

# SCIENTIFIC REPORTS

There are amendments to this paper

OPEN

## Downstream Products are Potent Inhibitors of the Heparan Sulfate 2-O-Sulfotransferase

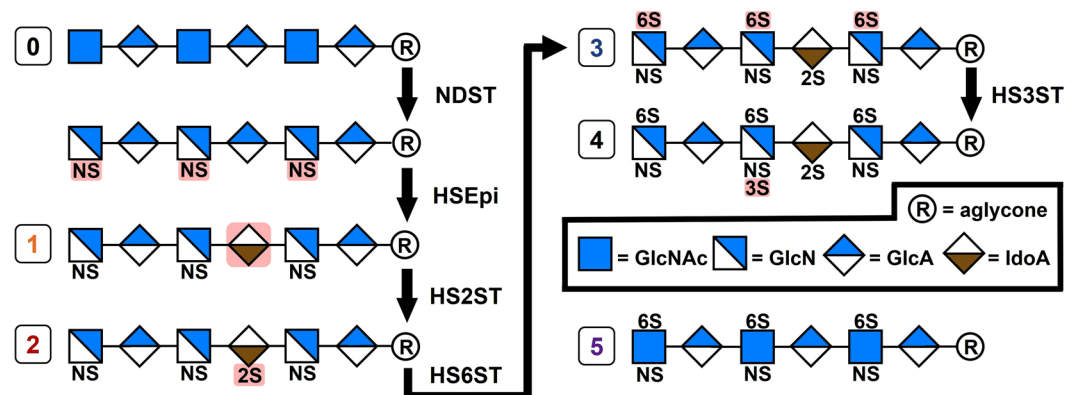
David F. Thieker<sup>1,2</sup>, Yongmei Xu<sup>3</sup>, Digantkumar Chapla<sup>2</sup>, Chelsea Nora<sup>4</sup>, Hong Qiu<sup>2</sup>, Thomas Felix<sup>1,2</sup>, Lianchun Wang<sup>1,2</sup>, Kelley W. Moremen<sup>1,2</sup>, Jian Liu<sup>3</sup>, Jeffrey D. Esko<sup>4</sup> & Robert J. Woods<sup>1,2</sup>

Heparan Sulfate (HS) is a cell signaling molecule linked to pathological processes ranging from cancer to viral entry, yet fundamental aspects of its biosynthesis remain incompletely understood. Here, the binding preferences of the uronyl 2-O-sulfotransferase (HS2ST) are examined with variably-sulfated hexasaccharides. Surprisingly, heavily sulfated oligosaccharides formed by later-acting sulfotransferases bind more tightly to HS2ST than those corresponding to its natural substrate or product. Inhibition assays also indicate that the IC<sub>50</sub> values correlate simply with degree of oligosaccharide sulfation. Structural analysis predicts a mode of inhibition in which 6-O-sulfate groups located on glucosamine residues present in highly-sulfated oligosaccharides occupy the canonical binding site of the nucleotide cofactor. The unexpected finding that oligosaccharides associated with later stages in HS biosynthesis inhibit HS2ST indicates that the enzyme must be separated temporally and/or spatially from downstream products during biosynthesis *in vivo*, and highlights a challenge for the enzymatic synthesis of lengthy HS chains *in vitro*.

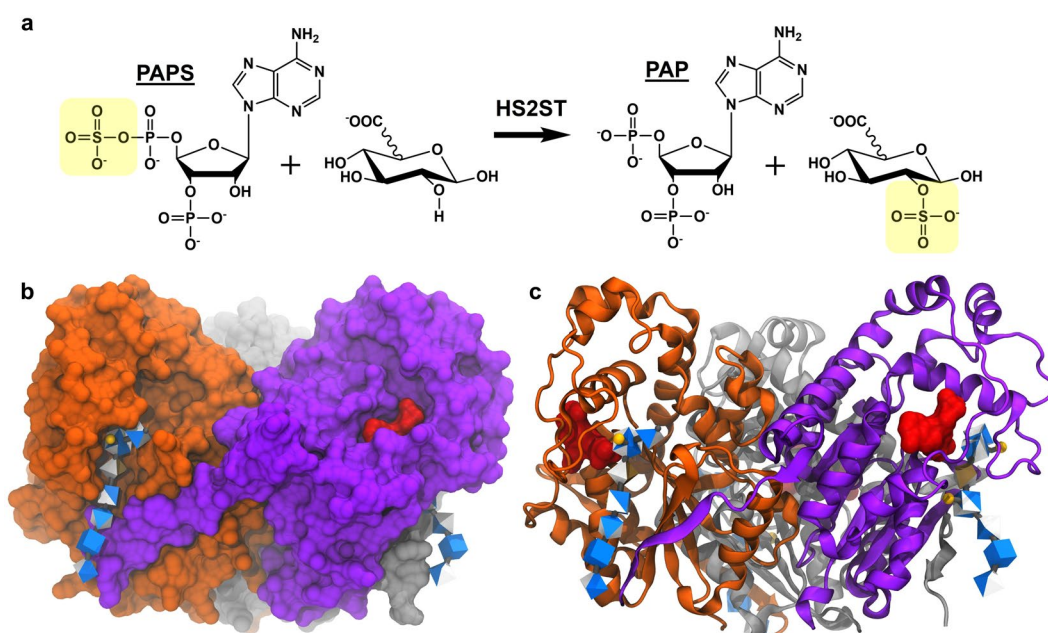
Glycosaminoglycans (GAGs) are a group of linear polysaccharides that consist of alternating hexosamine and hexuronic acid units. With the exception of hyaluronan, GAGs are variably sulfated, and are categorized into heparan sulfate (HS), chondroitin sulfate, and dermatan sulfate based on the composition and linkage of the monosaccharide residues<sup>1</sup>. HS has received considerable attention due to the broad functionality it displays in key cellular signaling events that are crucial for proper development<sup>2</sup>, such as vasculogenesis<sup>3</sup>, axon guidance<sup>4</sup>, and cellular differentiation<sup>5</sup>. The critical role for HS within animals, as well as its presence on cell surfaces, has provided pathogenic organisms with an accessible route for attachment and entry into the cell. Families of viruses including *Herpesviridae*<sup>6</sup>, *Flaviviridae*<sup>7</sup>, and *Filoviridae*<sup>8</sup> each exploit HS for cell adhesion. HS has also been implicated in various forms of cancer<sup>9</sup> and may either promote or inhibit metastasis depending on the sulfation pattern along the carbohydrate chain<sup>10</sup>. Efforts to characterize HS-protein interactions indicate that some binding modes recognize specific sulfation patterns, while others are non-specific and appear to be dependent primarily on negative charge density<sup>11,12</sup>. The prototypical example of highly specific HS binding is that between antithrombin and heparin (a more highly sulfated form of HS bearing antithrombin binding sites) whose affinity increases more than 1000-fold upon 3-O-sulfation of a key hexosamine residue<sup>13-16</sup>. There is growing interest in the therapeutic potential of HS mimetics due to the diverse functionality and possibility for highly specific interactions<sup>17-19</sup>.

Biosynthesis of HS begins with the assembly of a core tetrasaccharide to specific serine residues of proteins<sup>20</sup>. Two enzymes then add alternating *N*-acetyl glucosamine (GlcNAc) and glucuronic acid (GlcA) residues until the chain reaches between 20 and 150 disaccharide units in length<sup>21</sup>. Assembly of the HS chain continues via a set of enzymes that modify the repeating disaccharide unit, including concomitant *N*-deacetylation and *N*-sulfation of GlcNAc units<sup>22</sup>, epimerization of GlcA to iduronic acid (IdoA), 2-O-sulfation of uronic acids and 6-O- and 3-O-sulfation of hexosamine units (Fig. 1). A sequential biosynthetic pathway was originally proposed based on label-chase studies of HS in mouse mastocytoma microsomal fractions<sup>23</sup>. In general, the enzymatic

<sup>1</sup>Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA, 30602, USA. <sup>2</sup>Complex Carbohydrate Research Center, University of Georgia, Athens, GA, 30602, USA. <sup>3</sup>Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina, Rm 1044, Genetic Medicine Building, Chapel Hill, USA. <sup>4</sup>Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, California, USA. Correspondence and requests for materials should be addressed to R.J.W. (email: [rwoods@ccrc.uga.edu](mailto:rwoods@ccrc.uga.edu))



**Figure 1.** Heparan Sulfate (HS) biosynthetic pathway. After formation of the linkage region and polymerization of the chain (not shown), the chains undergo five modification steps catalyzed by corresponding enzymes, including the GlcNAc N-Deacetylase-N-Sulfotransferase (NDST), GlcA C5 Epimerase (HSEpi), 2-O-Sulfotransferase (HS2ST), 6-O-Sulfotransferase (HS6ST), and 3-O-Sulfotransferase (HS3ST). Differences from the products of the individual enzymes are highlighted in red. Five hexasaccharides were synthesized for this study and are numbered accordingly.



**Figure 2.** Structure of the 2-O-Sulfotransferase (HS2ST). **(a)** Reaction scheme for the HS2ST enzyme. The sulfate (yellow) from PAPS is transferred to the 2-position of a uronic acid (UA), resulting in PAP as a byproduct. **(b)** The co-crystal of the trimeric HS2ST in complex with PAP (red) and a HS heptasaccharide (3D-Symbol Nomenclature for Glycans<sup>76</sup>), sulfate depicted with yellow spheres). **(c)** The protein is depicted in a ribbon structure to convey the three symmetrical active sites that contain both PAP and HS residues.

preferences for upstream substrates and the inactivity of downstream products for each enzyme support a sequential pathway<sup>24</sup>, but cannot explain the irregular sulfation patterns observed *in vivo*. A method of regulating enzyme-substrate interactions has been proposed to resolve this discrepancy, in which the HS biosynthetic enzymes form a macro-complex termed the GAGosome<sup>1,20</sup>. Although a GAGosome involving all of the biosynthetic enzymes has not been detected directly, the epimerase and 2-O-Sulfotransferase (HS2ST) interact *in vitro*<sup>25</sup> and co-localize in the medial Golgi *in vivo*<sup>26</sup>.

Only one isoform of the HS2ST enzyme has been identified<sup>27</sup>, in contrast to the multiple isoforms with varying specificities that exist for all other HS sulfotransferases<sup>20</sup>. Bullock *et al.*<sup>28</sup> have shown that HS2ST is critical for normal development in mice. HS2ST catalyzes the transfer of a sulfate moiety from the co-factor 3'-phosphoadenosine-5'-phosphosulfate (PAPS) to the second hydroxyl position (2-sulfation) of a uronic acid (UA) residue (Fig. 2a); a motif that is required for HS recognition by proteins such as fibroblast growth factor-2<sup>29</sup>. Although HS2ST is capable of transferring sulfate to either GlcA or IdoA residues *in vitro*, GlcA2S residues are

sparse *in vivo* in contrast to IdoA2S<sup>27,30</sup>. The scarcity of GlcA2S may be due to co-localization of the epimerase and the HS2ST enzymes<sup>26</sup>, and the preference of the enzyme for IdoA (2.5-fold enhancement towards IdoA)<sup>27</sup>. The HS2ST enzyme exists as a homotrimeric complex, in which the C-terminus of one monomer extends to the substrate binding site in the neighboring monomer<sup>27</sup> (Fig. 2b,c). Although -IdoA2S-GlcNAc- units were detected within HS and heparin samples<sup>31,32</sup>, the enzyme strictly required a -GlcNS-UA-GlcNS- motif for catalysis *in vitro*<sup>33</sup>. Absence of either *N*-sulfate (NS) abolished activity, as did the presence of a sulfate at the 6-position (6S) on the reducing side glucosamine unit. Additionally, the minimum oligosaccharide size for activity was determined to be a pentasaccharide.

A lack of activity for HS2ST with 6-*O*-sulfated substrates is consistent with a sequential mechanism of assembly of the chain; therefore, little attention has been paid to the affinity or inhibitory ability of downstream biosynthetic products. Here, we present equilibrium binding data for hexasaccharides with varying sulfation patterns that correspond to steps within the HS biosynthetic pathway (Fig. 1 and Supplementary Table 1). Unexpectedly, binding and enzyme activity inhibition assays led to the observation that downstream products of HS biosynthesis are effective inhibitors and bind with high affinity to HS2ST. Binding studies confirmed the enhanced binding of recombinant HS2ST in the biological context of full length HS as well. These observations raised several questions, including why downstream products do not inhibit the enzyme *in vivo*, and by what mechanism they bind to and inhibit HS2ST. We explored the latter question through targeted point mutagenesis, as well as docking and molecular dynamics (MD) simulations. All downstream oligosaccharides were predicted to be able to fit into the HS2ST active site and remained bound to the protein over the course of the 200 ns MD simulations. Moreover, the inactivity of the 6-*O*-sulfated oligosaccharides as substrates was predicted to arise from the positioning of the 6-*O*-sulfate in the PAPS sulfate-binding site, presumably preventing the co-factor from binding.

## Results

**Binding analyses indicate sulfation-dependent HS2ST binding preferences.** HS-protein binding studies often utilize biologically-derived oligosaccharides of heterogeneous length and sulfation<sup>34,35</sup>; however, the synthesis of oligosaccharides with defined length and specific sulfation patterns has recently become possible<sup>36</sup>. We employed biolayer interferometry (BLI) to determine equilibrium surface binding constants ( $K_D$ ) for HS2ST with synthetic HS hexasaccharides (Fig. 1 and Supplementary Table 1). The HS compounds were attached to the BLI biosensor surface to conserve sample and to ensure a significant signal upon protein binding. The HS compounds contained a biotinylated PEG4 linker at the reducing terminus, enhancing the likelihood of a consistent presentation of the linear glycan on the sensor surface, in contrast to immobilizations via intra-chain modifications<sup>37</sup>.

Binding constants were collected for two isoforms of the enzyme; a glycosylated human variant (glyc-HS2ST) expressed in mammalian cell lines, and a highly homologous (96% identity) non-glycosylated chicken isoform expressed in *E. coli* (non-glyc-HS2ST). Glyc-HS2ST demonstrated modestly enhanced binding responses (Fig. 3a,b) and lower  $K_D$  values (Table 1) than the non-glyc-HS2ST construct. The high sequence identity suggested that the differences in binding primarily resulted from glycosylation of the enzyme. Glycosylation generally dampens protein dynamics<sup>38–40</sup> and has been shown to impair enzyme activity<sup>38,40</sup>. Thus, the enhanced ligand affinity seen in HS2ST was unexpected and may relate to the specific position of the glycosylation site (N108)<sup>41</sup>. This site is proximal to the HS binding site (Fig. 3c), and MD simulations suggest the glycan may stabilize HS binding by capping the substrate in the binding site.

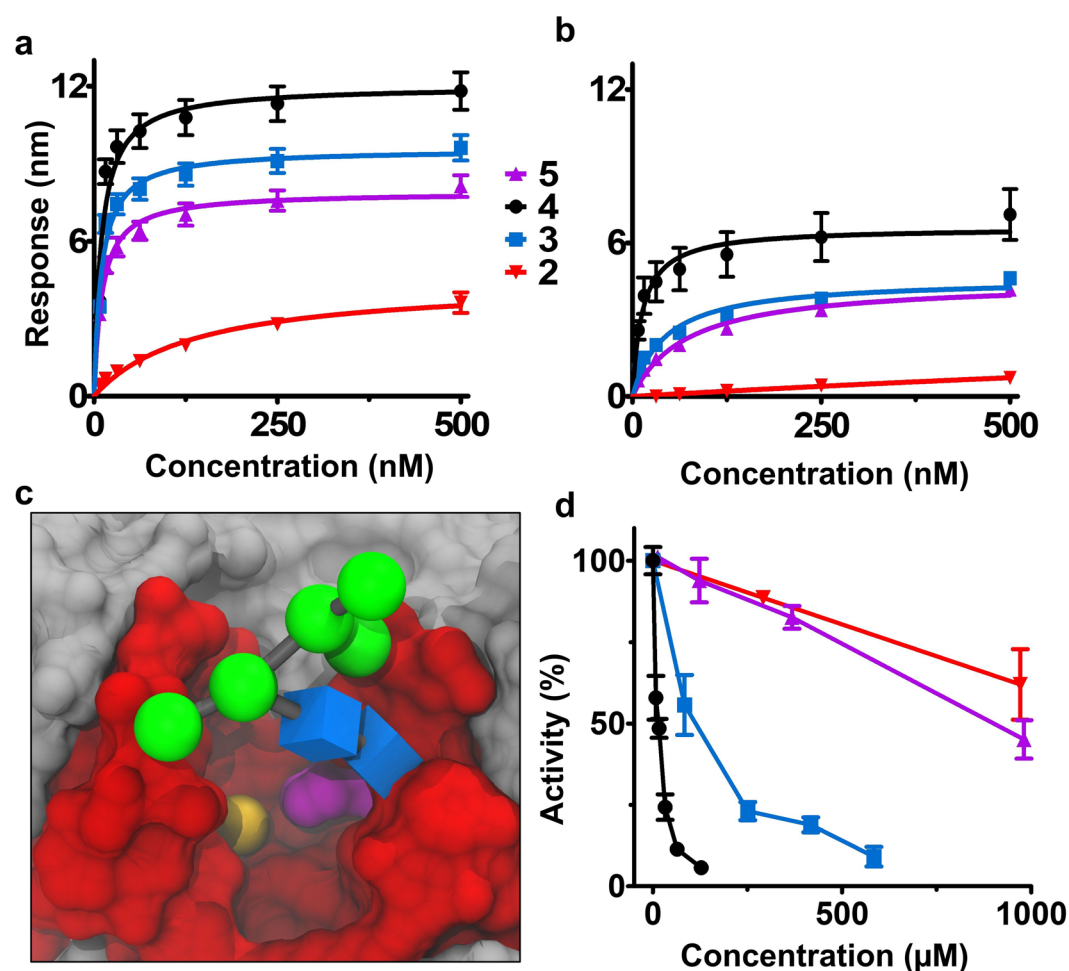
Interestingly, the affinity of the hexasaccharide that resembles the natural substrate of the reaction (1) was too weak to measure for either HS2ST isoform, but both displayed measurable affinities for sulfated molecules despite the lack of enzyme activity associated with these compounds (Table 1). The non-glyc-HS2ST demonstrated a clear trend in which the affinity increased according to the degree of HS sulfation. High affinities were observed for 6-*O*-sulfated compounds (5, 3, and 4) with glyc-HS2ST, but the similarity of their values (Table 1) likely reflects the limit of the sensitivity of the BLI assay (7.8 nM was the lowest concentration of protein whose binding was detectable). Since glycosylation of the enzyme did not appear to affect the binding preferences of the enzyme, and bacterial expression was more convenient, the remainder of the study utilized non-glyc-HS2ST.

**HS2ST binding preferences reflect inhibitory potential.** The tight binding of the higher-sulfated compounds suggested that they might also act as inhibitors of the HS2ST enzyme. Inhibition assays were thus performed with non-glyc-HS2ST and all compounds except for 1, the enzymatic substrate. As in the case of the affinities, the  $IC_{50}$  values were found to depend strongly upon the degree of sulfation of the hexasaccharide (Fig. 3d and Table 2). Compound 5 appeared to inhibit at least as well as the product of the reaction (2); however, the  $IC_{50}$  was not reached even at the maximum substrate concentration (1 mM) for either 2 or 5. The addition of a single sulfate moiety enhanced the potency of the glycan as an inhibitor by ~10-fold (compare 3 and 4). The presence of both 2- and 3-*O*-sulfate moieties in the hexasaccharide (4) enhanced its inhibitory potential by approximately 100-fold relative to 5. These *in vitro* inhibition assays indicate that downstream products of the HS biosynthetic pathway could interfere with HS2ST catalysis *in vivo*.

**Downstream HS inhibitors bind within the HS2ST active site.** To assess whether the inhibitory compounds interact with the active site, point mutations were introduced in non-glyc-HS2ST (Fig. 4a). Four mutants of the catalytic residue (H142A/E/R/K) were created, along with a mutation of the residue that facilitates 2-*O*-sulfation of IdoA residues (R189A)<sup>27</sup>. Additionally, a negative control (NC) was generated containing multiple mutations of positively charged residues (R189A, R80E, and R184E) that reduced the net charge of each protein monomer by 5. The extent that the mutations affected affinity provided an indication that the ligand had been bound in the active site (Table 1 and Supplementary Fig. 1). While R189A markedly diminished the affinity for all compounds, the NC mutations reduced affinity to immeasurable levels. Removal of the catalytic side chain

| Compound | WT                    | WT             | H142A    | H142K    | H142R    | H142E    | R189A     |
|----------|-----------------------|----------------|----------|----------|----------|----------|-----------|
|          | glyc-HS2ST            | non-glyc-HS2ST |          |          |          |          |           |
| 2        | 138 ± 19 <sup>b</sup> | — <sup>c</sup> | 460 ± 91 | 198 ± 34 | 136 ± 20 | —        | —         |
| 3        | 10 ± 1                | 42 ± 5         | 20 ± 3   | 15 ± 2   | 12 ± 1   | 41 ± 4   | 448 ± 124 |
| 4        | 10 ± 2                | 13 ± 3         | 10 ± 2   | 11 ± 2   | 10 ± 1   | 19 ± 3   | 63 ± 5    |
| 5        | 11 ± 1                | 72 ± 8         | 25 ± 2   | 22 ± 2   | 13 ± 1   | 140 ± 11 | —         |

**Table 1.** Equilibrium  $K_D$  values (nM)<sup>a</sup> for the interaction of HS2ST constructs with varying HS compounds, as calculated from saturation binding curves obtained with BLI. <sup>a</sup>n = 3. <sup>b</sup>Standard errors of best fit parameters determined by nonlinear regression. <sup>c</sup>Response too weak to fit.

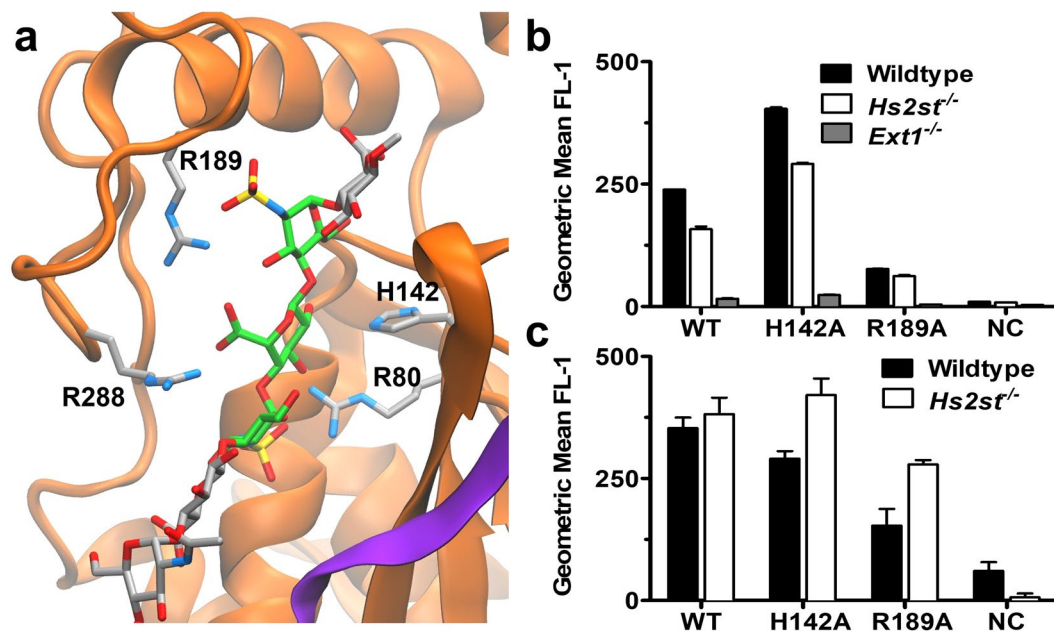


**Figure 3.** HS2ST WT binding and inhibition assays. (a) Equilibrium binding curves (n = 3) for glyc-HS2ST with compounds 2, 3, 4, and 5 as red, blue, black, and purple, respectively. (b) Equilibrium binding curves (n = 3) for non-glyc-HS2ST (c) A model of glyc-HS2ST was created by adding a short, high-mannose glycan (green and blue) to N108 of the crystal structure (PDB ID: 4NDZ). The trimeric protein is depicted in gray, except for amino acids within 10 Å of the acceptor IdoA residue (red) and the catalytic H142 (magenta). The modeled sulfate from PAPS is also depicted as a yellow sphere. (d) Inhibition assays in which a single concentration of non-glyc-HS2ST and substrate was incubated with varying concentrations of HS hexasaccharides. Errors are reported as the standard deviations of the mean.

(H142A) enhanced the affinity of each compound, compared to the WT, by approximately 2–3-fold. Addition of a positive charge at this location (H142K/R) further enhanced the affinity for each compound, so much so that a  $K_D$  value could be measured even for 2. The H142E mutation diminished the affinity of 5 by half compared to WT, but its effects were less significant for binding to 3 or 4. The basic character of the catalytic residue (H142) in the WT enzyme may be significant enough to be approximately comparable to that of a glutamate in this specific microenvironment.

| Compound | non-glyc-HS2ST     |
|----------|--------------------|
| 2        | >1000 <sup>a</sup> |
| 3        | 14 <sup>b</sup>    |
| 4        | 114 <sup>b</sup>   |
| 5        | >1000 <sup>a</sup> |

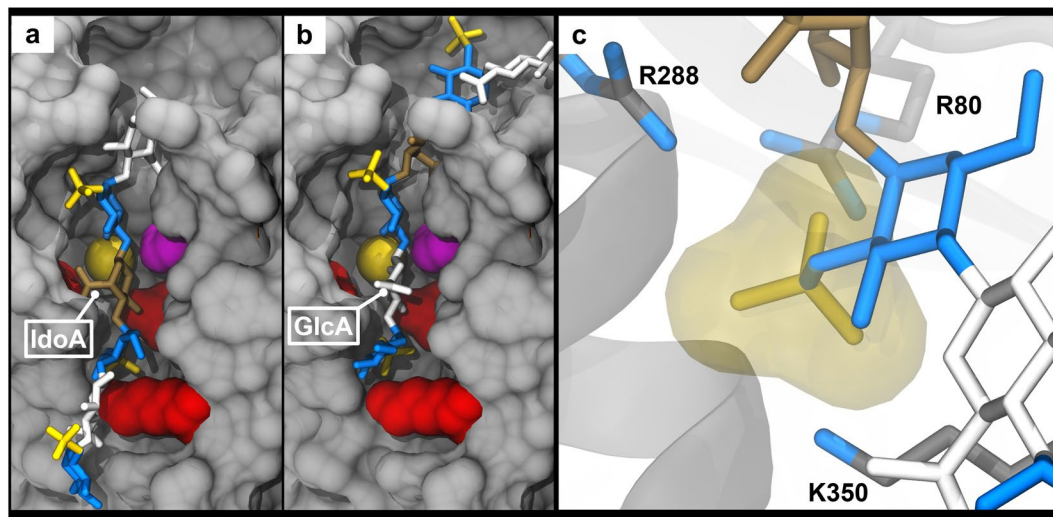
**Table 2.** IC<sub>50</sub> values (nM) of the various HS hexasaccharides, as determined from inhibition assays of the reaction between HS2ST and completely de-O-sulfated *N*-sulfated heparin. <sup>a</sup>n = 2. <sup>b</sup>n = 3.



**Figure 4.** HS2ST active site mutations. (a) The crystal structure of the co-complex (PDB ID: 4NDZ) is depicted without PAP, for clarity. HS2ST monomers are distinguished by orange and purple ribbons, and mutated amino acids side-chains are depicted in gray. The glycan is depicted in gray, except for the acceptor (IdoA) and flanking GlcNS residues, which are highlighted in green. (b) Flow cytometry of non-glyc-HS2ST constructs with various CHO cell lines (n = 2). (c) Flow cytometry of non-glyc-HS2ST constructs with various MLE cell lines (n = 3). Errors are reported as the standard deviations of the mean.

**HS2ST mutants display differential binding to HS-producing cells.** The interaction of WT HS2ST and mutants with biologically relevant full-length HS was examined using flow cytometry with binding being detected by fluorescently-labeled antibody to the maltose binding protein tag on the recombinant HS2ST. Chinese hamster ovary (CHO) cells lack the HS3ST enzyme<sup>42</sup>; therefore, the cell surface HS chains do not contain compound 4. Non-glyc-HS2ST constructs were incubated with wild-type CHO cells, as well as with CHO cells bearing mutations in different biosynthetic enzymes (Fig. 4b). Cell-surface binding reproduced the general trends in affinities observed with synthetic oligosaccharides. Specifically, the H142A mutant showed the highest affinity for Wildtype CHO cells, and diminished in affinity in the order of WT, R189A, and NC. The HS produced by the *Hs2st*<sup>-/-</sup> cells has no detectable 2-O-sulfate groups, and this loss is accompanied by an increase in the amount of both *N*- and 6-O-sulfate groups<sup>43</sup>. Thus, the composition of HS produced by *Hs2st*<sup>-/-</sup> is increased for 1 and 5 and decreased for 2 and 3 compared to Wildtype cells. Since the HS2ST constructs demonstrated a tighter affinity for 2 and 3 compared to 1 and 5 in the previous BLI studies, the lower binding signal for the *Hs2st*<sup>-/-</sup> cells is understandable. Negligible binding for each of the HS2ST constructs was observed in *Ext1*<sup>-/-</sup> cells, which lack HS. This finding suggests that the HS2ST does not interact with other glycosaminoglycans, such as chondroitin sulfate, which are up-regulated in this cell line<sup>44</sup>. Collectively, the cell binding data are consistent with the oligosaccharide binding assays and reinforce the hypothesis that the products of the early-acting HS sulfotransferases must be temporally and/or spatially separated from the products of the later enzymes in the biosynthetic pathway.

A separate flow cytometry assay was conducted with Wildtype and *Hs2st*<sup>-/-</sup> mouse lung endothelial (MLE) cells (Fig. 4c), which produce HS3ST. The *Hs2st*<sup>-/-</sup> MLE cells demonstrated a similar trend in affinities towards the HS2ST constructs as was found with the CHO cells and BLI experiments; namely, enhancing affinity in the order of NC, R189A, WT, and H142A. Interestingly, HS2ST constructs demonstrated enhanced binding with the *Hs2st*<sup>-/-</sup> MLE cells compared to Wildtype. This difference could suggest a role for 3-O-sulfate in HS2ST binding; however, we were unable to directly compare the relative contributions of each moiety to HS2ST binding as



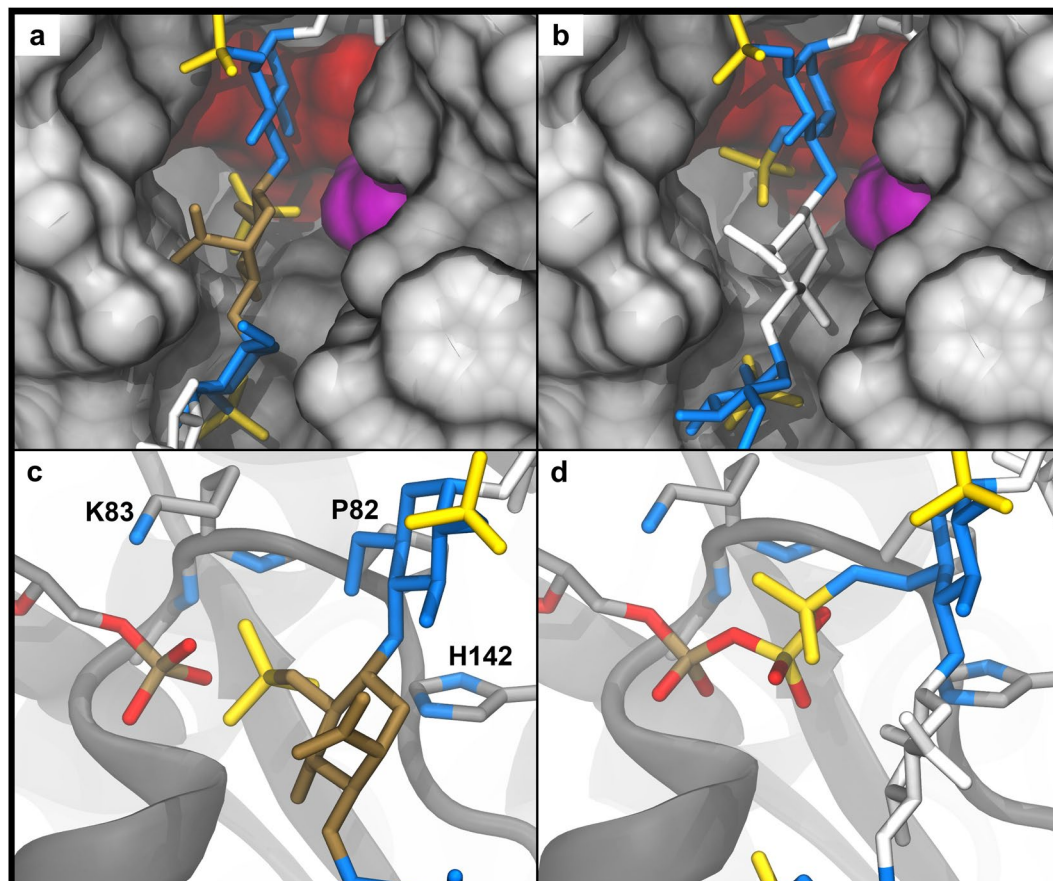
**Figure 5.** Predicted structures of HS2ST with compound 1. The same glycan color scheme is used as Fig. 4. (a,b) Amino acids that form the *N*-sulfate binding pocket (R288, R80, and K350) are colored red and the catalytic H142 is magenta. The sulfur atom modeled onto PAP is displayed for context (yellow sphere). (a) Docked model of **1** that matches the crystal structure. (b) Docked model of **1** that is translocated one disaccharide unit, resulting in GlcA as the acceptor residue. (c) A close-up view of the *N*-sulfate binding pocket from the crystal structure (PDB ID: 4NDZ) with the three amino acids coordinating the *N*-sulfate. The transparent, yellow surface represents the VDW radii of the sulfate moiety.

our hexasaccharide library did not include a compound that both lacked 2-*O*-sulfate and included 3-*O*-sulfate. Alternatively, enhanced binding of *Hs2st*<sup>-/-</sup> compared to WT cells may result from a difference in total charge of the HS chains on the cell surface.

**Predicted structures of HS2ST-inhibitor complexes.** Docking and MD simulations were performed to structurally characterize the binding mechanism of HS2ST with each of the hexasaccharides. The non-glyc-HS2ST enzyme has been crystallized in complex with PAP, a byproduct of the reaction, and a heptasaccharide substrate (GlcA-GlcNAc-GlcA-GlcNS-IdoA-GlcNS-GlcA-PNP) that is nearly identical to **1** (GlcNS-GlcA-GlcNS-IdoA-GlcNS-GlcA-PNP), except for the presence of a non-reducing terminal GlcA-GlcNAc unit in the former (PDB ID 4NDZ<sup>33</sup>). The hexasaccharides were docked to unliganded (apo) HS2ST using Vina-Carb<sup>45</sup>. As a positive control, the equivalent hexasaccharide (lacking the non-reducing terminal GlcA) was created with GLYCAM-Web ([www.glycam.org/gag](http://www.glycam.org/gag)) and docked into the active site of HS2ST, with the same level of flexibility as employed for the other compounds. The pose of the top-ranked model matched that of the ligand in the co-crystal (RMSD of 1.0 Å) (Supplementary Fig. 2a), validating the docking protocol. The remaining compounds were created similarly and docked into the HS2ST active site.

Results from the docking of **1** indicated two potential modes of binding (Fig. 5). In the top ranked pose, the entire glycan was translocated one disaccharide unit within the binding groove, into a position where a GlcA was poised to receive the sulfate from PAPS, rather than IdoA. The second binding mode matched that seen in the co-crystal structure (RMSD of 1.1 Å). Translocation of the oligosaccharide appeared to be promoted by the formation of electrostatic interactions between the *N*-sulfate and three positively charged protein residues (R288, R80, and K350). Indeed, these same amino acids coordinate with the sulfate moiety of the GlcNS that flanks the acceptor residue on the non-reducing side in the co-crystal structure (Fig. 5c). We refer to this site as the *N*-sulfate binding pocket. Notably, flanking GlcNS residues are required for HS2ST activity<sup>33</sup>; a feature that is now explained in terms of the enhanced stability provided by the interactions between the NS sulfate and R288/R80/K350.

The top-ranked model from docking the product of the reaction (**2**) was in essentially the same position as was found in the co-crystal of HS2ST with substrate (Fig. 6a,c). The 2S moiety was positioned where the sulfate from PAPS would reside, in the canonical phospho-sulfate binding (PSB) site that is shared amongst all sulfotransferases<sup>46</sup>. This pocket is favorable for a sulfate due to interactions with backbone amides (T84 and K83) that are oriented towards the PAPS phospho-sulfate moiety, along with the side chain of K83. Although **5** was not expected to bind due to a previously predicted steric clash between its 6-*O*-sulfate and P82/Y173<sup>33</sup>, the top-ranked model swiveled the 6-*O*-sulfate into the PSB site (Fig. 6b,d) enabling **5** to bind. The model was also translocated by one disaccharide unit, similar to the result for **1**. In the case of **5**, the 6-*O*-sulfate adopted a position equivalent to that occupied by the 2-*O*-sulfate in the modeled complex of **2**. Results from **3** indicate that compounds containing both 2S and 6S moieties are capable of binding in a similar orientation as in **2** (Supplementary Fig. 2b). Interestingly, addition of a 3S moiety to the non-reducing flanking GlcNS residue (**4**) appeared to prevent that residue from positioning the *N*-sulfate moiety in the *N*-sulfate binding pocket, resulting in a top-ranked model that was again translocated by one disaccharide unit (Supplementary Fig. 2c,d).



**Figure 6.** Predicted structures of HS2ST with non-substrate HS sequences that occupy the Phospho-Sulfate Binding (PSB) loop. **(a,c)** Docked model of compound 2. **(b,d)** Docked model of compound 5. **(a,b)** The trimeric protein is depicted in gray, except for the catalytic residue (H142A; magenta) and the PSB loop (T84, K83, and P82; red). **(c,d)** The catalytic residue and residues from the PSB loop are depicted in dark gray. Residues 180–191 are hidden for clarity. PAP or PAPS is displayed for context (gray, left), but was not present during the docking procedure.

Docking also generated poses in which the oligosaccharide was placed in a reversed orientation (flipped end-for-end) relative to the co-crystal structure (Supplementary Fig. 2e). This binding mode only occurred with 2-sulfated compounds (**2**, **3**, **4**), placing the 2S in the *N*-sulfate binding pocket and the NS in the PSB pocket. While plausible for binding, these modes are not appropriate for catalytic activity.

A 200 ns MD simulation of the trimeric apo-protein indicated an unexpected mobility of the protein loops proximal to the active site (Supplementary Fig. 3). By the end of the simulation, each of the three active sites of the trimer adopted topologically distinct conformations; one domain maintained a conformation similar to the crystal structure, another opened to expose the PAPS and substrate-binding sites, while the third closed over the substrate-binding site. A simulation of the non-sulfated HS-precursor (**0**) bound to HS2ST was performed to determine whether the complex was stable, given that HS2ST does not act on this oligosaccharide. Simulations of each of the other compounds were also performed, including a translocated model of **1**. All ligands remained bound during the simulation, although in many instances the active site loops showed motions as seen in the apo-HS2ST simulations. The ligand mobility (RMSF) and interaction energy for each compound was collected to determine the stability of the complexes (Table 3). HS precursor **0** demonstrated significantly enhanced positional fluctuations than the sulfated compounds, as well as showed the weakest interaction energy. Surprisingly, both starting positions of **1** appeared equally stable during the trajectory. The interaction energies of HS compounds largely corresponded to the experimentally observed trends in affinities, except for compound **5**, whose binding energy was predicted to be lower than that of **2**; possibly suggesting that the bound orientation of **5** differs from the pose selected from docking.

## Discussion

This study of the HS2ST binding specificity presents the surprising discovery that downstream HS products demonstrate higher affinity than both the enzymatic substrate and product of the HS2ST reaction, presenting an interesting challenge for biosynthesis. Considering the efficiency of 2-O-sulfation *in vivo*<sup>47</sup>, a mechanism for separating HS2ST from fully processed HS chains must exist within the cell. Chase-sulfation experiments of microsomal fractions indicated a stepwise process in which a limited amount of substrate was successively modified

|                | RMSF (Å)               | $\Delta G_{\text{bind}}$ (kcal/mol) |
|----------------|------------------------|-------------------------------------|
| 0 Aligned      | 4.6 ± 3.2 <sup>b</sup> | -169 ± 11 <sup>b</sup>              |
| 1 Aligned      | 1.4 ± 0.5              | -263 ± 11                           |
| 1 Translocated | 1.4 ± 0.7              | -268 ± 10                           |
| 2 Aligned      | 1.0 ± 0.5              | -327 ± 12                           |
| 3 Aligned      | 1.2 ± 0.6              | -343 ± 13                           |
| 4 Translocated | 1.2 ± 0.6              | -375 ± 14                           |
| 5 Translocated | 1.6 ± 1.0              | -310 ± 13                           |

**Table 3.** Root mean square fluctuations (RMSF) and interaction energies ( $\Delta G_{\text{bind}}$ )<sup>a</sup> obtained from MD simulations of models of compounds 0–5 docked into the HS2ST active site. Docked models are described as either in-register (Aligned) with regards to the structure of hexasaccharide-bound HS2ST (PDB ID: 4NDZ), or out-of-register (Translocated). <sup>a</sup>Values were calculated for the latter 50 ns of the 200 ns trajectory and do not include entropy contributions. <sup>b</sup>Standard deviation.

until the entire pool of oligosaccharides was consumed<sup>47</sup>, and a supramolecular complex of each of the sulfotransferases (GAGosome) was proposed as a mechanism for coupling the enzymatic reactions<sup>23</sup>. However, this model requires a regulatory process to maintain a unidirectional pathway and thereby avoid inhibition by downstream products. We propose a system in which small complexes of HS biosynthetic enzymes are separated into distinct Golgi stacks. This model is supported by previous studies that indicate epimerase localization is dependent on HS2ST expression<sup>26</sup>, whereas HS6ST isoforms localize independently of HS2ST<sup>48</sup>. Although discrete localization within the Golgi has not been reported for HS sulfotransferases, the functionally similar GlcNAc6ST family of enzymes is distributed in an isoform-specific manner<sup>49</sup>. Additional studies will be necessary to determine the cellular mechanism for regulating exposure of HS2ST to downstream products.

It was previously shown that HS2ST was inactive against substrates if glucosamine residues flanking uronic acid were 6-*O*-sulfated, and it was hypothesized that steric collisions involving the 6-*O*-sulfate groups prevented these sequences from binding in the active site of HS2ST<sup>33</sup>. Our results indicate that oligosaccharides containing the 6-*O*-sulfate motif bind stably in the active site, and molecular modeling suggests a mechanism in which the 6-*O*-sulfate moiety blocks co-factor binding by occupying the PAPS binding site, thereby preventing catalysis. The PSB pocket is a fundamental feature of sulfotransferases, and is structurally similar to the P-loop motif associated with nucleotide-binding proteins<sup>50</sup>. Heparin inhibition of members of the kinase family has been demonstrated previously<sup>51,52</sup>; however, Hathway, *et al.* also identified kinases that were unaffected by the highly sulfated GAG. Therefore, while this nucleotide-binding motif may be a general binding site for sulfate moieties, it is not sufficient for heparin inhibition.

The slightly higher affinity for product than substrate may simply be due to electrostatic interactions between the 2-*O*-sulfate and PSB loop. Although PAPS was not included in the binding assays, the cofactor was present during the inhibition assays and appeared unable to outcompete the HS fragments. An inability to release the product of the reaction would impede overall activity, so we propose a mechanism for destabilizing the product following 2-*O*-sulfation that is based upon known conformational preferences of the IdoA ring. The co-crystal structure contains an IdoA unit in the <sup>4</sup>C<sub>1</sub> conformation, which is unusual as this ring typically demonstrates a strong preference for the <sup>2</sup>S<sub>0</sub> or <sup>1</sup>C<sub>4</sub> conformations in solution<sup>53</sup>. This contrasts with GlcA, which strongly prefers the <sup>4</sup>C<sub>1</sub> conformation<sup>54</sup>. In the IdoA co-crystal structure, the <sup>4</sup>C<sub>1</sub> conformation is stabilized by a salt bridge between the IdoA carboxylate group and the side chain of R189. This stabilization is lost in the R189A mutation, leading to a preferential impairment of the activity HS2ST toward IdoA-containing substrates, relative to substrates containing GlcA<sup>27</sup>. Upon 2-*O*-sulfation, the <sup>4</sup>C<sub>1</sub> conformation of IdoA is further destabilized<sup>53,55</sup>, which we hypothesize leads to a disruption of its interaction with R189, ultimately enhancing the rate of product release. The higher activity of the wild type enzyme for IdoA, relative to GlcA, is then explained by the fact that both GlcA and GlcA2S are stable in the <sup>4</sup>C<sub>1</sub> conformation<sup>54</sup>, and thus would be predicted to be released from the enzyme more slowly.

## Methods

**Non-Glyc-HS2ST expression, purification, and mutagenesis.** The chicken HS2ST-Maltose Binding Protein (MBP) fusion protein was expressed and purified according to previous studies<sup>27</sup>. The fusion protein was transformed into *E. coli* Origami B (DE3) cells, cultured at 37 °C until OD<sub>600</sub> reached 0.5, then induced with isopropyl-β-D-thiogalactopyranoside and allowed to grow overnight at 20 °C. Cells were sonicated, followed by a two-step purification procedure. Whole-cell supernatant was passed through an amylose column with Tris buffer (25 mM Tris, 500 mM NaCl, pH 7.5). An elution buffer consisted of Tris Buffer with 40 mM Maltose. The amylose column eluent was concentrated and loaded on a Superdex 200 Increase 10/300 GL with PBS buffer (10 mM sodium phosphate, 150 mM NaCl, pH 7.8) to separate trimeric protein from aggregate and independent MBP (Supplementary Fig. 4). All reported mutations formed homotrimeric complexes according to the SEC profiles, including the negative control (NC = R189A/R80E/R184E). Protein concentration was assessed via UV at 280 nm (MW of trimer ~210 kDa, extinction coefficient ~119,000 M<sup>-1</sup> cm<sup>-1</sup> according to ExPASy ProtParam webtool<sup>56</sup>). The trimeric protein preparation was adjusted to 5% glycerol and stored at -20 °C until use. Additional non-glyc-HS2ST constructs were created with the Q5<sup>®</sup> Mutagenesis kit (New England Biolabs Inc.; cat. # E0554S) and sequenced at the Georgia Genomics Facility at the University of Georgia.



**Glyc-HS2ST expression and purification.** An expression construct was generated encoding the truncated catalytic domain of human HS2ST as an NH<sub>2</sub>-terminal fusion protein in the pGen2 expression vector essentially as described in prior studies<sup>57</sup>. The fusion protein coding region was comprised of a 25-amino acid signal sequence, an His<sub>6</sub> tag, AviTag, the “superfolder” GFP coding region, the 7-amino acid recognition sequence of the tobacco etch virus (TEV) protease followed by the human HS2ST catalytic domain region comprising of 328 amino acid residues. The recombinant human HS2ST was expressed as a soluble secreted protein by transient transfection of suspension culture HEK293-F cells (FreeStyle™ 293-Fcells, Thermo Fisher Scientific, Waltham MA) maintained at 0.5–3.0 × 10<sup>6</sup> cells/ml in a humidified CO<sub>2</sub> platform shaker incubator at 37 °C with 55% humidity. Transient transfection was performed using HEK293-F cells in the expression medium comprised of a 9:1 ratio of Freestyle™293 expression medium (Thermo Fisher Scientific, Waltham MA) and EX-Cell expression medium including Glutmax (Sigma-Aldrich). Transfection was initiated by the addition of plasmid DNA and polyethyleneimine as transfection reagent (linear 25-kDa polyethyleneimine, Polysciences, Inc., Warrington, PA) as previously described<sup>58,59</sup>. Twenty-four hours post-transfection the cell cultures were diluted with an equal volume of fresh media supplemented with valproic acid (2.2 mM final concentration) and protein production was continued for an additional 4–5 days at 37 °C. The cell cultures were harvested, clarified by sequential centrifugation at 1200 rpm for 10 min and 3500 rpm for 15 min at 4 °C, and passed through a 0.8 μm filter (Millipore, Billerica, MA). The enzyme preparation was adjusted to contain 20 mM HEPES, 20 mM imidazole, 300 mM NaCl, pH 7.5, and subjected to Ni<sup>2+</sup>-NTA Superflow (Qiagen, Valencia, CA) chromatography using a column pre-equilibrated with 20 mM HEPES, 300 mM NaCl, 20 mM imidazole, pH 7.5 (Buffer I). Following loading of the sample the column was washed with 3 column volumes of Buffer I followed by 3 column volumes of Buffer I containing 50 mM imidazole, and eluted with Buffer I containing 300 mM imidazole at pH 7.0. The protein was concentrated to approximately 3 mg/ml using an ultrafiltration pressure cell (Millipore, Billerica, MA) with a 10-kDa molecular mass cutoff membrane. The NH<sub>2</sub>-terminal GFP fusion sequences were removed by *in vitro* cleavage with TEV protease, which cleaves at the boundary of the HS2ST and the fusion sequences. His/GFP-tagged TEV protease<sup>60,61</sup> expressed in *E. coli* was added to the concentrated protein sample in the Ni<sup>2+</sup>-NTA elution buffer at a ratio of 1:5 relative to the fusion protein and incubated at 4 °C for 24 h. The cleaved HS2ST was further separated from the fusion tag and protease by Ni<sup>2+</sup>-NTA chromatography and gel filtration on a Superdex G-75 column (GE Healthcare) preconditioned with a buffer containing 20 mM HEPES, 100 mM NaCl, 0.05% Sodium azide, pH 7.0. Peak fractions of HS2ST were collected, buffer-exchanged into PBS and adjusted to 5% glycerol (Supplementary Fig. 5). The final protein preparation was stored at –20 °C until use. Protein concentration was determined via UV at 280 nm (MW of trimer ~117 kDa, extinction coefficient ~53,000 M<sup>-1</sup> cm<sup>-1</sup> as predicted by the ExpASY ProtParam webtool<sup>56</sup>).

**Bio-layer interferometry (BLI).** The Octet® RED96 system (FortéBio Inc.) was employed to detect the affinity of the enzymes for specific HS targets. Due to the small size of the glycan and the deep HS2ST binding cavity, the glycan and protein were treated as ligand and analyte, respectively. Varyingly sulfated HS fragments conjugated to biotin linkers were purchased from Glycan Therapeutics Inc. with purities greater than 93% according to HPLC (sequences located in Supplementary Table 1). The glycans were immobilized onto streptavidin (SA) labeled biosensors (FortéBio; cat. #18–5019) at 0.5 μg/mL. Experiments were performed at 25 °C in PBS Buffer (10 mM sodium phosphate, 150 mM NaCl, pH 7.8) with a final concentration of 1X Pall Kinetics buffer (FortéBio Inc; cat. #18–1092). Fresh biosensors and protein samples were used for each measurement (no regeneration step).

**Multi-sensor protocol.** The kinetics protocol began with a custom step (120 sec), followed by glycan loading (720 sec). A baseline (120 sec) was then established before the association and dissociation steps (420 sec each).

**Single-cycle protocol.** Another BLI assay format was employed to reduce the number of biosensors required to collect maximum response values for the creation of equilibrium binding curves. In this format, the same sensor is exposed to increasing concentrations of the protein without a dissociation step (Supplementary Fig. 6). Each step matched the timescale described for the multi-sensor protocol.

**Data analysis.** Results were processed with the Octet Data Analysis software, version 9.0. A reference sensor consisting of a glycan-coated biosensor and blank association wells (PBS-only) was included in each experiment and subtracted from the maximum response values. The non-specific binding signals between the biosensors and HS2ST were low for the 500 nM range of protein concentrations (Supplementary Fig. 7a); however, parallel reference sensors were implemented for the high concentration range (2 μM) employed with the glyc-HS2ST and compound 2. Equilibrium binding curves and affinity determinations were completed with Graphpad Prism V7.0 (Graphpad Prism Software Inc.) according to the one-site specific binding equation:  $Y = B_{max} * X / (Kd + X)$ .

**Inhibition assay.** The same glycan sequences were synthesized<sup>62</sup> for the inhibition assay as were used for the binding studies (Supplementary Table 1), except these terminated in p-nitrophenol rather than the PEG4-Biotin linker. The reaction mixture contained Tris pH 7.5 50 mM, 0.3 μg non-glyc-HS2ST WT, 0.08 nmol <sup>35</sup>S-labeled 3'-phosphoadenosine 5'-phosphosulfate ([<sup>35</sup>S] PAPS) (3 × 10<sup>4</sup> cpm), completely de-O-sulfated N-sulfated heparin (CDNS heparin) 1 μg, and either compound 2, 3, 4 (0, 1, 2, 4, 8, and 16 μg) or 5 (2, 20, 60, and 160 μg) in a total volume of 0.1 ml. After incubation at 37 °C for 1 h, the reaction was quenched by adding 1 mL of 0.01% Triton X-100 buffer at pH 5.0 containing 150 mM NaCl, 50 mM NaOAc, 3 M urea and 1 mM ethylenediaminetetraacetic acid (EDTA). The samples were loaded on a small DEAE column (200 μl in size). The column was then washed five times with the same buffer, each time 1 ml. The column was further washed five times with 0.25 M NaCl in

0.001% Triton X-100 buffer each time 1 ml. The column was eluted with 1 ml of 1 M NaCl in 0.001% Triton X-100 buffer. The  $^{35}\text{S}$  cpm of the eluted samples were tested by a scintillation counter.

**Flow cytometry.** *Chinese hamster ovary (CHO) cells.* Wildtype CHO-K1, pgsD-677 (*Ext1*<sup>-/-</sup>), and pgs-F17 cells (*Hs2st*<sup>-/-</sup>)<sup>43</sup> were grown as monolayers. Single cell suspensions were obtained by lifting the cells with Cell Dissociation Buffer (Life Technologies; cat. #13151014), sedimented by centrifugation, resuspended in ice cold PBS containing 0.1% BSA (Life Technologies; cat. # 14190250, Sigma; cat. # A3059–100G) and  $3 \times 10^5$  cells were aliquoted to each well of a 96-well clear v-bottom plate (Fisher; cat. # 7200108). After centrifugation at  $500 \times g$  at 4 °C for 5 min, the cells were resuspended in 200  $\mu\text{L}$  of in PBS/BSA containing 30 nM of non-glyc-HS2ST WT, mutants H142A, R189A, or the negative control (R189A/R80E/R184E). After 1 hr at 4 °C, the cells were centrifuged, washed twice with cold PBS/BSA, and resuspended in buffer containing 1  $\mu\text{g}/\text{ml}$  of Maltose Binding Protein Monoclonal Antibody (R29) (Fisher; cat. # MA5–14122). After 1 hr at 4 °C, the cells were washed twice, resuspended in a solution containing Alexa Fluor 488-tagged goat anti-mouse IgG (1.25  $\mu\text{g}/\text{mL}$ ) (Fisher; cat. # A11001) for 30 min. The cells were washed with ice cold PBS and analyzed by flow cytometry (Benton Dickinson FACSCalibur cytometer). Data was analyzed using FlowJo and plotted with Graphpad Prism V7.0 (Graphpad Prism Software Inc.).

*Mouse lung endothelial (MLE) cells.* The *Hs2st*<sup>-/-</sup> cell lines were generated by deriving immortalized primary MLE Cells from conditionally targeted (floxed) *HS2st* mice with SV40 large T-antigen and deleting the floxed gene by expressing Cre-recombinase. HS disaccharide composition analysis determined that the Cre-recombinase-mediated *Hs2st* deletion completely diminished 2-O-sulfation accompanied by increased N- and 6-O-sulfation (manuscript under review to report the generation and characterization of these two cell lines). Wildtype and *Hs2st*<sup>-/-</sup> cells were cultured in DMEM high glucose medium supplemented with 10% FBS, penicillin (100 U/mL), and streptomycin (100  $\mu\text{g}/\text{mL}$ ). The cells were detached by PBS without Calcium and Magnesium supplemented with 2 mM EDTA and 0.5% BSA (PBS-E). Wildtype and mutant cells ( $5 \times 10^5$  cells) were re-suspended in 100  $\mu\text{L}$  cold PBS-E and wildtype and mutant non-glyc-HS2ST protein were added to a final concentration of 1  $\mu\text{M}$ . The cells were washed with PBS-E buffer three times after incubation on ice for 45 min and re-suspended in 100  $\mu\text{L}$  cold PBS-E buffer. The anti-MBP antibody (MA5–14122, Invitrogen) was added to the cells (1:100) and incubated for 45 min. Following, the cells were washed three times with PBS-E buffer. The washed cells were incubated with Goat anti-mouse IgG secondary antibody conjugated with Alexa Fluor 488 (A11001, Invitrogen) for 45 min in 100  $\mu\text{L}$  PBS-E buffer. The cells were washed with PBS-E buffer three times and re-suspended in 300  $\mu\text{L}$  cold PBS-E buffer for flow cytometry (LSRII cytometer). Data was analyzed using FlowJo (Treestar) and plotted with Graphpad Prism V7.0 (Graphpad Prism Software Inc.).

**Molecular docking.** Vina-Carb<sup>45</sup>, a modified version of Autodock Vina<sup>63</sup>, was used for docking. The starting structure of the HS-PAP-HS2ST complex was obtained from PDB ID: 4NDZ<sup>33</sup>. Heparan sulfate hexasaccharide structures were built with the GLYCAM Web tool ([www.glycam.org](http://www.glycam.org)). The resulting PDB files were processed via AutodockTools<sup>64</sup>. All glycosidic linkages, hydroxyl groups, and most sulfate torsions could rotate during the model generation procedure. The only explicitly restrained torsion angles were N–C, N–S, and O6–S. Ring conformations are not capable of transitioning by the program. The gridbox was centered on the catalytic histidine residue (H142), which corresponds to an x, y, and z coordinate of 2.465, 41.242, and 77.155, respectively. The dimensions of the x, y, and z lengths were set to 50, 40, and 20, respectively (Supplementary Fig. 8). The exhaustiveness was set to 80, 10x higher than the default, due to the number of rotatable bonds and size of the gridbox. The chi\_coefficient and chi\_cutoff were set to 1 and 2, respectively. RMSD values for docked structures were computed for the ring atoms of the acceptor and flanking residues, as depicted in Fig. 4a, with Chimera<sup>65</sup>.

**Molecular dynamics (MD) simulations.** Chimera was employed to create a co-complex of the trimer bound to three ligands by aligning the docked model (with protein) to each of the remaining binding sites. MD simulations were performed with the pmemd.cuda version of AMBER14<sup>66,67</sup>. Amino acid and carbohydrate residues were parameterized with the FF12SB and GLYCAM06 (J-1) force fields, respectively<sup>67,68</sup>. The systems were neutralized and solvated with TIP3P in a truncated octahedral box with 15 Å distance from solute to the unit cell boundary. The Particle-Mesh Ewald algorithm accounted for electrostatic interactions<sup>69</sup> and the non-bonded cutoff was set to 8 Å. SHAKE was employed to permit an integration time step of 2 fs. Restraints were imposed in specific situations and were enforced with a 10 kcal/mol Å<sup>2</sup> energy barrier in each case. Three minimization steps were performed, each consisting of 1000 cycles of the steepest descent and 24,000 cycles of the conjugate gradient methods. The first restrained all solute atoms, the second restrained the protein backbone (C $\alpha$ ) and carbohydrate ring atoms (C1, C2, C3, C4, C5, O5), and the last was completely unrestrained. The systems were then heated to 300 K under NVT conditions over 60 ps, employing the Berendsen thermostat with a 1 ps coupling constant. The subsequent simulations were performed under NPT conditions for 200 ns.

**Computational analysis.** Analysis was performed on the latter 50 ns of the 200 ns simulations. Root mean square fluctuations (RMSF) were generated using the cpptraj module of AmberTools14<sup>70</sup>. Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) calculations were employed to quantify the total  $\Delta G_{\text{bind}}$  between the HS2ST receptor and HS ligand<sup>71,72</sup>. The igb values and internal dielectric (idel) constants were set to 2 and 4, respectively, according to a previous study of sulfated ligands<sup>73</sup>. Images were created using the Visual Molecular Dynamics (VMD) program<sup>74</sup>. Cartoon representations of glycans follow the SNFG format<sup>75</sup>, and the 3D shapes were created with the 3D-SNFG plugin for VMD<sup>76</sup>.

## References

- Victor, X. V. *et al.* Investigating the elusive mechanism of glycosaminoglycan biosynthesis. *Journal of Biological Chemistry* **284**, 25842–25853 (2009).
- Perrimon, N. & Bernfield, M. Specificities of heparan sulphate proteoglycans in developmental processes. *Nature* **404**, 725–728 (2000).
- Ferreras, C. *et al.* Endothelial heparan sulfate 6-O-sulfation levels regulate angiogenic responses of endothelial cells to fibroblast growth factor 2 and vascular endothelial growth factor. *Journal of Biological Chemistry* **287**, 36132–36146 (2012).
- Irie, A., Yates, E., Turnbull, J. & Holt, C. Specific heparan sulfate structures involved in retinal axon targeting. *Development (Cambridge, England)* **129**, 61–70 (2002).
- Forsberg, M. *et al.* Undersulfation of heparan sulfate restricts differentiation potential of mouse embryonic stem cells. *Journal of Biological Chemistry* **287**, 10853–10862 (2012).
- Shukla, D. *et al.* A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell* **99**, 13–22 (1999).
- Barth, H. *et al.* Cellular binding of hepatitis C virus envelope glycoprotein E2 requires cell surface heparan sulfate. *Journal of Biological Chemistry* **278**, 41003–41012 (2003).
- Salvador, B. *et al.* Filoviruses utilize glycosaminoglycans for their attachment to target cells. *Journal of virology* **87**, 3295–3304 (2013).
- Knelson, E. H., Nee, J. C. & Blobbe, G. C. Heparan sulfate signaling in cancer. *Trends in biochemical sciences* **39**, 277–288 (2014).
- Liu, D., Shriver, Z., Venkataraman, G., El Shabrawi, Y. & Sasisekharan, R. Tumor cell surface heparan sulfate as cryptic promoters or inhibitors of tumor growth and metastasis. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 568–573, <https://doi.org/10.1073/pnas.012578299> (2002).
- Lindahl, U. & Li, J. p. Interactions between heparan sulfate and proteins—design and functional implications. *International review of cell and molecular biology* **276**, 105–159 (2009).
- Singh, A. *et al.* The interaction of heparin tetrasaccharides with chemokine CCL5 is modulated by sulfation pattern and pH. *Journal of Biological Chemistry* **290**, 15421–15436 (2015).
- Lindahl, U., Bäckström, G., Thunberg, L. & Leder, I. G. Evidence for a 3-O-sulfated D-glucosamine residue in the antithrombin-binding sequence of heparin. *Proceedings of the National Academy of Sciences* **77**, 6551–6555 (1980).
- Hernaiz, M., Liu, J., Rosenberg, R. D. & Linhardt, R. J. Enzymatic modification of heparan sulfate on a biochip promotes its interaction with antithrombin III. *Biochemical and biophysical research communications* **276**, 292–297 (2000).
- Atha, D., Lormeau, J., Petitou, M. & Rosenberg, R. Contribution of 3-O- and 6-O-sulfated glucosamine residues in the heparin-induced conformational change in antithrombin III. *Biochemistry*, <https://doi.org/10.1021/bi00394a024> (1987).
- Atha, D. H., Lormeau, J. C., Petitou, M., Rosenberg, R. D. & Choay, J. Contribution of monosaccharide residues in heparin binding to antithrombin III. *Biochemistry* **24**, 6723–6729 (1985).
- Coombe, D. & Kett, W. Heparan sulfate-protein interactions: therapeutic potential through structure-function insights. *Cellular and molecular life sciences* **62**, 410–424 (2005).
- Zhou, H. *et al.* M402, a novel heparan sulfate mimetic, targets multiple pathways implicated in tumor progression and metastasis. *PLoS one* **6**, e21106 (2011).
- Schwörer, R., Zubkova, O. V., Turnbull, J. E. & Tyler, P. C. Synthesis of a Targeted Library of Heparan Sulfate Hexa- to Dodecasaccharides as Inhibitors of  $\beta$ -Secretase: Potential Therapeutics for Alzheimer's Disease. *Chemistry-A European Journal* **19**, 6817–6823 (2013).
- Esko, J. D. & Selleck, S. B. Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annual review of biochemistry* **71**, 435–471 (2002).
- Sarrazin, S., Lamanna, W. C. & Esko, J. D. Heparan sulfate proteoglycans. *Cold Spring Harbor perspectives in biology* **3**, a004952 (2011).
- Carlsson, P., Presto, J., Spillmann, D., Lindahl, U. & Kjellén, L. Heparin/heparan sulfate biosynthesis Processive formation of N-sulfated domains. *Journal of Biological Chemistry* **283**, 20008–20014 (2008).
- Jacobsson, I. & Lindahl, U. Biosynthesis of heparin. Concerted action of late polymer-modification reactions. *Journal of Biological Chemistry* **255**, 5094–5100 (1980).
- Liu, J., Moon, A. F., Sheng, J. & Pedersen, L. C. Understanding the substrate specificity of the heparan sulfate sulfotransferases by an integrated biosynthetic and crystallographic approach. *Current opinion in structural biology* **22**, 550–557 (2012).
- Préchoux, A., Halimi, C., Simorre, J.-P., Lortat-Jacob, H. & Laguri, C. d. C5-epimerase and 2-O-sulfotransferase associate *in vitro* to generate contiguous epimerized and 2-O-sulfated heparan sulfate domains. *ACS chemical biology* **10**, 1064–1071 (2015).
- Pinhal, M. A. *et al.* Enzyme interactions in heparan sulfate biosynthesis: uronosyl 5-epimerase and 2-O-sulfotransferase interact *in vivo*. *Proceedings of the National Academy of Sciences* **98**, 12984–12989 (2001).
- Bethea, H., Xu, D., Liu, J. & Pedersen, L. Redirecting the substrate specificity of heparan sulfate 2-O-sulfotransferase by structurally guided mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 18724–18729, <https://doi.org/10.1073/pnas.0806975105> (2008).
- Bullock, S. L., Fletcher, J. M., Beddington, R. S. & Wilson, V. A. Renal agenesis in mice homozygous for a gene trap mutation in the gene encoding heparan sulfate 2-sulfotransferase. *Genes & development* **12**, 1894–1906 (1998).
- Guglieri, S. *et al.* Minimum FGF2 Binding Structural Requirements of Heparin and Heparan Sulfate Oligosaccharides As Determined by NMR Spectroscopy†. *Biochemistry* **47**, 13862–13869 (2008).
- Lindahl, B., Eriksson, L. & Lindahl, U. Structure of heparan sulphate from human brain, with special regard to Alzheimer's disease. *Biochemical Journal* **306**, 177–184 (1995).
- Gerdy, B. *et al.* Detection of 2-O-sulfated iduronate and N-acetylglucosamine units in heparan sulfate by an antibody selected against acharan sulfate (IdoA2S-GlcNAc) n. *Journal of Biological Chemistry* **279**, 38346–38352 (2004).
- Tsuda, H. *et al.* Structures of Five Sulfated Hexasaccharides Prepared from Porcine Intestinal Heparin Using Bacterial Heparinase structural variants with apparent biosynthetic precursor-product relationships for the antithrombin III-binding site. *Journal of Biological Chemistry* **271**, 10495–10502 (1996).
- Liu, C. *et al.* Molecular mechanism of substrate specificity for heparan sulfate 2-O-sulfotransferase. *Journal of Biological Chemistry* **289**, 13407–13418 (2014).
- Ricard-Blum, S., Beraud, M., Raynal, N., Farndale, R. W. & Ruggiero, F. Structural requirements for heparin/heparan sulfate binding to type V collagen. *Journal of Biological Chemistry* **281**, 25195–25204 (2006).
- Muñoz, E. *et al.* Affinity, kinetic, and structural study of the interaction of 3-O-sulfotransferase isoform 1 with heparan sulfate. *Biochemistry* **45**, 5122–5128, <https://doi.org/10.1021/bi052403n> (2006).
- Xu, Y. *et al.* Chemoenzymatic synthesis of homogeneous ultralow molecular weight heparins. *Science* **334**, 498–501 (2011).
- Osmond, R. I., Kett, W. C., Skett, S. E. & Coombe, D. R. Protein–heparin interactions measured by BIAcore 2000 are affected by the method of heparin immobilization. *Analytical biochemistry* **310**, 199–207 (2002).
- Solá, R. J. & Griebenow, K. Chemical glycosylation: New insights on the interrelation between protein structural mobility, thermodynamic stability, and catalysis. *FEBS letters* **580**, 1685–1690 (2006).
- Hashimoto, Y. *et al.* Effects of glycosylation on the structure and dynamics of eel calcitonin in micelles and lipid bilayers determined by nuclear magnetic resonance spectroscopy. *Biochemistry* **38**, 8377–8384, <https://doi.org/10.1021/bi983018j> (1999).
- Rudd, P. M. *et al.* The effects of variable glycosylation on the functional activities of ribonuclease, plasminogen and tissue plasminogen activator. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* **1248**, 1–10 (1995).

41. Zhou, X., O'Leary, T. R., Xu, Y., Sheng, J. & Liu, J. Chemoenzymatic synthesis of heparan sulfate and heparin. *Biotransformation* **30**, 296–308 (2012).
42. Zhang, L. *et al.* The effect of precursor structures on the action of glucosaminyl 3-O-sulfotransferase-1 and the biosynthesis of anticoagulant heparan sulfate. *Journal of Biological Chemistry* **276**, 28806–28813 (2001).
43. Zhang, L., Lawrence, R., Frazier, B. A. & Esko, J. D. CHO glycosylation mutants: proteoglycans. *Methods in enzymology* **416**, 205–221 (2006).
44. Lidholt, K. *et al.* A single mutation affects both N-acetylglucosaminyltransferase and glucuronosyltransferase activities in a Chinese hamster ovary cell mutant defective in heparan sulfate biosynthesis. *Proceedings of the National Academy of Sciences* **89**, 2267–2271 (1992).
45. Anita Nivedha, D. T., Robert Woods. *Vina-Carb: Implementation of a novel scoring function for carbohydrate docking* (2013).
46. Negishi, M. *et al.* Structure and function of sulfotransferases. *Archives of biochemistry and biophysics* **390**, 149–157 (2001).
47. Höök, M., Lindahl, U., Hallen, A. & Bäckström, G. Biosynthesis of heparin. Studies on the microsomal sulfation process. *Journal of Biological Chemistry* **250**, 6065–6071 (1975).
48. Nagai, N., Habuchi, H., Esko, J. D. & Kimata, K. Stem domains of heparan sulfate 6-O-sulfotransferase are required for Golgi localization, oligomer formation and enzyme activity. *Journal of Cell Science* **117**, 3331–3341 (2004).
49. de Graffenried, C. L. & Bertozzi, C. R. Golgi localization of carbohydrate sulfotransferases is a determinant of L-selectin ligand biosynthesis. *Journal of Biological Chemistry* **278**, 40282–40295 (2003).
50. Via, A., Ferrè, F., Brannetti, B., Valencia, A. & Helmer-Citterich, M. Three-dimensional view of the surface motif associated with the P-loop structure: cis and trans cases of convergent evolution. *Journal of molecular biology* **303**, 455–465 (2000).
51. Hathaway, G. M., Lubben, T. H. & Traugh, J. A. Inhibition of casein kinase II by heparin. *Journal of Biological Chemistry* **255**, 8038–8041 (1980).
52. Herbert, J.-M., Clowes, M., Lea, H. J., Pascal, M. & Clowes, A. W. Protein kinase C  $\alpha$  expression is required for heparin inhibition of rat smooth muscle cell proliferation *in vitro* and *in vivo*. *Journal of Biological Chemistry* **271**, 25928–25935 (1996).
53. Hsieh, P.-H., Thieker, D. F., Guerrini, M., Woods, R. J. & Liu, J. Uncovering the Relationship between Sulphation Patterns and Conformation of Iduronic Acid in Heparan Sulphate. *Scientific reports* **6** (2016).
54. Hsieh, P.-H., Xu, Y., Keire, D. A. & Liu, J. Chemoenzymatic synthesis and structural characterization of 2-O-sulfated glucuronic acid-containing heparan sulfate hexasaccharides. *Glycobiology* **24**, 681–692 (2014).
55. Sattelle, B. M., Hansen, S. U., Gardiner, J. & Almond, A. Free energy landscapes of iduronic acid and related monosaccharides. *Journal of the American Chemical Society* **132**, 13132–13134 (2010).
56. Gasteiger, E. *et al.* *Protein identification and analysis tools on the ExPASy server.* (Springer, 2005).
57. Subedi, G. P., Johnson, R. W., Moniz, H. A., Moremen, K. W. & Barb, A. High Yield Expression of Recombinant Human Proteins with the Transient Transfection of HEK293 Cells in Suspension. *Journal of visualized experiments: JoVE*, e53568–e53568 (2015).
58. Meng, L. *et al.* Enzymatic basis for N-glycan sialylation: structure of ratalpha2,6-sialyltransferase (ST6GAL1) reveals conserved and unique features for glycan sialylation. *J Biol Chem* **288**, 34680–34698, <https://doi.org/10.1074/jbc.M113.519041> (2013).
59. Subedi, G. P., Johnson, R. W., Moniz, H. A., Moremen, K. W. & Barb, A. High Yield Expression of Recombinant Human Proteins with the Transient Transfection of HEK293 Cells in Suspension. *J Vis Exp*, e53568, <https://doi.org/10.3791/53568> (2015).
60. Wu, X. *et al.* A novel method for high-level production of TEV protease by superfolder GFP tag. *Journal of biomedicine & biotechnology* **2009**, 591923, <https://doi.org/10.1155/2009/591923> (2009).
61. van den Berg, S., Lofdahl, P. A., Hard, T. & Berglund, H. Improved solubility of TEV protease by directed evolution. *J Biotechnol* **121**, 291–298, <https://doi.org/10.1016/j.jbiotec.2005.08.006> (2006).
62. DeAngelis, P., Liu, J. & Linhardt, R. Chemoenzymatic synthesis of glycosaminoglycans: re-creating, re-modeling and re-designing nature's longest or most complex carbohydrate chains. *Glycobiology* **23**, 764–777, <https://doi.org/10.1093/glycob/cwt016> (2013).
63. Oleg, T. & Arthur, J. O. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, <https://doi.org/10.1002/jcc.21334> (2009).
64. Sanner, M. F. Python: a programming language for software integration and development. *J Mol Graph Model* **17**, 57–61 (1999).
65. Pettersen, E. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25**, 1605–1612, <https://doi.org/10.1002/jcc.20084> (2004).
66. Götz, A. *et al.* Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *Journal of chemical theory and computation* **8**, 1542–1555, <https://doi.org/10.1021/ct200909j> (2012).
67. Case, D. *et al.* The Amber biomolecular simulation programs. *Journal of computational chemistry* **26**, 1668–1688, <https://doi.org/10.1002/jcc.20290> (2005).
68. Kirschner, K. *et al.* GLYCAM06: a generalizable biomolecular force field. *Carbohydrates. Journal of computational chemistry* **29**, 622–655, <https://doi.org/10.1002/jcc.20820> (2008).
69. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N $\cdot$  log (N) method for Ewald sums in large systems. *The Journal of chemical physics* **98**, 10089–10092 (1993).
70. Roe, D. R. & Cheatham, T. E. III PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *Journal of chemical theory and computation* **9**, 3084–3095 (2013).
71. Kollman, P. A. *et al.* Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research* **33**, 889–897 (2000).
72. Xu, L., Sun, H., Li, Y., Wang, J. & Hou, T. Assessing the Performance of MM/PBSA and MM/GBSA Methods. 3. The Impact of Force Fields and Ligand Charge Models. *The journal of physical chemistry. B* **117**, 8408–8421, <https://doi.org/10.1021/jp404160y> (2013).
73. Krishnamurthy, V. R. *et al.* Glycopeptide analogues of PSGL-1 inhibit P-selectin *in vitro* and *in vivo*. *Nature communications* **6** (2015).
74. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *Journal of molecular graphics* **14**, 33–38 (1996).
75. Varki, A. *et al.* Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **25**, 1323–1324 (2015).
76. Thieker, D. F., Hadden, J. A., Schulten, K. & Woods, R. J. 3D implementation of the symbol nomenclature for graphical representation of glycans. *Glycobiology* **26**, 786–787 (2016).

## Acknowledgements

The authors thank the National Institutes of Health for support U01 CA207824 (to R.J.W.), P41 GM103390 (to R.J.W.), P41 GM107012 (to K.W.M.), RO1 HL093339 (to L.W.) and GM33063 (to J.D.E.).

## Author Contributions

D.F.T., L.W., K.W.M., J.L., J.D.E., and R.J.W. designed the study; D.F.T. and T.F. generated non-glyc-HS2ST protein and mutants; D.F.T. performed BLI assays, docking calculations, and molecular dynamics; D.K. generated glyco-HS2ST; Y.X. performed the inhibition assays; C.N. produced CHO cells; H.Q. produced MLEC cells; C.N. and H.Q. performed flow cytometry experiments; all authors analyzed the data; and D.F.T. and R.J.W. wrote the manuscript, which was reviewed by all co-authors.

### Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-29602-4>.

**Competing Interests:** J. L. is a founder of Glycan Therapeutics. D.F.T., Y.X., D.K., C.N., T.F., H.Q., L.W., K.W.M., J.D.E., and R.J.W. declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018