Structure-based mutagenic analysis of mechanism and substrate specificity in mammalian glycosyltransferases: Porcine ST3Gal-I

Bojana Rakitić1,2, Francesco V Rao1,2,3,4,5, Karen Freimann2,3, Warren Wakarchuk6, Natalie C J Strynadka4,5, and Stephen G Withers1,2,3

2Department of Chemistry; 3The Centre for High-throughput Biology (CHiBi); 4The Department of Biochemistry and Molecular Biology; 5The Centre for Blood Research, University of British Columbia, Vancouver, BC, Canada V6T 1Z3; and 6The Department of Molecular and Cellular Biology, University of Guelph, Guelph, ON, Canada N1G 2W1

Received on September 10, 2012; revised on December 11, 2012; accepted on January 2, 2013

Introduction

Sialyltransferases (STs) play essential roles in signaling and in the cellular recognition processes of mammalian cells by selectively installing cell-surface sialic acids in an appropriate manner both temporally and organ-specifically. The availability of the first three-dimensional structure of a mammalian (GT29) sialyltransferase has, for the first time, allowed quantitative structure/function analyses to be performed, thereby providing reliable insights into the roles of key active site amino acids. Kinetic analyses of mutants of ST3Gal-I, in conjunction with structural studies, have confirmed the mechanistic roles of His302 and His319 as general acid and base catalysts, respectively, and have quantitated other interactions with the cytosine monophosphate-α-acetyl-β-neuraminic acid donor substrate. The contributions of side chains that provide key interactions with the acceptor substrate, defining its specificity, have also been quantitated. Particularly important transition-state interactions of 2.5 and 2.7 kcal mol⁻¹ are found between the acceptor axial 4-hydroxyl and the conserved side chains of Gln108 and Tyr269, respectively. These results provide a basis for the engineering of mammalian STs to accommodate non-natural substrate analogs that should prove valuable as chemical biological probes of sialyltransferase function.

Keywords: kinetics / mammalian sialyltransferase / mutagenesis / ST3Gal-I / substrate specificity

© The Author 2013. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
asialo core 2 oligosaccharides present in healthy tissues to carrying sialylated core 1 in diseased tissues.

STs are localized in the Golgi and are Type II membrane proteins. They have a short N-terminal cytoplasmic tail that tethers them to the Golgi membrane, a hydrophobic transmembrane domain, a long-stem domain and a large catalytic C-terminal domain in the lumen of the Golgi. They are inverting GTs, so called because the sugar anomeric center is inverted upon transfer, and follow a direct displacement SN2-like mechanism in which the base deprotonates the hydroxyl group of the acceptor during attack on the anomeric centre of the CMP-Neu5Ac donor, with reaction occurring via an oxocarbenium-ion like transition state. In ST3Gal-I, which catalyzes the transfer of neuraminic acid from CMP-Neu5Ac to the 3-hydroxyl residue of galactose in the Galβ1,3GalNAc moiety attached to Ser/Thr O-glycans or ganglio-series glycolipid acceptors, the 3-hydroxyl of the galactose moiety directly attacks the C2 of CMP-Neu5Ac to form the Neu5Ac2,3Gal linkage (Scheme 1).

The first (and only) crystal structure of a mammalian sialyltransferase to date is that of porcine ST3Gal-I (pST3Gal-I), which has 85% sequence identity to the human ST3Gal-I (Rao et al. 2009). The enzyme was crystallized in the apo form, as well as in a binary complex with the Galβ1,3GalNAcα-PhNO2 acceptor and in a ternary complex with the CMP product and the disaccharide acceptor. The crystal structure of pST3Gal-I reveals a mixed αβ fold with seven twisted β-strands flanked by 12 α-helices, and adopts a GT-A like fold. However, because of its many differences from other GT-A proteins that utilize nucleotide-activated donor sugars, it is classified as GT-A variant 2. The GT42 ST Cst-II, which shares a similar β-sheet structure with pST3Gal-I, represents variant 1 of the GT-A fold (Chiu et al. 2004). pST3Gal-I lacks the DxD motif characteristic of GTs that require divalent cations for their activity, consistent with the fact that no metal was observed in the crystal structure of pST3Gal-I nor was the activity of the enzyme enhanced in the presence of metal cations. The crystal structure with the bound disaccharide acceptor and the CMP product revealed the locations and structures of the conserved peptide sequences of the catalytic domain known as sialylmotifs large (L), small (S) and very small (VS). Three disulfide bonds are observed in the pST3Gal-I crystal structure, two of which are located in the loop region near the N terminus, Cys62–67 (C1) and Cys65–142 (C2) and the third Cys145–284 (C3) connects β-sheet 1 in the L sialylmotif, with α-helix 11 in the S sialylmotif. Sialylmotifs L and S each contain one invariant cysteine residue, separated by ~150 amino acids, which form the intramolecular disulfide bridge between the L and S motifs that are essential for proper protein folding and activity (Datta et al. 2001). In the pST3Gal-I structure, this disulfide bridge brings the end of the helix α1 from sialylmotif S to the cysteine in β1 from sialylmotif L (Cys145–284). The segment of residues from 305 to 316 is disordered in the native and ligand-bound crystal structures. A shallow acceptor binding pocket is located on the enzyme surface with the galactosyl moiety of Galβ1,3GalNAcα-PhNO2 substrate facing in and the rest of the acceptor moiety facing out toward the missing loop and the solvent. Very few differences in the protein positions are seen between the crystal structures of the apo pST3Gal-I and the pST3Gal-I ternary complex with the disaccharide acceptor Galβ1,3GalNAcα-PhNO2, and CMP product, implying that

Scheme 1. Schematic of reaction mechanism and active site.
the acceptor binding site is largely preformed. This is consistent with the random-order kinetic mechanism determined earlier for pST3Gal-I (Rearick et al. 1979).

The acceptor binding sites of different mammalian STs must vary in structure in order for them to accommodate different carbohydrate moieties such as galactose in ST3Gal and ST6Gal, N-acetylgalactosamine in ST6GalNAc or sialic acid in ST8Sia. Likewise different regiochemical outcomes must be encoded; thus, ST3Gal and ST6Gal need to bind the galactosyl moiety in different orientations relative to CMP-Neu5Ac in order to make α2,3 and α2,6 linkages, respectively. Previous mutagenesis studies revealed several conserved peptide sequences within each subgroup family, located between the L and S sialylmotifs, that were postulated to contribute to the different linkages and acceptor specificity (Patel and Balaji 2006). Two such submotifs (residues 187–196 and 212–219) are observed in the case of pST3Gal-I.

Here, we report a kinetic analysis of pST3Gal-I variants in which key residues have been mutated in order to further investigate the role of the homologous motifs as well as to probe the enzyme donor and acceptor substrate specificity. In this study, we worked with the Δ45N-truncated version of pST3Gal-I, having an N-terminal hexahistidine (His) tag and a maltose binding protein (MBP) for purification and solubility considerations, respectively. The enzyme lacks 45 amino acids from the N-terminal anchoring site and the stem region, but it contains the entire catalytic domain. Although this stem region varies in amino acid composition and length among STs (from no stem region in ST6GalNAc-III to 200 amino acids long in ST6GalNAc), we and others have shown that its deletion has no effect on the enzyme activity in vitro. However, the stem region is sensitive to proteolysis and might be involved in retention of the enzyme in the Golgi and in the in vivo acceptor specificity (Datta 2009). The stem region appears to have little secondary structure and it often contains cysteine residues and several N- and O-glycosylation sites. pST3Gal-I is N-glycosylated at four asparagine residues in vivo with one glycosylation site on the N-terminal stem and three others on the protein surface. Our pST3Gal-I is expressed in Escherichia coli cells and lacks these posttranslational modifications. Although this deficiency does not affect in vitro protein folding or activity, glycosylation presumably has a more complex role in vivo related to protein stabilization and trafficking.

The availability of a three-dimensional structure for a mammalian sialyltransferase makes it possible to directly interrogate the roles of active site residues in substrate binding and catalysis and to interpret results of previous such studies of a structural basis. In this report, we investigate the key interactions in the donor and acceptor site and quantitate the importance of each such interaction in ground- and/or transition-state binding through kinetic analysis of purified mutants. Analysis of these contributions in terms of typical strengths of hydrogen bonds between two neutral partners (1–1.5 kcal mol⁻¹) or between a charged and a neutral partner (3–4.5 kcal mol⁻¹) agrees well with the structures determined or inferred, greatly strengthening our confidence in the structures and models that have been derived (Fersht et al. 1985; Street et al. 1986). Supporting evidence for two histidines playing roles as acid and base catalysts is provided, as well as insight into the mechanistic role of a flexible active site loop.

**Results**

**Choice of mutants, protein production and circular dichroism spectroscopy**

As shown in Scheme 1 and Figure 1, a suite of well-positioned amino acid side chains provides hydrogen bonding and other interactions for recognition of the donor and acceptor substrates, as well as presumably for optimal binding and stabilization of the reaction transition state. Among these, two histidine residues are thought to play particularly important direct roles in catalysis. His319 is well positioned to function as the general base catalyst by interacting with the 3'-hydroxyl of the acceptor galactose moiety (Figure 1). Likewise, His302 interacts with the phosphate moiety of CMP-Neu5Ac and likely serves either as a formal Brønsted acid catalyst or as a strong hydrogen bonding or electrostatic partner for the departing CMP (Figure 1). Consequently, this suite of residues was targeted for mutagenesis and kinetic analysis of the resulting variants. In addition, mutations were targeted at three positions (Phe313, Arg314 and Lys315) within a segment of polypeptide loop that is not observed crystallographically, but is proposed to form a closed lid over the active site in the ternary complex (Rao et al. 2009). The three specific mutation sites were chosen on the basis of the similar closed lid domain seen in the Cst-II crystal structure in complex with the stable analog CMP3F-Neu5Ac (Chiu et al. 2004), where these three residues interact with the sialic acid moiety.

Mutations were made within the N-terminally deleted (Δ45) construct to which had been appended an N-terminal His tag for improved purification purposes, as well as a maltose binding tag for solubilization and stabilization of the protein, as described previously (Rao et al. 2009). Kinetic studies were performed on this full-length protein, since comparison of kinetic parameters with those of the crystallographically solved cleaved version lacking MBP and His tags, revealed no significant differences. The full length version is more stable and soluble, thus is much easier to work with. All mutants were purified using nickel affinity chromatography, then concentrated and dialyzed into standard Tris/NaCl buffer, pH 7.5. Correct folding of these mutant proteins was confirmed by circular dichroism (CD) spectroscopy and identical CD spectra were seen for all mutants studied (Figure 2).

**Kinetic analysis of wild-type and mutant forms**

Kinetic parameters for the wild-type and each of the mutants (kcat, KM and kcat/KM) were determined for both the donor substrate CMP-Neu5Ac (at a fixed 1 mM concentration of acceptor), and for the acceptor substrate at a fixed (0.6 mM) concentration of donor (Table 1). The majority of the kinetic studies were performed with the Galβ1,3GalNAc-α-OBN acceptor, based upon previous studies (Rearick et al. 1979) that showed Galβ1,3GalNAc as the optimal acceptor for this enzyme in the context of mucin glycoproteins. Kinetic parameters for a limited, and relevant, subset of mutants were also
determined with an alternate acceptor Galβ1,6GalNAc-α-OBn (Table II). Only the $k_{cat}/K_M$ value could be determined for Galβ1,3GlcNAc-α-OBn since saturation kinetics could not be observed at the substrate concentrations achievable. Values of $\Delta\Delta G^\ddagger$ were calculated based upon changes in the $k_{cat}/K_M$ value for donor or acceptor relative to wild-type enzyme, as described in Materials and methods. $k_{cat}/K_M$ values reflect the first irreversible step which, under these initial rate conditions, is likely to be the chemical step of bond cleavage. The $\Delta\Delta G^\ddagger$ values, therefore, reflect the consequences of that mutation on the transition state and thus are a measure of the contribution of the interactions with that residue to catalysis.

**Discussion**

**Wild-type enzyme**

Kinetic parameters for the wild-type enzyme, using Galβ1,3GalNAc-α-OBn as acceptor, are in good agreement with those determined previously for the proteolytically cleaved and crystallographically defined version of this protein missing the MBP and His tags (Rao et al. 2009). Interestingly, while indeed the Galβ1,3GalNAc containing substrate is the best acceptor tested, as concluded by Rearick et al. (1979) previously, the $k_{cat}/K_M$ value is only three times higher than that of its 1,6-isomer. This contrasts with the >100-fold difference in relative rates ($k_{cat}$ and $K_M$ values were not determined...
Table I. Kinetic parameters using CMP-Neu5Ac as donor substrate and Galβ1,3GalNAc-α-OBn as acceptor substrate

<table>
<thead>
<tr>
<th>pST3Gal-I</th>
<th>Donor CMP-Neu5Ac</th>
<th>Acceptor Galβ1,3GalNAc-α-OBn</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_m$ (µM)</td>
<td>$k_{cat}$ (min$^{-1}$)</td>
</tr>
<tr>
<td>WT</td>
<td>123 ± 11</td>
<td>170 ± 5</td>
</tr>
<tr>
<td>H302A</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>H319A</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>W300A</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Y194F</td>
<td>460 ± 30</td>
<td>40 ± 2</td>
</tr>
<tr>
<td>Y269F</td>
<td>480 ± 60</td>
<td>250 ± 16</td>
</tr>
<tr>
<td>T272A</td>
<td>260 ± 50</td>
<td>620</td>
</tr>
<tr>
<td>Q108A</td>
<td>390 ± 40</td>
<td>80 ± 4</td>
</tr>
<tr>
<td>WT</td>
<td>670 ± 100</td>
<td>70 ± 5</td>
</tr>
</tbody>
</table>

Table II. Comparison of selected kinetic parameters for Galβ1,3GalNAc-α-OBn and Galβ1,6GalNAc-α-OBn

<table>
<thead>
<tr>
<th>pST3Gal-I</th>
<th>Acceptor Galβ1,3GalNAc-α-OBn</th>
<th>Acceptor Galβ1,6GalNAc-α-OBn</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_m$ (µM)</td>
<td>$k_{cat}$ (min$^{-1}$)</td>
</tr>
<tr>
<td>WT</td>
<td>70 ± 10</td>
<td>172 ± 7</td>
</tr>
<tr>
<td>Y269F</td>
<td>2600 ± 200</td>
<td>81 ± 3</td>
</tr>
<tr>
<td>Q108A</td>
<td>2340 ± 240</td>
<td>100 ± 5</td>
</tr>
<tr>
<td>Y233A</td>
<td>260 ± 60</td>
<td>105</td>
</tr>
<tr>
<td>Y269F</td>
<td>890 ± 130</td>
<td>67 ± 4</td>
</tr>
<tr>
<td>Q108A</td>
<td>390 ± 40</td>
<td>80 ± 4</td>
</tr>
<tr>
<td>Y233A</td>
<td>390 ± 40</td>
<td>80 ± 4</td>
</tr>
</tbody>
</table>

*Galβ1,3GlcNAc-α-OBn is a very poor acceptor ($k_{cat}/K_m = 5$ mM$^{-1}$ min$^{-1}$) with the wild-type enzyme. Galactose does not function as an acceptor.

in this case) measured by Rearick et al. This difference may well have its origins in the additional presence of an α-linked aromatic substituent at the anomeric centre. In order to probe this a 3D model of Galβ1,6GalNAc was generated and superimposed onto the Galβ1,3GlcNAc model. These studies suggest that this substituent may stack well with Tyr233, thereby providing additional binding interactions.

Also of interest is our demonstration that removal of the axial C-4 hydroxyl of the N-acetylhexosamine sugar moiety of the acceptor severely disrupts interactions. The $k_{cat}/K_m$ values for Galβ1,3GlcNAc-α-OBn is some 500-fold lower than that of its GalNAc isomer (Table II), corresponding to a loss of ≈4 kcal mol$^{-1}$ in transition-state binding affinity. This is fully consistent with earlier measurements by Rearick, who measured $V_{max}/K_m$ values that differ by ≈650-fold for the two isomers. Inspection of the structure (Scheme 1, Figure 1) reveals H-bonding interactions at that position with Tyr269 and Gln108. Presumably, these two interactions together contribute this sum of 4 kcal mol$^{-1}$ transition-state stabilization and serve as key discriminating ligands to the acceptor substrate. Indeed, as will be discussed in more detail later, it is interesting to note that individual mutations of these residues suggest they provide a total of (2.5 + 2.7) = 5.2 kcal mol$^{-1}$ of transition-state stabilization. The majority (4 kcal mol$^{-1}$) of this can be attributed to their interactions with the axial 4-hydroxyl. Unsurprisingly, galactose alone is not found to be an acceptor.

**His302 and His319 as acid/base catalysts**

The postulated roles for these two histidine residues are reinforced by this study. Replacement of the highly sequence-conserved His319 by Ala, whose methyl group cannot possibly function as a base catalyst, results in a completely inactive mutant enzyme (Table I). Although the mutants were not fully purified and quantitated in those cases, similarly inactive mutants at that position were found previously for hST3Gal-I and for the homologous histidine in ST8 Sia-II and -IV (Kitazume-Kawaguchi et al. 2001; Jeanneau et al. 2004). Particularly, strong support for such a base catalytic role for histidine residues in STs was obtained previously for hST3Gal-I and for the homologous histidine in ST8 Sia-II and -IV ((Kitazume-Kawaguchi et al. 2001; Jeanneau et al. 2004). Particularly, strong support for such a base catalytic role for histidine residues in STs was obtained previously for His188 in Cst-II (Chiu et al. 2004; Chan et al. 2009). 13C-nuclear magnetic resonance spectroscopy was used to measure the intrinsic pKₐ of selectively 13C-labelled His188. The measured pKₐ value of 6.6 agreed well with the pKₐ derived from the pH dependence of $k_{cat}/K_m$ for the whole enzyme, pKₐ = 6.5, consistent with its role as base.

Similarly, replacement of His302, which interacts with the phosphate moiety of CMP-Neu5Ac (Figure 1), by Ala also results in a completely inactive mutant enzyme (Table I). This residue is conserved in all ST3 enzymes and a previous kinetic analysis of mutants of the human ST3Gal-I had suggested that replacement by Ala abolishes all activity (Jeanneau et al. 2004). Here, using fully purified mutants, we clearly establish the importance of that residue in a structurally defined enzyme. This imidazole side chain likely either functions as a formal acid catalyst and transfers a proton to the CMP phosphate moiety as sialyl transfer occurs, or retains the proton and electrostatically stabilizes the developing dianion.

**Contributions of side-chain interactions with CMP-Neu5Ac**

All attempts to date to capture a complex of ST3Gal-I with donor CMP-Neu5Ac or analogs thereof have been

540
unsuccessful. However, the structure of ST3Gal-I with CMP and acceptor reveals the nucleotide bound in a cleft adjacent to the Rossman fold-containing β-core in a very similar manner to that in which CMP3F-Neu5Ac is bound to the GT42 enzyme Cst-II (Chiu et al. 2004). This allowed a model to be built of CMP3F-Neu5Ac bound to ST3Gal-I by superposition (Figure 1B). This modeled structure places the 3-hydroxyl of the acceptor galactose moiety ~2.8 Å from the anomeric centre of CMP3F-Neu5Ac, with the phosphate group oriented axially for departure. As noted, the phosphate group interacts with His302 as well as two asparagines (Asn150 and Asn173), and this histidine plays a crucial role in the mechanism, consistent with its presence as a highly conserved residue within sialylmotif 3 (H/y)(Y/W/F/h)(D/E/q/g). In pST3Gal-I, the specific sequence is HYWE.

Trp300 is located in a position from which it could readily move to form stacking interactions with the cytidine ring upon binding of CMP-Neu5Ac, as seen with some bacterial STs. This would explain our observation that mutation of Trp300 to alanine completely abrogated activity, even though the protein folds correctly, as shown by CD (Table I and Figure 2). Such aromatic stacking interactions are often found in nucleotide-binding sites. Indeed aromatic side chains of Phe155, Phe170 and Tyr156 play a similar role in the structurally related GT42 enzymes Cst-I and -II (Chiu et al. 2004, 2007; Lee et al. 2009, 2011). However, no π stacking with aromatic side chains is seen in the GT 80 Pasteurella multocida structure (Ni et al. 2007) (GT-B fold) or in the recently determined GT-B fold GT52 sialyltransferase from Neisseria meningitidis (Lin et al. 2011); here the alkyl side chains of Pro281 and Ile299 sandwich the base. Within ST3Gal-I and -II enzymes, Trp300 is fully conserved and precedes the aforementioned sialylmotif 3. Apparent key interactions with the sialic acid moiety that are predicted crystallographically are those of Ser271 and Thr272 with the C-1 carboxylate moiety (Figure 1). These residues form part of sialylmotif S and sit at the end of helix 11. Ser271 is fully conserved in ST3Gal-I and -II and is conservatively substituted by Thr in the remaining members of the ST3Gal group. The hydroxyl side chain performs a classical helix capping function, forming a strong hydrogen bond with the main chain nitrogen at position 274. Thr272 is also highly conserved across all ST3Gal enzymes except ST3Gal-II and -V, where it is replaced by Leu and Ile, respectively. Unfortunately, all our attempts to express the Ser271Ala mutant failed—suggesting that its helix capping function is indeed important to the folding and stability of the enzyme. In contrast, the Thr272Ala mutant expressed well, though surprisingly this mutation had very little effect on kinetic parameters for the donor CMP-Neu5Ac—simply a doubling of the $K_M$ (Table I). Small effects were, however, observed in the 4-fold increase of the acceptor $K_M$ value, possibly reflecting loss of the water-mediated hydrogen bonds between this threonine hydroxyl and the 3 and 4 hydroxyls of the galactose moiety. The minor effect of this substitution on kinetic parameters for CMP-Neu5Ac suggests that either the proposed interaction of Thr272 with the donor carboxylic acid is not important or that in the mutant a water molecule is able to substitute for the threonine hydroxyl. Indeed in retrospect the fact that similar substitutions are seen in ST3Gal-III and -V, which of course use exactly the same donor substrate, is completely consistent (Harduin-Lepers et al. 2005).

The final interaction with the donor sugar that was probed is that of Tyr194 with the 7-hydroxyl of CMP-Neu5Ac. Indeed removal of that hydrogen-bonding interaction via mutation to Ala results in ~15-fold loss in activity, corresponding to a decrease in transition-state stabilization of 1.7 kcal mol$^{-1}$ and fully consistent with the removal of a single neutral hydrogen bond (Persh et al. 1985; Street et al. 1986). Much of this effect is likely upon $K_M$—since saturation kinetics were not observed at 0.7 mM donor (the highest donor concentration achievable due to limits placed by spontaneous hydrolysis). The mutation also affects the acceptor $K_M$ value, increasing it 7-fold, though this remains an estimate of the true effect since, for this variant, the value was determined using sub-saturating concentrations of donor due to its high $K_M$ value. An effect on the acceptor would be anticipated in any case since Tyr194 forms stacking interactions with the disaccharide (Figure 1).

**Contributions of side-chain interactions with the acceptor disaccharide**

A complex network of possible hydrogen bonding interactions is seen with the acceptor disaccharide as shown in Scheme 1, where dashed lines indicate groups within 4 Å, although this full complement of interactions may not be present at all times. Gal OH-4 hydrogen bonds with His319, the hydroxyl group of the Tyr269 side chain and also with ordered water molecules. Gal OH-6 forms several interactions: It hydrogen bonds with the nitrogen in the Gln108 side-chain residue, with the side-chain hydroxyl group of Tyr233, and interacts via water molecules with the side chains of Glu196, Lys213 and Asp216 (Figure 1A). Finally, the galactose ring in Galβ1,3GalNAc-α-PhNO$_2$ interacts with Tyr194 via stacking interactions.

Within the N-acetylgalactosamine moiety of Galβ1,3GalNAc-α-PhNO$_2$ the axial C4-hydroxyl forms hydrogen bond interactions with the hydroxyl group in the Tyr269 side chain, with the amide nitrogen in the Gln108 side chain, and via an ordered water molecule with the guanidinium side chain of Arg268, and the backbone carbonyl in Leu107. The C6-hydroxyl forms stacking interactions with Tyr233 and establishes water-mediated hydrogen bonds. Finally, the crystal structure shows water-assisted hydrogen bond interactions between the N-acetyl group and Arg268 (Figure 1A). The preference of ST3Gal-I toward Galβ1,3GalNAc over related acceptors suggests other possible interactions of the N-acetamide group in the GalNAc moiety with amino acids in the distorted lid domain.

One of the most important amino side chains in providing acceptor affinity and specificity would appear to be Tyr269 (Figure 1). Its interactions with both the C-4 axial hydroxyls likely align the acceptor for attack and provide the basis for discrimination of Galβ1,3GalNAc- over Galβ1,3GlcNAc-containing acceptors, since GlcNAc has an equatorial hydroxyl at this position. Indeed Tyr269 is fully conserved in ST3Gal-I and -II enzymes, both of which predominantly use Galβ1,3GalNAc as acceptor, while other ST3Gal enzymes that prefer acceptors with Glc or GlcNAc contain different amino acids in place of the tyrosine. Tyr269 also forms a
bifurcated hydrogen-bonding interaction, as donor, with the endocyclic oxygen of Gal and the glycosidic oxygen. Indeed, mutation of Tyr269 to Phe resulted in a very large (80-fold) decrease of the $k_{cat}/K_M$ value for the acceptor, corresponding to a 2.7 kcal mol$^{-1}$ loss in transition-state binding affinity (Table I). Such a decrease is completely consistent with loss of 2–3 hydrogen bonds between neutral partners (Fersht et al. 1985; Street et al. 1986), as would be predicted from Scheme 1. Most of this decrease is reflected in the $K_M$ value indicating similar loss of interactions in the ground and transition states. A somewhat smaller deleterious effect (25-fold) was seen for transfer to Galβ1,6GalNAc-α-OBn by this mutant (Table II), consistent with the different binding mode anticipated. The modeled structure indicates that the GalNAc moiety flips in the active site such that the axial C4 hydroxyl orients away from Tyr269, with the former hydroxide bond being replaced by an interaction with an ordered water molecule. Since this latter interaction is likely to be of less consequence, a smaller decrease in $k_{cat}/K_M$ upon mutation is reasonable.

Tyr233 is another fully conserved residue in ST3Gal-I and -II and forms hydrogen bonds with the 6-hydroxyl of the Gal moiety as well as stacking interactions with the GalNAc. Mutation of Tyr233 to Ala should remove both classes of interaction and indeed kinetic analysis reveals a 32-fold decrease in $k_{cat}/K_M$, corresponding to a 2.1 kcal mol$^{-1}$ decrease in transition-state stabilization (Table I). Most of this loss is seen in the 12-fold increase in $K_M$ indicating similar interactions formed at the ground and transition states with their net contribution being entirely consistent with a neutral hydrogen bond plus stacking interactions. Interestingly, the effects of this mutation on transfer to the Galβ1,6Gal-α-OBn acceptor are somewhat larger and result in a 115-fold decrease in $k_{cat}/K_M$, corresponding to a 2.9 kcal mol$^{-1}$ increase in activation free energy (Table II). Interactions with the 6-hydroxyl of the galactosyl residue should be maintained; as shown in the model. Presumably either improved stacking interactions, or additional binding interactions not detected by modeling account for this small additional contribution to transition-state stabilization for this acceptor.

A third residue that appears to be particularly important for acceptor binding is Gin108, which also interacts with the axial C4 hydroxyl group of GalNAc and the C6 hydroxyl of Gal (Scheme 1). Accordingly, mutation to Ala yields a mutant with a 60-fold lower $k_{cat}/K_M$ for the acceptor, corresponding to a 2.5 kcal mol$^{-1}$ decrease in transition-state affinity, with much of the loss being seen in the ground state (from $K_M$ values) (Table I). The similarity of this effect to that seen for Tyr269 is striking and again consistent with two neutral hydrogen bond interactions. The effects of this mutation on transfer to the Galβ1,6GlcNAc substrate are very similar.

The missing loop and its interactions

By analogy with Cst-II, it seems probable that structuring of the missing region of polypeptide from residues 305–316 in ST3Gal-I upon binding of CMP-Neu5Ac stabilizes binding of the donor. This structuring helps exclude hydrolytic water, and form part of the acceptor binding site. Individual mutations of the residues in the loop Phe313, Arg314 and Lys315 to Ala in each case created mutants of reduced, but not extinguished, catalytic activity (Table I). Values of $k_{cat}$ for donor and acceptor were affected very little in each case, indeed being increased in one case (Lys315 Ala), while $K_M$ values for donor and acceptor both increased by $\sim$2.5-fold (Table I). These modest, individual effects indicate contributing, but not essential roles in catalysis, with particular importance for the binding of both the donor and acceptor. Further, it is tempting to ascribe the 1.5-fold increase in $k_{cat}$ for Lys315 Ala to increased mobility of the loop in this mutant, accelerating on and off rates. However, this conclusion will require future dynamics studies of this or related enzymes.

Conclusions

Kinetic analyses of active site mutants of mammalian ST3Gal-I, in conjunction with structural studies, have confirmed the roles of a number of the residues in catalysis and quantitated their net contributions to transition-state stabilization. A number of the interactions formed with the donor sugar proved to be essential, since their removal abrogated activity. Further, the mutations in the acceptor site generally had relatively little effect on $k_{cat}$, but instead increased $K_M$ and thus reduced $k_{cat}/K_M$. This finding requires that the removed interactions play an approximately equivalent role in stabilizing the binding of the ground state and of the transition state. Such behavior is consistent with studies of glycan binding sites on other oligosaccharide-processing enzymes, which typically contribute to catalysis by stabilizing ground and transition states equally. This observation reflects the difficulties of creating relatively remote interactions that preferentially stabilize transition states over ground states (Wicki, Schloegl et al. 2007; Wicki, Williams et al. 2007). A consequence of this is that any new interactions provided in the acceptor site that improve binding will themselves improve $k_{cat}/K_M$ values and thus specificity, as was dramatically seen for the evolved Phe91 Tyr mutant of Cst-2 (Aharoni et al. 2006).

These studies have clarified and quantified the active site interactions found within the important, yet under-studied GT29 family of mammalian STs. These data provide a basis for the selective engineering of mammalian sialytransferases to modify their specificities and thereby permit the use of tailored substrates as chemical biological probes of sialyltransferase function in vivo.

Materials and methods

General

Disaccharides Galβ1,3GalNAc-α-OBn and Galβ1,6GalNAc-α-OBn were obtained from Toronto Research Chemicals. Pyruvate kinase/lactate dehydrogenase (PK/LDH), NADH and phosphoenolpyruvate were obtained from Sigma-Aldrich. NMPK was obtained from Roche (Germany). Ni-NTA/His tag columns for protein purification were obtained from GE Healthcare (Montreal, Canada). DNA purification kits were obtained from Qiagen (Mississauga, Ontario, Canada).

Cloning, expression, and purification of pST3Gal-I Δ45 wild type and mutants

A truncated pST3Gal-I Δ45 (45 residue N-terminal deletion) was cloned into pET-28 vector (Novagen) with an N-terminal His tag and a maltose-binding tag and subsequently
transformed into chemically competent *E. coli* cells [origami 2(DE3), Novagen (San Diego, California, USA)] for expression. Site-directed mutagenesis was performed either with Pfu hotstart polymerase or with Phusion DNA Polymerase (New England Biolabs, Pickering, Ontario, Canada) using pST3Gal-I A45 as a template. When using Phusion DNA Polymerase, the thermocycling conditions consisted of a hot start at 98°C followed by 17 cycles of 10 s denaturation at 98°C, 30 s annealing at 60°C and 8 min extension at 72°C. When using Pfu hotstart, the polymerase thermocycler program was: 95°C for 2 min, followed by 18 cycles of 30 s denaturation at 95°C, 1 min annealing at 55°C and 12 min extension at 68°C.

Polymerase chain reaction (PCR) products were purified using the PCR purification kit, followed by incubation with DpnI (FastDigest) for 40 min to remove the template. Following the second PCR kit purification, PCR products were transformed into DH5α chemically competent cells. The mutants were confirmed by DNA sequencing (Genewiz, South Plainfield, NJ, USA).

The following primers (Integrated DNA Technologies) were used to generate the specific mutants: ST3Gal-I H302A-F (5′-cgattcaaaaaggaaactggcatgcctattgggaaaataatcc-3′) and ST3Gal-I H302A-R (5′-cggtattttctcaaatgcagcctgtcctctttgaatgcg-3′) were used to generate the H319A mutant. ST3Gal-I Q108A-F (5′-gctttcgaaatctccatcagcaactccagttttacgaaatgcgc-3′) and ST3Gal-I Q108A-R (5′-ggcaagtttcgcgaaaggattcaggaaatacaaaatggtgggttg-3′) were used to generate the K315A mutant. ST3Gal-I Y194F-F (5′-ggcaggtacaggaacggcggtgggaaatgc-3′) and ST3Gal-I Y194F-R (5′-ggattattttcccaataggcatgccagtttccttttgaatcg-3′) were used to generate the F313A mutant. ST3Gal-I Y233A-F (5′-gctttcgaaatctccatcagcaactccagttttacgaaatgcgc-3′) and ST3Gal-I Y233A-R (5′-ggcaagtttcgcgaaaggattcaggaaatacaaaatggtgggttg-3′) were used to generate the W300A mutant. ST3Gal-I Y269F-F (5′-cagctggctccgcctggcccgAAAAACAACCC-3) and ST3Gal-I Y269F-R (5′-ggattattttcccaataggcatgccagtttccttttgaatcg-3′) were used to generate the Y269F mutant. ST3Gal-I Y269F-F (5′-gctttcgaaatctccatcagcaactccagttttacgaaatgcgc-3′) and ST3Gal-I Y269F-R (5′-ggcaagtttcgcgaaaggattcaggaaatacaaaatggtgggttg-3′) were used to generate the H319A mutant. ST3Gal-I Q108A-F (5′-GTGGCTCCGCTTGCCCGTGAACCAACCC-3′) and ST3Gal-I Q108A-R (5′-gctttcgaaatctccatcagcaactccagttttacgaaatgcgc-3′) were used to generate the Q108A mutant. ST3Gal-I Y233A-F (5′-gctttcgaaatctccatcagcaactccagttttacgaaatgcgc-3′) and ST3Gal-I Y233A-R (5′-ggcaagtttcgcgaaaggattcaggaaatacaaaatggtgggttg-3′) were used to generate the F313A mutant. ST3Gal-I Y194F-F (5′-ggcaggtacaggaacggcggtgggaaatgc-3′) and ST3Gal-I Y194F-R (5′-ggcaagtttcgcgaaaggattcaggaaatacaaaatggtgggttg-3′) were used to generate the Y269F mutant.

ST3Gal-I H302A-F (5′-ccatctacgaaAGGTgggatggaaaacggccgtgacc-3′) and ST3Gal-I H302A-R (5′-ccattttgtatttcctgaatcctttcgcgaacttgcc-3′) were used to generate the H319A mutant. ST3Gal-I Q108A-F (5′-ccatcatgaactccagttgcacgaaatgcgc-3′) and ST3Gal-I Q108A-R (5′-ccatcatgaactccagttgcacgaaatgcgc-3′) were used to generate the Q108A mutant. ST3Gal-I Y233A-F (5′-ccatcatgaactccagttgcacgaaatgcgc-3′) and ST3Gal-I Y233A-R (5′-ccatcatgaactccagttgcacgaaatgcgc-3′) were used to generate the F313A mutant. ST3Gal-I K315A-F (5′-ccatcatgaactccagttgcacgaaatgcgc-3′) and ST3Gal-I K315A-R (5′-ccatcatgaactccagttgcacgaaatgcgc-3′) were used to generate the K315A mutant. ST3Gal-I Y194F-F (5′-ccatcatgaactccagttgcacgaaatgcgc-3′) and ST3Gal-I Y194F-R (5′-ccatcatgaactccagttgcacgaaatgcgc-3′) were used to generate the Y269F mutant.

The lysozyme was centrifuged at 4°C and 15,000 rpm for 30 min and the resulting supernatant was filtered through a 5 mm filter and loaded onto a Ni-NTA column equilibrated in a buffer (50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 50 mM imidazole, supplemented with a Roche ethylenediaminetetraacetic acid-free protease pill, benzonase nuclease, and lysozyme). Cells were lysed using a French press. The lysate was centrifuged at 37°C and 15,000 rpm for 30 min and the resulting supernatant was filtered through a 5 mm filter and loaded onto a Ni-NTA column equilibrated in a buffer (50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 50 mM imidazole, supplemented with a Roche ethylenediaminetetraacetic acid-free protease pill, benzonase nuclease, and lysozyme). Cells were lysed using a French press. The lysate was centrifuged at 37°C and 15,000 rpm for 30 min and the resulting supernatant was filtered through a 5 mm filter and loaded onto a Ni-NTA column equilibrated in a buffer (50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 50 mM imidazole, supplemented with a Roche ethylenediaminetetraacetic acid-free protease pill, benzonase nuclease, and lysozyme). Cells were lysed using a French press. The lysate was centrifuged at 37°C and 15,000 rpm for 30 min and the resulting supernatant was filtered through a 5 mm filter and loaded onto a Ni-NTA column equilibrated in a buffer (50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 50 mM imidazole, supplemented with a Roche ethylenediaminetetraacetic acid-free protease pill, benzonase nuclease, and lysozyme). Cells were lysed using a French press. The lysate was centrifuged at 37°C and 15,000 rpm for 30 min and the resulting supernatant was filtered through a 5 mm filter and loaded onto a Ni-NTA column equilibrated in a buffer (50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 50 mM imidazole, supplemented with a Roche ethylenediaminetetraacetic acid-free protease pill, benzonase nuclease, and lysozyme). Cells were lysed using a French press. The lysate was centrifuged at 37°C and 15,000 rpm for 30 min and the resulting supernatant was filtered through a 5 mm filter and loaded onto a Ni-NTA column equilibrated in a buffer (50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 50 mM imidazole, supplemented with a Roche ethylenediaminetetraacetic acid-free protease pill, benzonase nuclease, and lysozyme). Cells were lysed using a French press. The lysate was centrifuged at 37°C and 15,000 rpm for 30 min and the resulting supernatant was filtered through a 5 mm filter and loaded onto a Ni-NTA column equilibrated in a buffer (50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 50 mM imidazole, supplemented with a Roche ethylenediaminetetraacetic acid-free protease pill, benzonase nuclease, and lysozyme). Cells were lysed using a French press. The lysate was centrifuged at 37°C and 15,000 rpm for 30 min and the resulting supernatant was filtered through a 5 mm filter and loaded onto a Ni-NTA column equilibrated in a buffer (50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 50 mM imidazole, supplemented with a Roche ethylenediaminetetraacetic acid-free protease pill, benzonase nuclease, and lysozyme). Cells were lysed using a French press.
Michaelis–Menten equation using nonlinear regression analysis in GraFit version 5.0.13 (Erithacus). \( \Delta \Delta G^\# \) values were calculated for both the donor and the acceptor substrates for each mutation using the following equation:

\[
\Delta \Delta G^\# = RT \log \left( \frac{k_{cat}/K_M} {k_{cat}/K_M}^{WT} \right) \frac{k_{cat}/K_M} {k_{cat}/K_M}^{mutant}
\]

**CD spectroscopy**

CD spectra were measured on a JASCO-J-815 spectropolarimeter, varying the wavelength from 190 to 280 nm at 200 nm min\(^{-1}\). The measurements were made at room temperature using a 1 mm path length quartz cell. Protein concentrations were \( ~1.5 \mu M \) in 5 mM Tris–HCl (pH 7.5) buffer. The CD spectra were recorded with a response of 1 s at a 1 nm bandwidth. The buffer spectrum was subtracted from the protein spectra and subsequently smoothed.

**Modeling studies**

The 3D (energy minimized) model of Gal\(1\),6GalNAc was generated using PRODRG and was superimposed onto the Gal\(1\),3GlcNAc model using least-squares fitting (COOT) of the Galactose moiety (Schüttelkopf and van Aalten 2004; Emsley et al. 2010).

**Funding**

This work was supported by the Canadian Institutes of Health Research. S.G.W. is supported by a Tier 1 Canada Research Chair

**Acknowledgements**

Bojana Rakic thanks the Natural Sciences and Engineering Research Council of Canada for a postdoctoral fellowship. Francesco Rao thanks EMBO and the Michael Smith Foundation for Health Research for a postdoctoral fellowship.

**Conflict of interest**

None declared.

**Abbreviations**

ATP, adenosine triphosphate; CAZy, Carbohydrate-Active enZymes; CD, circular dichroism; CMP-Neu5Ac, cytosine monophosphate N-acetyl \( \beta \) neuraminic acid; GT, glycosyltransferase; His, hexahistidine; L, large; MBP, maltose binding protein; NADH, nicotinamide adenine dinucleotide; NMPK, nucleoside monophosphate kinase; PCR, polymerase chain reaction; pST3Gal-I, porcine ST3Gal-I; STs, sialyltransferases.

**References**


