Protein engineering of α2,3/2,6-sialyltransferase to improve the yield and productivity of in vitro sialyllactose synthesis

Yun Hee Choi2, Jong Hoon Kim2, Joon Ho Park3, Nahum Lee3, Dae-Hee Kim5, Kyoung-Soon Jang2, IL-Hyang Park2, and Byung-Gee Kim1,2,3,4

1Interdisciplinary Program for Biochemical Engineering and Biotechnology, Institute of Molecular Biology and Genetics; 2School of Chemical and Biological Engineering, and 3Institute of Bioengineering, Seoul National University, 151-742, Seoul, South Korea; and 4GeneChem Inc., Daejeon, 302-804, South Korea

Received on May 6, 2013; revised on October 10, 2013; accepted on October 13, 2013

In the large-quantity production of α2,3- and α2,6-sialyllactose (Neu5Ac(α2,3)Galβ1,4Glc (3′-SL) and Neu5Ac(α2,6)Galβ1,4Glc (6′-SL)) using sialyltransferases (STs), there are major hurdles to overcome for further improvement in yield and productivity of the enzyme reactions. Specifically, Pasteurella multocida α2,3-sialyltransferase (α2,3PST) forms a by-product to a certain extent, owing to its multifunctional activity at pH below 7.0, and Photobacterium damselae α2,6-sialyltransferase (α2,6PdST) shows relatively low ST activity. In this study, α2,3PST and α2,6PdST were successfully engineered using a hybrid approach that combines rational design with site-saturation mutagenesis. Narrowly focused on the substrate-binding pocket of the STs, putative functional residues were selected by multiple sequence alignment and alanine scanning, and subsequently subjected to site-saturation mutagenesis. In the case of α2,3PST, R313N single mutation improved its activity slightly (by a factor of 1.5), and further improvement was obtained by making the double mutants (R313N/T265S and R313H/T265S) resulting in an overall 2-fold improvement in its specific α2,3 ST activity, which is mainly caused by the increase in $k_{cat}$. It was revealed that the R313 mutations to N, D, Y, H or T greatly reduced the α2,6 ST side-reaction activity of α2,3PST at below pH 7.0. In the case of α2,6PdST, single-mutation L433S/T and double-mutation 1411T/L433T exhibited 3- and 5-fold enhancement of the α2,6 ST-specific activity compared with the wild-type, respectively, via increase in $k_{cat}$ values. Our results show a very good model system for enhancing ST activity and demonstrate that the generated mutants could be used efficiently for the mass production of 3′-SL and 6′-SL with enhanced productivity and yield.

Keywords: hybrid approach / protein engineering / sialyllactose / sialyltransferase

Introduction

Sialyloligosaccharides are known to have various important functions in human health and disease. Particularly, sialyllactose (SL) moieties are found in gangliosides in neuron and brain tissues, which are a component of cell plasma membrane that modulates cell signal transduction events such as brain development and memory formation (Antoine et al. 2003; Wang 2009). Sialyl Lewis X/A molecules containing sialic acid moiety are well-known carbohydrate ligand molecules recognized by selectins expressed on endothelial cells, and their bindings result in inflammation to recruit leucocytes by helping the adhesion of the leucocytes to the inner wall of a blood vessel (Lowe 2003). SL is known as one of the major components of human milk oligosaccharides (HMO) and can also function as a neutralizing reagent of toxins produced by enteric bacteria (Idota et al. 1995; Sinclair et al. 2008). They are also known as potent inhibitors of bacterial or viral adhesion to the epithelial surface in the initial stages of the infection process (Kunz et al. 2000; Lehmans et al. 2006). 3′-SL and 6′-SL containing Neu5Ac(α2,3)Gal and Neu5Ac(α2,6)Gal sequences are differentially recognized by avian and human influenza viruses, respectively (Lehmans et al. 2006; McCullough et al., 2012; Hideshima et al. 2013), through interaction with viral hemagglutinins. Synthetic SLs can be used for various applications, such as functional food ingredients, nutraceutical medicine and virus adsorbent materials, so that their large-scale enzymatic production is in great demand (Endo et al. 2000; Priem et al. 2002; Drouillard et al. 2010).

α2,3/α2,6-sialyllactose (3′-SL/6′-SL) can be synthesized by sialyltransferases (STs) (EC 2.4.99) which transfer N-acetylneuraminic acid (Neu5Ac) from cytidine monophosphate N-acetylneuraminic acid (CMP-Neu5Ac) to a lactose acceptor substrate, the galactose moiety of lactose more specifically (Scheme 1). Previously, we successfully accomplished an enzymatic process for CMP-Neu5Ac synthesis using a five-enzyme system (GlcNAc-2-epimerase, acetate kinase, CMP kinase, Neu5Ac aldolase and CMP-Neu5Ac synthetase), which could efficiently and cheaply supply the donor substrate CMP-Neu5Ac for ST reaction (Lee et al. 2002). As the next step, the large-scale synthesis of 3′-SL and 6′-SL has been studied using a combination of the enzymatic synthesis of CMP-Neu5Ac and ST reaction. For the STs, α2,3ST from...
By three times (Banas et al. 2006). Wu et al. also focused on the interface of the domains of a monooxygenase, identifying amino acid residues as functional sites and conducting saturation mutagenesis (Wu et al. 2010).

YH Choi et al.

**Photobacterium damselae** (α2,3PST) (Yu et al. 2005) and α2,6ST from *Photobacterium damselae* (α2,6PdST) (Sun et al. 2008; Cheng et al. 2010), which belong to glycosyltransferase (GT) family 80 (Yamamoto et al. 2008), were selected. Each enzyme has its own advantages and disadvantages for the enzymatic synthesis of 3′-SL and 6′-SL using the STs. In general, α2,3PST has shown approximately six times higher specificity than α2,6PdST, but it has pH-dependent multifunctional substrate regio-specificity, so that 6′-SL is produced as a by-product at below pH 7.0 (Yu et al. 2005). Since the optimum pH of the α2,3PST reaction is 8.5, the operation of the α2,3PST reaction under this condition can greatly reduce the generation of 6′-SL by-product (Sugiarto et al. 2011; Yu et al. 2005). However, since the optimum pH of multistep enzyme reactions for CMP-Neu5Ac synthesis is 7.0, combining the two steps of the synthesis of CMP-Neu5Ac and ST reaction at around pH 7.0 is more desirable to make α2,3PST with reduced α2,6ST activity. In the case of α2,6PdST, its natural low activity hampers the large-scale mass production of 6′-SL (Sun et al. 2008), but no additional by-products caused by poor regio-selectivity are observed, unlike with α2,3PST.

To meet such demands in the changes in ST properties, a hybrid approach combining alanine scanning based on a protein structure model with site-saturated mutagenesis was attempted. The hybrid approach called “semi-rational design” was developed to generate a smaller but higher-quality library to investigate a rather narrowly focused region by utilizing protein sequences, structure information and computational modeling (Chica et al. 2005; Lutz 2010). Banas et al. carried out a small-library site-saturation mutagenesis near the substrate access tunnels in haloalkane dehalogenase, showing that the mutations restricting solvent water accessibility to the hydrophobic substrate-protein complex can increase its $k_{cat}/K_m$ value by three times (Banas et al. 2006). Wu et al. also focused on the interface of the flavin adenine dinucleotide (FAD) and nicotinamide adenine dinucleotide phosphate (NADP)-binding domains of a monooxygenase, identifying two amino acid residues as functional sites and conducted saturation mutagenesis (Wu et al. 2010).

We focused on the residues contacting with CMP-Neu5Ac and lactose within a distance of 5 Å from the boundary surface of the substrate molecules in the substrate-binding pocket of each ST, and selected functional residues based on multiple sequence alignments and alanine scanning. Considering the conserved and functional residues deduced from multiple sequence alignment analysis (Yeom et al. 2008; Prabhu et al. 2010) and the neutral drift of some mutations, saturation mutagenesis was performed on the selected functional residues. Using this hybrid approach, herein we demonstrated that the ST activities of α2,3PST and α2,6PdST were increased 2-fold and 5-fold, respectively, and in the case of α2,3PST, the 6′-SL by-product synthesis activity was greatly reduced at low pH (<pH 7.0). In addition, detailed characterizations of the screened mutants were performed, and the mutation effects on the activity of the STs were investigated at molecular level based on a computer model. This study is the first report thus far showing that ST activity toward CMP-Neu5Ac and lactose can be improved by protein engineering for the synthesis of 3′-SL and 6′-SL. The mutants of α2,3PST and α2,6PdST can be efficiently applied to one-pot reactions for the large-scale synthesis of 3′-SL and 6′-SL.

**Results**

**Selection of target region and alanine scanning**

Prior to the selection of target residues for alanine scanning, the 2IHZ (PDB ID, α2,3PST-CMP-3F(α)Neu5Ac-lactose) crystal structure of α2,3PST (Ni et al. 2007) was used to identify substrate-binding pocket regions. For the homology modeling of α2,6PdST, 2Z4T was used as a template, whose structural data were acquired with a protein-ligand complex of CMP and lactose (Kakuta et al. 2008). When the model structure of α2,6PdST was superimposed onto the 2IHZ structure of α2,3PST, its structure was very similar to that of 2IHZ due to the 58% amino acid similarity. Both α2,3PST and α2,6PdST have GT-B-fold structures consisting of two separate Rossmann-fold domains. According to the 2IHZ α2,3PST crystal structure, the catalytic key residue that functions as a
general base to deprotonate from the reactive hydroxyl group of lactose acceptor was identified as Asp141 (Ni et al. 2007; Kim et al. 2008), which corresponded to Asp215 of α2,6PdST. The active site of the α2,3PST and α2,6PdST is located in the deep cleft between the two Rossmann domains. In the case of α2,3PST, upon the binding of CMP-Neu5Ac to the active site, its closed conformation (i.e. active form) is induced by the movement of α12α/α12β helices in the C-terminal domain, which is known to help define the subsequent lactose binding site (Ni et al. 2006).

According to the 2IHZ α2,3PST crystal structure, 76 amino acid residues were identified within a sphere of 12 Å radius from the Cα of Asp141. Among them, when the two substrate-binding pocket regions were examined in detail following the catalytic mechanism, 10 and 30 residues were identified to interact with the two substrates within 5 Å distance from the boundary surface of lactose (specifically, 5 Å distance from the oxygen atoms of hydroxyl groups) and CMP-Neu5Ac (specifically, 5 Å distance from the oxygen/nitrogen atoms of hydroxyl groups/amino groups), respectively. In the case of the α2,6PdST homology model, 78 residues were identified within a sphere of 12 Å radius from the Cα of Asp215, and similarly, 10 and 22 residues were identified in the binding sites of the lactose and CMP-Neu5Ac, respectively.

To minimize the experiments for alanine scanning, conserved sequences among STs were analyzed and eliminated. Multiple sequence alignments for α2,3PST and α2,6PdST were done with the sequences of the GT family 80 retrieved by Position-Specific Iterated-Basic local alignment search tool (PSI-BLAST), such as Pasteurella multocida gi 90109232, Pasteurella dagmantis gi 260914426, Photobacterium leiognathi gi 267844791, Photobacterium damselae gi 269102115, Photobacterium phosphoreum gi 148537333, Photobacterium sp. JT-ISH-224 gi 159147903, Vibrioaceae photobacterium gi 178847432, Vibrio sp. JT-FAJ-16 gi 158455298, Haemophilus ducreyi gi 33151326 and Shewanella piezotolerans gi 212637736 (Yamamoto et al. 2008). A total of 27 conserved amino acid residues including catalytic key residues such as Asp141, H311, S355 and S356 in α2,3PST (Asp215, H387, S431 and S432 in α2,6PdST) were found (Supplementary data, Figure S1). Considering that the conserved residues in a protein structure have high possibilities of showing functional or structural importance for its catalytic activity, they were were eliminated in the alanine scanning experiments. For example, H311, S355 and S356 in α2,3PST are conserved for the stabilization of CMP-Neu5Ac by holding the phosphate oxygen with hydrogen bonds (Ni et al. 2007; Kim et al. 2008). Similarly, the corresponding H387, S431 and S432 in α2,6PdST were identified as conserved amino acid residues that have the same function. Among the 10 and 30 residues interacting with the lactose and CMP-Neu5Ac in the binding pocket of α2,3PST, 5 (P34, N85, H112, M144 and R313 for lactose) and 12 (L40, S143, V147, T265, T267, T268, T269, P312, I335, S336, L357 and Y388 for CMP-Neu5Ac) were nonconserved residues were selected for alanine scanning, respectively. Similarly, 5 (H106, K135, R136, N157 and W347 for lactose) and 6 (S108, L112, T344, I411, S412 and L433 for CMP-Neu5Ac) for α2,6PdST were selected for alanine scanning.

### Table I. Relative activities and relative specific activities of alanine-substituted mutant enzymes for α2,3PST/α2,6PdST

<table>
<thead>
<tr>
<th>(A)</th>
<th>Relative activity (%)</th>
<th>(B)</th>
<th>Relative specific activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>α2,3PST mutants</td>
<td>WT</td>
<td>100</td>
<td>H106A</td>
</tr>
<tr>
<td></td>
<td>P54A</td>
<td>61</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>L40A</td>
<td>77</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>N85A</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td></td>
<td>H112A</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>S143A</td>
<td>39</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>M144A</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>V147A</td>
<td>49</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>T265A</td>
<td>48</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>T267A</td>
<td>26</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>T268A</td>
<td>102</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>T269A</td>
<td>48</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>P312A</td>
<td>48</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>R313A</td>
<td>107</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>I335A</td>
<td>43</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>S336A</td>
<td>17</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>L357A</td>
<td>23</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Y388A</td>
<td>30</td>
<td>27</td>
</tr>
</tbody>
</table>

Relative activity *(a)* for alanine-substituted mutants was determined by U/mL, and relative specific activity *(b)* was defined as (U/mL of mutant/level of soluble protein expression of mutant relative to the WT)/(U/mL of WT/Relative level of soluble protein expression of WT). The mean values are shown from duplicate experiments, n.d.: not detected.

### Material and methods

Alanine scanning for all the selected residues was performed to determine the functional residues for site-saturation mutagenesis as the next step. The relative ST activity was compared using the pH-based colorimetric method mentioned in the “Material and methods” section. Since the pKα of cresol red and phenol red are 8.32 and 7.5, respectively, cresol red was better suited for α2,3PST assay, whose optimal pH was 8.5, and phenol red for α2,6PdST assay, whose optimal pH was 8.0. The relative activities (U/mL) and relative specific activities (i.e. relative activity/level of soluble protein expression relative to the wild-type, (Supplementary data, Table S1 and Supplementary data, Figure S2)) for Ala-substituted α2,3PST mutants are given in Table 1A.

The soluble enzyme expression levels of the most α2,3PST mutants were about the same as that of wild-type α2,3PST, except for N85A and L357A mutants, which showed <30% of the wild-type levels (Supplementary data, Table S1 and Supplementary data, Figure S2). The N85A mutant appeared to produce improperly folded enzymes, and lost most of the α2,3 ST activity. In addition, despite having a reasonable expression level, H112A almost lost its activity, yielding 6% of the activity of the wild-type enzyme. Such remarkable losses in α2,3 ST activity suggest that N85 is an important key residue for proper protein folding and substrate binding, and H112 is a key residue involved in substrate binding and/or significantly modulating the catalytic efficiency of the enzyme. According to the structure of the α2,3PST-CMP-3F(α)Neu5Ac-lactose ternary complex, N85 appears to make water-mediated hydrogen (H)-bonding interactions with the O1’ and O5’ of the galactose moiety of the lactose, and Nε2 of His112 is likely to
generate H-bonds with O3’ and O4’ of the galactose moiety. Thus, N85 and H112 were eliminated in the following mutation study. The L357 site was also excluded from the candidates of the study owing to a 29% of low soluble protein expression level compared with the wild-type despite its relative specific activity of 78%. In the result, only the mutation sites yielding proper soluble expression with high relative specific activities were selected. Among the remaining mutation sites, the sites showing over 60% of the relative specific activity compared with the wild-type were chosen from the perspective of neutral drift, with P34, T265, T268 and R313 remaining for the sites for saturation mutagenesis.

Likewise, among the alanine-scanned α2,6PdST mutants, H106A, S108A and W347A mutants showed <30% soluble expression levels compared with the wild-type α2,6PdST, and H106A and S108A showed no activities (Table IB). The H106 residue is situated within an H-bonding distance from O4’ and O5’ of the galactose moiety of the lactose, and appears to contribute to the binding of the lactose. The relative specific activities of L112A, N157A and T344A mutants were <50%, even though their soluble protein expression levels were quite comparable with or higher than that of the wild-type α2,6PdST. Among them, the N157 appears to make an H-bond with the O1 of the glucose moiety of the lactose, which corresponds to N85 in α2,3PST interacting with the galactose moiety of the lactose. As a result, I411 and L433 were chosen for the sites for site-saturation mutagenesis, since I411A and L433A mutants showed 173 and 95% of its relative specific activities, respectively, without changing their soluble protein expression levels.

Saturation mutagenesis for functional residues
P34, T265, T268 and R313 residues in α2,3PST were individually subjected to saturation mutagenesis, and the mutants showing high activities were screened at pH 8.5 using the same colorimetric assay. In the case of R313 located on the loop in the binding site of the glucose moiety of lactose, R313A and R313G replaced with small amino acids appeared to show essentially the same ST activity as the wild-type, but the mutants replaced with hydrophilic residues like Ser, Thr, Tyr, Asp, His and Asn showed at least 1.5-times higher activities than the wild-type (Supplementary data, Table SII). The relative specific activities of the R313T, R313Y, R313D, R313H and R313N mutants were 129, 125, 108, 146 and 231% compared with the wild-type, respectively, indicating that the R313N mutant showed the highest specific activity (Table IIA). In the case of T265 located on the loop at a 4.7 Å distance from the N atom of NH2 of cytidine in CMP-Neu5Ac, the T265G, T265S and T265N mutants were selected (Supplementary data, Table SII), and their relative specific activities were 94, 116 and 126%, respectively (Table IIA). Despite a slightly low specific activity of T265G, the mutant showed the highest overall activity (i.e., 168%) due to its higher expression level. On the other hand, no mutant at P34 and T268 sites showed higher specific activities than its wild-type (below 90%), even though the protein expression levels of the mutants were similar to that of the wild-type (data not shown).

In the case of α2,6PdST, I411 and L433 were individually subjected to saturation mutagenesis. According to the modeling structure of α2,6PdST, I411 is at an H-bonding distance of 3.4 Å from the O2 of the cytosine ring, and L433 is located in a deep cleft region for the CMP binding site by making an H-bond with an O3’ atom of the ribose ring of CMP. Among the L433 variants, L433S and L433T mutants showed ~3-fold improvement in the specific activities (Table IIIB). Notably, among the I411 variants screened, the I411T mutant showed not only 2-fold improvement in the specific activity, but also an enhanced expression level of its soluble protein compared with the wild-type α2,6PdST, which was also confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (data not shown).

Characterization of α2,3PST mutants
To understand the underlying mechanisms of the individual mutations of α2,3PST to the changes in specific activities, the kinetic parameters of the single mutants and double mutants were further evaluated using the purified enzymes (Table IIIA and Supplementary data, Figure S3). R313N and R313H single mutations enhanced $k_{cat}$ values for CMP-Neu5Ac (~2-fold increase) and lactose (~40–70% increase), whereas the mutations slightly decreased their substrate-binding affinities for CMP-Neu5Ac (~40–80% $K_m$ value increase) and lactose (~10% $K_m$ increase). In the results, the $k_{cat}/K_m$ values for CMP-Neu5Ac of the R313N and R313H mutations were increased by ~32 and 20%, and those for lactose were enhanced by 55 and 33%, respectively. In the case of T265S mutation, the $k_{cat}/K_m$ value was increased by ~30% for CMP-Neu5Ac due to the increase in the $k_{cat}$ value, but no significant changes were shown in $k_{cat}$ and $K_m$ values for lactose. The kinetic parameters of T265N single mutant were not determined, because the mutation showed relatively lower activity than T265S when combined with R313H/Y/N mutations in the double mutants as shown in Supplementary data, Table SIII.

To identify any synergistic combinations of the single mutations, a total of seven double mutants were constructed by site-
directed mutagenesis based on the high specific activities of R313N, R313H, R313Y and R313T, as well as T265N and T265S variants (the R313T/T265N double mutant was not constructed, but its activity would be lower than that of the R313T/T265S mutant according to the data shown in Supplementary data, Table SIII). The protein expression levels of all the double mutants were quite similar to that of the wild-type, and notably, only the relative activities of R313H/T265S and R313N/T265S were higher than those of the other double mutants (Supplementary data, Table SIII). The specific activities of R313H/T265S and R313N/T265S mutants were 237 and 216%, respectively, compared with that of the wild-type (Table II). In general, both double mutants had significantly improved $k_{cat}$ values for both CMP-Neu5Ac ($\approx3$–3.5-fold increase) and lactose ($\approx50$% increase). However, their $K_m$ values for CMP-Neu5Ac and lactose were somewhat increased and decreased, respectively, indicating the reduced binding affinity for CMP-Neu5Ac donor, but the enhanced binding affinity for lactose acceptor. Overall, by varying the concentration of CMP-Neu5Ac or lactose, the $k_{cat}/K_m$ values of both the R313N/T265S and R313H/T265S mutants were identified as increased by a factor of 1.7 and 1.5 for CMP-Neu5Ac, whereas 1.7- and 2-fold increases were observed for lactose compared with that of the wild-type, respectively.

At the same time, the 6'-SL synthesis activities of α2,3PST for R313 single and double mutants were compared with that of the wild-type at below pH 7.0 (Figure 1). Interestingly, the R313 single mutants (R313N, R313T, R313Y, R313H and R313D) and the double mutants R313N/T265S and R313H/T265S showed very low α2,6 activity at pH 6.0 and pH 6.5, respectively, although wild-type α2,3PST was known to produce 3–4% 6'-SL among the overall products. When the reactions were conducted with the purified enzymes (2.25 μg, 0.05 U) for 30 min at pH 6.0, the production yields of 6'-SL of the R313 single and double mutants decreased by a factor of 4–30 compared with that of the wild-type. Notably, all the R313 single and double mutants did not produce 6'-SL at all at pH 6.5, except for R313Y. The result demonstrates that R313 is a key amino acid residue for reducing the α2,6 ST side-reaction activity of α2,3PST.

**Characterization of α2,6PdST mutants**

The kinetic parameters of the single mutants I411T, L433S and L433T of α2,6PdST were determined, as given in Table IIIIB. All the single mutants showed significantly enhanced $k_{cat}$ values for CMP-Neu5Ac ($\approx2.5$–4.7-fold increase) and lactose ($\approx5.9$–20-fold increase), whereas the mutations decreased their substrate-binding affinities for CMP-Neu5Ac ($\approx3$–40% increase).
increase in $K_m$ value) and lactose ($\sim$3.6–8-fold increase in $K_m$ value). However, unlike others, only the L433T mutation increased the substrate-binding affinity for CMP-Neu5Ac $\sim$2.5-fold. In the results, the $k_{cat}/K_m$ values for CMP-Neu5Ac of the I411T, L433S and L433T mutants were increased 2.4, 3.5, and 6.4-fold, and those for lactose were also enhanced 1.6-, 2.5- and 2.2-fold, respectively. To identify any synergistic combinations of the single mutations, two double mutants, I411T/L433S and I411T/L433T, were constructed. Both of the double mutants had increased $k_{cat}$ values for CMP-Neu5Ac 3.9–4.4-fold, whereas the CMP-Neu5Ac-binding affinity of the I411T/L433S mutation decreased 1.8-fold, but the I411T/L433T mutation had increased CMP-Neu5Ac binding affinity by $\sim$57%. In all, I411T/L433S and I411T/L433T had increased $k_{cat}/K_m$ values for CMP-Neu5Ac 2- and 7.5-fold, respectively. Surprisingly, the $k_{cat}$ values for lactose of the I411T/L433S and I411T/L433T mutations were 24- and 26-fold higher than those of the wild-type, while their $K_m$ values for lactose increased 9.9- and 8.1-fold, respectively. The $k_{cat}/K_m$ values for lactose of I411T/L433S and I411T/L433T increased 2.3- and 3.1-fold relative to those of the wild-type. The I411T/L433T double mutation displayed further improvement in activity than single mutations mainly by increasing the $k_{cat}$ values, which resulted in an increase ($\sim$20–40%) of the $k_{cat}/K_m$ values for both substrates compared with its single mutant L433T.

Discussion

We have attempted to engineer two STs, $\alpha$2,3PST and $\alpha$2,6PdST, using a hybrid approach to increase catalytic activity for the production of 3′-SL and 6′-SL. Functional residues contributing to the catalytic activity and the expression level of soluble protein of the two STs were selected by combining computer modeling of the ST structures and alanine scanning experiments, after excluding the conserved sequences extracted from multiple alignments. Through the alanine scanning, the mutations displaying complete loss of ST activity or low ST activity were eliminated, and the functional candidate residues for site-saturation mutagenesis were identified.

In the case of $\alpha$2,3PST, only the R313, which is present in the loop region making an H-bond with glucose in the 2IHZ structure ($\alpha$2,3PST-CMP-3F(a)Neu5Ac-lactose) (Ni et al. 2007), was able to generate interesting and highly active mutants. Its replacements with hydrophobic residues such as T/Y/D/H/N rather than hydrophilic or small amino acid residues displayed somewhat higher ST activity. The 2IV structure ($\alpha$2,3PST-CM-lactose ternary structure) superimposed with 2IHK ($\alpha$2,3PST-CMP-3F(e)Neu5Ac binary structure) gives useful information about Neu5Ac and lactose coordinates for the $\alpha$2,3 ST reaction, such as indicating a very close positioning of the galactose O3 of the lactose to the anomeric carbon (C2) of sialic acid (Figure 2A) (Ni et al. 2007). According to the structure, N12 and N12 of R63 located at 3 Å from the carboxylic acid of sialic acid helps to anchor the Neu5Ac during its distortion along the reaction coordinate. While the N12 of R313 is 4.9 Å away from the N12 of R63, R313N mutation decreases the distance (corresponding to the distance from N82 of N313 to N12 of R63) to 2.63 Å. In the R313H mutant, N2 of H313 could also make H-bonds with the N12 atoms of R63. This calculation suggests that a more stable structure of Neu5Ac is formed by R313N and R313H mutations by generating more reactive conformation of $\alpha$2,3PST through anchoring the Neu5Ac in close proximity to the lactose. In addition, the reoriented hydroxyl group generated by T265S mutation is likely to result in H-bonding (2.87 Å) to the N4 atom of the cytosine ring of CMP, whereas the distance between O2 of T265S and CMP is 4.87 Å in the 2IV structure of closed conformation (Figure 2A). This interaction might help to stabilize the CMP in more active conformation, and trigger the movement of the loop near Neu5Ac like a movement through interaction between G266 and N4 of CMP (Ni et al. 2007). The movement could help to define the acceptor binding pocket by W270 located on the loop. This addtive T265S mutation to R313N/H significantly increased the $k_{cat}$ values for CMP-Neu5Ac, and slightly increased the binding affinity of lactose (Table IIIA).

In terms of the $\alpha$2,6 ST side reaction activities of $\alpha$2,3PST, all the R313 single variants (N, T, Y, H and D) and double variants (R313N/T265S and R313H/T265S) showed significantly low levels of $\alpha$2,6 ST activities compared with that of the wild-type (Figure 1). Notably, none of the mutants showed $\alpha$2,6 ST side-reaction activity at pH 6.5, except for R313Y. The 2IH structure ($\alpha$2,3PST-CMP-3F(a)Neu5Ac-lactose ternary structure) indicates alternative lactose-binding conformation for $\alpha$2,6 ST activity at low pH (Ni et al. 2007). In the structure, R313 can make an H-bond to the O6′ of galactose with a distance of 2.84 Å (Figure 2B), suggesting that the H-bonding would help the O6′ atom of the galactose to situate toward the C2 of sialic acid. When the R313N/T/Y/H/D variants were compared with the wild-type $\alpha$2,3PST, the distances from O6′ of the galactose to the replaced amino acids were much longer, and the longest distance (6.4 Å) was found with R313T, which shows the lowest $\alpha$2,6ST activity. R313 residue also appears to bind the O2 and O3 of the glucose moiety of the lactose in the 2IH structure, leading the O6′ of the galactose moiety of lactose to come closer to the anomeric carbon (C2) of Neu5Ac (Figure 2B). In the case of other R313 mutants, the distance between the replaced residue at the 313 site and glucose becomes much longer, suggesting that R313 is the key determinant residue to generate the side reaction of $\alpha$2,6 ST activity. When $\alpha$2,3PST and $\alpha$2,6PdST were superimposed, their overall structures were quite similar, with minor conformational differences, including the lactose orientation. The R313 in $\alpha$2,3PST appeared to correspond to the R136 in $\alpha$2,6PdST. In the results on $\alpha$2,6PdST, the relative specific activity of the R136A mutant was only 30% that of its wild-type (Table IB), indicating that the residue might contribute to the orientation of the lactose acceptor required for its $\alpha$2,6 ST activity. Interestingly, the R313Y mutant was recently shown to decrease the sialidase activity of $\alpha$2,3PST at pH 5.5 (Sugianto et al. 2011). Additionally, we have measured the $\alpha$2,3 sialidase activities of R313 single and of double mutants for 3′-SL with the excess amount of the purified enzymes (45 μg, which is 100-fold higher concentration of the enzyme used in the $\alpha$2,3 ST reaction) at pH 5.5 (50 mM MES buffer). Under this condition, $\alpha$2,3 sialidase activities of all the mutants decreased compared with that of the wild-type. For examples, R313N, R313H, R313Y, R313D, R313T, R313N + T265S and R313H + T265S mutants produced 13.1, 7.8, 3.3, 3.6, 13.5, 1.4 and 1.9% of...
sialic acid from 1 mM of 3′SL in 1 h, respectively, whereas 19.6% of sialic acid was produced by the wild-type. This result indicates that α2,3 sialidase activity of the mutants decreased significantly, and 7–69% of their original sialidase activities remained intact in the mutants. In the case of R313N + T265S mutant, 93% of its α2,3 sialidase activity decreased, suggesting that the decrease in the sialidase activity would be one of the major causes of high ST specific activity (2-fold, shown at Table II) of the mutant. However, the measurement of the reduced α2,3 sialidase activities was performed at a different reaction condition compared with α2,3 ST reaction. The enzyme concentration was 100-fold higher and the reaction time was six times longer than those of α2,3 ST activity measurement, reaction was proceeded with ~0.45 μg of the purified enzyme in 5 mM Tris–HCl buffer at pH 8.5 for 10 min. All together, the α2,3 sialidase activity of R313N + T265S mutant was ~600 times lower than its α2,3 ST activity. Therefore, the effect of the decrease in α2,3 sialidase activity of the mutants on the increase in α2,3 ST activity was negligible as we have expected. Nevertheless, our experimental results showed that R313 residue is also somewhat involved in controlling the α2,3 sialidase activity of wild-type α2,3ST, suggesting that R313 is a key residue to control the multifunction of α2,3ST.

All of the high-performing R313 mutants of α2,3ST, such as R313N, R313N/T265S and R313H/T265S, showed no changes in their optimum pH of the α2,3 ST activity, displaying 50–60% relative activity at pH 7.0 compared with those at pH 8.5. The same fold changes in their α2,3ST activities at pH 7.0 remained in all the mutants, resulting in 2-fold higher activity for the synthesis of 3′SL than that of wild-type α2,3PST without 6′SL by-product formation.

In the case of α2,6PdST, the I411 located on a bent loop toward CMP is within H-bonding distance (3.4 Å) of the N4 atom of the cytosine ring of CMP (Figure 2C). When I411 was replaced by Thr, a hydrophilic amino acid, the H-bond distance between the carbonyl oxygen of T411 and the cytosine ring was not changed, but a possible H-bond was newly generated between P408 and T411. Therefore, I411T mutation may reduce the flexibility of the bent loop containing P408 and T411 by H-bonding, and bring A409/K410 onto the rigid loop and cytosine ring closer, which allows less room for

![Protein engineering of α2,3/2,6-sialyltransferase](https://academic.oup.com/glycob/article-abstract/24/2/159/1988423)
CMP-Neu5Ac to move around in the active site, and perhaps pushes CMP-Neu5Ac toward the lactose, resulting in a 2-fold increase in the $k_{cat}/K_m$ of the enzyme.

L433 of $\alpha_2,6$PDST is located on a short alpha-helix very close to the CMP-Neu5Ac binding site in the deep cleft of the active site and generates an H-bond (3.77 Å) between the main chain of L433 and O3 of CMP. When L433 was replaced with S or T, the distance of the H-bond was not changed much, but the enzyme surface for the CMP binding site was changed and the substrate access channel became more open (Figure 2D). When the L433 residue and CMP-Neu5Ac were represented using a sphere form, structural collisions occurred between L433 and the cytidine ring, whereas L433S or L433T mutation showed no collision with CMP-Neu5Ac, and the enlarged space for the CMP binding would allow easy access of CMP-Neu5Ac, resulting in an increased $k_{cat}$ value ($\sim$2.5–4.7-fold) for CMP-Neu5Ac. Therefore, the double mutation I411T/L433T showed an additive effect by enhancing the $k_{cat}/K_m$ value 7.5-fold relative to the wild-type for CMP-Neu5Ac.

According to the current model, however, both I411 and L433 are situated inside the substrate-binding pocket to make more direct contact with CMP-Neu5Ac rather than lactose, so that the two residues appear not to affect the direct lactose binding. Therefore, in the case of lactose, it is quite hard to explain the mechanisms using the current model of the 25-fold increase in the $k_{cat}$ value for I411T/L433S and I411T/L433T, and the dramatic increases in $K_m$ ($\sim$3.5-fold and 8-fold in I411T and L433S/T mutants, respectively).

In summary, this study provides insight into the protein engineering strategy to improve the activity of STs belonging to GT-B-fold enzymes, which can be utilized for the production of 3'-SL and 6'-SL. Our approach is a hybrid approach that combines rational design based on computer modeling, alanine scanning and saturation mutagenesis, and it can generally be used with any other proteins. The results have demonstrated that the catalytic activity of ST can be efficiently improved by the mutagenesis mainly focused on its active site, on the surrounding regions of the binding sites of the two substrates, i.e., glycan donor (CMP-Neu5Ac) and glycan acceptor (lactose). The alanine scanning and saturation mutation could generate dramatic changes in the ST expression level as a soluble form, and in its specific activities. Although most of them are still quite speculative and unknown, the detailed mechanisms to make such changes in protein solubility and activity could be explained to a certain extent by computer modeling and measurement of the kinetic parameters. In the case of $\alpha_2,3$PST/ $\alpha_2,6$PDST, our results suggested that the introduction of a few H-bonds might be able to move the anomeric carbon of the CMP-Neu5Ac donor toward the hydroxyl group of the acceptor lactose, and/or to stabilize the binding of the two substrates in the active site, and that some mutations to reduce the steric hindrance to the access of substrates are key to the improvement of their catalytic activities. In addition, since the 2.6 ST side-reaction activity of $\alpha_2,3$PST appears to be caused by erroneous positioning of the two substrates in the active sites (especially the lactose acceptor), its abolishment was quite effective with even only one site-directed mutation, such as the mutation at R313, yielding quite fascinating results. Another interesting observation in the mutation study was that one or two amino acid changes such as L433S and I411T/L433T of $\alpha_2,6$PDST could make dramatic changes in the $k_{cat}$ value for lactose.

For further studies, one-pot reaction combining multiple enzyme steps for the large-scale synthesis of CMP-Neu5Ac and the SLs can be carried out with enhanced yield and productivity without generating the by-product using the improved ST mutants. Furthermore, the costs of the synthesis of 3'-SL and 6'-SL as well as their purification process originating from by-product formation (6'-SL in the case of 3'-SL synthesis) would be greatly reduced, and the mutants could be widely used in the synthesis of various other 3'- and 6'-sialyloligosaccharides.

Materials and methods

Chemicals and preparation of STs

CMP-Neu5Ac sodium salt, 3'-SL and 6'-SL were obtained from GeneChem, Inc. (Daejeon, Korea). All of the other chemicals were purchased from Sigma-Aldrich, Inc. (St. Louis, MO) and Junsei Chemical Co. Ltd. (Tokyo, Japan). $\alpha_2,3$PST truncated with 24 amino acids at the N-terminus (Yu et al. 2005) was cloned in pET23a, and $\alpha_2,6$PDST truncated with 15 amino acids at the N-terminus (Sun et al. 2008) was cloned in both pET15b and pET28a. The truncated $\alpha_2,3$PST and $\alpha_2,6$PDST were used as wild-type. Each plasmid was transformed into Escherichia coli BW25113 (DE3), and the cloned ST was induced with 0.5 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG), when OD$_{600}$ of the culture reached 0.5–0.8. The induced cells were further incubated in a rotary shaker at 200 rpm and 30°C for 5 h in the case of $\alpha_2,3$PST and at 18°C for 18 h in the case of $\alpha_2,6$PDST. The cells were harvested from 50 mL of the culture broth by centrifugation at 4000 rpm for 10 min at 4°C, suspended in 5 mL of 10 mM potassium phosphate buffer (pH 7.0) containing 2 mM ethylenediaminetetraacetic acid (EDTA), 1 mM phenylmethanesulfonyl fluoride and 0.01% 2-mercaptoethanol, and lysed by sonication on ice. Cell lysate (soluble fraction) was obtained by centrifugation at 15,000 rpm for 30 min at 4°C.

Sialyltransferase assay

ST activity assays were performed in a total volume of 100 μL of 50 mM Tris–HCl buffer at pH 8.5 for $\alpha_2,3$PST (Yu et al. 2005) and at pH 8.0 (Yamamoto et al. 2007) for $\alpha_2,6$PDST. The reaction mixtures contained 10 mM CMP-Neu5Ac, 5 mM lactose and appropriate amounts of a specified enzyme. The reactions took place for 1 h for $\alpha_2,3$PST and 2 h for $\alpha_2,6$PDST at room temperature. The reactions were stopped by heating at 95°C for 40 s, after which the reaction mixture was centrifuged, and its supernatant was analyzed quantitatively using a Bio-LC DX-300 (Dionex Co., Sunnyvale, CA). The supernatant of the reaction mixture was diluted with distilled water, and injected into a Bio-LC DX-300 high-performance anion-exchange chromatography with a pulsed amperometric detector (Dionex Co., Sunnyvale, CA). To separate the lactose, 3'-SL and 6'-SL in the reaction mixture using a CarboPac PA-100 column (Dionex Co., Sunnyvale, CA), 100 mM NaOH solution was used at a flow rate of 1 mL/min. ST activities were also quantitatively determined by a colorimetric method using cresol red.
(pK_a = 8.32) for α2,3PST (Park et al. 2009) and phenol red (pK_a = 7.5) for α2,6PdST as pH indicators. The method is based upon the color and absorbance changes of the pH indicators according to the proton release during the glycosidic bond formation of the products. The ST activity assays were conducted in a 96-well plate with 100 μL of 5 mM Tris–HCl buffer solution (pH 8.5 for α2,3PST, and pH 8.0 for α2,6PdST) containing 4 mM CMP-Neu5Ac, 4 mM lactose, 0.2 mM pH indicator and a specified enzyme. The increase in absorbance at 405 nm for cresol red and the decrease in absorbance at 560 nm for phenol red were recorded for each sample at 30 s or 1-min intervals for 5–10 min using a spectrophotometer (Spectrostar nano, BMG LABTECH, Germany). The reactions were started by adding the substrate mixtures to the enzyme reaction mixtures, and then the initial reaction rate was calculated by measuring the changes in absorbance in the reaction mixture. The control for each sample containing the same reaction mixture without lactose substrate was used, and the difference in optical density (ΔOD) was calculated between the control and reaction samples. The released proton concentration during the reaction was calculated based on an HCl calibration curve. To generate a calibration curve of protons, 0–1.2 mM HCl were mixed into the 100 μL of 5 mM Tris-HCl buffer with the same reaction condition containing either 4 mM CMP-Neu5Ac or 4 mM lactose, with 0.2 mM pH indicator and appropriate amounts of enzyme (10% v/v). The enzyme activity was defined as the amount of enzyme required to produce 1 μmol of proton released during glycosidic bond formation per minute at room temperature. Correlation coefficient between the colorimetric and Bio-LC assays was 0.95 (Supplementary data, Figure S4).

Homology modeling and sequence alignment

The three-dimensional (3D) homology modeling for α2,6PdST was done with the MODELLER program (Sali and Blundell 1993) using 2Z4T (protein data bank (PDB) accession code of α2,6-ST from Photobacterium sp. JT-ISH-224) as a template sequence (Kakuta et al. 2008), which has 76% protein sequence similarity with α2,6PdST. The alignment between α2,6PdST and 2Z4T was done using ClustalW1.8. The residues within 5 Å distance from the boundary surface of the substrate molecules in the binding pocket were selected using Swiss PDB Viewer v.3.7. PSI-BLAST (Altschul et al. 1997) searches for α2,3PST and α2,6PdST were done with E-value better than the threshold (E-value < 0.005), and then the called protein sequences were retrieved from the National Center for Biotechnology Information database. The multiple sequence alignments of α2,3PST and α2,6PdST were performed using clustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/). The alignment was examined using SeaView (ver. 4.3.0) to identify the conserved and nonconserved sequences.

Alanine scanning and saturation mutagenesis

For alanine scanning, all the nonconserved amino acid residues in the substrate-binding pocket of the target STs were individually mutated to alanine using GCG codon. Soluble fractions of the enzymes were obtained after cell sonication, and then ST activities were measured with the cell-extracts. The levels of the soluble protein expression for alanine-substituted enzymes relative to the wild-type were calculated by the image analysis of the soluble fraction bands on SDS–PAGE using a GeneTools analysis software equipped with Genetools (Syngene, UK). Saturation mutagenesis was carried out to generate various single mutants for the selected residues from alanine scanning. The α2,3PST/pET23a and α2,6PdST/pET15b plasmids were used as DNA templates, and N = G or A or T or C, K = G or T nucleotide (NNK) codon was used for saturation mutagenesis. The polymerase chain reaction (PCR) for ST gene amplification was conducted with an initial denaturation at 95°C for 5 min, followed by 15 cycles of 95°C for 30 s, 55°C for 1 min, 68°C for 12 min and 68°C for 12 min for a final extension. The PCR product was subsequently treated with DpnI for 1 h at 37°C. 10 μL of the PCR product was transformed into E. coli DH5α, and the harvested plasmids from the cells were transformed into E. coli BW25113 (DE3).

Mutants screening

In order to screen all the single amino acid-exchange mutants, a sufficient number of colonies (>200 colonies at one site) were examined to satisfy the condition P_f > 0.99, where P_f = 1 − (1 − F_i)^j, where F_i represents the frequency at which sequence i is present in the NNK library (i.e. since NNK contains 32 codons, F_i becomes 1/32), and P_f is the probability such that the sequence i is detected among the T transformants generated from the library (where T is the number of colonies examined). Individual colonies were inoculated into 96-deep-well microplates, with each well containing 500 μL of Luria broth (LB) medium supplemented with 100 μg mL⁻¹ ampicillin. After sealing the plates with Axyseal sealing film (Axygen Scientific, Inc., Union City, CA), cells were grown in a shaking incubator at 500 rpm and 37°C for 18 h. Twenty-five microliters of each culture was transferred to a fresh deep-well plate containing 500 μL of LB medium supplemented with 100 μg/mL ampicillin and 0.5 mM IPTG. The freshly inoculated plates were incubated for 18 h at 30°C for α2,3PST and 40 h at 18°C for α2,6PdST. The cells were harvested by centrifugation at 3390 × g for 10 min at 4°C. In the case of α2,3PST, cell pellets were thoroughly resuspended in 100 μL of 5 mM Tris–HCl buffer (pH 8.5), and then 10 μL of the whole cells were used for screen mutants. On the other hand, in the case of α2,6PdST, to break out the harvested cells, the cell pellets were thoroughly resuspended with 50 μL of BugBuster protein extraction reagent (Novagen, San Diego, CA), followed by incubation for 20 min at room temperature. The cell debris was subsequently removed by centrifugation. Positive variants were identified from 96-well plates based on their colors and corresponding absorbance changes. For the colorimetric assay, 10 μL of whole cells or clear supernatants were transferred to a new microplate containing 90 μL of substrate solution (5 mM Tris–HCl, pH 5.5 or pH 8.0, 4 mM lactose, 4 mM CMP-Neu5Ac and 0.2 mM pH indicator), and then the absorbance changes were measured at several points in time. The positive mutants were cultured once again in a 250-mL shake flask (working volume 50 mL), and the initial ST activity (U/mL) of the mutants was confirmed by the same colorimetric assay.

Purification and kinetic assays

Wild-type α2,3PST, α2,6PdST, and the screened mutants expressed in E. coli BW25113 (DE3) cells were harvested and...
purified by nickel nitrilotriacetic acid column (QIAGEN, Inc., Valencia, CA). First, a protein mixture was loaded onto a column equilibrated with a 50 mM Tris–HCl buffer (pH 8.0) containing 5 mM imidazole and 300 mM NaCl. The ST fraction was eluted from the column by applying a 50–250 mM imidazole buffer containing 300 mM NaCl. The active fractions were then collected and desalted by VIVASPIN (Sartorius Stedim Biotech GmbH, Germany). The protein concentration was determined using a Bradford protein assay kit (Bio-Rad Laboratories, Hercules, CA). Kinetic analysis of ST variants for acceptor and donor substrates was carried out using the same colorimetric method. The kinetic parameters of $k_{\text{cat}}$ and $K_m$ were obtained by nonlinear regression analysis of the Michaelis–Menten equation using SigmaPlot 10.0.

### Supplementary data

Supplementary data for this article is available online at [http://glycob.oxfordjournals.org/](http://glycob.oxfordjournals.org/).

### Funding

This research was supported by the Converging Research Center Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2013K000234), and the WCU (World Class University) Program through the NRF grant funded by the Ministry of Education, Science and Technology (R322012000102130).

### Conflict of interest

None declared.

### Abbreviations

CMP-Neu5Ac, cytidine monophosphate N-acetylneuraminic acid; EDTA, ethylenediaminetetraacetic acid; GT, glycosyltransferase; HMO, human milk oligosaccharide; IPTG, isopropyl-1-thio-β-D-galactopyranoside; LB, Luria broth; LC, liquid chromatography; Neu5Ac, N-acetylneuraminic acid; NNK, N = G or A or T or C, K = G or T nucleotide; PCR, polymerase chain reaction; PDB, protein data bank; PSI-BLAST, Position-Specific Iterated-Local alignment search tool; SDS–PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; SL, sialylactose; STs, sialyltransferases; α2,3PST, Pasteurella multocida α2,3-sialyltransferase; α2,6PdST, Photobacterium damselae α2,6-sialyltransferase.

### References


Sun M, Li Y, Chokhawala HA, Henning R, Chen X. 2008. N-Terminal 112 amino acid residues are not required for the sialyltransferase activity of...


