Crystal Structures of *Pasteurella multocida* Sialyltransferase Complexes with Acceptor and Donor Analogues Reveal Substrate Binding Sites and Catalytic Mechanism[†,‡]

Lisheng Ni,[§] Harshal A. Chokhawala,[§] Hongzhi Cao,[§] Ryan Henning,[§] Laura Ng,[§] Shengshu Huang,[§] Hai Yu,[§] Xi Chen,*[,§] and Andrew J. Fisher*[,§,||]

Department of Chemistry and Section of Molecular and Cellular Biology, University of California, One Shields Avenue, Davis, California 95616

Received February 19, 2007; Revised Manuscript Received March 27, 2007

ABSTRACT: Sialyltransferases are key enzymes involved in the biosynthesis of biologically and pathologically important sialic acid-containing molecules in nature. Binary X-ray crystal structures of a multifunctional *Pasteurella multocida* sialyltransferase (Δ24PmST1) with a donor analogue CMP-3F(*a*)Neu5Ac or CMP-3F(*e*)Neu5Ac were determined at 2.0 and 1.9 Å resolutions, respectively. Ternary X-ray structures of the protein in complex with CMP or a donor analogue CMP-3F(*a*)Neu5Ac and an acceptor lactose have been determined at 2.0 and 2.27 Å resolutions, respectively. This represents the first sialyltransferase structure and the first GT-B-type glycosyltransferase structure that is bound to both a donor analogue and an acceptor simultaneously. The four structures presented here reveal that binding of the nucleotide-activated donor sugar causes a buried tryptophan to flip out of the protein core to interact with the donor sugar and helps define the acceptor sugar binding site. Additionally, key amino acid residues involved in the catalysis have been identified. Structural and kinetic data support a direct displacement mechanism involving an oxocarbenium ion-like transition state assisted with Asp141 serving as a general base to activate the acceptor hydroxyl group.

Terminal sialic acid (*N*-acetylneuraminic acid or Neu5Ac)[1] residues on the carbohydrate moieties of higher animal glycoconjugates are directly involved in many biologically and pathologically important processes such as molecular recognition, cell−cell interaction, bacterial and viral infection, and tumor metastasis (*1*, *2*). Sialic acids have also been found in many pathogenic bacteria such as *Campylobacter jejuni*, *Neisseria meningitidis*, *Neisseria gonorrheae*, *Escherichia coli* K1, group B *Streptococci*, *Haemophilus ducreyi*, and *Haemophilus influenzae* as components of their lipooligosaccharides (LOS) or capsular polysaccharides, and their presence is often associated with virulence (*1*−*9*). The importance of sialic acids has made sialyltransferases, the key enzymes involved in the biosynthesis of sialic acid-containing structures, potential therapeutic targets. It is, therefore, crucial to understand the catalytic mechanism of sialyltransferases.

Δ24PmST1 (*Pasteurella multocida* sialyltransferase 1 with an N-terminal 24 amino acid residue deletion) is a multifunctional sialyltransferase from *P. multocida* strain P-1059. The major function of the enzyme is an α2,3-sialyltransferase that catalyzes the transfer of a sialic acid residue from its activated sugar nucleotide donor, cytidine 5′-monophosphate sialic acid (CMP-Neu5Ac), to its acceptor, a structure terminated with a galactose residue (*10*). The crystal structures of the enzyme have been determined previously in native form and in complex with CMP (*11*). The overall structure of Δ24PmST1 belongs to the glycosyltransferase B (GT-B) fold and consists of two separate Rossmann domains with a deep nucleotide-binding cleft between the two domains (*11*). It is, thus, different from the GT-A-like fold of the only other known sialyltransferase CstII structure (*12*, *13*). In order to determine the key amino acid residues that interact with the Neu5Ac in the CMP-Neu5Ac donor and the carbohydrate acceptor, CMP-3FNeu5Ac compounds

[†] This work was supported by start-up funds to X.C. from the Regents of the University of California. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.

[‡] Protein coordinates have been deposited in the Protein Data Bank (http://www.rcsb.org/pdb/): accession codes 2IHJ and 2IHK for binary structures of Δ24PmST1−CMP-3F(*a*)Neu5Ac and Δ24PmST1−CMP-3F(*e*)Neu5Ac, respectively; accession codes 2IHZ and 2ILV for ternary structures of Δ24PmST1−CMP-3F(*a*)Neu5Ac−lactose and Δ24PmST1−CMP−lactose, respectively.

* Corresponding authors. A.J.F.: tel, (530) 754-6180; fax, (530) 752-8995; e-mail, fisher@chem.ucdavis.edu. X.C.: tel, (530) 754-6037; fax, (530) 752-8995; e-mail, chen@chem.ucdavis.edu.

[§] Department of Chemistry, University of California.

[||] Section of Molecular and Cellular Biology, University of California.

[1] Abbreviations: ASU, asymmetric unit; CMP, cytidine 5′-monophosphate; CMP-3F(*a*)Neu5Ac, cytidine 5′-monophosphate 3-(*axial*)-fluoro-*N*-acetylneuraminic acid; CMP-3F(*e*)Neu5Ac, cytidine 5′-monophosphate 3-(*equatorial*)fluoro-*N*-acetylneuraminic acid; GT, glycosyltransferase; LacMU, 4-methylumbelliferyl-β-D-lactoside; ManNAc, *N*-acetylmannosamine; MME, monomethyl ether; Neu5Ac, *N*-acetylneuraminic acid; rmsd, root mean square deviation; SSRL, Stanford Synchrotron Radiation Laboratory.

10.1021/bi700346w CCC: $37.00
Published on Web 05/08/2007

(inert CMP-Neu5Ac donor analogues with a fluorine atom in either the axial or equatorial position at C-3 of the Neu5Ac) were synthesized and used in the cocrystallization studies. Binary structures of the protein in the presence of CMP-3F(*a*)Neu5Ac or CMP-3F(*e*)Neu5Ac have been obtained. In addition, ternary structures of the protein in the presence of CMP-3F(*a*)Neu5Ac or CMP and lactose have also been determined. These structures identified substrate binding pockets in Δ24PmST1 for both the donor and acceptor and the key amino acid residues that are involved in catalysis. Site-directed mutagenesis, kinetics, and crystallographic studies support a direct displacement mechanism involving an oxocarbenium ion-like transition state as predicted for inverting glycosyltransferases (*14*−*16*). Structural and kinetic data point to Asp141 as the general base that deprotonates the hydroxyl group (3′-OH and possibly 6′-OH) in the terminal galactose residue of the acceptor lactose, to activate it as a nucleophile for the direct displacement mechanism. The current study reported here affords a significant understanding of the mechanism of action for inverting GT-B-type glycosyltransferases. The structural and mechanistic knowledge gained from this study will further aid in the discovery of potential therapeutics that work through the inhibition of sialyltransferases.

EXPERIMENTAL PROCEDURES

Synthesis of CMP-Neu5Ac, CMP-3F(a)Neu5Ac, and CMP-3F(e)Neu5Ac. CMP-Neu5Ac was synthesized from *N*-acetylmannosamine (ManNAc), sodium pyruvate, and CTP using a recombinant sialic acid aldolase and a recombinant *N. meningitidis* CMP-sialic acid synthetase (NmCSS) as reported (*17*). Both 3F(*e*)Neu5Ac and 3F(*a*)Neu5Ac were formed from ManNAc and 3-fluoropyruvate (5.0 equiv) by a sialic acid aldolase (*17*) catalyzed reaction. They were separated by silica gel chromatography and used in the NmCSS (*17*) catalyzed synthesis of CMP-3F(*e*)Neu5Ac and CMP-3F(*a*)Neu5Ac, respectively. The detailed synthesis and the characterization of the compounds will be reported separately.

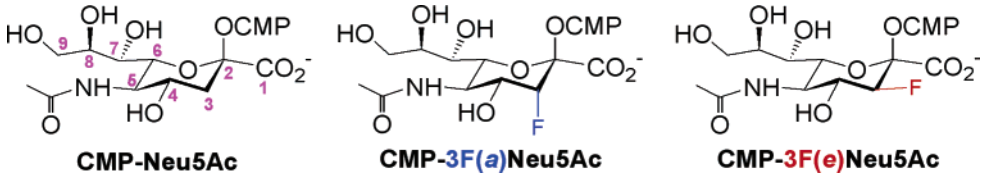
Crystallization. Native sialyltransferase from *P. multocida* is likely anchored into the membrane by a predicted membrane-spanning domain consisting of the N-terminal 24-residue stretch of basic and hydrophobic residues (*11*). A soluble construct was made by deleting the first 24 amino acid residues and replacing Ser25 with the Met start codon (Δ24PmST1), which was expressed as described previously (*11*). Δ24PmST1 in 20 mM Tris buffer (pH = 7.5) was concentrated to 10 mg mL^{-1}, and CMP-3F(*a*)Neu5Ac or CMP-3F(*e*)Neu5Ac was added to a final concentration of 2 mM. Binary CMP-3FNeu5Ac crystals were grown by hanging drop with 3 μL of the sample mixed with an equal volume of reservoir buffer [30% polyethylene glycol monomethyl ether (PEG MME) 2000, 200 mM $MgCl_2$, 0.4% Triton X-100]. Ternary crystals with both CMP-3FNeu5Ac and lactose bound were obtained by soaking binary crystals in buffer containing 33% PEG MME 2000, 50 mM $MgCl_2$, 0.4% Triton X-100, and 10 mM lactose for 1 h. All crystals were transferred to Paratone-N and frozen in a steam of nitrogen to 100 K for data collection (*18*).

Data Collection, Model Building, and Refinement. X-ray diffraction data were collected on a variety of Stanford Synchrotron Radiation Laboratory (SSRL) beam lines and an in-house rotating anode (Table 2). The SSRL collected data were processed with MOSFLM and scaled with SCALA (*19*). Data collected in-house on the rotating anode were collected on the Proteum CCD (Bruker-AXS, Madison, WI) and processed with SAINT and scaled with PROSCALE (Bruker-AXS, Madison, WI). All structures were solved by molecular replacement using the previously solved structure of Δ24PmST1 with CMP bound (PDB ID 2EX1) employing the program AMoRe (*20*). The atomic model building was carried out with the molecular graphics program COOT (*21*). All structures were refined with the program REFMAC (*22*, *23*) using 95% of the measured data as a target function. The final *R*-factor and *R*-free along with the quality of the models based on PROCHECK (*24*) are listed in Table 2. Only Ala219 falls in the disallowed region in the Ramachandran plot for all four structures presented here, as well as the two previously published structures (*11*). The electron density map clearly defines the same conformation in all models ($\varphi \approx 45°$, $\psi \approx 120°$). The reason for the unfavorable main-chain angles is that Ala219 falls in a tight turn between α6b and β7, which then proceeds into the C-terminal domain.

Site-Directed Mutagenesis. Site-directed mutagenesis was carried out using the commercially available QuikChange multi-site-directed mutagenesis kit from Statagene according to manufacturer's direction. The primers used were PmST1_H112A_R (5′CGAATTAATTGAACGGAAGCAGCAATATTTAAG3′), PmST1_D141A_R (5′CCATTGAGCCAGCGTCATAAAG3′), PmST1_D141N_R (5′CCATTGAGCCATTGTCATAAAG3′), PmST1_W270A_F (5′CCGGCACGACAACTGCGGAAGGAAATACC3′), PmST1_W270F_F (5′CCGGCACGACAACTTTTGAAGGAAATACC3′), and PmST1_H311A_F (5′CTACTTTAAAGGGGCTCCTAGAGGTGGTG3′). The double mutant H112A/D141A was generated using the plasmid of the H112A mutant as a template and the mutation primer of D141A as a primer in the polymerase chain reaction for mutagenesis. The expression and purification of mutant proteins were carried out as previously reported for the wild-type enzyme (*10*, *11*).

Sialyltransferase Kinetic Assays. Unless specified, all kinetic assays were carried out in triplicate in a total volume of 50 μL in Tris-HCl buffer (100 mM, pH 8.5) containing CMP-Neu5Ac, LacMU, and Δ24PmST1 or its mutant. Reactions were allowed to proceed for 10 min at 37 °C before they were quenched by adding ice-cold 12% acetonitrile (150 μL). The 0.5 mL microcentrifuge tubes containing assay mixtures were kept on ice for 10 min before being centrifuged for 5 min at 13000 rpm to remove precipitated protein. The formation of product was measured by injecting 10 μL of the supernatant into a Shimadzu LC-2010A HPLC system equipped with a membrane on-line degasser, a temperature control unit (maintained at 30 °C throughout the experiment), and a RF-10A fluorescence detector. A reverse-phase Premier C18 column (250 × 4.6 mm i.d., 5 μm particle size) protected with a C18 guard column cartridge was used. The mobile phase was 12% acetonitrile aqueous solution with a flow rate of 1 $mL \cdot min^{-1}$. The fluorescent compounds LacMU and Neu5Acα2,3LacMU were detected by excitation at 325 nm and emission at 372 nm (*10*, *25*). The system was controlled by Shimadzu EZStart v7.2 SP1 chromatography software. Apparent kinetic parameters for wild-type Δ24PmST1 and its mutants were determined by varying the CMP-Neu5Ac

Scheme 1: Structures of CMP-Neu5Ac, CMP-3F(*a*)Neu5Ac, and CMP-3F(*e*)Neu5Ac



concentration from 0.1 to 6.0 mM and a fixed concentration of LacMU (1 mM) or by a fixed concentration of CMP-Neu5Ac (1 mM) and varying the LacMU concentration from 0.5 to 7.0 mM. Grafit 5.0 was used to calculate kinetic parameters by fitting the initial velocity data into the Michaelis−Menten equation. Under assay conditions, no spontaneous hydrolysis of CMP-Neu5Ac was observed.

CMP-3F(a)Neu5Ac Inhibition Assays. Inhibition assays were carried out similarly to that described for the kinetic assays. The typical condition was 37 °C in Tris-HCl buffer (100 mM, pH 8.5) containing CMP-Neu5Ac (0.1−5.0 mM), LacMU (1 mM), CMP-3F(*a*)Neu5Ac inhibitor (10 μM and 50 μM), and Δ24PmST1 enzyme (41 ng). The type of inhibition and the K_i value were obtained using Grafit 5.0 by fitting the initial velocity data into respective equations.

RESULTS

Overall Structures. Here we present four crystal structures of Δ24PmST1 with donor analogues bound in the presence or the absence of lactose acceptor. Previously, our laboratories determined the crystal structures of Δ24PmST1 in the absence of any substrates and in the presence of CMP (*11*). The CMP structure was actually generated by cocrystallizing Δ24PmST1 with its natural sugar nucleotide donor CMP-Neu5Ac. However, only the CMP moiety was observed binding in the active site, suggesting that the enzyme hydrolyzed the Neu5Ac moiety from CMP-Neu5Ac and only CMP remained bound to the enzyme. This result was similar to that reported for CstII sialyltransferase, for which crystallization of the sialyltransferase in the presence of CMP-Neu5Ac also resulted in observing only CMP bound to the enzyme (*12*). Therefore, to identify amino acid residues important for binding to the Neu5Ac moiety in the donor and catalytic activity of sialyltransferase, nonhydrolyzable analogues of the natural donor substrate CMP-Neu5Ac with an electronegative fluorine attached to carbon-3 next to the anomeric carbon C2 of the sialic acid moiety (CMP-3FNeu5Ac) were synthesized and used in crystal structure studies (*12*, *26*, *27*). Two CMP-3FNeu5Ac compounds were used. As shown in Scheme 1, CMP-3F(*a*)Neu5Ac is the donor analogue with the fluorine in the axial position at C-3 of the Neu5Ac pyranose ring and CMP-3F(*e*)Neu5Ac is the analogue with the fluorine in the equatorial position at C-3 of the Neu5Ac.

As predicted, the introduction of an electron-withdrawing fluoride group at the carbon-3 next to the anomeric carbon-2 of the Neu5Ac in the CMP-Neu5Ac sialyltransferase donor slows down the enzyme turnover rate by destabilizing the presumed oxocarbenium ion-like transition state involved in the sialyltransferase-catalyzed reaction (*12*, *26*). Kinetic data obtained indicate that CMP-3F(*a*)Neu5Ac is a competitive inhibitor of donor CMP-Neu5Ac for the α2,3-sialyltransferase activity of Δ24PmST1 with a K_i value of 25.7 μM. The K_i value (25.7 μM) of CMP-3F(*a*)Neu5Ac is more than 16-fold less than the K_m (0.43 mM) of CMP-Neu5Ac, suggesting that CMP-3F(*a*)Neu5Ac binds to Δ24PmST1 stronger than CMP-Neu5Ac. However, with higher enzyme concentrations and extended incubation times, Δ24PmST1 was able to catalyze the transfer of fluorinated Neu5Ac from donor analogues to the acceptor. This was confirmed by donor substrate specificity assays using CMP-Neu5Ac, CMP-3F(*e*)Neu5Ac, and CMP-3F(*a*)Neu5Ac as donor substrates (Table 1). While both CMP-3F(*e*)Neu5Ac and CMP-3F(*a*)-Neu5Ac can still be used as donor substrates by Δ24PmST1 for the formation of α2,3-sialylated products, both compounds displayed around 4 orders of magnitude less activity than the native donor CMP-Neu5Ac (Table 1). Nevertheless, the introduction of the C-3 fluorine atom in either the equatorial or axial position in the sugar of CMP-Neu5Ac overcomes the problem of the hydrolysis of the donor sugar by Δ24PmST1 during cocrystallization. Therefore, cocrystallization of Δ24PmST1 with either CMP-3F(*e*)Neu5Ac or CMP-3F(*a*)Neu5Ac produced two binary structures corresponding to the enzyme bound with the respective sugar donor analogues. The introduction of the C-3 fluorine atom in the Neu5Ac of CMP-Neu5Ac also helps to stabilize the complex formation with acceptor lactose. Previously, soaking the CMP-bound Δ24PmST1 binary complex with 10 mM lactose for about 1 h before data collection did not produce a ternary structure (data not shown). In comparison, performing the similar procedure by soaking the binary CMP-3F(*a*)Neu5Ac complex with 10 mM lactose resulted in a ternary complex structure with intact donor analogue CMP-3F(*a*)-Neu5Ac along with acceptor lactose bound to the enzyme. A similar procedure by soaking the binary CMP-3F(*e*)-Neu5Ac complex with 10 mM lactose also resulted in a ternary structure. However, no electron density for the sialic acid moiety of the donor analogue was observed in this structure, resulting in a ternary structure of the protein complexed with CMP along with acceptor lactose bound. This observation indicates that the donor analogue CMP-3F(*e*)Neu5Ac can still be utilized as a substrate for the enzyme. Indeed, substrate specificity assays confirmed that CMP-3F(*e*)Neu5Ac displayed more than three times the activity compared to that of CMP-3F(*a*)Neu5Ac (Table 1).

Table 1: Donor Substrate Specificity Assays of Δ24PmST1 Using CMP-Neu5Ac, CMP-3F(*a*)NeuAc, and CMP-3F(*e*)Neu5Ac

donor substrate	relative activity[a]
CMP-Neu5Ac	8779
CMP-3F(*e*)Neu5Ac	1
CMP-3F(*a*)Neu5Ac	0.37

[a] Relative activity was determined in duplicate by single point readings. The enzymatic assays were performed at 37 °C in a total volume of 50 μL in Tris-HCl buffer (100 mM, pH 8.5) containing CMP-Neu5Ac or its fluorinated analogue (1.5 mM), LacMU (1 mM), and the recombinant enzyme [180 ng for CMP-Neu5Ac and 36 μg for both CMP-3F(*e*)Neu5Ac and CMP-3F(*a*)Neu5Ac].

Table 2: Data Collection, Phasing, and Refinement Statistics

	Δ24PmST1 with CMP-3F(*a*)Neu5Ac	Δ24PmST1 with CMP-3F(*e*)Neu5Ac	Δ24PmST1 with CMP-3F(*a*)Neu5Ac + lactose	Δ24PmST1 with CMP + lactose
data collection statistics				
X-ray source	SSRL BL9-2	SSRL BL1-5	Cu Kα	SSRL BL11-1
wavelength (Å)	0.979	0.979	1.54	0.979
resolution (Å)	2.00 (2.05−2.00)	1.90 (1.95−1.90)	2.00 (2.05−2.00)	2.27 (2.33−2.27)
space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
cell parameters				
a (Å)	61.8	60.9	60.5	61.9
b (Å)	65.3	65.0	64.5	63.6
c (Å)	66.8	64.8	64.4	63.8
β (deg)	98.14	98.56	99.03	98.36
monomers/ASU	1	1	1	1
V_M (Å^3/Da), solvent (%)	3.0, 59	2.9, 57	2.8, 56	2.8, 56
no. of reflections	112190	70437	199896	56832
no. of unique reflections	33718	35617	33274	25036
R_{merge}[a] (%)	5.1 (31.0)	4.8 (20.4)	4.6 (21.1)	8.3 (38.1)
mean $(I)/\sigma(I)$	15.3 (2.4)	13.5 (3.8)	10.3 (2.1)	9.6 (2.1)
completeness (%)	98.63 (99.31)	90.9 (87.9)	99.92 (99.10)	92.8 (95.7)
refinement statistics				
resolution (Å)	66.08−2.00	64.02−1.90	63.63−2.00	63.12−2.27
no. of reflections ($F \geq 0$)	33495 (2603)	33809 (2537)	31548 (2409)	20108 (1593)
R-factor[b] (%)	19.7 (30.4)	17.7 (21.7)	21.1 (22.9)	21.7 (28.1)
R-free[b] (%)	22.1 (38.7)	20.4 (31.2)	25.2 (26.5)	25.7 (35.6)
rmsd bond lengths (Å)	0.011	0.013	0.010	0.012
rmsd bond angles (deg)	1.281	1.302	1.333	1.302
mean *B* values (Å^2)				
protein (no. of atoms)	31.1 (3113)	28.1 (3096)	27.4 (3099)	32.2 (3108)
donor analogue (no. of atoms)	25.0 (42)	19.8 (42)	23.4 (42)	23.6 (21)
lactose (no. of atoms)			40.3 (23)	32.5 (23)
water (no. of atoms)	36.5 (422)	35.1 (320)	33.5 (359)	36.5 (228)
Ramachandran plot statistics				
no. of residues[c]	342	340	343	342
no. of waters	423	320	382	237
most favorable region (%)	93.0	93.5	90.7	90.1
allowed region (%)	6.7	6.2	9.0	9.6
generously allowed region (%)	0.0	0.0	0.0	0.0
disallowed region (%)	0.3	0.3	0.3	0.3
PDB ID	2IHJ	2IHK	2IHZ	2ILV

[a] $R_{merge} = [\sum_h \sum_i |I_h - I_{hi}| / \sum_h \sum_i I_{hi}]$, where I_h is the mean of I_{hi} observations of reflection *h*. Numbers in parentheses represent the highest resolution shell. [b] *R*-factor and *R*-free $= \sum ||F_s| - |F_c|| / \sum |F_o| \times 100$ for 95% of recorded data (*R*-factor) or 5% of data (*R*-free). [c] Number of non-proline and non-glycine residues used for calculation.

All crystal structures are isomorphous with very similar crystal lattice parameters (Table 2). The Δ24PmST1 structure belongs to the GT-B structural group (*28*−*33*). The enzyme is made up of two α/*β*/α Rossmann domains (*34*). A deep cleft between the two domains (Figure 1A) comprises the active site (see below). We previously showed that the two domains closed ∼23° around the active site cleft upon binding to donor substrate (*11*). All four structures reported here resemble the closed conformation with an rmsd of ∼0.6 Å for 374 equivalent α-carbons to the previously reported closed CMP-bound structure. Electron density clearly defines residues 25−412 except for residues 374−382, which correspond to the loop between strand *β*12 and helix α12a as well as part of the helix α12a. This loop lies in the C-terminal domain near the active site and is disordered in all structures except for the ligand-free apo structure (*11*). While it appears that this disordered loop is not involved in binding sialyltransferase acceptors, its exact role both in sialyltransferase and in sialidase functions remains to be determined.

Δ24PmST1−*CMP-3FNeu5Ac Binary Structures.* Original molecular-replacement-phased electron density maps clearly defined the CMP-3FNeu5Ac binding sites (Figure 2). The structures of the enzyme in both binary structures Δ24PmST1−CMP-3F(*a*)Neu5Ac and Δ24PmST1−CMP-3F(*e*)Neu5Ac are virtually identical and can be superimposed onto each other with an rmsd of 0.47 Å for 376 equivalent α-carbons. Additionally, the conformations of two CMP-3FNeu5Ac compounds observed in these two binary structures are very similar except for the positions of the fluorine on carbon-3 (one is in axial and the other is in equatorial) and a slight rotation of the sialic acid carboxylate group orientation caused by the difference of the fluorine positions.

The CMP nucleotide moiety in both structures interacts with the enzyme in a manner very similar to that in the previously reported CMP-bound binary structure (*11*). CMP makes extensive hydrogen bonds with only the C-terminal nucleotide-binding domain, and no contacts are made to the N-terminal domain. Eleven of the 12 hydrogen bonds between the CMP moiety and the protein are conserved between the CMP binary structure previously reported and the two CMP-3FNeu5Ac binary structures reported here. The only hydrogen bond that is absent in the CMP-3FNeu5Ac binary structures is the one observed between O*γ* of Ser355 and the O3P oxygen of the phosphate in the CMP binary structure. The tethering of the sialic acid moiety to the phosphate in the CMP-3FNeu5Ac causes the phosphate to rotate and move away from Ser355. However, the hydrogen

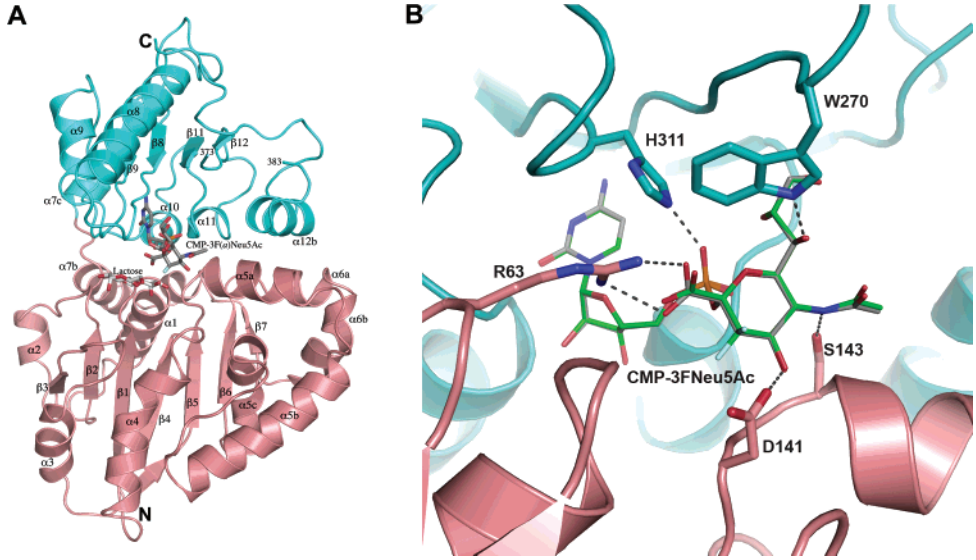


FIGURE 1: (A) Overall ribbon drawing of Δ24PmST1. The N- and C-terminal domains are colored in salmon and teal, respectively. Bound CMP-3F(*a*)Neu5Ac and lactose are rendered as sticks with gray- and white-colored carbons, respectively. Secondary structures elements are labeled. (B) Shown is the active site of Δ24PmST1 with the activated sugar nucleotide donor analogue CMP-3F(*a*)Neu5Ac bound, drawn in thin sticks with gray-colored carbon atoms. Superimposed is the CMP-3F(*e*)Neu5Ac from the binary structure drawn in thin sticks with green-colored carbon atoms. Fluorine atoms are colored light blue. Residue side chains that hydrogen bond to the phosphate and Neu5Ac moiety are drawn in sticks, with potential hydrogen bonds in gray dashes.

bonds between the phosphate oxygen O2P and Ser356 (Oγ) as well as those between the O1P and His311 (Nϵ2) are still maintained in all structures.

The sialic acid sugar is held in place through interactions with amino acid residues from both N- and C-terminal domains of the protein (Figure 1B). The indole ring nitrogen of Trp270 in the loop between strand β8 and helix α8 of the C-terminal domain forms a hydrogen bond to oxygen O7 of sialic acid. This Trp270 side chain is disordered in the previous CMP binary structure but is ordered in the apo structure where it swings ~180° around and aligns next to helix α8. Comparing to the apo structure (*11*), the Trp270-containing loop moves substantially upon binding to CMP-3FNeu5Ac. The Trp270 side chain in the new location helps to define the acceptor binding pocket by a stacking interaction (see below).

Arg63, Asp141, and Ser143 from the N-terminal domain also interact with the sialic acid moiety of the CMP-3FNeu5Ac molecules (Figure 1). Arg63 ion pairs with the sialyl carboxylic acid group, Asp141 hydrogen bonds with sialic acid oxygen O4, and Ser143 hydrogen bonds with the amide nitrogen of the *N*-acetyl group. All of these interactions are conserved in both binary structures with protein bound to CMP-3F(*a*)Neu5Ac or CMP-3F(*e*)Neu5Ac.

Δ24PmST1−CMP-3F(a)Neu5Ac−Lactose Ternary Structure. Crystals of the CMP-3F(*a*)Neu5Ac binary complex were soaked with lactose for 1 h prior to freezing and data collection. The resultant electron density map clearly defines the galactopyranose moiety of the disaccharide acceptor while the glucopyranose is less defined although it can still be modeled unambiguously (Figure 2D). The lactose has higher temperature (*B*) values (40.3 Å^2) compared to CMP-3F(*a*)Neu5Ac (23.4 Å^2) (Table 2), suggesting that it does not bind as tightly. The more buried galactopyranose moiety contacts five amino acid residues in the active site, while the glucopyranose is solvent exposed and interacts with only one amino acid residue (Figure 3A). Trp270, whose indole ring nitrogen hydrogen bonds to the hydroxyl oxygen O7 of the sugar in the donor, forms a van der Waals stacking interaction to the "hydrophobic" surface of the galactopyranose ring of the lactose. Residues His112 and Asp141 each make two hydrogen bonds to O3′ and O4′ of the galactopyranose moiety. Nϵ2 of His112 makes a bifurcated hydrogen bond to these two oxygens, while Oδ1 and Oδ2 of Asp141 make a pair of parallel bidentate hydrogen bonds to oxygens O3′ and O4′ of the galactopyranose moiety, respectively. Nη2 of Arg63, which ion pairs to the donor sialic acid carboxyl group, also hydrogen bonds to O6 of the glucopyranose in the acceptor. This O6 oxygen is also in hydrogen-bonding distance to the guanidino group of Arg313. Meanwhile, Arg313 forms hydrogen bonds with O2 and O3 of the hydroxyl groups on C2 and C3 of glucopyranose. These are the only interactions found between the glucopyranose in the acceptor and Δ24PmST1 and explain the observed weak electron density for this monosaccharide residue.

Δ24PmST1−CMP−Lactose Ternary Structure. In the ternary crystal structure above, the distance between the O3′ hydroxyl of the galactopyranose moiety and the anomeric carbon C2 of the sialic acid is 5.3 Å, which is too great a distance for a productive nucleophilic attack to generate the α2,3 linkage. This suggests that the galactopyranose moiety, as observed in the above crystal structure, may not be bound in a conformation that leads to a productive α2,3-sialyltransferase activity. This is likely due to the presence of the fluorine atom on carbon C3 of the donor, which is in the axial position and lies in between the donor sugar and the acceptor lactose. The fluorine is 3.08, 3.37, and 3.35 Å away from carbons C4′ and C3′ and oxygen O3′ of the galactopyranose, respectively. Therefore, the axial position of the fluorine prevents a closer interaction between the oxygen O3′ of the galactopyranose in the acceptor (the nucleophile) and the anomeric carbon of the sugar in the donor, which

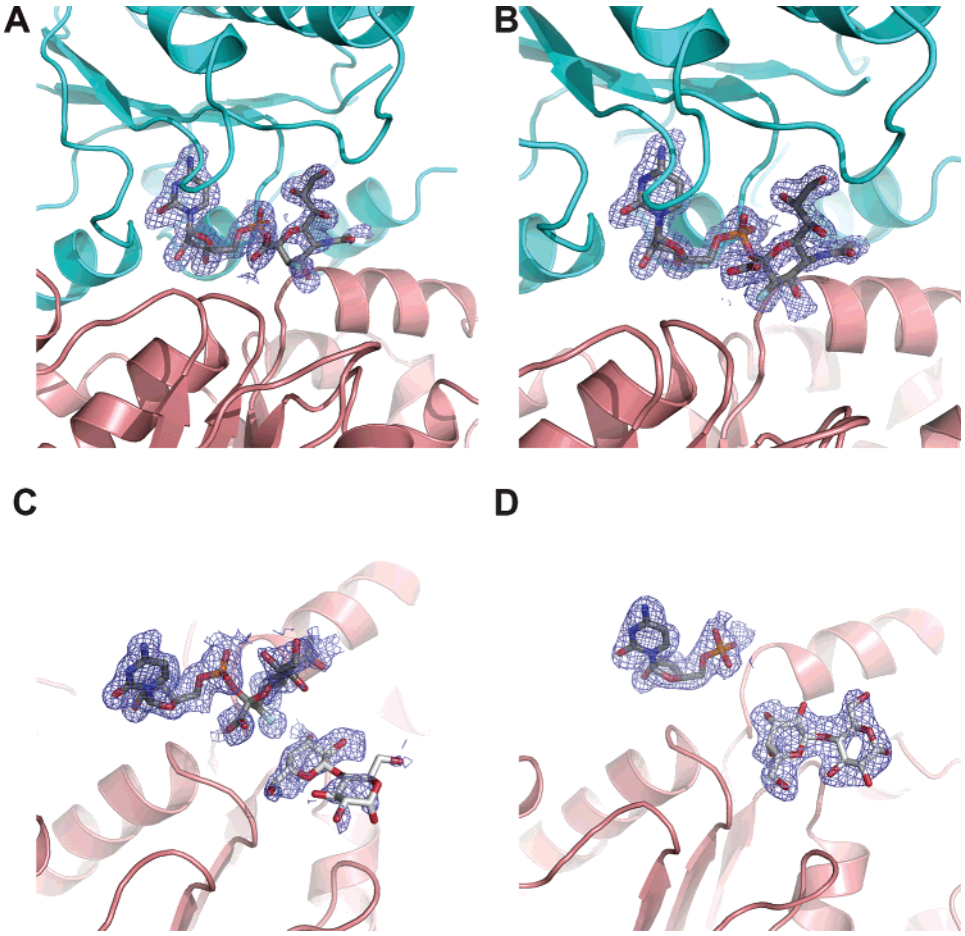


FIGURE 2: Original experimental omit maps calculated using phases from the molecular replacement solution, which contained protein but no substrates or water molecules. Maps represent original $F_o - F_c$ electron density maps contoured at 2.5σ prior to any protein refinement. (A) and (B) represent the binary structures CMP-3F(*a*)Neu5Ac and CMP-3F(*e*)Neu5Ac, respectively, drawn with the final refined model. (C) and (D) represent the ternary structures CMP-3F(*a*)Neu5Ac + lactose and CMP + lactose, respectively. The C-terminal domain is omitted for clarity in (C) and (D). Electron density improved for all ligands upon refinement.

would be required for a productive α2,3-sialyl-transfer reaction.

Therefore, to determine if the fluorine in the axial position was affecting the acceptor lactose binding location, a donor CMP-Neu5Ac analogue CMP-3F(*e*)Neu5Ac with the fluorine in the equatorial position off carbon C3 of Neu5Ac was synthesized. This analogue, when bound to the enzyme, would not position the fluorine between the nucleophilic hydroxyl group in the acceptor sugar and the anomeric carbon of the Neu5Ac in the donor. The binary structure of Δ24PmST1 with this donor sugar was described above, which is essentially identical to the binary complex with the donor analogue containing the fluorine in the axial position of C3 in Neu5Ac. However, when the Δ24PmST1−CMP-3F(*e*)Neu5Ac binary complex crystals were soaked for 1 h with lactose prior to freezing and data collection, the resultant electron density map clearly defined the acceptor lactose binding position and the CMP nucleotide, but no density was observed for the sialic acid moiety (Figure 2D).

The electron density that defines the lactose is strong and better defined than that of the lactose bound to the CMP-3F(*a*)Neu5Ac ternary complex (Figure 2D). This is surprising considering that the lactose in the CMP ternary complex makes fewer contacts to the protein than the lactose in the CMP-3F(*a*)Neu5Ac ternary complex (Figure 3B). Of significance though is the observation that, in the CMP ternary complex, the lactose rotates and shifts a little deeper into the active site pocket compared to the CMP-3F(*a*)Neu5Ac ternary complex. The galactopyranose ring, the ring that accepts the donor sugar, shifts 1.1 Å (ring center to ring center) toward what would have been the sialic acid, if it was not enzymatically transferred during soaking procedures (Figure 3B). The glucopyranose ring center rotates and shifts 5.0 Å down toward Met144 compared to the other lactose-bound ternary structure.

The shift of the lactose in this structure results in breaking some contacts and making a couple new interactions to the enzyme. Asp141 still makes a bidentate interaction with galactopyranosyl hydroxyl oxygens O3′ and O4′, but His112 now only hydrogen bonds to hydroxyl oxygen O4′, not hydroxyl oxygen O3′ anymore (Figure 3B). Both Arg63 and Arg313 no longer interact with the lactose. Trp270 still maintains its hydrophobic stacking interaction with the lactose galactopyranose ring, resulting in Trp270 shifting ∼1.5 Å down (as viewed in Figure 3B) and pulling the β8−α8 loop with it, compared to the other structures. The galactopyranose ring is now sandwiched between Trp270 and Met144, where Met144 makes new van der Waals contact with the bottom of the ring (as viewed in Figure 3B). Finally, a new hydrogen bond develops between Asn85 (Oδ1) and the hydroxyl oxygen O6′ of the galactose in the lactose acceptor. While in the CMP-3F(*a*)Neu5Ac−lactose ternary

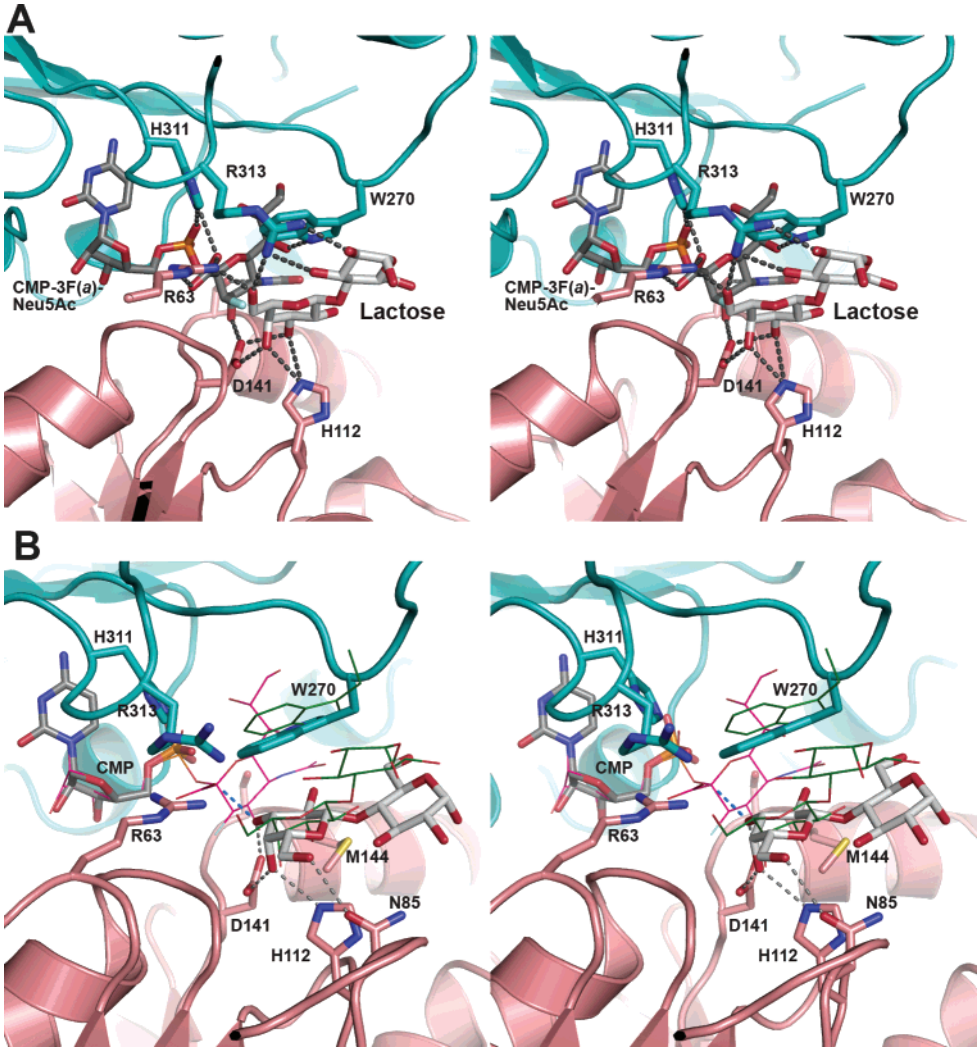


FIGURE 3: Active site structures of ternary complexes. (A) Stereoview of the active site ternary complex with the activated sugar donor analogue CMP-3F(*a*)Neu5Ac plus acceptor lactose bound. The activated sugar nucleotide donor is drawn with gray-colored carbon atoms, and the lactose acceptor is drawn with white-colored carbon atoms. Only side-chain residues that interact with the sialic acid moiety and lactose are drawn in stick representation with hydrogen bonds drawn in gray dashes. N-Terminal and C-terminal domains of Δ24PmST1 are colored in salmon and teal, respectively. (B) Stereoview of the active site ternary complex with CMP plus lactose bound. Lactose and Trp270, as observed in the CMP-3F(*a*)Neu5Ac-bound ternary complex, are drawn as lines with green-colored carbons. The activated sugar donor analogue CMP-3F(*e*)Neu5Ac, as observed in its binary structure, is drawn in line representation with magenta-colored carbons. Hydrogen bond interactions to bound lactose are drawn in gray dashes. The hypothetical distance between the acceptor lactose hydroxyl oxygen O3 (as seen in the CMP−lactose ternary structure) is 2.7 Å away (blue dashed line) from the donor anomeric carbon C2 in CMP-3F(*e*)Neu5Ac (as observed in its binary structure).

structure, Asn85 makes a water-mediated interaction to galactopyranosyl oxygens O1′ and O5′ of the lactose (not shown).

Although the sialic acid moiety is absent in this structure, the structural position of lactose bound here is likely in its productive conformation for α2,3-sialyltransferase activity. This is because the O3′ hydroxyl has now moved closer to a reasonable distance from what would be the anomeric carbon C2 position. By superimposing the CMP-3F(*e*)-Neu5Ac analogue (as observed in its binary structure) onto the CMP in the ternary structure, one can get a reasonably accurate position of the sialic acid (drawn as lines in Figure 3B). In this superposition, the O3′ hydroxyl of lactose is now 2.7 Å from the anomeric carbon C2 of sialic acid (blue dashed line in Figure 3B), an ideal distance for a productive transfer.

Comparison to Other Sialyltransferase Structures. The only other sialyltransferase whose structure is known at atomic detail is CstII from the pathogenic bacterium *C. jejuni*. CstII structures have been obtained in the presence of CMP or the donor analogue CMP-3F(*a*)Neu5Ac (*12*). Unlike Δ24PmST1 structures, which belong to the GT-B fold, CstII structures fold into the GT-A-like single Rossmann domain. These two types of sialyltransferase structures have two uniquely distinct binding sites and donor recognition motifs. In CstII, the donor analogue binds between the C-terminal end of the central β-sheet and a small lid domain that covers the active site. The ribose lies at the N-terminal end of helix αF where the 2′-OH and 3′-OH hydrogens bond to the main-chain atoms. The nucleotide moiety lies deep in a crevice between the central β-sheet and helix αF, with the lid domain covering the nucleotide. CstII main-chain atoms mediate all hydrogen bonds between the nucleoside and the enzyme. Additionally, the cytosine base π-stacks against Tyr156 in CstII. In comparison, in Δ24PmST1, the donor analogue lies in a deep cleft between the two Rossmann domains. The

Table 3: Apparent Kinetic Parameters for the α2,3-Sialyltransferase Activity of Δ24PmST1 and Its Mutants

	CMP-Neu5Ac[a]			LacMU[a]		
Δ24PmST1	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (mM^{-1} s^{-1})	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (mM^{-1} s^{-1})
WT	31.6 ± 1.45	0.43 ± 0.05	72.5	46.6 ± 1.46	1.41 ± 0.09	33.0
D141A[b]	ND	ND	ND	ND	ND	ND
D141N[b]	ND	ND	ND	ND	ND	ND
H112A/D141A[b]	ND	ND	ND	ND	ND	ND
H112A	0.91 ± 0.10	7.25 ± 1.42	0.12	0.62 ± 0.05	6.57 ± 0.79	0.09
H311A	2.34 ± 0.13	0.90 ± 0.17	2.58	1.76 ± 0.17	1.42 ± 0.24	1.24
W270F	1.07 ± 0.08	1.92 ± 0.57	0.56	1.31 ± 0.06	0.89 ± 0.14	1.46
W270A	0.51 ± 0.03	1.36 ± 0.27	0.37	1.37 ± 0.03	2.27 ± 0.13	0.60

[a] Apparent kinetic parameters for CMP-Neu5Ac were determined at a fixed concentration (1 mM) of LacMU and for LacMU at a fixed concentration (1 mM) of CMP-Neu5Ac. [b] The activity was too low to accurately measure the kinetics parameters (ND, not determined). However, measurements by single point readings indicated more than 20000-fold reduction in α2,3-sialyltransferase activity of the mutant as compared to that of the wild-type Δ24PmST1. The enzymatic assays were performed in a total volume of 50 μL in Tris-HCl buffer (100 mM, pH 8.5) containing CMP-Neu5Ac (1 mM), LacMU (1 mM), and the recombinant enzyme (30 μg for D141N, 58 μg for H112A/D141A, and 79 μg for D141A Δ24PmST1 mutants) for 1 h at 37 °C. Less than 2% conversion was observed for the D141A and D141N single mutants as well as the H112A/D141A double mutant under these conditions. (In comparison, assays for wild-type enzyme carried out at 37 °C for 10 min using 41 ng of enzyme resulted in about 10% conversion.)

2′-OH and 3′-OH hydrogens bond to Glu338 and Ser336 and make no contact with any main-chain atoms. The nucleoside makes hydrogen bonds to atoms in both the main chain and side chain. The base does not π-stack with any aromatic side chains.

Comparing the active sites of these two types of sialyltransferase structures also reveals key differences in the donor substrate conformations. The donor substrate analogue, CMP-3F(*a*)Neu5Ac, is more collapsed in the CstII-bound form compared to that in the Δ24PmST1-bound form (which is in a more extended conformation). The collapsed CMP-3F(*a*)Neu5Ac in the CstII structure is due to the extensive hydrogen-bonding network between the Neu5Ac moiety and the enzyme. In particular, Tyr156, which π-stacks against the cytosine ring, hydrogen bonds to two phosphate oxygens, causing phosphate to rotate around the C5′−O5′ bond into an ∼90° dihedral angle. This causes the carboxylate oxygen in Neu5Ac to form a short hydrogen bond (2.58 Å) to the 3′-OH of the ribose, and the Neu5Ac carboxylate group does not ion pair with any basic residue. In Δ24PmST1, the C4′−C5′−O5′−PA dihedral angle is ∼180°, causing the Neu5Ac to extend away from the nucleoside where the carboxylate group ion pairs with Arg63 and is 6.3 Å away from the 3′-OH of ribose. Another difference between the two donor binding motifs is at the C8 and C9 hydroxyl groups of sialic acid. In Δ24PmST1, both hydroxyl groups are solvent exposed where they hydrogen bond to ordered water molecules, while in CstII, the hydroxyls are buried in the active site where they hydrogen bond to side chains.

The different interactions of CMP-3F(*a*)Neu5Ac to Δ24PmST1 and CstII provide the first insights into the design of a specific inhibitor for one or a list of sialyltransferases. For example, the C8-OH in the Neu5Ac moiety of CstII forms two potential intermolecular hydrogen bonds with Gln32 and one intramolecular bond with a phosphate oxygen. The same hydroxyl in Neu5Ac binding to PmST1 makes no hydrogen bonds and actually points toward a small cavity above the ribose ring. At the opposite end of this cavity are two main-chain carbonyl oxygens. This suggests that adding a functional group that can donate hydrogen bonds could increase both binding and selectivity to PmST1 but preclude binding in CstII ST. We are in the process of synthesizing CMP-3FNeu5Ac with an alkyl group at the 8-OH group in 3FNeu5Ac. The specificity of these compounds as potential inhibitors particularly to Δ24PmST1 will be tested.

Sialic Acid Conformation and Interactions. The sialic acid ring in all Δ24PmST1−CMP-3FNeu5Ac-bound protein complexes is in a $^2C_5$ chair conformation. However, the sialic acid observed in both CMP-3F(*a*)Neu5Ac-bound binary and ternary structures is slightly more planar than that is observed in the CMP-3F(*e*)Neu5Ac binary structure. For the CMP-3F(*a*)Neu5Ac structures bound in both binary and ternary complexes, the oxygens in the carboxylate group, C1, C2, and O6 of the 3F(*a*)Neu5Ac moiety are roughly coplanar. But in the CMP-3F(*e*)Neu5Ac-bound binary complex, the carboxylate group of the 3F(*e*)Neu5Ac rotates out of the plane due to the equatorial position of the fluorine.

Overall, the interactions between the sialic acid carboxylate group and the enzyme differ from those in the other sialyltransferase CstII structures. For CstII, the carboxylate group interacts with neutral side chains but lies near the positive dipole of a nearby helix (*12*). In comparison, for Δ24PmST1, the carboxylate group of the sialic acid ion pairs with Arg63 and lies near the dipole of helix α1. This also differs from the carboxylate coordination commonly observed in sialidase and *trans*-sialidase structures where a conserved arginine triad interacts with the carboxylate group of Neu5Ac (*35*−*42*).

Kinetic Analysis. The structures presented here have provided invaluable insight in identifying potential residues important for catalytic activity. Asp141, which is hydrogen bonded to hydroxyl oxygen O3′ of the galactopyranose in the lactose, may serve as a catalytic base to abstract a proton from the hydroxyl group at carbon-3′ of the galactopyranose to make O3′ a better nucleophile to attack the anomeric carbon C2 of the CMP-Neu5Ac for α2,3-sialyltransferase activity. His112, which is near Asp141 and also interacts with the acceptor sugar, may play a role in activating or regulating the pK_a of Asp141. To test these hypotheses, we mutated these residues and measured their effect on the α2,3-sialyltransferase activity of Δ24PmST1. Both single mutants D141A and D141N as well as a double mutant D141A/H112A had such a low activity that accurate kinetic parameters could not be determined. However, single point

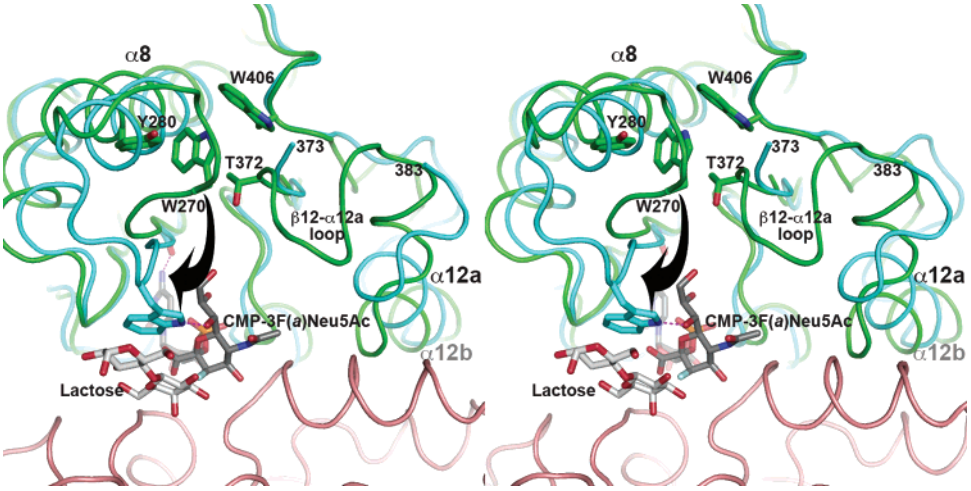


FIGURE 4: Stereoview superposition between the previously determined apo Δ24PmST1 structure and the CMP-3F(*a*)Neu5Ac ternary structure showing the movement of Trp270 upon binding to donor. Only the C-terminal domain of the apo Δ24PmST1 structure is shown in green. The N- and C-terminal domains of the ternary structure are shown in salmon and teal, respectively. CMP-3F(*a*)Neu5Ac and lactose are rendered in sticks with gray- and white-colored carbon atoms, respectively. Additionally, critical amino acid side chains are also shown in sticks. The side chain of Trp270 moves >13 Å upon binding to the activated donor sugar (black arrow). The β12−α12a loop is shown in the apo structure (green), which becomes disordered upon binding donor sugar.

readings with higher concentrations of mutant enzyme and a longer reaction time (1 h) revealed that these mutants had a more than 20000-fold reduction of the α2,3-sialyltransferase activity compared to that of the wild-type enzyme (Table 3). The H112A mutant was more active than the Asp141 mutants, but the activity (k_{cat}/K_m) was reduced 350−700-fold compared to that of the wild-type enzyme.

His311, which hydrogen bonds to the oxygen on the phosphate in the sugar nucleotide donor and may serve to stabilize the CMP phosphate leaving group, was mutated to Ala. The H311A mutant resulted in a less than 30-fold decrease in efficiency, suggesting that it also makes a minor contribution to the α2,3-sialyltransferase catalytic mechanism. Trp270, which stacks against the galactopyranose ring in the lactose acceptor where it makes van der Waals interactions, has also been mutated. Both W270F and W270A mutants had similar effects with more than a 35-fold decrease in k_{cat}, suggesting the important role of Trp270 in the α2,3-sialyltransferase activity of Δ24PmST1 (Table 3).

DISCUSSION

The structure of Δ24PmST1 consisting of two Rossmann-like domains belongs to the glycosyltransferase B (GT-B) structural group (*32*). This differs from the only other known sialyltransferase structure of CstII, a sialyltransferase from *C. jejuni* (*12*), which belongs to the single Rossmann-domain-containing glycosyltransferase A (GT-A) group. The GT-B family glycosyltransferases typically have more diversity in their substrates and products compared to those from the GT-A family, which may help to explain the four catalytic activities measured for Δ24PmST1 (*10*).

Four crystal structures of Δ24PmST1 illustrated here with different donor analogues in the prescence or the absence of acceptor substrate indicate that while the activated sugar nucleotide donor binds in a defined site with a comparable conformation, there are possibly two distinct lactose binding orientations.

The lactose position seen in the Δ24PmST1−CMP−lactose ternary structure (Figure 3B) is likely the productive orientation for the α2,3-sialyltransferase activity because of the close proximity of the galactopyranose oxygen O3′ to the potential anomeric carbon of sialic acid. The lactose location in the Δ24PmST1−CMP-3F(*a*)Neu5Ac−lactose ternary structure (Figure 3A) represents an alternative lactose binding conformation, which might be required for the α2,6-sialyltransferase activity [favored at low pH (*10*)]. In this lactose conformation, oxygen O6′ is pointed away from the anomeric carbon C2 where it hydrogen bonds to Arg63 and Arg313 (Figure 3A). However, rotating the torsion angle around the C5−C6 bond can move the O6′ oxygen from 6.6 to 5.1 Å away from the anomeric carbon. While this distance is still significantly large, the lactose disposition might shift slightly at low pH. Furthermore, the absence of the fluorine atom in the natural donor may bring the O6′ oxygen closer to the anomeric carbon to form the α2,6-sialyl linkage. The O6′ oxygen of the lactose galactopyranose ring in the CMP ternary structure is further away (8.3 Å) from the anomeric carbon-2 (Figure 3B). This corroborates that the alternative lactose binding orientation seen in the CMP-3F(*a*)Neu5Ac−lactose ternary complex might represent the bound conformation for α2,6-sialyltransferase activity. It might also represent a sugar binding conformation for the sialidase and/or *trans*-sialidase activity of the enzyme (*10*). But additional experiments are required to establish this.

The observation of the intact CMP-3F(*e*)Neu5Ac in the binary complex indicates that the CMP-3F(*e*)Neu5Ac is not hydrolyzable by the enzyme. Soaking the crystals of this binary complex with a sialyltransferase acceptor lactose resulted in the removal of the 3F(e)Neu5Ac moiety from the crystal. This demonstrates that the sugar moiety 3F(*e*)-Neu5Ac in the donor analogue can be transferred to the acceptor lactose in the crystals (this has been confirmed by the synthesis of fluorinated sialosides using Δ24PmST1, which will be reported separately), the sialylated product can then diffuse out of the pocket, and excess lactose from the soaking solution can bind to the acceptor binding pocket of the CMP-bound enzyme. This also suggests that the closed conformation of the protein, similar to that of the CMP- or

CMP-3FNeu5Ac-bound ones, is the one involved in the sialyltransferase-catalyzed reaction of Δ24PmST1. Additionally, it implies that the CMP product may not be released from the enzyme until the two domains open up as seen in the ligand-free apo structure.

The transition state of sialic acid in the sialyltransferase-catalyzed reaction has been believed to have an sp^2 oxocarbenium ion-like character at the anomeric carbon (*26*). In all three sialic acid-containing structures presented here, the pyranose ring conformations of the sialic acids are in a similar $^2C_5$ chair conformation. This indicates that the ring must be distorted to generate the transition state where the O6, C2, C1, C3, and C4 are all coplanar, required for the oxocarbenium ion formation. Arg63 is coordinated to the sialic acid carboxylate, which helps to anchor the substrate during its distortion along the reaction coordinate. This arginine anchoring is reminiscent to that in the sialidase and *trans*-sialidase arginine triad but is absent from CstII.

Trp270 in the β8−α8 loop plays an important role as a gatekeeper in the α2,3-sialyltransferase activity of Δ24PmST1. Trp270 swings down only after the CMP-activated sugar nucleotide donor binds. This movement is likely facilitated by the cytosine ring hydrogen bonding to the β8−α8 loop. A likely scenario is that, in the absence of CMP-activated donor sugar, Trp270 is packed in the C-terminal domain protein core between Tyr280, Trp406, and Thr372 (Figure 4). Upon binding to CMP-activated donor sugar, the hydrogen bond formed between the nitrogen N4 of the cytosine ring and the carbonyl oxygen of Gly266 in the β8−α8 loop triggers the movement of the β8−α8 loop, which allows the Trp270 residue in the loop to flip around >13 Å to form a hydrogen bond between the indole nitrogen and O7 of sialic acid (Figure 4). The new location of the Trp270 side chain now helps to define the acceptor sugar binding pocket by forming a stacking interaction with the galactopyranose ring. This type of stacking interaction is common for many other carbohydrate binding proteins (*43*−*46*). In the CMP binary structure previously reported, the side chain of Trp270 is disordered because there is no sialic acid to interact with. It seems that Trp270 does not only help to form the acceptor binding pocket upon binding to the sugar nucleotide donor but also helps to stabilize the donor itself. Mutating this residue had significant consequences in the α2,3-sialyltransferase activity. The mutants had a greater effect on k_{cat} (∼35-fold decrease) than K_m (less than 5-fold different) for both donor and acceptor substrates (Table 3). Surprisingly, the W270F mutant actually decreased the K_m for acceptor and the W270A mutant only increased the K_m by a trivial amount (1.6-fold). This suggests that while Trp270 does contribute to some extent to the binding of acceptor and donor, it has a larger effect on catalytic turnover, possibly by helping to maintain a productive orientation of both substrates. Since Trp270 forms similar interactions with both conformations of bound lactose, it may play a similar role in the α2,6-sialyltransferase activity. It remains to be tested if Trp270 plays a similar role in sialidase/*trans*-sialidase activity.

In the Δ24PmST1 binary and ternary structures presented here, movement of two adjacent loops is observed upon binding to CMP-3FNeu5Ac donor substrate analogues. The β8−α8 loop shifts upon binding to the donor analogues and causes Trp270 in this loop to swing down. This Trp270 movement likely triggers the shifting of the adjacent β12−α12a loop (residues 374−382), which becomes disordered (Figure 4). Comparable loop movements are observed near the active sites in virtually all glycosyltransferases (*47*), including the CstII sialyltransferase structure (*12*). The function of this loop is currently unknown. It remains to be determined if this β12−α12a loop plays a functional role in any of the four measured activates of this enzyme.

The structures presented here have identified Asp141 as a potential catalytic base in the α2,3-sialyltransferase activity, with His112 potentially activating or modulating the Asp141 activity. This hypothesis has been tested by site-directed mutagenesis. The D141A mutant decreases the α2,3-sialyltransferase activity by more than 20000-fold, while the H112A mutant decreases the activity more than 350-fold but less than 1000-fold. This suggests that Asp141 is the catalytic base and abstracts the proton from the O3′ hydroxyl of the Gal in the acceptor to make the negatively charged oxygen a better nucleophile to attack the anomeric carbon of the sialic acid in the donor. This is consistent with the mechanism of inverting glycosyltransferases. The D141A/H112A double mutant has kinetics similar to that of the D141A single mutant. This suggests that there is no concerted effect of Asp141 and His112.

The location of His311 suggests that its function may be to stabilize the phosphate leaving group. Therefore, the H311A mutant was constructed to test its role in stabilizing the negatively charged CMP when it leaves. The kinetic data (Table 3), however, suggest that this residue alone is not critical in assisting the departure of the phosphate. In this case, Ser356 along with His311 may help in the departure of the CMP leaving group as it is also found to donate a hydrogen bond to the phosphate oxygen O2P. This may be similar to that of CstII wherein the concerted effort of two tyrosine residues is required to stabilize the leaving phosphate group (*12*). It may also be analogous to the catalytic architecture of retaining glycosidases wherein histidine/serine/tyrosine residues assist in the hydrolysis of the glycosyl enzyme intermediate by stabilizing the released carboxylate (*48*, *49*).

REFERENCES

1. Angata, T., and Varki, A. (2002) Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective, *Chem. Rev. 102*, 439−469.
2. Schauer, R. (2000) Achievements and challenges of sialic acid research, *Glycoconjugate J. 17*, 485−499.
3. Vimr, E., and Lichtensteiger, C. (2002) To sialylate, or not to sialylate: that is the question, *Trends Microbiol. 10*, 254−257.
4. Gilbert, M., Brisson, J. R., Karwaski, M. F., Michniewicz, J., Cunningham, A. M., Wu, Y., Young, N. M., and Wakarchuk, W. W. (2000) Biosynthesis of ganglioside mimics in *Campylobacter jejuni* OH4384. Identification of the glycosyltransferase genes, enzymatic synthesis of model compounds, and characterization of nanomole amounts by 600-mHz (1)H and (13)C NMR analysis, *J. Biol. Chem. 275*, 3896−3906.
5. Smith, H., Parsons, N. J., and Cole, J. A. (1995) Sialylation of neisserial lipopolysaccharide: a major influence on pathogenicity, *Microb. Pathog. 19*, 365−377.
6. Mandrell, R. E., McLaughlin, R., Aba Kwaik, Y., Lesse, A., Yamasaki, R., Gibson, B., Spinola, S. M., and Apicella, M. A. (1992) Lipooligosaccharides (LOS) of some *Haemophilus* species mimic human glycosphingolipids, and some LOS are sialylated, *Infect. Immun. 60*, 1322−1328.
7. Hood, D. W., Cox, A. D., Gilbert, M., Makepeace, K., Walsh, S., Deadman, M. E., Cody, A., Martin, A., Mansson, M., Schweda, E. K., Brisson, J. R., Richards, J. C., Moxon, E. R., and Wakarchuk, W. W. (2001) Identification of a lipopolysaccharide

alpha-2,3-sialyltransferase from *Haemophilus influenzae*, *Mol. Microbiol. 39*, 341−350.
8. Hood, D. W., Makepeace, K., Deadman, M. E., Rest, R. F., Thibault, P., Martin, A., Richards, J. C., and Moxon, E. R. (1999) Sialic acid in the lipopolysaccharide of *Haemophilus influenzae*: strain distribution, influence on serum resistance and structural characterization, *Mol. Microbiol. 33*, 679−692.
9. Bouchet, V., Hood, D. W., Li, J., Brisson, J. R., Randle, G. A., Martin, A., Li, Z., Goldstein, R., Schweda, E. K., Pelton, S. I., Richards, J. C., and Moxon, E. R. (2003) Host-derived sialic acid is incorporated into *Haemophilus influenzae* lipopolysaccharide and is a major virulence factor in experimental otitis media, *Proc. Natl. Acad. Sci. U.S.A. 100*, 8898−8903.
10. Yu, H., Chokhawala, H., Karpel, R., Wu, B., Zhang, J., Zhang, Y., Jia, Q., and Chen, X. (2005) A multifunctional *Pasteurella multocida* sialyltransferase: a powerful tool for the synthesis of sialoside libraries, *J. Am. Chem. Soc. 127*, 17618−17619.
11. Ni, L., Sun, M., Yu, H., Chokhawala, H., Chen, X., and Fisher, A. J. (2006) Cytidine 5′-monophosphate (CMP)-induced structural changes in a multifunctional sialyltransferase from *Pasteurella multocida*, *Biochemistry 45*, 2139−2148.
12. Chiu, C. P., Watts, A. G., Lairson, L. L., Gilbert, M., Lim, D., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. (2004) Structural analysis of the sialyltransferase CstII from *Campylobacter jejuni* in complex with a substrate analog, *Nat. Struct. Mol. Biol. 11*, 163−170.
13. Breton, C., Snajdrova, L., Jeanneau, C., Koca, J., and Imberty, A. (2006) Structures and mechanisms of glycosyltransferases, *Glycobiology 16*, 29R−37R.
14. Charnock, S. J., and Davies, G. J. (1999) Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms, *Biochemistry 38*, 6380−6385.
15. Murray, B. W., Takayama, S., Schultz, J., and Wong, C. H. (1996) Mechanism and specificity of human alpha-1,3-fucosyltransferase V, *Biochemistry 35*, 11183−11195.
16. Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A., and Negishi, M. (2000) Heparan/chondroitin sulfate biosynthesis. Structure and mechanism of human glucuronyltransferase I, *J. Biol. Chem. 275*, 34580−34585.
17. Yu, H., Yu, H., Karpel, R., and Chen, X. (2004) Chemoenzymatic synthesis of CMP-sialic acid derivatives by a one-pot two-enzyme system: comparison of substrate flexibility of three microbial CMP-sialic acid synthetases, *Bioorg. Med. Chem. 12*, 6427−6435.
18. Hope, H. (1988) Cryocrystallography of biological macromolecules: a generally applicable method, *Acta Crystallogr. B 44* (Part 1), 22−26.
19. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 50*, 760−763.
20. Navaza, J. (1993) On the computation of the fast rotation function, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 49*, 588−591.
21. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126−2132.
22. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 57*, 122−133.
23. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 53*, 240−255.
24. Vaguine, A. A., Richelle, J., and Wodak, S. J. (1999) SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55*, 191−205.
25. Kajihara, Y., Kamiyama, D., Yamamoto, N., Sakakibara, T., Izumi, M., and Hashimoto, H. (2004) Synthesis of 2-[(2-pyridyl)amino]-ethyl beta-D-lactosaminide and evaluation of its acceptor ability for sialyltransferase: a comparison with 4-methylumbelliferyl and dansyl beta-D-lactosaminide, *Carbohydr. Res. 339*, 1545−1550.
26. Burkart, M. D., Vincent, S. P., Duffels, A., Murray, B. W., Ley, S. V., and Wong, C. H. (2000) Chemo-enzymatic synthesis of fluorinated sugar nucleotide: useful mechanistic probes for glycosyltransferases, *Bioorg. Med. Chem. 8*, 1937−1946.
27. Williams, S. J., and Withers, S. G. (2000) Glycosyl fluorides in enzymatic reactions, *Carbohydr. Res. 327*, 27−46.
28. Hu, Y., and Walker, S. (2002) Remarkable structural similarities between diverse glycosyltransferases, *Chem. Biol. 9*, 1287−1296.
29. Vrielink, A., Ruger, W., Driessen, H. P., and Freemont, P. S. (1994) Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose, *EMBO J. 13*, 3413−3422.
30. Ha, S., Walker, D., Shi, Y., and Walker, S. (2000) The 1.9 Å crystal structure of *Escherichia coli* MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis, *Protein Sci. 9*, 1045−1052.
31. Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R., and Davies, G. J. (2002) Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA, *Chem. Biol. 9*, 1337−1346.
32. Bourne, Y., and Henrissat, B. (2001) Glycoside hydrolases and glycosyltransferases: families and functional modules, *Curr. Opin. Struct. Biol. 11*, 593−600.
33. Unligil, U. M., and Rini, J. M. (2000) Glycosyltransferase structure and mechanism, *Curr. Opin. Struct. Biol. 10*, 510−517.
34. Rossmann, M. G., Moras, D., and Olsen, K. W. (1974) Chemical and biological evolution of nucleotide-binding protein, *Nature 250*, 194−199.
35. Crennell, S., Garman, E., Laver, G., Vimr, E., and Taylor, G. (1994) Crystal structure of *Vibrio cholerae* neuraminidase reveals dual lectin-like domains in addition to the catalytic domain, *Structure 2*, 535−544.
36. Crennell, S., Takimoto, T., Portner, A., and Taylor, G. (2000) Crystal structure of the multifunctional paramyxovirus hemagglutinin-neuraminidase, *Nat. Struct. Biol. 7*, 1068−1074.
37. Crennell, S., Garman, E., Philippon, C., Vasella, A., Laver, W., Vimr, E., and Taylor, G. (1996) The structures of *Salmonella typhimurium* LT2 neuraminidase and its complexes with three inhibitors at high resolution, *J. Mol. Biol. 259*, 264−280.
38. Burmeister, W., Ruigrok, R., and Cusack, S. (1992) The 2.2 Å resolution crystal structure of influenza B neuraminidase and its complex with sialic acid., *EMBO J. 11*, 49−56.
39. Gaskell, A., Crennell, S., and Taylor, G. (1995) The three domains of a bacterial sialidase: a beta-propeller, an immunoglobulin module and a galactose-binding jelly-roll, *Structure 3*, 1197−1205.
40. Luo, Y., Li, S., Chou, M., Li, Y., and Luo, M. (1998) The crystal structure of an intramolecular trans-sialidase with a NeuAc alpha2→3Gal specificity, *Structure 6*, 521−530.
41. Luo, Y., Li, S., Li, Y., and Luo, M. (1999) The 1.8 Å structures of leech intramolecular trans-sialidase complexes: evidence of its enzymatic mechanism, *J. Mol. Biol. 285*, 323−332.
42. Amaya, M., Buschiazzo, A., Nguyen, T., and Alzari, P. (2003) The high resolution structures of free and inhibitor-bound *Trypanosoma rangeli* sialidase and its comparison with *T. cruzi* trans-sialidase, *J. Mol. Biol. 325*, 773−784.
43. Bray, M. R., Johnson, P. E., Gilkes, N. R., McIntosh, L. P., Kilburn, D. G., and Warren, R. A. (1996) Probing the role of tryptophan residues in a cellulose-binding domain by chemical modification, *Protein Sci. 5*, 2311−2318.
44. Din, N., Forsythe, I. J., Burtnick, L. D., Gilkes, N. R., Miller, R. C., Jr., Warren, R. A., and Kilburn, D. G. (1994) The cellulose-binding domain of endoglucanase A (CenA) from *Cellulomonas fimi*: evidence for the involvement of tryptophan residues in binding, *Mol. Microbiol. 11*, 747−755.
45. Lehtio, J., Sugiyama, J., Gustavsson, M., Fransson, L., Linder, M., and Teeri, T. T. (2003) The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules, *Proc. Natl. Acad. Sci. U.S.A. 100*, 484−489.
46. Williamson, M. P., Le Gal-Coeffet, M. F., Sorimachi, K., Furniss, C. S., Archer, D. B., and Williamson, G. (1997) Function of conserved tryptophans in the *Aspergillus niger* glucoamylase 1 starch binding domain, *Biochemistry 36*, 7535−7539.
47. Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2005) Substrate-induced conformational changes in glycosyltransferases, *Trends Biochem. Sci. 30*, 53−62.
48. Gebler, J. C., Trimbur, D. E., Warren, A. J., Aebersold, R., Namchuk, M., and Withers, S. G. (1995) Substrate-induced inactivation of a crippled beta-glucosidase mutant: identification of the labeled amino acid and mutagenic analysis of its role, *Biochemistry 34*, 14547−14553.
49. White, A., Tull, D., Johns, K., Withers, S. G., and Rose, D. R. (1996) Crystallographic observation of a covalent catalytic intermediate in a beta-glycosidase, *Nat. Struct. Biol. 3*, 149−154.

BI700346W