solved by rigid body refinement using Refmac and the refined model of hGlce, followed by rounds of restrained refinement. The substrate and product were fitted into unbiased Fo-Fc difference electron density maps calculated after rigid body refinement. The final structures of hGlce and the Tyr578Phe-dp10 and -dp8 complexes comprise 511/516 residues (Leu102-Asn617) for each subunit in the dimer (rmsd value, ~0.8 Å for 510 C $\alpha$  atoms). High temperature factors and weak electron densities are associated with a disordered Lys243-Pro247 segment at the tip of surface loop Asp241-Asp249. Structural comparison of the apo and substrate/product-bound proteins reveals that substrate/product binding causes slight reorientations only of side chains (Tyr158, Gln213, Arg575) (rmsd value, 0.35 Å for 511 C $\alpha$  atoms). Stereochemistry of the structures was analyzed with internal modules of Coot and the Molprobity server (19). Electrostatic surface potentials were calculated using APBS (20) with the PyMOL APBS tools. Figures 2-4 were generated using PyMOL (21). The ring conformations of the bound substrate and product were analyzed using the Cremer-Pople parameter calculator (22).



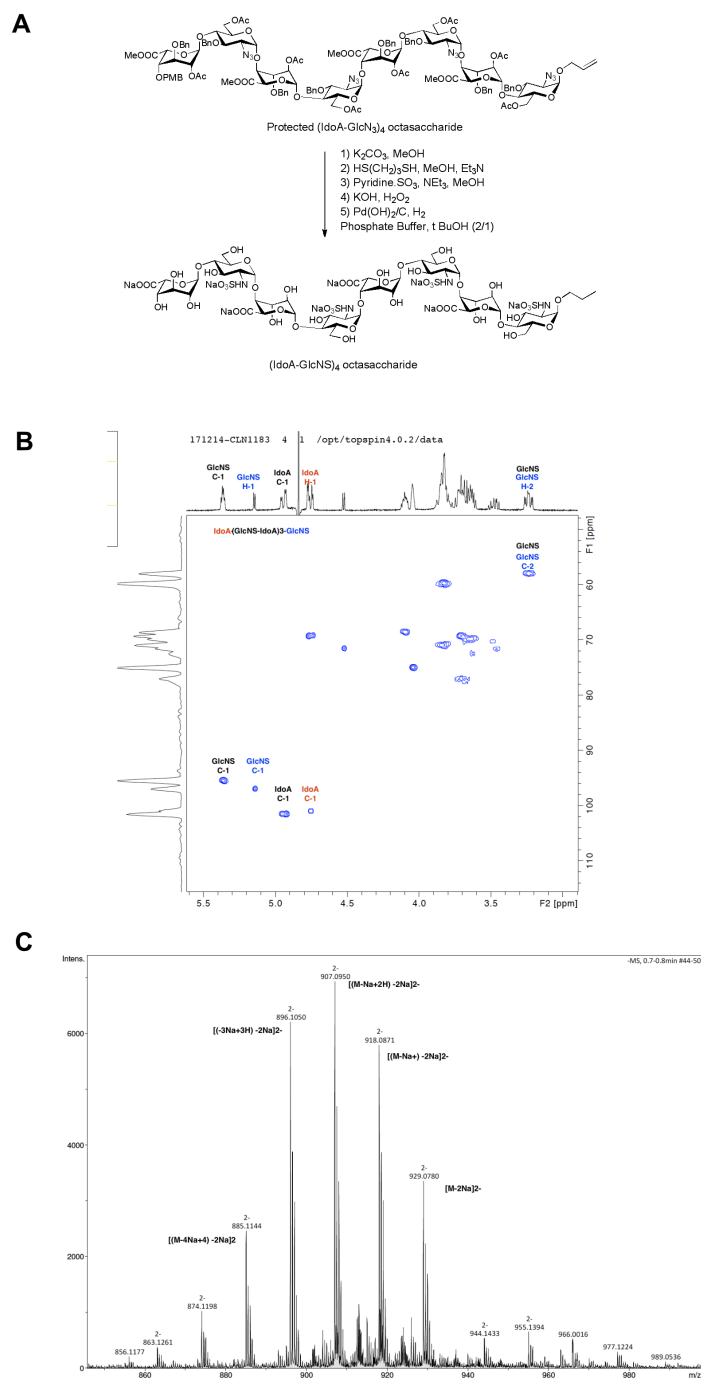
**Fig. S1: Structural comparison.** Comparison of the overall fold, viewed in a similar orientation, of **(A)** the  $\beta$ -sandwich domain of hGlce (left, in violet and green) with representative members of CBM29 (center, in cyan; accession code 1oh3) and CBM22 (right, in orange; 2wys), and **(B)** the  $(\alpha/\alpha)_4$  barrel of hGlce (left, colored blue and green for the two non-continuous sequences) with the  $(\alpha/\alpha)_6$  barrel of a cellulose 2-epimerase (center, in yellow; 3wkg) and a GH8 family xylanase (right, in violet; 2bf4). The location of the concave face side (CVS) in CBMs involved in oligosaccharide binding is indicated. In hGlce, the large insertion (in green) in the  $\beta$ -sandwich domain participates to inter-domain stabilization. The common locations of the  $\text{Ca}^{2+}$ -binding site in the  $\beta$ -sandwich domain and the substrate binding cleft at the bottom of the barrel are evident.



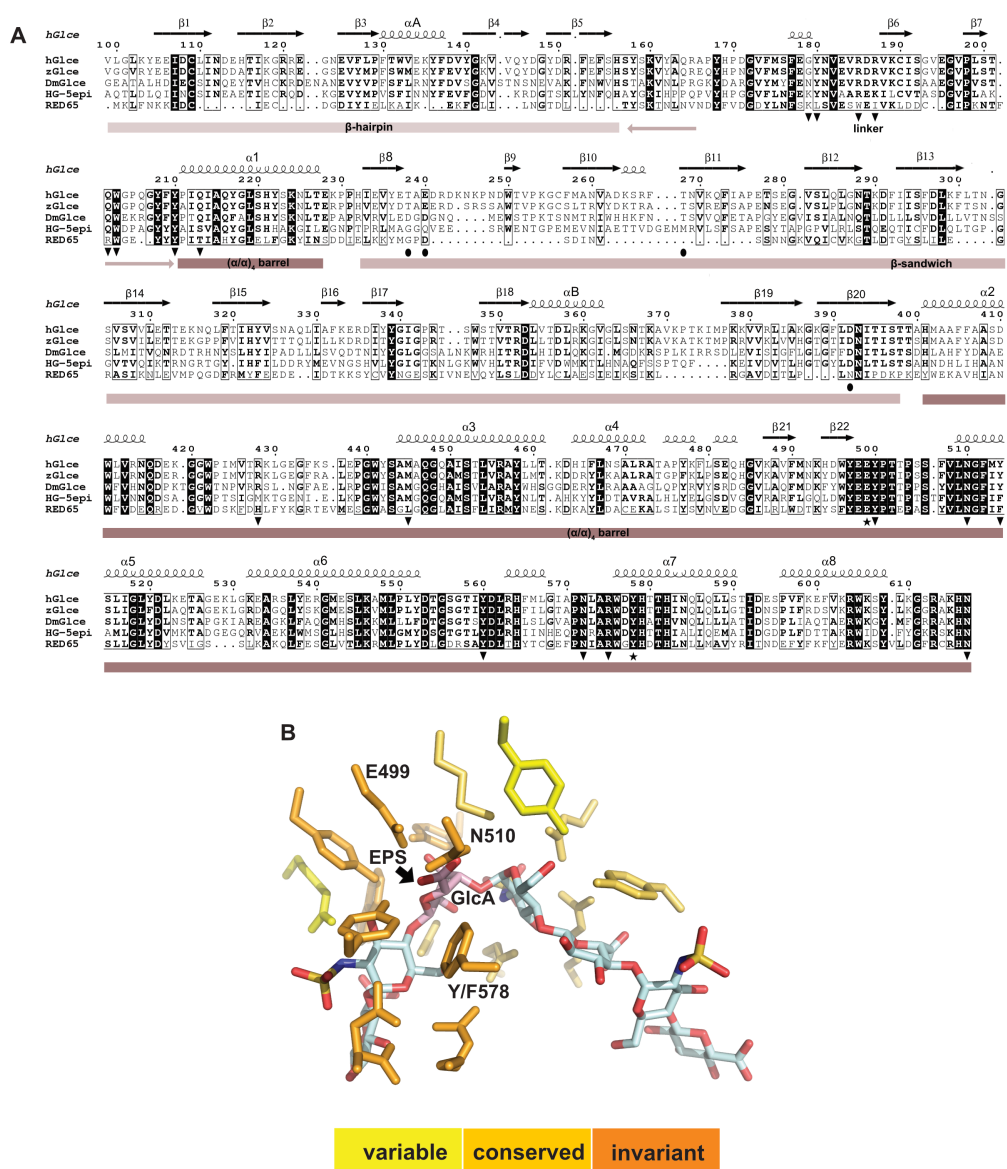


**Fig. S2: Preparation and characterization of *N*-sulfated heparosan-derived deca-saccharide (dp10).** (A). Elution profiles of dp2 to dp10 oligosaccharides separated by FPLC using two serial Superdex peptide 10/300GL columns. The dp10-containing fractions are indicated by dotted lines. (B, top spectrum) Characteristic  $^1\text{H}$  NMR chemical shifts of purified dp10 (intra-chain GlcNAc  $\delta_{\text{C}_1\text{H}_1} = 5.45$ ; reducing-end

$\alpha$ -form GlcNAc  $\delta_{C1H1} = 5.16$ ; hexuronic acid  $\delta_{C1H1} = 5.28$  and  $\delta_{C4H4} = 5.88$ ; glucuronic acid  $\delta_{C2H2} = 3.43$ ). **(B)**, bottom spectrum)  $^1\text{H}$  NMR analysis of chemically de-*N*-acetylated and re-*N*-sulfated dp10 reveals characteristic resonances of GlcNS C1H1 and C2H2 ( $\delta_{C1H1} = 5.65$  and  $\delta_{C2H2} = 3.32$ ). **(C)** SAX analysis digestion profile of chemically *N*-sulfated dp10 (plain line) and heparosan-derived oligosaccharide (dotted line). Commercial disaccharide standards were used for the qualitative and quantitative calibration.

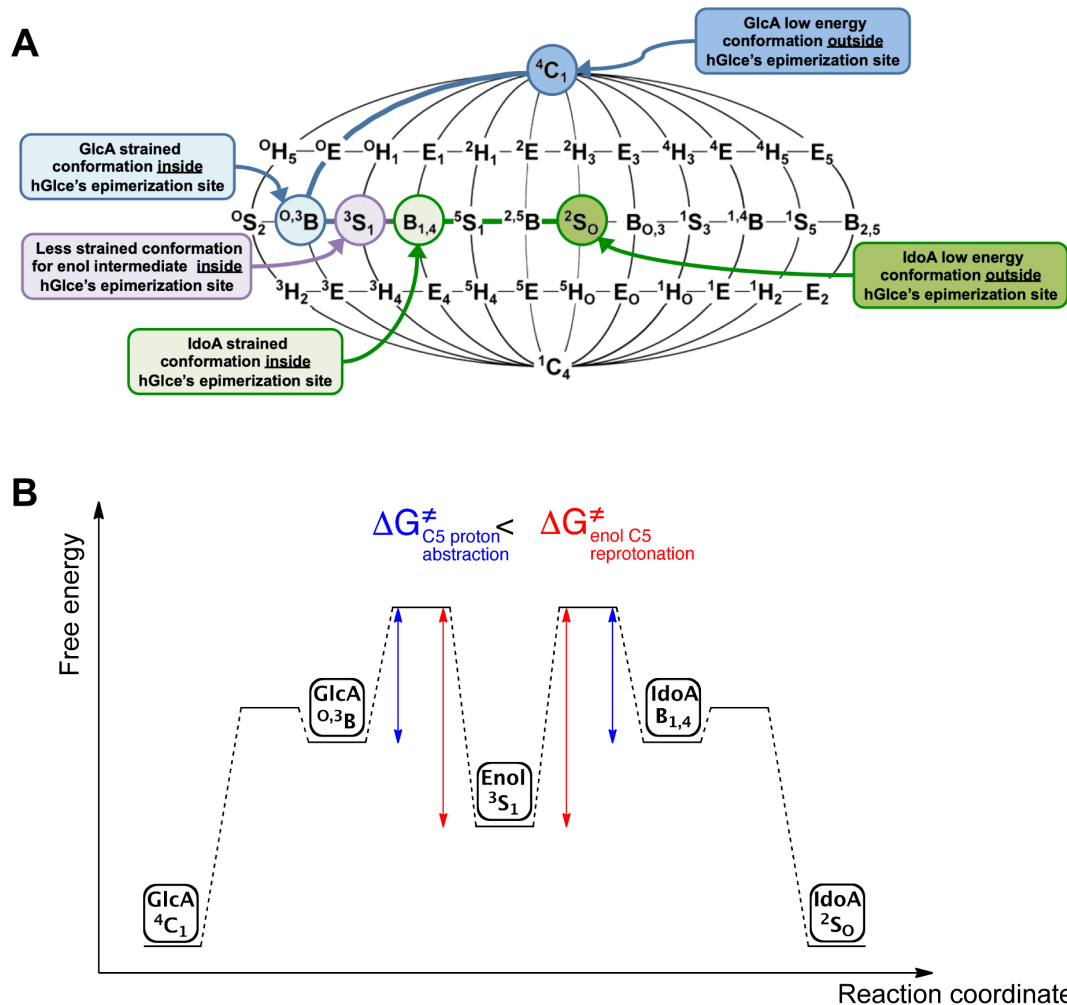


**Fig. S3: Preparation and characterization of *N*-sulfated IdoA-containing octasaccharide (dp8).** (A) Protected (IdoA-GlcN<sub>3</sub>)<sub>4</sub> octasaccharide was used as a template to synthesize the (IdoA-GlcNS)<sub>4</sub> octasaccharide in 5 steps, which was characterized by (B) NMR (<sup>1</sup>H, HSQC) spectroscopy and (C) mass spectrometry. NMR analysis reveals characteristic chemical shifts of GlcNS ( $\delta$  H-2<sup>GlcNS</sup> = 3.35-3.26 ppm).



**Fig. S4. Multiple sequence alignment of Glce with homologous proteins.** (A) The alignment of the sequences of hG1ce, close homologs from Zebrafish (zG1ce) and *Drosophila melanogaster* (DmG1ce) and related members from the giant African snail *Achatina fulica* (HG-5epi) and the marine bacterium *Bermanella marisrubri* (RED65) was generated with MUSCLE (23) and displayed with ESPrpt (24). Identical residues are indicated with a black background while similar residues are boxed. Catalytic residues and other functional residues involved in  $\text{Ca}^{2+}$  or substrate/product binding in hG1ce are indicated by an asterisk, filled circles and filled triangles, respectively.

Secondary structure elements and domain boundaries are indicated above and below the alignment. **(B)** Conservation of active site residues, shown as sticks, in Glce; those residues that interact with the central GlcNS-1-GlcA/IdoA+1-GlcNS+2 trisaccharide motif at the epimerization site are largely conserved.



**Fig. S5: Proposed conformational itinerary and Gibbs free energy of the epimerization reaction catalyzed by Glce**

(A) Pseudorotational itinerary for the interconversion of sugar ring conformations highlighting the ring puckering signatures and proposed transition states employed by Glce (two-dimensional projection of the CP sphere adapted from (25)). The Cremer-Pople sphere shows the 38 canonical puckering designations and the letter designs the type of pucker (C, chair; E, envelope; H, half-chair; S, skew; B, boat). The thick blue and green lines illustrate the proposed conformational itineraries linking  ${}^4C_1$  GlcA to  ${}^2S_0$  IdoA along a reaction coordinate involving the sole improper C5-O5-C6-C4 dihedral angle (**Fig. 3A**), *ie.* the sole motion of the C5 hydrogen atom. In heparinase I,

the corresponding IdoA unit adopts a standard and stable  ${}^2S_0$  conformation, suggesting that ring geometry distortion is not required as a driving force for the lytic process. **(B)** Gibbs free energy of the hGlcE-catalyzed epimerization process along the reaction pathway. The  ${}^{0,3}B$  and  $B_{1,4}$  conformations lie on a conformational itinerary interconverting  ${}^4C_1$  GlcA to  ${}^2S_0$  IdoA through a  ${}^3S_1$  skew boat, whose conformational strain is likely lower than those from the two distorted  ${}^{0,3}B$  and  $B_{1,4}$  conformers, supporting a rate-limiting step for C5 re-protonation in the epimerization reaction as shown by kinetic isotope effects (26).

**Table S1. Quality of the fit and kinetic parameters**

Protein	$x_{IdoA}^0$	$x_{IdoA}^\infty$	$k$ (min <sup>-1</sup> )	$\chi^2$	$V_0$ (min <sup>-1</sup> )
<i>wt</i> hGlcE	0.009 ± 0.005	0.387 ± 0.01	0.034 ± 0.001	0.00111765	0.013
Tyr500Phe	0.006 ± 0.01	0.297 ± 0.02	0.027 ± 0.002	0.00169619	0.008



**Table S2. Data collection and refinement statistics**

	substrate-free	Y578F-dp10	Y578F-dp8
<b>Data collection</b>			
Beamline	SOLEIL PX1	ESRF ID30B	SOLEIL PX1
Space group	P6 <sub>1</sub>	P6 <sub>1</sub>	P6 <sub>1</sub>
Cell dimensions:			
A=b, c (Å)	99.8,262.9	99.7,258.9	99.8,260.5
Resolution (Å)	49-2.52 (2.67-2.52)	48.95-2.1 (2.15-2.1)	49.0-2.45 (2.5-2.45)
R <sub>merge</sub>	0.11 (0.73)	0.083 (1.55)	0.14 (1.63)
R <sub>pim</sub>	-	0.038 (0.71)	0.04 (0.44)
CC(1/2)	0.99 (0.86)	0.99 (0.42)	0.99 (0.73)
I / $\sigma I$	12.2 (1.9)	12.6 (1.1)	17.4 (1.6)
Completeness (%)	99.5 (97.2)	99.9 (100)	100 (100)
Redundancy	5.7 (5.7)	5.7 (5.7)	14.3 (14.5)
Wilson B (Å <sup>2</sup> )	50.9	39.3	41.6
<b>Refinement</b>			
Resolution (Å)	43.83-2.52	40.3-2.1	49-2.45
No. reflections	47276	80072	51003
R <sub>work</sub>	17.6 (33.8)	16.2 (29.5)	17.4 (30.8)
R <sub>free</sub>	22.6 (35.1)	19.5 (31.5)	21.7 (33.1)
No. atoms			
Protein/Water	8357/372	8402/564	8335/362
Carbohydrate/ion	252/4	303/2	267/2
Ligand/others	-/14	188/88	203/97
B-factors (Å <sup>2</sup> )			
Protein	52.1/54.9	47.3/49.7	51.5/54.3
Carbohydrate/Water	103.7/44.6	91.2/52.0	98.0/48.5
Ligand/others	-/67.1	70.8/63.3	71.2/77.0
R.m.s. deviations			
Bond lengths (Å)	0.012	0.013	0.012
Bond angles (°)	1.58	1.2	1.68
Ramachandran			
Favoured (%)	96.9	96.5	96.1
Outliers (%)	0.2	0.1	0.2
<b>PDB accession code</b>	<b>6HZZ</b>	<b>6I01</b>	<b>6I02</b>

\*One single crystal was used for each data collection. #Values in parentheses are for the highest-resolution shell.

## References

1. Flint J, et al. (2004) Ligand-mediated dimerization of a carbohydrate-binding molecule reveals a novel mechanism for protein-carbohydrate recognition. *J Mol Biol* 337:417–426.
2. Najmudin S, et al. (2010) Putting an N-terminal end to the *Clostridium thermocellum* xylanase Xyn10B story: crystal structure of the CBM22-1-GH10 modules complexed with xylohexaose. *J Struct Biol* 172:353–362.
3. Schmidt A, Schlacher A, Steiner W, Schwab H, Kratky C (1998) Structure of the xylanase from *Penicillium simplicissimum*. *Protein Sci* 7:2081–2088.
4. Fujiwara T, et al. (2013) Crystal structure of *Ruminococcus albus* cellobiose 2-epimerase: structural insights into epimerization of unmodified sugar. *FEBS Lett* 587:840–846.
5. Attigiani A, et al. (2016) The crystal structure of the endoglucanase Cel10, a family 8 glycosyl hydrolase from *Klebsiella pneumoniae*. *Acta Crystallogr Sect F, Struct Biol Commun* 72:870–876.
6. Nakamichi Y, Mikami B, Murata K, Hashimoto W (2014) Crystal structure of a bacterial unsaturated glucuronyl hydrolase with specificity for heparin. *J Biol Chem* 289:4787–4797.
7. Abbott DW, van Bueren AL (2014) Using structure to inform carbohydrate binding module function. *Curr Opin Struct Biol* 28:32–40.
8. Loignon M, et al. (2008) Stable high volumetric production of glycosylated human recombinant IFN $\alpha$ 2b in HEK293 cells. *BMC Biotechnol* 8:65.
9. Barreteau H, Richard E, Drouillard S, Samain E, Priem B (2012) Production of intracellular heparosan and derived oligosaccharides by lyase expression in metabolically engineered *E. coli* K-12. *Carbohydr Res* 360:19–24.
10. Hijmans RS, et al. (2017) High sodium diet converts renal proteoglycans into pro-inflammatory mediators in rats. *PLoS One* 12:e0178940.
11. Lubineau A, Lortat-Jacob H, Gavard O, Sarrazin S, Bonnaffé D (2004) Synthesis of tailor-made glycoconjugate mimetics of heparan sulfate that bind IFN- $\gamma$  in the nanomolar range. *Chemistry* 10:4265–4282.
12. Laguri C, et al. (2011)  $^{13}\text{C}$ -labeled heparan sulfate analogue as a tool to study protein/heparan sulfate interactions by NMR spectroscopy: application to the CXCL12 $\alpha$  chemokine. *J Am Chem Soc* 1339:642–9645.
13. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66:133–144.
14. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235–242.
15. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
16. Qin Y, et al. (2015) Structural and Functional Study of D-Glucuronyl C5-epimerase. *J Biol Chem* 290:4620–4630.
17. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501.
18. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr D Biol Crystallogr* 53:240–255.
19. Chen VB, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66:12–21.
20. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* 98:10037–10041.
21. The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC (2010).
22. Cremer D, Pople JA (1975) General definition of ring puckering coordinates. *J Am Chem Soc* 97:1354–1358.
23. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high

- throughput. *Nucleic Acids Res* 32:1792–1797.
24. Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42:W320-324.
  25. Mayes HB, Broadbelt LJ, Beckham GT (2014) How sugars pucker: electronic structure calculations map the kinetic landscape of five biologically paramount monosaccharides and their implications for enzymatic catalysis. *J Am Chem Soc* 136:1008–1022.
  26. Hagner-Mcwhirter A, Lindahl U, Li J p. (2000) Biosynthesis of heparin/heparan sulphate: mechanism of epimerization of glucuronyl C-5. *Biochem J* 347:69–75.