2016

Computational characterisation of the interactions between human ST6Gal I and transition-state analogue inhibitors: insights for inhibitor design

Andrew Montgomery
University of Wollongong, apm977@uowmail.edu.au

Remi Szabo
University of Wollongong, rs424@uowmail.edu.au

Danielle Skropeta
University of Wollongong, skropeta@uow.edu.au

Haibo Yu
University of Wollongong, hyu@uow.edu.au

Publication Details

Montgomery, A., Szabo, R., Skropeta, D. & Yu, H. (2016). Computational characterisation of the interactions between human ST6Gal I and transition-state analogue inhibitors: insights for inhibitor design. Journal of Molecular Recognition, 29 (5), 210-222.

Computational characterisation of the interactions between human ST6Gal I and transition-state analogue inhibitors: insights for inhibitor design

Abstract

Human β-galactoside α-2,6-sialyltransferase I (hST6Gal I) catalyses the synthesis of sialylated glycoconjugates involved in cell-cell interactions. Overexpression of hST6Gal I is observed in many different types of cancers, where it promotes metastasis through altered cell surface sialylation. A wide range of sialyltransferase (ST) inhibitors have been developed based on the natural donor, cytidine 5′-monophosphate N-acetylneuraminic acid (CMP-Neu5Ac). Of these, analogues that are structurally similar to the transition state exhibit the highest inhibitory activity. In order to design inhibitors that are readily accessible synthetically and with favourable pharmacokinetic properties, an investigation of the replacement of the charged phosphodiester-linker, present in many ST inhibitors, with a potential neutral isostere such as a carbamate or a 1,2,3-triazole has been undertaken. To investigate this, molecular docking and molecular dynamics simulations were performed. These simulations provided an insight into the binding mode of previously reported phosphodiester-linked ST inhibitors and demonstrated that targeting the proposed sialyl acceptor site is a viable option for producing selective inhibitors. The potential for a carbamate- or triazole-linker as an isosteric replacement for the phosphodiester in transition-state analogue ST inhibitors was established using molecular docking. Molecular dynamics simulations of carbamate- and phosphodiester-linked compounds revealed that both classes exhibit consistent interactions with hST6Gal I. Overall, the results obtained from this study provide a rationale for synthetic and biological evaluation of triazole- and carbamate-linked transition-state analogue ST inhibitors as potential new antimetastatic agents.

Disciplines

Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Montgomery, A., Szabo, R., Skropeta, D. & Yu, H. (2016). Computational characterisation of the interactions between human ST6Gal I and transition-state analogue inhibitors: insights for inhibitor design. Journal of Molecular Recognition, 29 (5), 210-222.

Computational characterisation of the interactions between human ST6Gal I and transition-state analogue inhibitors: Insights for inhibitor design

Short title: Molecular modelling of human ST6Gal I-inhibitor complexes

Authors: Andrew Montgomery,[a] Rémi Szabo,[a] Danielle Skropeta[a,b] and Haibo Yu[a,b]

Affiliation: [a] School of Chemistry, Faculty of Science, Medicine and Health, University of Wollongong, Wollongong NSW 2522, Australia

[b] Centre for Medical and Molecular Bioscience, University of Wollongong, Wollongong NSW 2522, Australia

Author to Check Proofs:

Haibo Yu, School of Chemistry, University of Wollongong, NSW 2522, Australia.

Email: hyu@uow.edu.au

Phone: +61 2 4221 4235

Fax: +61 2 4221 4287

Abstract:

Human β-galactoside α-2,6-sialyltransferase I (hST6Gal I) catalyses the synthesis of sialylated glycoconjugates involved in cell-cell interactions. Overexpression of hST6Gal I is observed in many different types of cancers, where it promotes metastasis through altered cell surface sialylation. A wide range of sialyltransferase (ST) inhibitors have been developed based on the natural donor, cytidine 5'-monophosphate *N*-acetylneuraminic acid (CMP-Neu5Ac). Of these, analogues that are structurally similar to the transition state exhibit the

highest inhibitory activity. In order to design inhibitors that are readily accessible synthetically and with favourable pharmacokinetic properties, an investigation of the replacement of the charged phosphodiester-linker, present in many ST inhibitors, with a potential neutral isostere such as a carbamate or a 1,2,3-triazole has been undertaken. To investigate this, molecular docking and molecular dynamics simulations were performed. These simulations provided an insight into the binding mode of previously reported phosphodiester-linked ST inhibitors and demonstrated that targeting the proposed sialyl acceptor site is a viable option for producing selective inhibitors. The potential for a carbamate- or triazole-linker as an isosteric replacement for the phosphodiester in transition-state analogue ST inhibitors was established using molecular docking. Molecular dynamics simulations of carbamate- and phosphodiester-linked compounds revealed that both classes exhibit consistent interactions with hST6Gal I. Overall, the results obtained from this study provide a rationale for synthetic and biological evaluation of triazole- and carbamate-linked transition-state analogue ST inhibitors as potential new antimetastatic agents.

Keywords: sialyltransferase; inhibitors; molecular docking; molecular dynamics simulations; transition-state analogues; metastasis

INTRODUCTION

Sialyltransferases (STs) are enzymes of the glycosyltransferase (GT) family that play an integral role in the biosynthesis of sialic acid (*N*-acetylneuraminic acid, Neu5Ac) containing oligosaccharides and glycoconjugates (Harduin-Lepers et al. 2001; Li and Chen 2012). STs are anchored within the Golgi apparatus' membrane, with a catalytic domain present in the lumen (Figure 1; Li and Chen 2012). Generally in glycosylation reactions catalysed by STs, the sugar nucleotide donor is cytidine 5'-monophosphate Neu5Ac (CMP-Neu5Ac) and the acceptor is an oligosaccharide or glycoconjugate terminated by a galactose (Gal), *N*-acetylgalactosamine (GalNAc), or other sialic acid residue (Chen and Varki 2010; Li and Chen 2012). STs are classified based on the position of the glycosyl acceptor that sialic acid is transferred. In humans these are ST3, ST6 and ST8, which form an α-glycosidic bond between the C2 atom of sialic acid and the 3'-, 6'- or 8'-hydroxyl group of the acceptor, respectively (Harduin-Lepers et al. 2001; Wang 2005; Li and Chen 2012).

The focus of this work, β-galactoside α-2,6-sialyltransferase I (ST6Gal I), is one of the two subfamilies of the ST6 family, that catalyse the transfer of a Neu5Ac residue to the 6'-hydroxyl group of the terminal Gal residues of the disaccharide Galβ[1,4]GalNAc. This disaccharide can be found free or as a terminal *N*-acetyllactosamine unit of *N*- or *O*- linked oligosaccharides (Harduin-Lepers et al. 2001; Audry et al. 2011). Human ST6Gal I (hST6Gal I; Figure 1) belongs to the GT29 family of GTs, which is comprised of eukaryotic and viral STs (http://www.cazy.org; Lombard et al. 2014).

The first sequence data of ST6Gal I was determined from a cDNA clone from rats in 1987 (Weinstein et al. 1987), with the sequence of hST6Gal I described three years later (Grundmann et al. 1990). The primary sequence of hST6Gal I, like all other eukaryotic STs, contains four characteristic regions: a short *N*-terminal cytoplasmic domain, followed by a

transmembrane domain that extends to a stem region and a larger *C*-terminal catalytic domain that faces the lumen of the Golgi apparatus (Figure 1; Kuhn et al. 2013). Within the catalytic domain there are four regions that show significant sequence conservation across eukaryotic STs (Figure 1), known as sialyl motifs L, S, III and VS (Jeanneau et al. 2004). Residues of motifs L (Trp181–Gly224; hST6Gal I numbering) and III (Tyr354–Gln357) are involved in binding the donor substrate (Datta et al. 2001); motif S residues (Pro321-Phe343) are involved in both donor and acceptor binding, while motif VS residues (His370–Glu375) participate in the catalytic reaction (Datta 2009). The first 80 residues of hST6Gal I, comprising the cytoplasmic domain, transmembrane domain and stem region, can be deleted without affecting activity (Legaigneur et al. 2001; Datta 2009).

There were two decades between the identification of the primary sequence and reporting of structural data of mammalian ST6Gal I, with ligand conformation and protein-ligand contact surfaces identified using nuclear magnetic resonance (NMR) appearing in 2007 and 2009 (Liu et al. 2007; Liu et al. 2009). The first mammalian ST crystal structure published was that of porcine ST3Gal I (Rao et al. 2009) in 2009, while the first mammalian ST6Gal I crystal structures that came from rat (Meng et al. 2013) and human (Protein Data Bank (PDB) id: 4JS1, 4JS2; Kuhn et al. 2013) were published in 2013. In addition to these, crystal structures of human ST8Sia III were published in 2015 (Volkers et al. 2015). All of the mammalian structures published are of the soluble catalytic domains with the cytoplasmic domain, transmembrane domain and a portion of the stem region deleted.

Human ST6Gal I adopts the GT-A (variant 2) fold, a seven-stranded central β-sheet flanked by multiple α-helices, and exhibits three disulphide bonds Cys142-Cys406, Cys184-Cys335 and Cys353-Cys364. The hST6Gal I crystal structure revealed that CMP binds with the ribose in a 3'-*endo* conformation created via interactions with the first and last residues of a SerSerGly sequence within the sialyl motif S (Figure S1). The disulphide bond Cys353-

Cys364 is contained within a well-defined loop, packing on top of the nucleoside. The glycan binding observed, showed the 6'-hydroxyl of the acceptor close to His370, a potential catalytic residue (Kuhn et al. 2013). Mutation carried out upon the equivalent histidine (His367) of rat ST6Gal I resulted in no detectable catalytic activity (Meng et al. 2013), which supports the conclusion that His370 is a catalytic residue in hST6Gal I.

It has been proposed that the mechanism of ST6Gal I (Figure 2) proceeds via nucleophilic attack on the Neu5Ac anomeric carbon by the deprotonated Gal 6'-hydroxyl of the disaccharide Galβ[1,4]GalNAc acceptor, facilitated by the catalytic His370 (Breton et al. 2012; Meng et al. 2013). This generates an oxocarbenium-like transition state, with the CMP portion acting as a leaving group. The β-configuration at the anomeric carbon of the donor is inverted in the final product (Breton et al. 2012). These aspects suggest that ST6Gal I functions as an inverting GT, utilising an S_N2-like direct displacement mechanism (Lairson et al. 2008; Audry et al. 2011). This is supported by a model of the Michaelis complex generated from the glycan binding mode observed in the crystal structure of hST6Gal I (Kuhn et al. 2013).

Depending on the substrate targeted by ST6Gal I, α-2,6-sialylation can modulate protein conformation, oligomerisation and/or receptor internalisation (Schultz et al. 2013). Overexpression of ST6Gal I is observed in many different types of cancers including lung, cervical, ovarian, breast and colon carcinoma (Cerami et al. 2012; Schultz et al. 2013). It has been suggested that the overexpression of ST6Gal I promotes cancer cell metastasis through altered sialylation patterns affecting the function of β1 integrin (Seales et al. 2005; Shaikh et al. 2008). ST6Gal I also has an important function in the regulation of galectins, via the sialylation of their galactose-containing substrates (Schultz et al. 2012; Schultz et al. 2013). For instance, α-2,6 sialylation of the galectin substrates, Fas and TNFR1 death receptors, has

been shown to hinder the apoptosis process, thus identifying ST6Gal I overexpression as a potential inhibitor of cell death pathways (Schultz et al. 2012; Schultz et al. 2013).

The overexpression of ST6Gal I in various cancers has lead to the hypothesis that it can serve as a target for the development of new anti-cancer therapeutics. This has resulted in the development of a diverse range of ST inhibitors (Drinnan et al. 2003; Jung et al. 2003; Wang et al. 2003). Of these, analogues that are structurally similar to the proposed oxocarbenium ion-like transition state (Figure 2) exhibit the highest affinity to STs (Schröder and Giannis 1999; Skropeta et al. 2004). Transition-state analogue inhibitors incorporating a 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid (Neu5Ac2en) moiety are amongst the most potent inhibitors, with the C2-C3 double bond mimicking the planar anomeric carbon of the oxocarbenium transition state (Amann et al. 1998). An extra carbon between the anomeric carbon and CMP leaving group oxygen was also incorporated to mimic the increased distance between these two atoms in the transition state (Amann et al. 1998; Skropeta et al. 2004). Derivatives with structural variations of the glycerol side chain of the Neu5Ac2en moiety produced the greatest inhibitory activity. Of these, phenoxy derivative **1** (Figure 3, K_i = 29 nM), showed the greatest inhibition of rat liver α-2,6-ST and is one of the most potent transition-state analogue ST inhibitors reported to date (Schwörer and Schmidt 2002). Derivatives that incorporate an aryl replacement of the Neu5Ac2en moiety have comparable activity to the phenoxy derivative **1** and are more synthetically accessible (Müller et al. 1998; Schröder and Giannis 1999; Skropeta et al. 2004). Of these compounds the (*R*)-isomer of the phenoxy derivative **2** was the most potent inhibitor of rat liver α-2,6-ST (Figure 3, K_i = 70 nM) and is used as the lead compound herein (Skropeta et al. 2004).

The charged phosphodiester linkage of transition-state analogue ST inhibitors, while contributing to potency (Skropeta et al. 2003), is generally thought to introduce pharmacokinetic issues such as: poor cellular permeability, low bioavailability and potential

loss of activity due to cleavage by phosphatases (Rye and Baell 2005; Kumar et al. 2013). In this study, to produce potent ST inhibitors that are readily accessible and with favourable pharmacokinetic properties, the replacement of the charged phosphodiester-linker of the lead compound **2** with a neutral-linker, such as a rigid 1,2,3-triazole or a flexible carbamate has been investigated. The 1,2,3-triazole motif is a suitable choice as it has previously been shown to be an applicable phosphodiester isostere in transition-state analogue ST inhibitors (Lee et al. 2006; Kumar et al. 2013), as well as being used as isosteres of pyrophosphate-linkers (Chen et al. 2010) and phosphodiester-linkers of oligonucleotides (El-Sagheer and Brown 2010). To date, there are a paucity of examples of carbamates being used as isosteres of phosphodiesters. Carbamates do, however, have potential to be an isosteric replacement, as this group has similar hydrogen bonding capabilities to a phosphodiester group and will maintain the three-atom distance between the cytidine moiety and Neu5Ac-mimic observed in reported phosphodiester-linked ST inhibitors. To investigate the potential binding effects of replacing the phosphodiester-linker in derivatives with triazoles (e.g. **3**; Figure 3) and carbamates (e.g. **4**; Figure 3), molecular docking and molecular dynamics simulations with a model of hST6Gal I were undertaken. Insight into this binding may suggest novel molecular structures and design strategies for the future development of clinically-relevant and selective ST inhibitors.

METHODS

Molecular docking setup: A selection of reported and newly designed inhibitors were docked into the binding site of the hST6Gal I crystal structure (PDB id: 4JS2; Kuhn et al. 2013) and snapshots sampled from every 20 ns of MD simulations (see below) with AutoDock Vina version 1.1.2. (Trott and Olson 2010). The structure of hST6Gal I was prepared for docking with AutoDockTools 1.5.6 (ADT; Sanner 1999). The three-dimensional structures of the inhibitors studied were prepared with ChemDraw 14.0 and Avogadro v1.1.1.

(Hanwell et al. 2012), with ADT used to assign both rigid and rotatable bonds and to remove non-polar hydrogen atoms. PROPKA 3.1 (Søndergaard et al. 2011) was used to predict the protonation state of active site residues and ionisable ligand groups at pH 7.0. Docking was performed in a 30 Å × 30 Å × 30 Å box centred at the active site highlighted in the hST6Gal I-CMP crystal structure (PDB id: 4JS2; Kuhn et al. 2013). The dimension was chosen to ensure it was big enough to accommodate the largest ligands and their interacting residues. In all the docking calculations, the receptor is kept rigid and no explicit waters have been included. The docking procedure was validated by re-docking of CMP to replicate the crystallographically determined hST6Gal I-CMP complex (Kuhn et al. 2013).

Molecular dynamics simulation setup: The crystal structure of hST6Gal I (PDB id: 4JS2) was taken as the initial structure for the simulations and was prepared using VMD version 1.9.1. (Humphrey et al. 1996). An overview of all simulations is given in Table 1. Simulation of the hST6Gal I-CMP complex, using the starting coordinates provided by the crystal structure was undertaken to validate the molecular dynamics simulation protocols used. To probe flexibility in the apo state, the hST6Gal I protein in the absence of a ligand was also simulated. Simulation of hST6Gal I-inhibitor complexes with reported phosphodiester-linked compounds **1** and **2** and the proposed carbamate-linked derivative ***(R)*-4** (Figure 3) was undertaken to determine if a carbamate-linker was a suitable replacement for the phosphodiester-linker used in previously reported inhibitors of hST6Gal I. The initial structure for the hST6Gal I-inhibitor complexes was taken from the docking studies detailed above.

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. Each box was neutralised with counter ions of Na^+, with the salt (NaCl) concentration set to 0.15 mol/L. Simulations were carried out using the NAMD 2.9 package (Phillips et al. 2005). The protein was

represented with the non-polarisable CHARMM PARAM36 force field (Brooks et al. 2009). Each ligand if not represented in this force field, had a parameter and topology file generated with the General Amber Force Fields (GAFF; Wang et al. 2004) using AmberTools13 (D.A. Case et al. 2012). All of the 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 (Martyna et al. 1994; Feller et al. 1995) for keeping the pressure constant at 1.0 bar. The electrostatic interactions were calculated using the Particle Mesh Ewald method (Darden et al. 1993). 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 (Andersen 1983). 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 systems were initially undertaken. During these simulations a harmonic restraint was placed upon all Cα atoms of hST6Gal I, with a decreasing force constant from 32.0 to 1.0 kcal/mol/Å^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/Å^2 was obtained (i.e. 32.0, 16.0 … 1.0 kcal/mol/Å^2). A force constant of 1.0 kcal/mol/Å^2 was then applied for the remainder of the 10 ns thermal equilibration. The simulations were further continued for 90 ns without any restraints. To perform trajectory analysis snapshots were saved every picosecond (i.e. every 1000 steps).

Analysis: The docking protocols outlined above were validated by replication of the experimentally determined binding mode of the mono-deprotonated CMP in the reported hST6Gal I crystal structure (Figure S2; Kuhn et al. 2013). To assess the ability of AutoDock Vina to rank potential hST6Gal I inhibitors in terms of their enzyme binding affinities, focused docking of selected reported inhibitors into the active site of hST6Gal I was undertaken.

To determine the optimal binding mode for each ligand investigated the top ranked models, based on the scoring function for binding affinities in AutoDock Vina were examined in VMD (Humphrey et al. 1996). Evaluation of the optimal binding mode was firstly based upon comparison to the position of CMP component of the inhibitors investigated in relation to residues that were highlighted as important for CMP binding in the crystal structure of hST6Gal I (Kuhn et al. 2013). This was followed by comparison of the Neu5Ac mimic position in relation to residues highlighted as interacting with the Neu5Ac component of CMP-3-F-Neu5Ac in a model of the Michaelis complex generated for hST6Gal I (Kuhn et al. 2013). For each compound the arithmetic mean binding affinity of the optimal binding mode was reported (Table 2-3). Two-tailed *t*-tests were used to determine whether differences in the mean binding affinities were significant, with significance defined as $P < 0.05$ (Table S1-S7).

To monitor the global properties of the structural evolution during MD simulations, 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 (Humphrey et al. 1996).

Hydrogen bond interactions in the hST6Gal I-inhibitor complex simulations were analysed according to a geometric criterion. A hydrogen bond was defined by a minimum donor-hydrogen-acceptor angle of 120º 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(Brooks et al. 2009) was utilised.

RESULTS AND DISCUSSION

Structural flexibility of the apo and holo form of hST6Gal I: To assess the structural stability and fluctuations of the recently reported hST6Gal I crystal structure (Kuhn et al. 2013) in both the apo and holo forms, MD simulations were performed. The Apo, CMP2 and CMP3 simulation RMSDs for carbon alpha (Cα) converged at 1.5 Å with CMP1 converging at 2.5 Å (Figure S3), indicating no significant change in the structure of the protein backbone. It should be noted that during the CMP simulations the CMP ligand remained tightly associated with the expected binding site. RMSFs for the Cα atoms, were comparable across the Apo and three CMP simulations (Figure S4). Each simulation showed consistently larger RMSF values for the flexible loop between Tyr355 and Tyr369.

Probing the interactions between reported inhibitors and hST6Gal I: Snapshots from the Apo and CMP1 simulations were used to dock reported inhibitors (**1-2** and **5-11**, Table 2) with similar results obtained with each condition. Results from CMP2 and CMP3 simulations were not included as no significant difference in binding affinity or binding mode was observed between the docking into the snapshots of each CMP simulation.

For both the Apo and CMP1 cases the AutoDock Vina scoring function for the more potent inhibitors (**1-2**, **5-8**) was unable to differentiate between compounds that exhibited similar K_i values (Table 2). For the less potent inhibitors (**9-11**) the AutoDock Vina scoring function was unable to differentiate between any of these compounds despite the large differences in K_i observed experimentally (Table 2). This suggests that obtaining a perfect ranking of docked compounds using this protocol is unlikely, which is consistent with the error reported in the systematic benchmarking studies for AutoDock Vina (Trott and Olson 2010). Encouragingly though, the scoring function was generally able to predict a significant difference between inhibitors of nanomolar, submicromolar, micromolar and submillimolar K_i (Table S1-S2), providing confidence that an indication of relative binding affinity between compounds could be obtained. In addition, experimentally reported trends were also observed

in the docking studies such as: no significant difference in binding affinity between the Neu5Ac2en containing **1** and its more synthetically accessible counterpart **2** being observed (Müller et al. 1998; Schröder and Giannis 1999; Skropeta et al. 2004); and a *m*-phenoxy substituted inhibitor (e.g. **2**) exhibiting greater binding affinity than the analogous unsubstituted inhibitor (e.g. **6**; Skropeta et al. 2004).

For all compounds except **1**, the binding affinity results obtained from docking into the Apo and CMP1 MD sampled structures were not significantly different (Table S3). However, in terms of the positioning of the inhibitors, docking using the MD sampled structures of the CMP1 simulation generally produced binding modes that were more consistent. This was particularly evident in the docking into the 80 ns and 100 ns snapshots, where CMP1 docking conditions produced binding modes where the CMP moiety of each inhibitor reflected the interactions highlighted within the hST6Gal I crystal structure (Kuhn et al. 2013), which was not consistently the case for the Apo docking conditions. This can be rationalised in that during the Apo simulation hST6Gal I relaxed back to the apo form, from the starting "induced-fit" form of the co-crystallised CMP-hST6Gal I structure.

Probing the interactions between 1,2,3-triazole- and carbamate-linked compounds and hST6Gal I: Based on the results discussed above, docking of the triazole- (**3**, **12-22**) and carbamate-linked compounds (**4**, **23-33**) was undertaken using MD snapshots from the CMP1 simulation (Table 3). Of the compounds docked there were six (***(R)*-3**, ***(R)*-4**, **12**, **16**, **23** and **27**) that only differed from the literature inhibitors examined by the replacement of the phosphodiester-linker with either a triazole or carbamate. The mean binding affinities for all analogous compounds, except for the phosphodiester-linked *m*-trifluoromethoxy ***(S)*-5** and the corresponding triazole ***(S)*-12**, were not significantly different (Table S4). For these triazole- and carbamate-linked compounds the position within the active site (Figure 4A) was also comparable to those obtained for their phosphodiester-linked counterparts, indicating

that these linkers may be feasible replacements of the reported phosphodiester-linker. This docking study also shows generally no significant difference between analogous carbamate- and triazole-linked compounds (Table S5). This suggests that in most cases there may be no preference for either a flexible carbamate- or rigid triazole-linker as an isostere for the phosphodiester-linker.

The triazole and carbamate compounds generally demonstrated no significant difference between the *(R)* and *(S)*-configured diastereomers of the same compound (Table S6-S7), consistent with earlier findings reported in the literature (Schwörer and Schmidt 2002; Skropeta et al. 2004). This counterintuitive result observed both experimentally and *in silico* appears to be related with the sufficient space available within the active site to accommodate the different stereochemistry, whilst maintaining proximity to residues that are expected to be key in binding of the inhibitors.

Different substituents upon the aryl portion of the triazole- and carbamate-linked compounds were also examined (Table 3 and S5-S7). The use of methoxy (**13** and **24**) and propoxy (**14** and **25**) *m*-substituents resulted in comparable predicted binding affinity to the corresponding unsubstituted triazole **16** and carbamate **27**. However, the use of a phenoxy (**3** and **4**) *m*-substituent exhibited a significant increase in predicted binding affinity when compared to the analogous unsubstituted triazole **16** and carbamate **27**, as observed experimentally for the corresponding phosphodiester-linked compounds (Skropeta et al. 2004). The other *m*-substituent examined was trifluoromethoxy. For both isomers of the *m*-trifluoromethoxy carbamate **23** there was a significant increase in predicted binding affinity compared to the corresponding unsubstituted carbamate **27**. In contrast the *(S)*-configured diastereomer of *m*-trifluoromethoxy triazole **12** had a significantly larger predicted binding affinity than the corresponding unsubstituted triazole **16**, which was not observed for the *(R)*-configured diastereomer of **12**. Comparing each of the *m*-substituents shows that for the carbamate-

linked inhibitors the *m*-phenoxy diastereomers (**4**) showed a significant increase in predicted binding affinity compared to *m*-methoxy (**24**) and *m*-propoxy (**25**) diastereomers, whereas no significant difference was observed for *m*-trifluoromethoxy diastereomers (**23**). These trends were not observed for the triazole-linked inhibitors with the only the *(R)*-configured diastereomer of the *m*-phenoxy triazole **3** shown to have significant improvement over the *(R)*-configured diastereomer of *m*-propoxy triazole **14**. The *p*-substituted methoxy triazole (**15**) and carbamate (**26**) showed comparable binding with the non-substituted counterparts (**16** and **27** respectively). It should be noted that this method is providing an indication of relative binding affinity between compounds and further exploration of substituents and substituent patterns needs to be undertaken before any definite conclusions can be drawn.

Compounds where a hydroxyl, methyl carbonate and methyl sulphonamide were trialled at the position of the phosphonate group of **16** and **27**, were also examined (Table 3 and S5-S7). For both the triazole (**17**-**19**) and carbamate (**28**-**30**) examples, no significant difference was observed between these compounds and the corresponding phosphodiester compounds, suggesting that these groups may also be potential isosteres. Although it should be noted that experimentally a charged group at this position, in particular a phosphonate, has been shown to produce the most potent inhibitors (Amann et al. 1998). Not being able to replicate this experimental trend can be attributed to the accuracy of the scoring function used in AutoDock Vina (Trott and Olson 2010).

Derivatives that are structurally different to the *m*-phenoxy structure of the lead compound (**2**) were also examined (Table 3 and S5-S7). For instance the triazole-linked indolone, *p*-tolylsulphonamide and 2,3-dihydro-1H-indenol derivatives **20**-**22** produced docking results that are comparable to the *m*-phenoxy structure **3**. Interestingly, the corresponding carbamate-linked **31**-**33** did not exhibit the same trend, with only the *(R)*-configured diastereomer of the indolone derivative **31**, the *(S)*-configured diastereomer of the *p*-tolylsulphonamide **32** and

2,3-dihydro-1H-indenol derivatives **33** having comparable results to the *m*-phenoxy structure **4**. This suggests that these structurally different derivatives benefit more from the rigidity provided by a triazole-linker.

Key features of hST6Gal I-inhibitor complexes: The general binding mode observed for all of the docked compounds exhibits a number of trends based on the different components of these structures. The CMP moiety of each compound reflected the interactions highlighted within the hST6Gal I crystal structure (Figure S1; Kuhn et al. 2013). The triazole- or carbamate-linker were generally in the proximity of residues such as Asn212, Tyr354 and His370, which is an analogous position to that observed for the phosphodiester-linker in compounds **1**-**2** and **5**-**9**. The groups that were mimicking the carboxylate of the natural donor, such as a phosphonate, were shown to be positioned in proximity to a central α-helix that has a positively polarised dipole (marked with an asterix in Figure 4A), that has been proposed to interact with the carboxylate (Kuhn et al. 2013; Meng et al. 2013). The different functionalities examined for the Neu5Ac mimic portion of the compounds were either located within or oriented towards one of two binding pockets (Figure 4B), both of which could be exploited in the future design of inhibitors of hST6Gal I. The first pocket, was occupied by compounds such as the unsubstituted benzyls (**16** and **27**) and those bearing substituents such as trifluoromethoxy (**5**, **12** and **23**) or methoxy (**13** and **24**). This pocket is thought to be one that the Neu5Ac of the sialyl donor could be accommodated within, based on the docking results obtained for CMP-3-F-Neu5Ac (**7**) and CMP-Neu5Ac (**8**). In compounds bearing larger substituents, such as a phenoxy (**1**-**4**) or a propoxy (**14** and **25**), these groups were located within or oriented towards the predicted sialyl acceptor site. This provides a rationale for the exploration of a potential route to selectivity, as each ST subfamily would be expected to have differences in this region to accommodate the different sialyl acceptors.

Molecular dynamics simulations of hST6Gal I-inhibitor complexes: To characterise important binding interactions between hST6Gal I and potential inhibitors, four all-atom MD simulations were performed in addition to those for CMP discussed previously. Two of these simulations utilised the reported potent phosphodiester-linked *(R)*-isomers, **1** (PL1) and **2** (PL2). The remaining two simulations were of the analogous carbamate-linked ***(R)*-4** (CAR1 and CAR2), with triazole-linked inhibitors to be examined in future studies. The PL1 and PL2 simulation RMSDs for Cα converged at 1.5 Å, with CAR1 and CAR2 converging at 2.0 Å (Figure S5), indicating no significant change in the structure of the protein backbone. It was noted that across all simulations each ligand remained tightly associated with the hST6Gal I active site. RMSFs for the Cα atoms, were consistent across all simulations (Figure S6). Each simulation showed consistently larger RMSF values for the flexible loop between Tyr355 and Tyr369 and the region that lies between Gln235 and Asn250. This region is adjacent to the portion of the active site that is expected to interact with the sialyl acceptor and the large fluctuation is most likely a result of the phenoxy being within the proposed sialyl acceptor site.

Interactions in the ligand-protein complexes analysed were characterised by monitoring the hydrogen bonds and hydrophobic contacts (Table S8-S15). The simulations with CMP in complex with hST6Gal I reproduced all hydrogen bonding interactions observed in the crystal structure (Figure S1; Kuhn et al. 2013). Additional interactions observed include Lys376 acting as a hydrogen bond donor to the pyridine like nitrogen of the cytosine ring and Glu375 acting as a hydrogen bond acceptor to the 2' hydroxyl group of the ribose. The water molecules shown to interact with CMP in the crystal structure were not observed, with two different bridging interactions from an oxygen of the phosphate group, to Ser189 and Tyr356 observed instead. Hydrophobic contacts that were observed were most significant with the cytosine ring, which was shown to interact with Ala190 and also have a proposed π-stacking

interaction with Tyr354. The ribose ring was shown to have hydrophobic interactions with Ser189, Phe211 and Leu372.

The PL1, PL2 and CAR simulations each revealed a number of consistent hydrogen bond and hydrophobic contacts with the hST6Gal I active site. These interactions have been schematically illustrated in Figure 5. The respective compounds of each simulation were shown to maintain the majority of the CMP moiety interactions that were observed in the CMP simulations, with a notable absence of the proposed π-stacking of the cytosine ring with Tyr354 in all cases. Additional water bridge interactions with the CMP moiety were also observed, with the interaction between the free amine of the cytosine ring and the backbone carbonyl of Val352 present in all simulations and the backbone NH of Gly324 and the 3'-hydroxyl of the ribose observed in all simulations except CAR1.

As well as the CMP moiety, each compound examined had two other conserved regions, the phosphonate and the *m*-phenoxy substituent. The conserved phosphonate component was shown to hydrogen bond with Tyr354 in all simulations and Asn233 in all simulations except PL2 (Figure 5). This does not correlate with the position generally observed in the docking studies, suggesting another potential position for the phosphonate has been captured. The *m*-phenoxy substituent for all compounds examined was shown to have hydrophobic contacts with Asn233, Gln235, Trp257, Pro259, Tyr275 and Tyr369 (Figure 5), which are present in the proposed sialyl acceptor binding site. This provides a further rationale for the exploration of *m*-phenoxy type substituents to target the proposed sialyl acceptor binding site as a potential route to the development of selective hST6Gal I inhibitors.

The two components of the compounds studied that were not consistent are the Neu5Ac mimic and the linking group. The Neu5Ac mimic in the PL1 simulation, Neu5Ac2en was shown to have limited interactions with hST6Gal I. Replacement of this moiety in the PL2 and CAR simulations with an aryl ring did not result in any significant change in interactions,

which is consistent with the comparable activity of compounds **1** and **2** that has been observed experimentally (Müller et al. 1998; Schröder and Giannis 1999; Skropeta et al. 2004). Consistent interactions were observed when comparing the phosphodiester-linker of the inhibitors in the PL1 and PL2 simulations to the carbamate-linker in the CAR simulations (Figure 5). These interactions include hydrogen bonding with Tyr354 and Asn212, which were consistent across all simulations and water bridged interactions with Ser323 and His370 that were consistent across PL1 and CAR simulations. The similarity in interactions observed for the carbamate-linker and the phosphodiester-linker in these simulations provides further evidence that a carbamate-linker may be a suitable isosteric replacement for the phosphodiester-linkers used in reported ST inhibitors.

Implication for inhibitor design: To further examine the promising trends that have been observed in this study, more rigorous calculations of binding free energy beyond that of molecular docking can be undertaken using techniques such as free energy perturbation (Hansen and van Gunsteren 2014). Relative binding free energies between the different linkers or different substitution patterns will provide insights into their effects on binding and guide the experimental optimisation process. This work is currently undertaken in our lab and will be reported in the future.

In previously published work, the utilisation of uncharged isosteres of phosphate in protein tyrosine phosphatase inhibitors, has shown promising results for increasing cell permeability and preventing cleavage by phosphatases (Rye and Baell 2005). Additionally, triazole-linkers have been shown to be an applicable phosphodiester isostere in transition-state analogue ST inhibitors (Lee et al. 2006; Kumar et al. 2013), suggesting that utilisation of uncharged isosteres may be a suitable strategy for overcoming these pharmacokinetic concerns. Recent work by Preidl et al. has demonstrated that derivatives of phosphodiester-linked transition-state analogue ST inhibitors bearing a fluorescent probe are able to cross the cellular

membrane via a vesicular uptake mechanism (Preidl et al. 2014), indicating that cell permeability may not be an issue for these compounds. However, *in vivo* it is expected that these phosphodiester-linked compounds would still be susceptible to phosphatase activity, which can be overcome with a neutral isostere. As demonstrated in this study both triazole and carbamate are potentially suitable isosteres for this purpose.

Improvements to the pharmacokinetic properties of transition-state analogue ST inhibitors that could be gained by the introduction of a neutral triazole- or carbamate-linker could logically be improved further by also replacing the charged phosphonate group with a neutral isostere. However, for phosphodiester-linked transition-state analogues it has been demonstrated experimentally that a charged group at this position, in particular a phosphonate, produces the most potent inhibitors (Amann et al. 1998). In order to maintain the phosphonate but still produce potential pharmacokinetic benefits, a neutral prodrug approach could be undertaken. The use of phosphonate prodrugs has been shown to be successful in producing several potent and selective nucleoside analogues that have clinical applications in the human immunodeficiency virus (HIV), hepatitis B, and hepatitis C virus fields (Pradere et al. 2014). A prodrug strategy as well as the incorporation of a neutral linker, as described in this study, could produce improvements to the pharmacokinetic properties of transition-state analogue ST inhibitors, and thus should be investigated further on a synthetic and biological level.

CONCLUSIONS

Molecular docking and molecular dynamics simulations were performed to investigate the replacement of the charged phosphodiester-linker used in reported transition-state analogue ST inhibitors with a neutral isostere, such as a carbamate or a triazole. With the molecular docking protocols used, eleven reported inhibitors and both the *(R)*- and *(S)*-configured

diastereomers of twelve of each triazole- and carbamate-linked derivatives were examined. This study successfully replicated experimental trends for reported inhibitors and furthermore demonstrated the potential for both the use of a carbamate- or triazole-linker as an isosteric replacement for the phosphodiester in transition-state analogue ST inhibitors. Molecular dynamics simulations of four different ligands complexed with hST6Gal I were performed and interactions in the complexes were characterised by monitoring hydrogen bonds and hydrophobic contacts over the course of the simulation. This analysis replicated all the hydrogen bonding interactions observed in the hST6Gal I-CMP complex crystal structure and demonstrated that in a dynamic system the phenoxy substituent of compounds **1**, **2** and ***(R)*-4** are oriented towards the proposed sialyl acceptor site, encouraging the targeting of this site as a potential route to design selective hST6Gal I inhibitors. Furthermore, it demonstrated that the carbamate-linker of ***(R)*-4** maintains the interactions with hST6Gal I observed for the phosphodiester-linker in the simulations of the reported inhibitors **1** and **2**. Overall, these simulations provide not only a rationale for the synthetic and biological evaluation of 1,2,3-triazole- and carbamate-linked transition-state analogue ST inhibitors, but also provide an insight into binding of these compounds, which can aid in the design of future ST inhibitors.

ACKNOWLEDGEMENTS

We wish to acknowledge the Australian Government for an Australian Postgraduate Award scholarship for A.M. and for an Australian Research Council Future Fellow for H.Y. (FT110100034); and Phil Clingan, Maxine Stewart and the Illawarra Cancer Carers for financial support and scholarship for R.S. We also wish to acknowledge that this research was undertaken with the assistance of resources provided at the NCI National Facility systems at the Australian National University through the National Computational Merit Allocation Scheme supported by the Australian Government.

REFERENCES

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

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

Audry M, Jeanneau C, Imberty A, Harduin-Lepers A, Delannoy P, Breton C. 2011. Current trends in the structure-activity relationships of sialyltransferases. *Glycobiology* **21**(6):716-726. doi: 10.1093/glycob/cwq189.

Breton C, Fournel-Gigleux S, Palcic MM. 2012. Recent structures, evolution and mechanisms of glycosyltransferases. *Curr. Opin. Struct. Biol.* **22**(5):540-549. doi: 10.1016/j.sbi.2012.06.007.

Brooks BR, Brooks Iii CL, Mackerell Jr AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S and others. 2009. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30**(10):1545-1614. doi: 10.1002/jcc.21287.

Burkart MD, Vincent SP, Düffels A, Murray BW, Ley SV, Wong CH. 2000. Chemo-enzymatic synthesis of fluorinated sugar nucleotide: Useful mechanistic Probes for glycosyltransferases. *Bioorg. Med. Chem.* **8**(8):1937-1946. doi: 10.1016/S0968-0896(00)00139-5.

Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E and others. 2012. The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**(5):401-404. doi: 10.1158/2159-8290.CD-12-0095.

Chen L, Wilson DJ, Xu Y, Aldrich CC, Felczak K, Sham YY, Pankiewicz KW. 2010. Triazole-linked inhibitors of inosine monophosphate dehydrogenase from human and mycobacterium tuberculosis. *J. Med. Chem.* **53**(12):4768-4778. doi: 10.1021/jm100424m.

Chen X, Varki A. 2010. Advances in the biology and chemistry of sialic acids. *ACS Chem. Biol.* **5**(2):163-176. doi: 10.1021/cb900266r.

D.A. Case TAD, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R., Luo RCW, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra,, J. Swails AWG, I. Kolossváry, K.F.Wong, F. Paesani, J. Vanicek, R.M.Wolf, J. Liu,, X. Wu SRB, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G., Cui DRR, D.H. Mathews, M.G. Seetin, R., Salomon-Ferrer, C. Sagui, V. Babin, T., Luchko SG, A. Kovalenko, and P.A. Kollman. 2012. AMBER 13: University of California, San Francisco.

Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**(12):10089-10092. doi: 10.1063/1.464397.

Datta AK, Chammas R, Paulson JC. 2001. Conserved Cysteines in the Sialyltransferase Sialylmotifs Form an Essential Disulfide Bond. *J. Biol. Chem.* **276**(18):15200-15207. doi: 10.1074/jbc.M010542200.

Datta AK. 2009. Comparative sequence analysis in the sialyltransferase protein family: Analysis of motifs. *Curr. Drug Targets* **10**(6):483-498. doi: 10.2174/138945009788488422.

Drinnan NB, Halliday J, Ramsdale T. 2003. Inhibitors of sialyltransferases: potential roles in tumor growth and metastasis. *Mini Rev. Med. Chem.* **3**(6):501-517. doi: 10.2174/1389557033487881.

El-Sagheer AH, Brown T. 2010. Click chemistry with DNA. *Chem. Soc. Rev.* **39**(4):1388-1405. doi: 10.1039/b901971p.

Feller SE, Zhang Y, Pastor RW, Brooks BR. 1995. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **103**(11):4613-4621. doi: 10.1063/1.470648.

Grundmann U, Nerlich C, Rein T, Zettlmeissl G. 1990. Complete cDNA sequence encoding human β-galactoside α-2,6-sialyltransferase. *Nucleic Acids Res.* **18**(3):667.

Hansen N, van Gunsteren WF. 2014. Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.* **10**(7):2632-2647. doi: 10.1021/ct500161f.

Hanwell MD, Curtis DE, Lonie DC, Vandermeerschd T, Zurek E, Hutchison GR. 2012. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **4**(8). doi: 10.1186/1758-2946-4-17.

Harduin-Lepers A, Vallejo-Ruiz V, Krzewinski-Recchi MA, Samyn-Petit B, Julien S, Delannoy P. 2001. The human sialyltransferase family. *Biochimie* **83**(8):727-737. doi: 10.1016/S0300-9084(01)01301-3.

Humphrey W, Dalke A, Schulten K. 1996. VMD: Visual molecular dynamics. *J. Mol. Graph. Model.* **14**(1):33-38. doi: 10.1016/0263-7855(96)00018-5.

Jeanneau C, Chazalet V, Augé C, Soumpasis DM, Harduin-Lepers A, Delannoy P, Imberty A, Breton C. 2004. Structure-function analysis of the human sialyltransferase ST3Gal I: Role of N-glycosylation and a novel conserved sialylmotif. *J. Biol. Chem.* **279**(14):13461-13468. doi: 10.1074/jbc.M311764200.

Jung KH, Schwörer R, Schmidt RR. 2003. Sialyltransferase Inhibitors. *Trends Glycosci. Glyc.* **15**(85):275-289. doi: 10.4052/tigg.15.275.

Kuhn B, Benz J, Greif M, Engel AM, Sobek H, Rudolph MG. 2013. The structure of human α-2,6-sialyltransferase reveals the binding mode of complex glycans. *Acta Crystallogr. D* **69**:1826-1828. doi: 10.1107/S0907444913015412.

Kumar R, Nasi R, Bhasin M, Khieu NH, Hsieh M, Gilbert M, Jarrell H, Zou W, Jennings HJ. 2013. Sialyltransferase inhibitors: Consideration of molecular shape and charge/hydrophobic interactions. *Carbohydr. Res.* **378**:45-55. doi: 10.1016/j.carres.2012.12.017.

Lairson LL, Henrissat B, Davies GJ, Withers SG. 2008. Glycosyl transferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**(1):521-555. doi: 10.1146/annurev.biochem.76.061005.092322.

Lee L, Chang K-H, Valiyev F, Liu H-J, Li W-S. 2006. Synthesis and Biological Evaluation of 5′-Triazole Nucleosides. *J. Chin. Chem. Soc.* **53**(6):1547-1555. doi: 10.1002/jccs.200600202.

Legaigneur P, Breton C, El Battari A, Guillemot JC, Augé C, Malissard M, Berger EG, Ronin C. 2001. Exploring the Acceptor Substrate Recognition of the Human β-Galactoside α2,6-Sialyltransferase. *J. Biol. Chem.* **276**(24):21608-21617. doi: 10.1074/jbc.M100860200.

Li Y, Chen X. 2012. Sialic acid metabolism and sialyltransferases: Natural functions and applications. *Appl. Microbiol. Biotechnol.* **94**(4):887-905. doi: 10.1007/s00253-012-4040-1.

Liu S, Meng L, Moremen KW, Prestegard JH. 2009. Nuclear magnetic resonance structural characterization of substrates bound to the α-2,6-sialyltransferase, ST6Gal-I. *Biochemistry* **48**(47):11211-11219. doi: 10.1021/bi9015154.

Liu S, Venot A, Meng L, Tian F, Moremen KW, Boons GJ, Prestegard JH. 2007. Spin-Labeled Analogs of CMP-NeuAc as NMR Probes of the α-2,6-Sialyltransferase ST6Gal I. *Chem. Biol.* **14**(4):409-418. doi: 10.1016/j.chembiol.2007.02.010.

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**(D1):D490-D495. doi: 10.1093/nar/gkt1178.

Martyna GJ, Tobias DJ, Klein ML. 1994. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101**(5):4177-4189. doi: 10.1063/1.467468.

Meng L, Forouhar F, Thieker D, Gao Z, Ramiah A, Moniz H, Xiang Y, Seetharaman J, Milaninia S, Su M and others. 2013. 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**(48):34680-34698. doi: 10.1074/jbc.M113.519041.

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

Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**(16):1781-1802. doi: 10.1002/jcc.20289.

Pradere U, Garnier-Amblard EC, Coats SJ, Amblard F, Schinazi RF. 2014. Synthesis of Nucleoside Phosphate and Phosphonate Prodrugs. *Chem. Rev.* **114**(18):9154-9218. doi: 10.1021/cr5002035.

Preidl JJ, Gnanapragassam VS, Lisurek M, Saupe J, Horstkorte R, Rademann J. 2014. Fluorescent mimetics of CMP-Neu5Ac are highly potent, cell-permeable polarization probes of eukaryotic and bacterial sialyltransferases and inhibit cellular sialylation. *Angew. Chem. Int. Ed.* **53**(22):5700-5705. doi: 10.1002/anie.201400394.

Rao FV, Rich JR, Raki B, Buddai S, Schwartz MF, Johnson K, Bowe C, Wakarchuk WW, Defrees S, Withers SG and others. 2009. Structural insight into mammalian sialyltransferases. *Nat. Struct. Mol. Biol.* **16**(11):1186-1188. doi: 10.1038/nsmb.1685.

Rye CS, Baell JB. 2005. Phosphate isosteres in medicinal chemistry. *Curr. Med. Chem.* **12**(26):3127-3141. doi: 10.2174/092986705774933452.

Sanner MF. 1999. Python: A programming language for software integration and development. *J. Mol. Graph. Model.* **17**(1):57-61.

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

Schultz MJ, Swindall AF, Bellis SL. 2012. Regulation of the metastatic cell phenotype by sialylated glycans. *Cancer Metastasis Rev.* **31**(3-4):501-518. doi: 10.1007/s10555-012-9359-7.

Schultz MJ, Swindall AF, Wright JW, Sztul ES, Landen CN, Bellis SL. 2013. ST6Gal-I sialyltransferase confers cisplatin resistance in ovarian tumor cells. *J. Ovarian Res.* **6**(1). doi: 10.1186/1757-2215-6-25.

Schwörer R, Schmidt RR. 2002. Efficient sialyltransferase inhibitors based on glycosides of N-acetylglucosamine. *J. Am. Chem. Soc.* **124**(8):1632-1637. doi: 10.1021/ja017370n.

Seales EC, Jurado GA, Brunson BA, Wakefield JK, Frost AR, Bellis SL. 2005. Hypersialylation of β 1 integrins, observed in colon adenocarcinoma, may contribute to cancer progression by up-regulating cell motility. *Cancer Res.* **65**(11):4645-4652. doi: 10.1158/0008-5472.CAN-04-3117.

Shaikh FM, Seales EC, Clem WC, Hennessy KM, Zhuo Y, Bellis SL. 2008. Tumor cell migration and invasion are regulated by expression of variant integrin glycoforms. *Exp. Cell Res.* **314**(16):2941-2950. doi: 10.1016/j.yexcr.2008.07.021.

Skropeta D, Schwörer R, Haag T, Schmidt RR. 2004. Asymmetric synthesis and affinity of potent sialyltransferase inhibitors based on transition-state analogues. *Glycoconjugate J.* **21**(5):205-219. doi: 10.1023/B:GLYC.0000045093.96413.62.

Skropeta D, Schwörer R, Schmidt RR. 2003. Stereoselective synthesis of phosphoramidate α(2-6)sialyltransferase transition-state analogue inhibitors. *Bioorg. Med. Chem. Lett.* **13**(19):3351-3354. doi: 10.1016/S0960-894X(03)00672-3.

Søndergaard CR, Olsson MHM, Rostkowski M, Jensen JH. 2011. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of p K a values. *J. Chem. Theory Comput.* **7**(7):2284-2295. doi: 10.1021/ct200133y.

Trott O, Olson AJ. 2010. Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**(2):455-461. doi: 10.1002/jcc.21334.

Volkers G, Worrall LJ, Kwan DH, Yu C-C, Baumann L, Lameignere E, Wasney GA, Scott NE, Wakarchuk W, Foster LJ and others. 2015. Structure of human ST8SiaIII sialyltransferase provides insight into cell-surface polysialylation. *Nat. Struct. Mol. Biol.* **22**(8):627-635. doi: 10.1038/nsmb.3060.

Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. 2004. Development and testing of a general amber force field. *J. Comput. Chem.* **25**(9):1157-1174. doi: 10.1002/jcc.20035.

Wang P-H. 2005. Altered Glycosylation in Cancer: Sialic Acids and Sialyltransferases. *J. Cancer Mol.* **1**(2):73-81.

Wang X, Zhang LH, Ye XS. 2003. Recent development in the design of sialyltransferase inhibitors. *Med. Res. Rev.* **23**(1):32-47. doi: 10.1002/med.10030.

Weinstein J, Lee EU, McEntee K, Lai PH, Paulson JC. 1987. Primary structure of β-galactoside α2,6-sialyltransferase. Conversion of membrane-bound enzyme to soluble forms by cleavage of the NH2-terminal signal anchor. *J. Biol. Chem.* **262**(36):17735-17743.

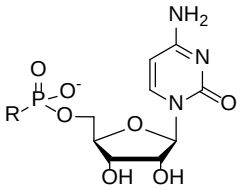
Whalen LJ, McEvoy KA, Halcomb RL. 2003. Synthesis and evaluation of phosphoramidate amino acid-based inhibitors of sialyltransferases. *Bioorg. Med. Chem. Lett.* **13**(2):301-304. doi: 10.1016/S0960-894X(02)00735-7.

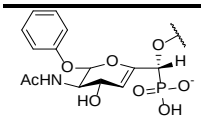
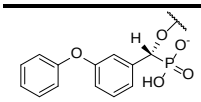
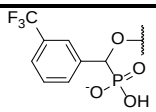
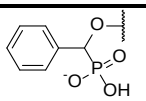
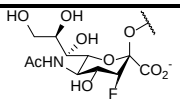
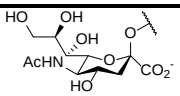
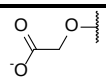
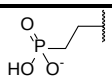
Table 1: Overview of all MD simulations[a]

Name	Ligand	Total Simulation Time (ns)	Equilibration (ns)	Free Simulation (ns)
Apo	-	100	10	90
CMP1	cytidine 5'-monophosphate (CMP)	100	10	90
CMP2	CMP	100	10	90
CMP3	CMP	100	10	90
PL1	Phosphodiester-linked inhibitor **1**	100	10	90
PL2	Phosphodiester-linked inhibitor **2**	100	10	90
CAR1	Carbamate-linked inhibitor ***(R)*-4**	100	10	90
CAR2	Carbamate-linked inhibitor ***(R)*-4**	100	10	90

[a]The systems consist of approximately 49 000 atoms with a box size of 81 Å × 81 Å × 81 Å

Table 2: Mean binding affinities[a] of the optimal binding modes for reported ST inhibitors based on AutoDock Vina



R	Cpd	Mean Binding Affinity Apo[b] (kcal/mol)	Mean Binding Affinity CMP[c] (kcal/mol)	Experimental rST6Gal I K_i (μM)	Experimental Estimate of Binding Affinity[d] (kcal/mol)
	***(R)*-1**	-10.0 ± 0.2	-9.6 ± 0.1	0.029[e]	-10.3
	***(R)*-2**	-10.0 ± 0.2	-9.7 ± 0.2	0.070[f]	-9.8
	***(R)*-5**	-9.7 ± 0.1	-9.4 ± 0.2	0.25[f]	-9.1
	***(S)*-5**	-9.3 ± 0.2	-9.3 ± 0.2	0.22[f]	-9.1
	***(R)*-6**	-9.1 ± 0.2	-9.1 ± 0.1	0.20[g]	-9.2
	***(S)*-6**	-8.9 ± 0.1	-8.8 ± 0.1	1.0[g]	-8.2
	**7**	-8.4 ± 0.1	-8.5 ± 0.1	5.7[h]	-7.2
	**8**	-8.6 ± 0.1	-8.6 ± 0.1	44[i]	-6.0
	**9**	-8.2 ± 0.1	-7.9 ± 0.1	370[j]	-4.7
	**10**	-8.0 ± 0.2	-7.8 ± 0.1	750[i]	-4.3
	**11**	-8.2 ± 0.1	-7.9 ± 0.1	2000[i]	-3.7

[a] Arithmetic mean binding affinity ± standard error of the mean (SEM) obtained from docking into six snapshots of MD simulation.

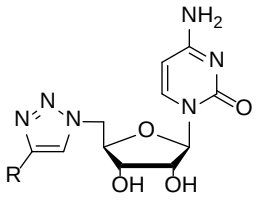
[b] Docking carried out on the snapshots from the Apo molecular dynamics simulations

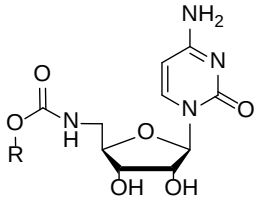
[c] Docking carried out on the snapshots from the CMP1 simulations.

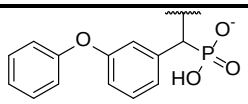
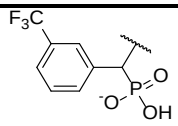
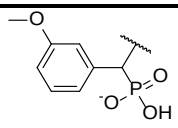
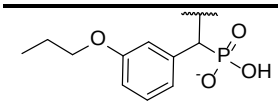
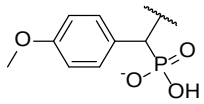
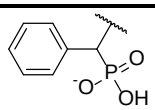
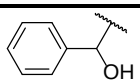
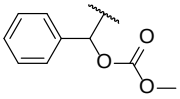
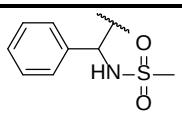
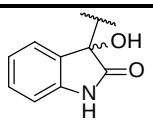
[d] Estimated using $\Delta G = RTln(K_i)$ where R = 1.9872 x 10^{-3} kcal/K/mol, T = 300 K and K_i values were M

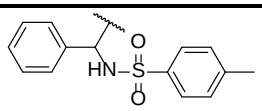
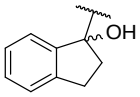
[e] (Schwörer and Schmidt 2002); [f] (Skropeta et al. 2004);[g] (Müller et al. 1998); [h] (Burkart et al. 2000); [i] (Amann et al. 1998), [j] (Whalen et al. 2003)

Table 3: Mean binding affinities of the optimal binding modes for 1,2,3-triazole- and carbamate-linked derivatives





R	Cpd	Mean Binding Affinity Triazole[a] (kcal/mol)	Cpd	Mean Binding Affinity Carbamate[a] (kcal/mol)
	***(R)*-3**	-9.7 ± 0.1	***(R)*-4**	-9.8 ± 0.2
	***(S)*-3**	-9.9 ± 0.1	***(S)*-4**	-9.8 ± 0.2
	***(R)*-12**	-9.7 ± 0.2	***(R)*-23**	-9.5 ± 0.1
	***(S)*-12**	-9.9 ± 0.1	***(S)*-23**	-9.7 ± 0.2
	***(R)*-13**	-9.4 ± 0.2	***(R)*-24**	-9.1 ± 0.2
	***(S)*-13**	-9.8 ± 0.2	***(S)*-24**	-9.1 ± 0.2
	***(R)*-14**	-9.1 ± 0.1	***(R)*-25**	-8.8 ± 0.1
	***(S)*-14**	-9.4 ± 0.2	***(S)*-25**	-9.0 ± 0.2
	***(R)*-15**	-9.1 ± 0.2	***(R)*-26**	-8.8 ± 0.1
	***(S)*-15**	-9.2 ± 0.2	***(S)*-26**	-8.9 ± 0.1
	***(R)*-16**	-9.3 ± 0.1	***(R)*-27**	-9.0 ± 0.2
	***(S)*-16**	-9.1 ± 0.2	***(S)*-27**	-8.9 ± 0.1
	***(R)*-17**	-9.1 ± 0.2	***(R)*-28**	-8.8 ± 0.1
	***(S)*-17**	-8.7 ± 0.1	***(S)*-28**	-8.6 ± 0.2
	***(R)*-18**	-8.9 ± 0.2	***(R)*-29**	-8.7 ± 0.2
	***(S)*-18**	-9.1 ± 0.2	***(S)*-29**	-8.7 ± 0.2
	***(R)*-19**	-9.0 ± 0.1	***(R)*-30**	-8.8 ± 0.1
	***(S)*-19**	-9.2 ± 0.1	***(S)*-30**	-9.0 ± 0.1
	***(R)*-20**	-9.7 ± 0.1	***(R)*-31**	-9.2 ± 0.3
	***(S)*-20**	-9.6 ± 0.2	***(S)*-31**	-9.1 ± 0.1

	***(R)*-21**	-9.8 ± 0.2	***(R)*-32**	-9.1 ± 0.2
	***(S)*-21**	-10.0 ± 0.3	***(S)*-32**	-9.6 ± 0.2
	***(R)*-22**	-9.7 ± 0.2	***(R)*-33**	-9.1 ± 0.2
	***(S)*-22**	-9.5 ± 0.2	***(S)*-33**	-9.2 ± 0.1

[a] Arithmetic mean of binding affinity ± SEM obtained from docking into six snapshots of CMP1 simulation.

Figure Legends

Figure 1. The structure of hST6Gal I (PDB id 4JS2; Kuhn et al. 2013) with sialyl motifs L (red), III (blue), S (yellow) and VS (orange) highlighted. The co-crystallised ligand cytidine 5'-monophosphate (CMP) is represented as pipes. The soluble catalytic domain of hST6Gal I extends into the lumen of the Golgi apparatus via a stem region from a transmembrane domain and a short *N*-terminal cytoplasmic domain.

Figure 2. Proposed reaction scheme for the catalytic action of hST6Gal I. R represents a glycan chain of a glycoconjugate.

Figure 3. Development of the phosphodiester-linked lead compound (**2**) and the proposed 1,2,3-triazole- (**3**) and carbamate- (**4**) linked derivatives. Reported K_i values expressed for compounds **1** (Schwörer and Schmidt 2002) and **2** (Skropeta et al. 2004) are determined against rat liver α-2,6-ST.

Figure 4. (A) Comparison of binding modes obtained using AutoDock Vina between the phosphodiester-linked lead compound **2** (purple) and the 1,2,3-triazole- ***(R)*-3** (green) and carbamate-linked ***(R)*-4** (orange) derivatives. Compounds pictured are the binding modes obtained from docking into the 40ns snapshot of the CMP1 simulation. The backbone of hST6Gal I is shown as grey ribbons and the central α-helix that has a positively polarised dipole is marked with an asterix. **(B)** Surface model of hST6Gal I highlighting the two

Figure 5. Schematic representation of conserved hydrogen bonding **(A)** and hydrophobic contacts **(B)** observed in the PL1, PL2 and CAR simulations. The consistent components of each inhibitor are shown as chemical structures. The red box represents the different linkers and the blue box represents the different Neu5Ac mimic. Dashed lines indicate the interacting atoms shown in Table S10-15 and "W" represents a single molecule of water.