SCIENTIFIC REP🅞RTS

OPEN

Transition state-based ST6Gal I inhibitors: Mimicking the phosphodiester linkage with a triazole or carbamate through an enthalpy-entropy compensation

Andrew P. Montgomery[1], Danielle Skropeta[1,2,3] & Haibo Yu [1,2,3]

Human β-galactoside α-2,6-sialyltransferase I (ST6Gal I) catalyses the synthesis of sialylated glycoconjugates. Overexpression of ST6Gal I is observed in many cancers, where it promotes metastasis through altered cell surface sialylation. A wide range of sialyltransferase inhibitors have been developed, with analogues structurally similar to the transition state exhibiting the highest inhibitory activity. To improve synthetic accessibility and pharmacokinetics of previously reported inhibitors, the replacement of the charged phosphodiester linker with a potential neutral isostere such as a carbamate or a 1,2,3-triazole has been investigated. Extensive molecular dynamics simulations have demonstrated that compounds with the alternate linkers could maintain key interactions with the human ST6Gal I active site, demonstrating the potential of a carbamate or a 1,2,3-triazole as a phosphodiester isostere. Free energy perturbation calculations provided energetic evidence suggesting that the carbamate and 1,2,3-triazole were slightly more favourable than the phosphodiester. Further exploration with free energy component, quasi-harmonic and cluster analysis suggested that there is an enthalpy-entropy compensation accounting for the replacement of the flexible charged phosphodiester with a neutral and rigid isostere. Overall, these simulations provide a strong rationale for the use of a carbamate or 1,2,3-triazole as a phosphodiester isostere in the development of novel inhibitors of human ST6Gal I.

Sialic acid (*N*-acetylneuraminic acid, Neu5Ac) is one of the body's most important sugars next to glucose[1]. Sialylation, the addition of sialic acid to cell surface molecules via the glycosyltransferase (GT) enzymes known as sialyltransferases (STs), is integral to cell function, governing numerous important biological processes including cell–cell recognition, protein targeting, cell adhesion, and fertilisation[2–4]. However, abnormal sialylation is strongly associated with cancer, with hypersialylation of up to 30–40% seen in many cancers, along with marked upregulation of ST activity. This has been directly correlated with an increased metastatic potential of tumours and poor patient prognosis[5,6]. Upregulation of STs has also been shown to produce cellular resistance in both paclitaxel and cisplatin treatment in human ovarian and colorectal cancer cell lines, along with reducing treatment efficacy[7]. Thus the critical role of STs in the process of tumour metastasis as well as drug resistance demonstrates that inhibitors of STs could be used either alone as anti-metastatic agents, or in combination with existing drugs to enhance their sensitivity against resistant tumours[8].

In humans, there are 20 different STs each of which share a common sugar nucleotide donor, cytidine 5′-monophosphate Neu5Ac (CMP-Neu5Ac)[2,9]. STs are classified into one of four groups based on the position and identity of the glycosyl acceptor (galactose, *N*-acetylgalactosamine or sialic acid) to which sialic acid is transferred: ST3Gal I-VI, ST6Gal I-II, ST6GalNAc I-VI, and ST8Sia I-VI[2,3]. Each ST subtype controls the synthesis

[1]School of Chemistry, Faculty of Science, Medicine and Health, University of Wollongong, Wollongong, NSW 2522, Australia. [2]Centre for Medical and Molecular Bioscience, University of Wollongong, Wollongong, NSW 2522, Australia. [3]Illawarra Health and Medical Research Institute, Wollongong, NSW 2522, Australia. Correspondence and requests for materials should be addressed to H.Y. (email: hyu@uow.edu.au)
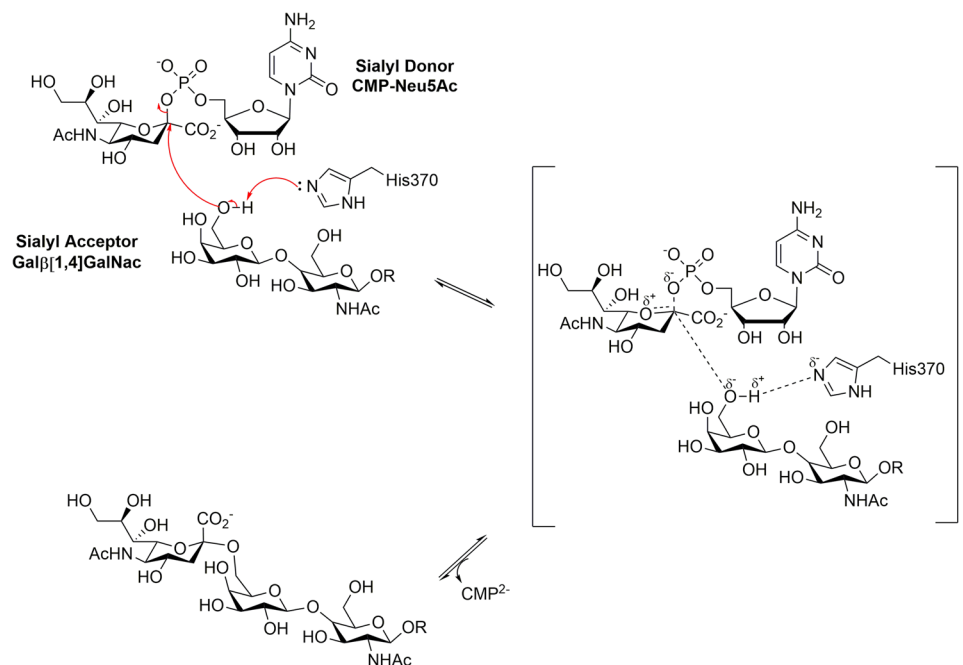
1

Figure 1. Proposed reaction scheme for the catalytic mechanism of human ST6Gal I. R represents a glycan chain of a glycoconjugate. Replicated from Montgomery, A., Szabo, R., Skropeta, D. & Yu, H. Computational characterisation of the interactions between human ST6Gal I and transition-state analogue inhibitors: insights for inhibitor design. *J. Mol. Recogn* **29**, 210–222, doi:10.1002/jmr.2520 (2016) with permission from Wiley.

of specific sialylated structures with unique biological roles, therefore it is crucial that selective inhibitors are developed.

Herein we focus on human β-galactoside α-2,6-sialyltransferase I (ST6Gal I), which is one of the two subfamilies of the ST6 family that belongs to the GT29 family (http://www.cazy.org)[10]. ST6Gal I catalyses the transfer of a Neu5Ac residue to the 6′-hydroxyl group of the terminal Gal residues of the disaccharide Galβ[1,4]GalNAc[3,11]. ST6Gal I mediated sialylation can modulate protein conformation, oligomerisation and/or receptor internalisation[12], depending on the substrate targeted. Overexpression of ST6Gal I is observed in several types of cancers including lung, cervical, ovarian, pancreatic, breast and colon carcinoma[13–16]. It has been recently reported that ST6Gal I promotes a cancer stem cell (CSC) phenotype, with the activity of ST6Gal I critical for CSC behaviours including tumour spheroid growth, chemoresistance and tumour initiating potential[16]. Additionally, overexpression of ST6Gal I has been suggested to promote cancer cell metastasis through altered sialylation patterns affecting the function of β1 integrin[17,18]. ST6Gal I also has an important role in galectin regulation, via the sialylation of their galactose-containing substrates[12,13]. For instance, α-2,6 sialylation of the galectin substrates Fas[19] and TNFR1[20] death receptors have been shown to hinder the apoptosis process, thus identifying ST6Gal I overexpression as an inhibitor of cell death pathways. It has also been recently demonstrated that ST6Gal I has a role in promoting tumour cell survival within serum-depleted environments, such as those within the hypovascularised regions of large, solid tumours[21].

The catalytic mechanism of human ST6Gal I (Fig. 1) follows a classical S_N2-like direct displacement mechanism for inverting GTs[11,22]. It is facilitated by the catalytic base His370, which promotes the nucleophilic attack on the sialyl donor anomeric carbon by deprotonating the galactose 6′-hydroxyl of the acceptor[23,24]. This generates an oxocarbenium-like transition state (TS), followed by CMP acting as a leaving group. The β-configuration at the anomeric carbon of the donor is inverted in the final product[24]. This is supported by a model of the Michaelis complex generated from the glycan binding mode observed in the crystal structure of human ST6Gal I[25].

The overexpression of ST6Gal I in various cancers has resulted in the development of a diverse range of ST inhibitors[26–28]. Of these, analogues mimicking the proposed oxocarbenium ion-like TS of the donor (Fig. 1) exhibit the highest affinity to STs[29,30]. TS analogue inhibitors incorporating a 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid (Neu5Ac2en) moiety are amongst the most potent inhibitors. The TS-like characteristics are captured with the C2-C3 double bond mimicking the planar anomeric carbon and an extra carbon between the anomeric carbon and CMP leaving group to mimic the elongated distance in the TS[29,31]. Various strategies have been adopted to increase the potencies of these inhibitors. Derivatives which explore variations of the glycerol side chain of the Neu5Ac2en moiety produced the greatest inhibitory activity. Of these, phenoxy derivative **1** (Fig. 2) is one of the most potent TS analogue inhibitors reported to date ($K_i = 29$ nM for rat liver α-2,6-ST)[32]. More synthetically assessable derivatives that incorporate an aryl replacement of the Neu5Ac2en moiety have comparable activity to the phenoxy derivative **1**[29,30,33]. Of these the (*R*)-isomer of the phenoxy derivative **2** (Fig. 2) was a potent inhibitor of rat liver α-2,6-ST ($K_i = 70$ nM)[29] and human ST6Gal I ($K_i = 19$ nM)[34] and is the lead compound herein.
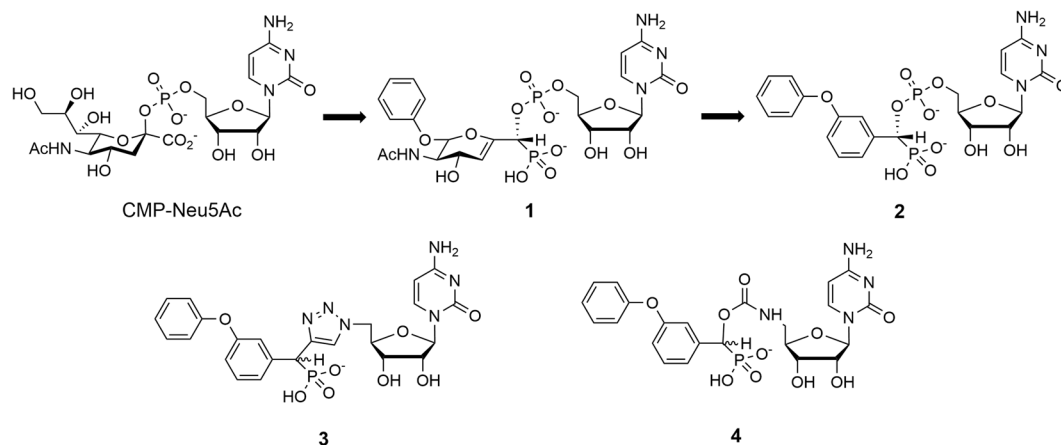
Figure 2. Development of the phosphodiester-linked lead compound (2) and the analogous 1,2,3-triazole- (3) and carbamate- (4) linked derivatives.

Recently, cyclopentyl[34] and amide[35] ST inhibitors have been reported with comparable nanomolar activity to the phenoxy lead derivative (**2**) against human ST6Gal I.

Many of the above inhibitors were designed with limited knowledge of the target ST structure. The recent characterisation of the mammalian ST crystal structures, in particular that of human ST6Gal I[25] and human ST8Sia III[36], has facilitated the use of computational modelling to aid in the design of new selective ST inhibitors. In this work, we use computational modelling to investigate the effect of replacing the charged phosphodiester linkage of TS analogue ST inhibitors. This is being investigated, as while contributing to potency[37] the phosphodiester linkage may potentially contribute to low bioavailability, loss of activity due to cleavage by phosphatases or instability in the ST active site[38,39]. To build upon our previous work[40,41], molecular dynamics (MD) simulations and binding free energy calculations via the free energy perturbation (FEP) method will be undertaken to investigate the effect of replacing the phosphodiester linker of our lead compound **2**, with either a 1,2,3-triazole (**3**; Fig. 2) or carbamate (**4**; Fig. 2), has on binding to human ST6Gal I. Force field parameters compatible with the CHARMM PARAM36 force field for the ligands will be developed[42]. Through equilibrium MD simulations the conserved interactions as well as potential differences between each compound will be compared. Furthermore, relative binding free energy calculations will be used to study the effect of the linker on binding affinity.

Methods

Force field development. Force fields generated from the CHARMM General Force Field (CGenFF)[43] using the paramchem interface (https://cgenff.paramchem.org) for the phosphodiester-linked **2** and the corresponding 1,2,3-triazole-linked **3** and carbamate-linked **4** derivatives (Fig. 2) were optimised using the GAAMP method[44]. The partial atomic charges (see Supplementary Fig. S1 and Table S1) and identified "soft" dihedrals, which were not already represented in the CHARMM and CGenFF force fields or that experienced a significant penalty score from paramchem, were optimised (see Supplementary Fig. S2).

Molecular dynamics simulations. All MD simulations were prepared using VMD version 1.9.2[45]. An overview of all simulations is given in Supplementary Table S2. Simulation of human ST6Gal I (PDB id: 4JS2)–inhibitor complexes with the reported phosphodiester-linked compound **2** and the proposed 1,2,3-triazole and carbamate-linked derivatives (**3** and **4** respectively; Fig. 2) were undertaken to determine if a carbamate or 1,2,3-triazole is a suitable replacement for the phosphodiester linker used in previously reported inhibitors. The initial structure for the ST6Gal I–inhibitor complexes was taken from previously reported docking studies[40]. Additionally, simulations of each inhibitor in a water box were also performed to determine the effect of linker flexibility. Prior to simulation being undertaken PROPKA 3.1[46] was used to predict the protonation state of ST6Gal I residues and ionisable ligand groups at pH 7.0.

Preparation of each simulation involved solvating the system in a box of water molecules, which extended approximately 9 Å from the surface of the protein, resulting in each system being solvated in an 82 Å × 82 Å × 82 Å cubic TIP3P[47] water box. Each box was neutralised with counter ions of Na$^+$ and Cl$^-$, with the salt (NaCl) concentration set to 0.15 mol/L. Simulations were carried out using the NAMD 2.9 package[48]. The protein was represented with the non-polarisable CHARMM PARAM36 force field[49]. The force field for each ligand were those optimised using the GAAMP method[44]. All systems were simulated in periodic boundary conditions using the Langevin algorithm for maintaining the temperature at 298.15 K, and the Langevin Piston Nose–Hoover method[50,51] for maintaining a constant pressure of 1.0 bar. The electrostatic interactions were calculated using the Particle Mesh Ewald method[52]. The van der Waals forces were treated with a cut-off of 12 Å and a smoothening function between 10 and 12 Å. All of the covalent bonds involving hydrogen were kept rigid with the RATTLE algorithm[53]. The integration time step was set to 1.0 fs, 2.0 fs and 4.0 fs for bonded, non-bonded and long-range electrostatics, respectively.

Equilibrium simulations for all human ST6Gal I-inhibitor complex systems were initially undertaken. During these simulations, a harmonic restraint was placed upon all Cα atoms of human ST6Gal I, with a decreasing
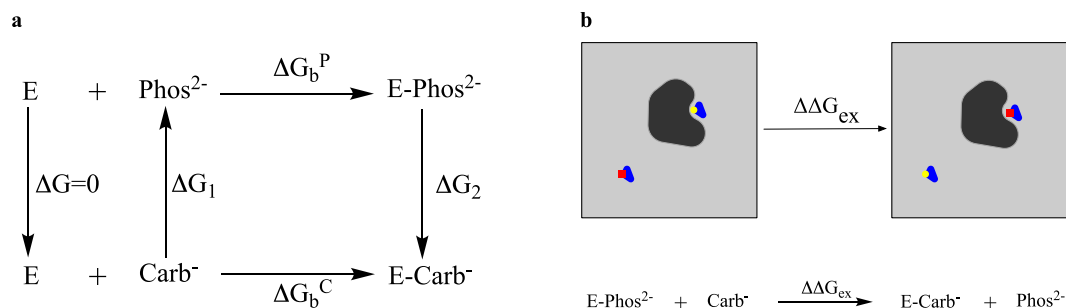
Figure 3. The thermodynamic cycle used (**a**) and a schematic representation of the perturbation performed (**b**) in the phosphodiester to carbamate FEP calculations. Phos and the yellow circle represent the phosphodiester-linked lead compound *(R)*-2, Carb and the red square represent the carbamate-linked derivative *(R)*-4, and E represents human ST6Gal I.

force constant from 32.0 to 1.0 kcal/mol/$\mathrm{\AA}^2$ over 10 ns. This decrease was achieved gradually by halving the force constant for each respective 1 ns interval until a force constant of 1.0 kcal/mol/$\mathrm{\AA}^2$ was obtained (i.e. 32.0, 16.0 … 1.0 kcal/mol/$\mathrm{\AA}^2$). A force constant of 1.0 kcal/mol/$\mathrm{\AA}^2$ was then applied for the remainder of the 10 ns thermal equilibration. The simulations were further continued for 90 ns without any restraints (see Supplementary Table S2). The inhibitor only simulations (see Supplementary Table S2) were performed for 20 ns with a harmonic restraint placed upon one atom of the corresponding ligand, with a force constant of 32 kcal/mol/$\mathrm{\AA}^2$. To perform trajectory analysis, snapshots were saved every ps (i.e. every 1,000 steps). Replicates of each simulation from different starting coordinates were generated by starting each simulation with three different initial minimisation steps (i.e. 2,000, 2,500 and 3,000 steps) and initiated with different velocities. In total, 2.2 μs of MD simulations were performed.

Simulation trajectory analysis. To monitor the global properties of the structural evolution during MD simulations, the atomic positional root-mean-square deviations (RMSD) with respect to a reference structure and the atomic positional root-mean-square fluctuations (RMSF) were calculated using VMD[45].

Hydrogen bond interactions were analysed according to a geometric criterion. A hydrogen bond was defined by a minimum donor–hydrogen–acceptor angle of 120° and a maximum heavy atom distance of 3 Å. Hydrophobic contacts were also analysed according to a geometric criterion of a maximum distance of 4 Å between hydrophobic atom pairs. To identify these interactions CHARMM version 38a1[54] was utilised.

The quasi-harmonic approximation[55–57] was used to examine the loss of configurational entropy of each ligand upon binding to human ST6Gal I. This analysis was performed, using CHARMM version 38a1[54], for trajectories encompassing 11–100 ns (11–55 ns and 56–100 ns were also performed to monitor convergence; See Supplementary Table S3) of each ligand in complex with human ST6Gal I and trajectories encompassing 1–20 ns (1–10 ns and 11–20 ns were also performed to monitor convergence; See Supplementary Table S3) of the ligand simulations. Prior to the analysis, the Cartesian coordinates of the ligands were aligned after least-squares fit superposition of the heavy atoms of all configurations onto the initial structure, to eliminate the overall translation and rotation. The change in configurational entropy (ΔS_{conf}) was determined based on the following relationship:

$$\Delta S_{conf} = S_{(complex)} - S_{(free)}. \tag{1}$$

Using the structures obtained from the MD simulations, conformational cluster analysis was performed for the ligands using MMTSB tools[58]. Clusters were identified using k-means clustering with an RMSD threshold for ligand heavy atoms of 1.6 Å. For each simulation, cluster analysis was based on combined trajectories of every 10 ps of triplicate simulations. This resulted in 27,000 structures used in the analysis of the complexed simulations and 6,000 structures for the ligand only simulations.

Free energy perturbation calculations. All free-energy perturbation (FEP) calculations described herein were carried out using the NAMD 2.10 package[48]. The simulation parameters were as per the MD simulation method described above. Depending on the calculation the starting structure was taken from the ST6Gal I-inhibitor simulations of the *(R)*-diastereomers of either the phosphodiester-linked lead compound **(R)-2** or the carbamate-linked derivative **(R)-4** and temperature was maintained at either 288.15 K or 298.15 K (see Supplementary Table S2).

Each calculation utilised a dual ligand approach, where a ligand is mutated to an analogue in the human ST6Gal I binding site and the reverse mutation is performed simultaneously on the analogue in bulk water in the same system (Fig. 3). This approach has been undertaken to minimise potential electrostatic finite-size effects posed by the change in charge state when the *(R)*-diastereomer of the phosphodiester-linked lead compound **2** (net charge of -2e) is perturbed to the *(R)*-diastereomer of either the 1,2,3-triazole- **3** or carbamate-linked **4** derivatives (net charge of -1e) (Fig. 3)[59,60]. In this approach, the overall charge of the system remains constant. This allows for the reduction of errors arising from the large solvation energy of charged ligands and potential

artefacts from the use of periodic boundary conditions. A similar approach has been successfully employed to handle changes in charge state in several different biomolecular systems[61].

Using the dual ligand approach, the change in the binding free energy due to linker mutation with perturbation from phosphodiester to carbamate as an example can be expressed as (Fig. 3)

$$
\begin{aligned}
\Delta\Delta G_b &= \Delta G_b^C - \Delta G_b^P \\
&= \Delta G_2 - \Delta G_1 \\
&= \Delta\Delta G_{ex} \\
&= \Sigma_i \Delta\Delta G_{ex,i}
\end{aligned}
\tag{2}
$$

In the FEP method, one introduces a hybrid Hamiltonian, $H(\lambda) = (1 - \lambda)H_0 + \lambda H_1$, where H_0 represents the Hamiltonian for the initial state and H_1 for the final state. The interval between 0 and 1 is divided into n subintervals (λ_i, where $i = 1, \ldots, n - 1$), and for each subinterval the free energy difference is calculated from the ensemble average, $\Delta\Delta G_{ex,i} = -k_B T \ln < \exp[-(H(\lambda_{i+1}) - H(\lambda_i))/k_B T >_{\lambda_i}$. The free energy difference between initial and final states is obtained from the sum, $\Delta\Delta G = \Sigma_i \Delta\Delta G_{ex,i}$. Alchemical transformations were carried out with 20 uniformly distributed λ values between 0 and 1 (0.0, 0.05, 0.1, …, 0.9, 0.95, 1.0), with the transformation simulated in reverse directions beginning from the $0. \times 5$ λ values (e.g. $0.0 \leftarrow 0.05 \rightarrow 0.1$). To prevent the occurrence of singularities at values of λ where an atom disappears/appears, a soft-core potential with $\alpha = 4.0$ Å^2 was used to reshape the Lennard-Jones potential into a form devoid of singularity when λ approaches either 0 or 1[62,63].

Each λ window underwent initial 5000 energy minimisation steps followed by 1 ns equilibration simulation to provide starting coordinates. Each window was then subsequently simulated for multiple ns with a harmonic restraint placed on the centre of mass of ST6Gal I. A harmonic restraint was also applied to one atom from the ligand in the bulk. These restraints were applied to prevent association of the ligand in the bulk and ST6Gal I and had no effects on the calculated binding free energy difference. Both restraints used a force constant of 32 kcal/mol/Å^2. In total 1.02 μs of FEP simulations were performed.

Results and Discussion

1,2,3-Triazole- and carbamate-linked *(R)*- and *(S)*-diastereoisomers maintain consistent binding site contacts with human ST6Gal I when compared to their phosphodiester derivative.

Hydrogen bond and hydrophobic contact analysis of phosphodiester- and carbamate-linked *(R)*-diastereomers has previously been published[40]. These simulations were performed using ligand force fields generated using the general AMBER force field (GAFF)[64] and with human ST6Gal I represented with the non-polarisable CHARMM PARAM36 force field[49]. In this work, a force field parameterisation compatible with the CHARMM PARAM36 force field was carried out for the ligands. We have undertaken simulation of both the *(R)*- and *(S)*-diastereoisomers of the phosphodiester- (**2**), 1,2,3-triazole- (**3**) and carbamate-linked (**4**) derivatives (Fig. 2) in complex with human ST6Gal I with the newly developed force field.

The MD simulations of each complex showed a stable trajectory with the positional RMSD of backbone Cα atoms being 1.5–2.0 Å for each simulation (see Supplementary Figs S3–S8). RMSFs for the Cα atoms, were generally consistent across all simulations (see Supplementary Figs S3–S8). Larger RMSF values, across each simulation, were observed for the flexible loop between Tyr355 and Tyr369. Large RMSF values were also observed for the regions between Gln235—Asn250 and Ser260—Ile295, which comprise the portion of the active site that is expected to interact with the sialyl acceptor. This can most likely be attributed to the phenoxy being within the proposed sialyl acceptor site. Across all simulations each ligand remained tightly associated with the human ST6Gal I active site. However, there was some variation in ligand position within this active site, which can be seen in the RMSD values that correspond to the heavy atoms of each ligand (see Supplementary Figs S3–S8). These fluctuations can be attributed to the *m*-phenoxy substituent, which depending upon the simulation either rotated to maintain contacts with the sialyl acceptor site or only weakly interacted with peripheral residues of this pocket, resulting in positional fluctuation.

Interactions in the human ST6Gal I-ligand complexes analysed were characterised by monitoring the hydrogen bonds and hydrophobic contacts (see Supplementary Tables S4–S15). Each simulation revealed similar interactions with the ST6Gal I active site, which have been schematically illustrated in Fig. 4.

The cytidine moiety of each simulation's respective compound maintained the hydrogen bonding interactions that were observed in the human ST6Gal I crystal structure[25]. An additional interaction observed in all simulations was Lys376 acting as a hydrogen bond donor to the pyridine like nitrogen of the cytosine ring. There was also Ser189 acting as a hydrogen bond donor with the oxygen embedded within the furan ring in the simulations of the *(S)*-diastereomers (Fig. 4b). The hydrophobic contacts that were observed were most significant with the cytosine ring, which was shown to interact with Ala190 and have a proposed π-stacking interaction with Tyr354. The ribose ring was generally shown to have hydrophobic interactions with Phe211 and Leu372. It should be noted that in the simulations of the *(R)*-diastereomers only (Fig. 4a), a water bridge interaction between the free amine of the cytosine ring and the backbone carbonyl of Val352 was also observed.

In addition to the CMP moiety, each compound had two other conserved regions, the phosphonate and the Neu5Ac mimic. For the simulation involving the *(R)*-diastereomers the phosphonate was shown to hydrogen bond with His370 in all simulations and, Ala368 and Tyr369 in all simulations except TC1 (see Supplementary Table S8). The phosphonate in the simulations of the *(S)*-diastereomers occupied a different orientation within the ST6Gal I active site and was shown to hydrogen bond with Tyr354 in all simulations. This difference between the diastereomers is to be expected as the phosphonate is a substituent at the only varied chiral centre. This could also be an explanation to the higher level of potency observed for *(R)*-**2** compared to *(S)*-**2** when tested against rat α-2,6 ST[29], which has been shown to have a high level of homology to the human equivalent[23].
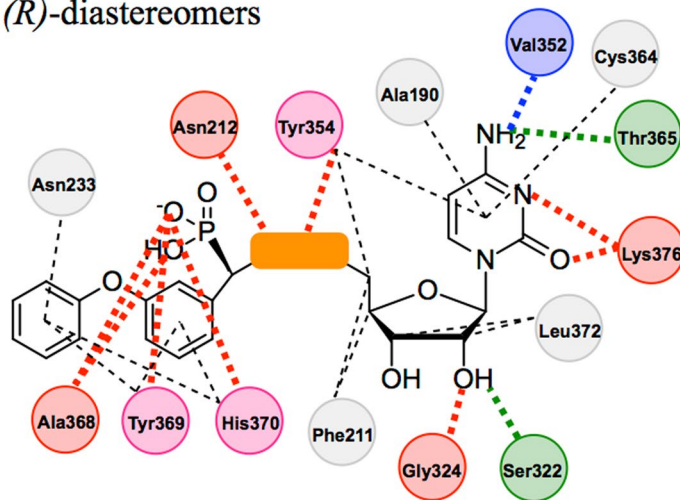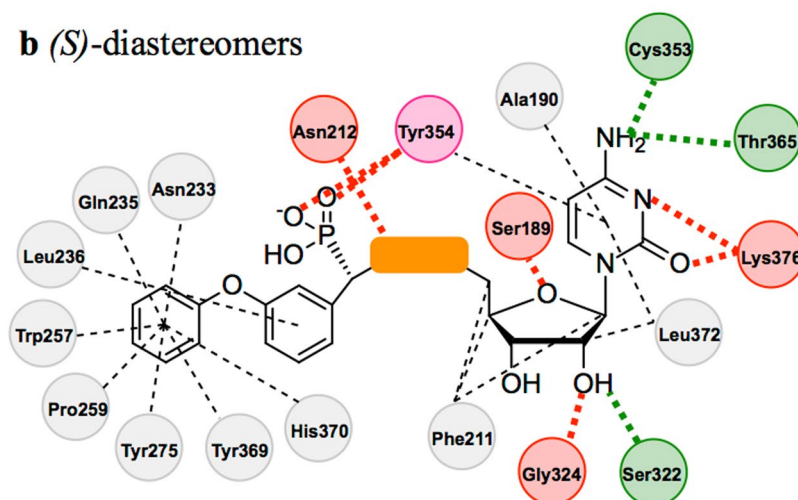
Figure 4. Schematic representation of conserved hydrogen bonds and hydrophobic contacts observed in the *(R)*-diastereomer- (**a**) and *(S)*-diastereomer-ST6Gal I complex (**b**) simulations. The consistent components of each inhibitor are shown as chemical structures. The orange box represents the different linkers. Interacting human ST6Gal I residues are represented as circles which are colour coded based on interaction type as follows: Red = hydrogen bond donor; Green = hydrogen bond acceptor; Blue = water bridged hydrogen bond; Grey = hydrophobic contact; Pink = hydrogen bond and hydrophobic contact. Dashed lines indicate hydrophobic contacts and dotted lines represent the hydrogen bonds shown in Supplementary Tables S4–S15.

The Neu5Ac mimic consists of two components, the aryl ring and its *m*-phenoxy substituent. The *m*-phenoxy substituent for the majority of the simulations was shown to be in, or oriented towards, the proposed sialyl acceptor site through observed hydrophobic contacts with Asn233, Gln235, Trp257, Pro259, Tyr275 and Tyr369. This was most evident for the *(S)*-diastereomers. More variation in this group's position was observed for the *(R)*-diastereomers, with the phenoxy generally orientated to interact only with the residues on the periphery of the pocket, being Asn233, Tyr369 and His370. Notably of the three different classes of linker observed, it was the 1,2,3-triazole that showed both the *(R)*- and *(S)*-diastereomers with the same *m*-phenoxy interactions (see Supplementary Tables S9 and S11), suggesting the added rigidity of the linker favours direction towards the proposed sialyl acceptor site. This is an important finding, as this site is expected to vary between the different ST subtypes, particularly those with different sialyl acceptors, and thus presents a potential route to the development of selective human ST6Gal I inhibitors.

The aryl ring of the Neu5Ac mimic was shown to have limited interactions with human ST6Gal I, with the *(R)*-diastereomers showing hydrophobic contacts with Tyr369 and His370 and the *(S)*-diastereomers with Leu236. Despite being limited these interactions are nevertheless important in directing the *m*-phenoxy towards the proposed sialyl acceptor site. This is clearly highlighted in the CL2 simulations (see Supplementary Table S2) where this aryl ring is shown to interact with Cys353, Tyr354 and Cys364 in a binding mode that does not see the *m*-phenoxy within the proposed sialyl acceptor site.

Alchemical Transformation	$\Delta\Delta G_b$		$\Delta\Delta H_b{}^c$	$-T\Delta\Delta S_b{}^c$
	288 K[a]	298 K[b]	—	—
$2 \to 4$	-1.1 ± 0.8	-1.7 ± 0.4	10 ± 30	-20 ± 30
$2 \to 3$	-0.3 ± 0.3	-1.4 ± 0.9	40 ± 30	-40 ± 30
$4 \to 3$	-0.7 ± 1.0	-0.5 ± 0.3	-10 ± 30	10 ± 30

Table 1. Difference in binding free energy to human ST6Gal I of proposed ST inhibitors. Running average of $\Delta\Delta G_b$ (kcal/mol) calculated at [a]288.15 K and [b]298.15 K, error reported is the SEM. [c]The $\Delta\Delta H_b$ (kcal/mol) and $-T\Delta\Delta S_b$ (kcal/mol) are calculated through finite-temperature difference from the $\Delta\Delta G_b$ values for both the simulations carried out at 288.15 K and 298.15 K. The values reported are those for 298.15 K, error reported is the SEM.

The only component of the potential ST inhibitors examined that varied was the group that links the CMP moiety to the Neu5Ac mimic. Consistent interactions were observed when comparing the phosphodiester linker to the 1,2,3-triazole- and carbamate linkers. In the simulations of the *(R)*-diastereomers, hydrogen bonds of the linkers with Asn212 and Tyr354 were observed. In the simulations of the *(S)*-diastereomers hydrogen bonds of the linkers with Asn212 were observed.

Analysis of the hydrogen bond and hydrophobic contacts of both the *(R)*- and *(S)*-diastereomers of the three different classes of compounds, has demonstrated that no matter which linker is used, key interactions of cytidine and the Neu5Ac mimic with the human ST6Gal I active site are able to be maintained. These results confirm our previous findings[40] and provide further evidence that a carbamate or a 1,2,3-triazole may be a suitable isosteric replacement for the phosphodiester linkers that have been used in reported ST inhibitors.

Binding free energy calculations confirm 1,2,3-triazole and carbamate as isosteres of the phosphodiester linker.
To build upon the results observed in the MD simulations on the thermodynamic level, FEP calculations were performed to study the relative binding affinities for compounds with three different linkers. These simulations involved the alchemical transformations of the phosphodiester-linked *(R)*-**2** to either the 1,2,3-triazole *(R)*-**3** or carbamate-linked *(R)*-**4** derivatives. To close the thermodynamic cycle in order to monitor the convergence, an additional transformation of the carbamate-linked *(R)*-**4** to the 1,2,3-triazole-linked *(R)*-**3** was also performed. The results of the FEP calculations are summarised in Table 1.

The $\Delta\Delta G_b$ values observed for the 288 K and 298 K calculations form a thermodynamic cycle which closes within a reasonable error of -1.5 kcal/mol and -0.8 kcal/mol respectively. These calculations indicate that there is only a slight difference in the free energy of binding between the carbamate- and 1,2,3-triazole-linked derivatives, with the 1,2,3-triazole favoured. It was also demonstrated that when the phosphodiester linker of the lead compound is perturbed to either a carbamate or a 1,2,3-triazole, a slightly more favourable binding to ST6Gal I is observed (-1.7 kcal/mol and -1.4 kcal/mol respectively at 298.15 K), with the carbamate being slightly more preferential.

This result is surprising as the charge of the phosphodiester linker has been previously thought to be important for binding, but to what extent is currently unknown[29,32,34,35,38]. It is expected that replacing the charged phosphodiester group with a neutral isostere could potentially result in the loss of favourable interactions with the binding site that are mediated by this charge. For example, in the MD simulations the phosphodiester linker was involved in three additional water mediated interactions that were not observed in those of the carbamate- and 1,2,3-triazole-linked derivatives (see Supplementary Tables S4, S8, S12). Therefore, one would expect an impact on the binding free energy (ΔG_b), particularly through enthalpy (ΔH_b), if the charged phosphodiester linker were to be replaced. Based on the relatively similar binding free energies for compounds with three different linkers, we hypothesise that the more rigid isosteric linkers might achieve comparable binding affinities through a favourable entropy compensation.

The error analyses for the FEP simulations were carried out by calculated the standard error of the mean (SEM) based on each ns simulation segment (Table 1) and all simulations converged with an error of <1.0 kcal/mol. RMSD of both protein backbone and complexed ligand were also monitored. RMSDs for Cα of ST6Gal I converged between 1.5 Å and 2.0 Å indicating that no significant change in the structure of the protein backbone occurred as one linker was perturbed to another. Across all simulations each ligand remained tightly associated with the ST6Gal I active site, with RMSD values that fluctuated between 1.0 Å and 3.0 Å observed. As with the MD simulations discussed above, this fluctuation was influenced by the *m*-phenoxy substituent, with it remaining fixed in the sialyl acceptor site for the lower RMSDs observed and with it orientated towards the sialyl acceptor site but not constantly located within it when larger RMSDs were observed.

Relative rigidity of the 1,2,3-triazole and carbamate linkers compensate for the enthalpic penalty that results from replacing the phosphodiester.
To test our enthalpy-entropy compensation hypothesis proposed based on the FEP calculations, we calculated $\Delta\Delta H_b$ and $\Delta\Delta S_b$ from the $\Delta\Delta G_b$ values observed at two different temperatures (Table 1)[65]. From this analysis, we were able to demonstrate that the replacement of the charged phosphodiester linker with either a neutral carbamate or 1,2,3-triazole indeed results in an enthalpic penalty (a positive $\Delta\Delta H_b$). In both cases this penalty was compensated for by a favourable entropy (a negative $\Delta\Delta S_b$). This supports our hypothesis of the resulting molecular rigidity created by the inclusion of a carbamate or 1,2,3-triazole in place of a phosphodiester compensates for the enthalpic penalty. For the

Cpd	ΔS_{conf} (kcal/mol/K)[a]	$-T\Delta S_{conf}$ (kcal/mol)[b]
(R)-2	-0.125 ± 0.009	37 ± 3
(S)-2	-0.14 ± 0.02	41 ± 6
(R)-3	-0.08 ± 0.01	23 ± 4
(S)-3	-0.09 ± 0.01	27 ± 4
(R)-4	-0.067 ± 0.007	20 ± 2
(S)-4	-0.09 ± 0.02	25 ± 6

Table 2. Configurational entropy change upon binding of phosphodiester-linked (2), 1,2,3-triazole-linked (3) and carbamate-linked (4) derivatives to human ST6Gal I. [a]Arithmetic mean $\Delta S_{conf} \pm$ standard error of the mean (SEM) obtained from triplicate simulations of each compound. [b]Arithmetic mean ΔS_{conf} multiplied by $-298.15\,K \pm$ SEM obtained from triplicate simulations of each compound.

perturbation from the carbamate to the 1,2,3-triazole, the opposite relationship was observed where changing to a triazole was enthalpically favourable which compensated for an entropic penalty.

The effect of the entropic component of binding free energy to human ST6Gal I has also been demonstrated through quasi-harmonic analysis of the trajectories obtained from the MD simulations performed (Table 2). This analysis has shown that for both the *(R)*- and *(S)*-diastereomers, the phosphodiester-linked lead compound **2** has a larger configurational entropy penalty upon binding to human ST6Gal I when compared to the 1,2,3-triazole-**3** and carbamate-linked **4** derivatives. Though the quasi-harmonic analysis only captures the configurational entropy of the ligands upon binding, while neglecting contributions from other factors, the relatively large difference observed supports the entropic compensation that was captured in the component analysis of the $\Delta\Delta G_b$ values (Table 1). It can also be seen that there is comparatively a very small difference (that is within error) between the $-T\Delta S_{conf}$ values for the carbamate and 1,2,3-triazole (Table 2). However, the 1,2,3-triazole does exhibit a slightly larger value for both the *(R)*- and *(S)*-diastereomers supporting the entropic penalty observed when a carbamate is perturbed to a 1,2,3-triazole (Table 1).

Comparing the diastereomers it can be seen while there are no major differences between the diastereomers, the *(S)*-diastereomer in all three cases has a slightly larger entropic penalty for binding to ST6Gal I. This result reflects the slight differences in the binding modes of the diastereomers observed during the analysis of the MD simulations.

Overall both the free energy component analysis and the quasi-harmonic analysis have demonstrated that the conformational restriction, and thus entropic favourability for binding, imparted by the 1,2,3-triazole and carbamate linkers when compared to the phosphodiester linker is enough to compensate the resulting enthalpic penalty. This result has provided a strong rationale for the use of these two groups as isosteres of the phosphodiester in the development of potential inhibitors of human ST6Gal I. Interestingly, while we have demonstrated the benefit of introducing conformational restriction to binding, there appears to be a limit to what can be tolerated. This is reflected in the carbamate being shown to be slightly more entropically favourable than the 1,2,3-triazole (Tables 1 and 2).

Conformational cluster analysis demonstrates different linker flexibilities. Additional clustering analysis was performed for the ligand conformations in the human ST6Gal I-ligand complex and in the aqueous solution. The most dominant cluster for each complex simulation (see Supplementary Fig. S9) showed similar positions for the cytidine and phenoxy substituents for both the *(R)*- and *(S)*-diastereomers. This supports the findings of the analysis of the protein-ligand interactions (Fig. 4). The conformations of these structures also provide an indication of the differing levels of flexibility of each linker, with the conformation of the 1,2,3-triazole derivatives being much more rigid and elongated when compared to the phosphodiester- and carbamate-linked counterparts. However, this conformational difference is much more obvious in the distribution of cluster population obtained from the ligand in water box simulations (see Supplementary Fig. S10). The distribution for the *(R)*-diastereomers shows the phosphodiester-linked ligand samples a much larger conformational space than both its carbamate- and 1,2,3-triazole-linked counterparts, as evidenced by the number of clusters identified. The distribution for the *(S)*-diastereomers shows a similar trend, however the difference between the phosphodiester and carbamate derivatives is not as large. The increased flexibility of the phosphodiester derivatives compared to the 1,2,3-triazole- and carbamate-linked derivatives provides further evidence to support the enthalpy-entropy compensation hypothesis.

It is worth noting that though the binding free energy differences are converged within 1.0 kcal/mol, a proper error propagation significantly increases the uncertainties in the enthalpic component ($\Delta\Delta H_b$) and entropic component ($-T\Delta\Delta S_b$) (Table 1). Nevertheless, our calculations have demonstrated that these three types of inhibitors with different linkers have similar potencies against human ST6Gal I. Complementary analyses based on binding entropy calculations, quasi-harmonic and conformational cluster analyses highlighted the potential role of entropy in the binding of transition state-based ST6Gal I inhibitors. Experimentally, we are synthesising both the 1,2,3-triazole- and carbamate-linked derivatives and their predicted potency will be validated with experimental inhibition rates and kinetic constants. The enthalpy-entropy compensation hypothesis can be then validated by examining the temperature-dependent data.

Conclusion

MD simulations and FEP calculations were performed to investigate the replacement of the charged phosphodiester linker used in reported TS analogue ST inhibitors with a neutral isostere, such as a carbamate or a 1,2,3-triazole. MD simulations of the *(R)-* and *(S)*-diastereomers of three ligands, which only differed in the linker, complexed with human ST6Gal I or in a water box were performed. Hydrogen bonds and hydrophobic contacts of the ligands with human ST6Gal I were monitored over the course of the simulations and have demonstrated that no matter which linker is used, important interactions with the active site are able to be maintained. These results confirmed the findings of our previous work[40], which was undertaken with a different set of force fields. Free energy perturbation calculations demonstrated that when the phosphodiester linker was perturbed to either a carbamate or a 1,2,3-triazole a slightly more favourable binding to human ST6Gal I was observed, with the carbamate being marginally more preferential. This result was a surprise considering the perceived importance of the phosphodiester linker. We rationalise this observation with a hypothesis of an enthalpy-entropy compensation, which is supported with free energy component analysis, quasi-harmonic and cluster analysis. These analyses demonstrated that the conformational restriction, and thus entropic favourability for binding, imparted by the 1,2,3-triazole and carbamate linkers when compared to the phosphodiester linker is enough to compensate the resulting enthalpic penalty. The results of these simulations have provided a strong rationale for the use of a carbamate or 1,2,3-triazole as an isostere of the phosphodiester. We are currently exploring both potential isosteres synthetically with the aim to develop novel inhibitors of human ST6Gal I, improve synthetic accessibility and address potential pharmacokinetic problems.

References

1. Angata, T. & Varki, A. Chemical diversity in the sialic acids and related α-keto acids: An evolutionary perspective. *Chem. Rev.* **102**, 439–469, https://doi.org/10.1021/cr000407m (2002).
2. Li, Y. & Chen, X. Sialic acid metabolism and sialyltransferases: Natural functions and applications. *Appl. Microbiol. Biotechnol.* **94**, 887–905, https://doi.org/10.1007/s00253-012-4040-1 (2012).
3. Harduin-Lepers, A. *et al*. The human sialyltransferase family. *Biochimie* **83**, 727–737, https://doi.org/10.1016/S0300-9084(01)01301-3 (2001).
4. Varki, A. Sialic acids as ligands in recognition phenomena. *FASEB J.* **11**, 248–255 (1997).
5. Bull, C., Stoel, M. A., den Brok, M. H. & Adema, G. J. Sialic acids sweeten a tumor's life. *Cancer Res.* **74**, 3199–3204, https://doi.org/10.1158/0008-5472.can-14-0728 (2014).
6. Pinho, S. S. & Reis, C. A. Glycosylation in cancer: mechanisms and clinical implications. *Nat Rev Cancer* **15**, 540–555, https://doi.org/10.1038/nrc3982 (2015).
7. Huang, S., Day, T. W., Choi, M.-R. & Safa, A. R. Human β-galactoside α-2,3-sialyltransferase (ST3Gal III) attenuated Taxol-induced apoptosis in ovarian cancer cells by downregulating caspase-8 activity. *Mol. Cell. Biochem.* **331**, 81–88, https://doi.org/10.1007/s11010-009-0147-9 (2009).
8. Szabo, R. & Skropeta, D. Advancement of Sialyltransferase Inhibitors: Therapeutic Challenges and Opportunities. *Med. Res. Rev.* **37**, 219–270, https://doi.org/10.1002/med.21407 (2017).
9. Chen, X. & Varki, A. Advances in the biology and chemistry of sialic acids. *ACS Chem. Biol.* **5**, 163–176, https://doi.org/10.1021/cb900266r (2010).
10. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–D495, https://doi.org/10.1093/nar/gkt1178 (2014).
11. Audry, M. *et al*. Current trends in the structure-activity relationships of sialyltransferases. *Glycobiology* **21**, 716–726, https://doi.org/10.1093/glycob/cwq189 (2011).
12. Schultz, M. J., Swindall, A. F. & Bellis, S. L. Regulation of the metastatic cell phenotype by sialylated glycans. *Cancer Metastasis Rev.* **31**, 501–518, https://doi.org/10.1007/s10555-012-9359-7 (2012).
13. Schultz, M. J. *et al*. ST6Gal-I sialyltransferase confers cisplatin resistance in ovarian tumor cells. *J. Ovarian Res*. **6**, https://doi.org/10.1186/1757-2215-6-25 (2013).
14. Cerami, E. *et al*. The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404, https://doi.org/10.1158/2159-8290.CD-12-0095 (2012).
15. Swindall, A. F. *et al*. ST6Gal-I protein expression is upregulated in human epithelial tumors and correlates with stem cell markers in normal tissues and colon cancer cell lines. *Cancer Res*. **73**, 2368–2378, https://doi.org/10.1158/0008-5472.CAN-12-3424 (2013).
16. Schultz, M. J. *et al*. The Tumor-Associated Glycosyltransferase ST6Gal-I Regulates Stem Cell Transcription Factors and Confers a Cancer Stem Cell Phenotype. *Cancer Res*. **76**, 3978–3988, https://doi.org/10.1158/0008-5472.can-15-2834 (2016).
17. Seales, E. C. *et al*. Hypersialylation of β 1 integrins, observed in colon adenocarcinoma, may contribute to cancer progression by up-regulating cell motility. *Cancer Res*. **65**, 4645–4652, https://doi.org/10.1158/0008-5472.CAN-04-3117 (2005).
18. Shaikh, F. M. *et al*. Tumor cell migration and invasion are regulated by expression of variant integrin glycoforms. *Exp. Cell Res*. **314**, 2941–2950, https://doi.org/10.1016/j.yexcr.2008.07.021 (2008).
19. Swindall, A. F. & Bellis, S. L. Sialylation of the Fas death receptor by ST6Gal-I provides protection against Fas-mediated apoptosis in colon carcinoma cells. *J. Biol. Chem.* **286**, 22982–22990, https://doi.org/10.1074/jbc.M110.211375 (2011).
20. Liu, Z. *et al*. ST6Gal-I regulates macrophage apoptosis via alpha2-6 sialylation of the TNFR1 death receptor. *J. Biol. Chem.* **286**, 39654–39662, https://doi.org/10.1074/jbc.M111.276063 (2011).
21. Britain, C. M., Dorsett, K. A. & Bellis, S. L. The Glycosyltransferase ST6Gal-I Protects Tumor Cells against Serum Growth Factor Withdrawal by Enhancing Survival Signaling and Proliferative Potential. *J. Biol. Chem.* **292**, 4663–4673, https://doi.org/10.1074/jbc.M116.763862 (2017).
22. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyl transferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521–555, https://doi.org/10.1146/annurev.biochem.76.061005.092322 (2008).
23. Meng, L. *et al*. Enzymatic basis for N-glycan sialylation: Structure of rat α2,6-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).
24. Breton, C., Fournel-Gigleux, S. & Palcic, M. M. Recent structures, evolution and mechanisms of glycosyltransferases. *Curr. Opin. Struct. Biol.* **22**, 540–549, https://doi.org/10.1016/j.sbi.2012.06.007 (2012).
25. Kuhn, B. *et al*. The structure of human α-2,6-sialyltransferase reveals the binding mode of complex glycans. *Acta Crystallogr. D* **69**, 1826–1828, https://doi.org/10.1107/S0907444913015412 (2013).
26. Jung, K. H., Schwörer, R. & Schmidt, R. R. Sialyltransferase Inhibitors. *Trends Glycosci. Glyc.* **15**, 275–289, https://doi.org/10.4052/tigg.15.275 (2003).
27. Wang, X., Zhang, L. H. & Ye, X. S. Recent development in the design of sialyltransferase inhibitors. *Med. Res. Rev.* **23**, 32–47, https://doi.org/10.1002/med.10030 (2003).

28. Drinnan, N. B., Halliday, J. & Ramsdale, T. Inhibitors of sialyltransferases: potential roles in tumor growth and metastasis. *Mini Rev. Med. Chem.* **3**, 501–517, https://doi.org/10.2174/1389557033487881 (2003).

29. Skropeta, D., Schwörer, R., Haag, T. & Schmidt, R. R. Asymmetric synthesis and affinity of potent sialyltransferase inhibitors based on transition-state analogues. *Glycoconjugate J.* **21**, 205–219, https://doi.org/10.1023/B:GLYC.0000045093.96413.62 (2004).

30. Schröder, P. N. & Giannis, A. From substrate to transition state analogues: The first potent inhibitor of sialyltransferases. *Angew. Chem. Int. Ed.* **38**, 1379–1380, https://doi.org/10.1002/(SICI)1521-3773(19990517)38:10<1379 (1999).

31. Amann, F., Schaub, C., Müller, B. & Schmidt, R. R. New potent sialyltransferase inhibitors - Synthesis of donor and of transition-state analogues of sialyl donor CMP-Neu5Ac. *Chem. Eur. J.* **4**, 1106–1115, https://doi.org/10.1002/(SICI)1521-3765(19980615)4:6<1106::AID-CHEM1106>3.0.CO;2-7 (1998).

32. Schwörer, R. & Schmidt, R. R. Efficient sialyltransferase inhibitors based on glycosides of N-acetylglucosamine. *J. Am. Chem. Soc.* **124**, 1632–1637, https://doi.org/10.1021/ja017370n (2002).

33. Müller, B., Schaub, C. & Schmidt, R. R. Efficient sialyltransferase inhibitors based on transition-state analogues of the sialyl donor. *Angew. Chem. Int. Ed.* **37**, 2893–2897, https://doi.org/10.1002/(SICI)1521-3773(19981102)37:20<2893::AID-ANIE2893>3.0.CO;2-W (1998).

34. Li, W., Niu, Y., Xiong, D.-C., Cao, X. & Ye, X.-S. Highly Substituted Cyclopentane–CMP Conjugates as Potent Sialyltransferase Inhibitors. *J. Med. Chem.* **58**, 7972–7990, https://doi.org/10.1021/acs.jmedchem.5b01181 (2015).

35. Guo, J., Li, W., Xue, W. & Ye, X.-S. Transition State-Based Sialyltransferase Inhibitors: Mimicking Oxocarbenium Ion by Simple Amide. *J. Med. Chem.* **60**, 2135–2141, https://doi.org/10.1021/acs.jmedchem.6b01644 (2017).

36. Volkers, G. *et al.* Structure of human ST8SiaIII sialyltransferase provides insight into cell-surface polysialylation. *Nat. Struct. Mol. Biol.* **22**, 627–635, https://doi.org/10.1038/nsmb.3060 (2015).

37. Skropeta, D., Schwörer, R. & Schmidt, R. R. Stereoselective synthesis of phosphoramidate α(2-6)sialyltransferase transition-state analogue inhibitors. *Bioorg. Med. Chem. Lett.* **13**, 3351–3354, https://doi.org/10.1016/S0960-894X(03)00672-3 (2003).

38. Kumar, R. *et al.* Sialyltransferase inhibitors: Consideration of molecular shape and charge/hydrophobic interactions. *Carbohydr. Res.* **378**, 45–55, https://doi.org/10.1016/j.carres.2012.12.017 (2013).

39. Rye, C. S. & Baell, J. B. Phosphate isosteres in medicinal chemistry. *Curr. Med. Chem.* **12**, 3127–3141, https://doi.org/10.2174/092986705774933452 (2005).

40. Montgomery, A., Szabo, R., Skropeta, D. & Yu, H. Computational characterisation of the interactions between human ST6Gal I and transition-state analogue inhibitors: insights for inhibitor design. *J. Mol. Recogn* **29**, 210–222, https://doi.org/10.1002/jmr.2520 (2016).

41. Montgomery, A. P., Xiao, K., Wang, X., Skropeta, D. & Yu, H. in *Advances in Protein Chemistry and Structural Biology* Vol. 109 (ed Karabencheva-Christova Tatyana) 25–76 (Academic Press, 2017).

42. Van Gunsteren, W. F. *et al.* Biomolecular modeling: Goals, problems, perspectives. *Angew. Chem., Int. Ed.* **45**, 4064–4092, https://doi.org/10.1002/anie.200502655 (2006).

43. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690, https://doi.org/10.1002/jcc.21367 (2010).

44. Huang, L. & Roux, B. Automated force field parameterization for nonpolarizable and polarizable atomic models based on ab initio target data. *J. Chem. Theory Comput.* **9**, 3543–3556, https://doi.org/10.1021/ct4003477 (2013).

45. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph. Model.* **14**, 33–38, https://doi.org/10.1016/0263-7855(96)00018-5 (1996).

46. Søndergaard, C. R., Olsson, M. H. M., Rostkowski, M. & Jensen, J. H. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. *J. Chem. Theory Comput.* **7**, 2284–2295, https://doi.org/10.1021/ct200133y (2011).

47. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys* **79**, 926–935, https://doi.org/10.1063/1.445869 (1983).

48. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802, https://doi.org/10.1002/jcc.20289 (2005).

49. MacKerell, A. D. *et al.* All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **102**, 3586–3616, https://doi.org/10.1021/jp973084f (1998).

50. Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys* **101**, 4177–4189, https://doi.org/10.1063/1.467468 (1994).

51. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys* **103**, 4613–4621, https://doi.org/10.1063/1.470648 (1995).

52. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys* **98**, 10089–10092, https://doi.org/10.1063/1.464397 (1993).

53. Andersen, H. C. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys* **52**, 24–34, https://doi.org/10.1016/0021-9991(83)90014-1 (1983).

54. Brooks, B. R. *et al.* CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614, https://doi.org/10.1002/jcc.21287 (2009).

55. Karplus, M. Method for estimating the configurational entropy of macromolecules. *Macromolecules* **14**, 325–332, https://doi.org/10.1021/ma50003a019 (1981).

56. Andricioaei, I. & Karplus, M. On the calculation of entropy from covariance matrices of the atomic fluctuations. *J. Chem. Phys* **115**, 6289–6292, https://doi.org/10.1063/1.1401821 (2001).

57. Schäfer, H. & Mark, A. E. & Gunsteren, W. F. v. Absolute entropies from molecular dynamics simulation trajectories. *J. Chem. Phys* **113**, 7809–7817, https://doi.org/10.1063/1.1309534 (2000).

58. Feig, M., Karanicolas, J. & Brooks, C. L. MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. *J. Mol. Graphics Modell.* **22**, 377–395, https://doi.org/10.1016/j.jmgm.2003.12.005 (2004).

59. Rocklin, G. J., Mobley, D. L., Dill, K. A. & Hünenberger, P. H. Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: An accurate correction scheme for electrostatic finite-size effects. *J. Chem. Phys* **139**, 184103, https://doi.org/10.1063/1.4826261 (2013).

60. Reif, M. M. & Oostenbrink, C. Net charge changes in the calculation of relative ligand-binding free energies via classical atomistic molecular dynamics simulation. *J. Comput. Chem.* **35**, 227–243, https://doi.org/10.1002/jcc.23490 (2014).

61. Heinzelmann, G., Baştuğ, T. & Kuyucak, S. Free Energy Simulations of Ligand Binding to the Aspartate Transporter GltPh. *Biophys. J.* **101**, 2380–2388, https://doi.org/10.1016/j.bpj.2011.10.010 (2011).

62. Zacharias, M., Straatsma, T. P. & McCammon, J. A. Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J. Chem. Phys* **100**, 9025–9031, https://doi.org/10.1063/1.466707 (1994).

63. Beutler, T. C., Mark, A. E., van Schaik, R. C., Gerber, P. R. & van Gunsteren, W. F. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett.* **222**, 529–539, https://doi.org/10.1016/0009-2614(94)00397-1 (1994).

64. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174, https://doi.org/10.1002/jcc.20035 (2004).

65. Peter, C., Oostenbrink, C., van Dorp, A. & van Gunsteren, W. F. Estimating entropies from molecular dynamics simulations. *J. Chem. Phys* **120**, 2652–2661, https://doi.org/10.1063/1.1636153 (2004).

Author Contributions

A.P.M., H.Y. & D.S. designed research. A.P.M. & H.Y. conducted experiments and analysed the data. A.P.M. wrote the main manuscript and prepared figures and tables. All authors have reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-14560-0.

Competing Interests: The authors declare that they have no competing interests.

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