Cytidine 5′-Monophosphate (CMP)-Induced Structural Changes in a Multifunctional Sialyltransferase from *Pasteurella multocida*[†,‡]

Lisheng Ni,[§] Mingchi Sun,[§] Hai Yu,[§] Harshal Chokhawala,[§] Xi Chen,*[,§] and Andrew J. Fisher*[,§,||]

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

Received November 23, 2005; Revised Manuscript Received December 19, 2005

ABSTRACT: Sialyltransferases catalyze reactions that transfer a sialic acid from CMP−sialic acid to an acceptor (a structure terminated with galactose, *N*-acetylgalactosamine, or sialic acid). They are key enzymes that catalyze the synthesis of sialic acid-containing oligosaccharides, polysaccharides, and glycoconjugates that play pivotal roles in many critical physiological and pathological processes. The structures of a truncated multifunctional *Pasteurella multocida* sialyltransferase (Δ24PmST1), in the absence and presence of CMP, have been determined by X-ray crystallography at 1.65 and 2.0 Å resolutions, respectively. The Δ24PmST1 exists as a monomer in solution and in crystals. Different from the reported crystal structure of a bifunctional sialyltransferase CstII that has only one Rossmann domain, the overall structure of the Δ24PmST1 consists of two separate Rossmann nucleotide-binding domains. The Δ24PmST1 structure, thus, represents the first sialyltransferase structure that belongs to the glycosyltransferase-B (GT-B) structural group. Unlike all other known GT-B structures, however, there is no C-terminal extension that interacts with the N-terminal domain in the Δ24PmST1 structure. The CMP binding site is located in the deep cleft between the two Rossmann domains. Nevertheless, the CMP only forms interactions with residues in the C-terminal domain. The binding of CMP to the protein causes a large closure movement of the N-terminal Rossmann domain toward the C-terminal nucleotide-binding domain. Ser 143 of the N-terminal domain moves up to hydrogen-bond to Tyr 388 of the C-terminal domain. Both Ser 143 and Tyr 388 form hydrogen bonds to a water molecule, which in turn hydrogen-bonds to the terminal phosphate oxygen of CMP. These interactions may trigger the closure between the two domains. Additionally, a short helix near the active site seen in the apo structure becomes disordered upon binding to CMP. This helix may swing down upon binding to donor CMP−sialic acid to form the binding pocket for an acceptor.

Sialic acids (2-keto-3-deoxynonulosonic acids) are negatively charged α-keto acids with a nine-carbon backbone. They are commonly found as terminal carbohydrate residues on cell surface glycoconjugates (glycoproteins and glycolipids) of higher animals. As the terminal carbohydrate residue, sialic acid is one of the first molecules encountered in cellular interactions and has been found to play important roles in cellular recognition and communication (*1*, *2*). Cell surface sialic acids have also been found in relatively few microorganisms, mainly pathogenic bacteria, and their presence is often associated with virulence (*1*−*4*). For example, sialic acid-containing capsular polysaccharides of *Neisseria meningitidis*, *Escherichia coli* K1, and group B *Streptococci* are important virulence factors (*5*, *6*). Sialic acids on lipooligosaccharides (LOS[1]) of several mucosal pathogens, including *N. meningitidis*, *Neisseria gonorrhoeae*, *Haemophilus ducreyi*, and *Haemophilus influenzae* strains (*7*−*15*) are considered important LOS components for molecular mimicry to allow them to evade the host's immune system (*8*, *15*−*17*) or as modulators of microbial interactions with host cells (*1*, *2*, *14*, *18*).

Sialyltransferases are key enzymes for the biosynthesis of sialic acid-containing structures. They catalyze reactions that transfer a sialic acid from its activated sugar nucleotide donor, cytidine 5′-monophosphate sialic acid (CMP-sialic acid), to its acceptor, usually a structure terminated with a galactose, an *N*-acetylgalactosamine, or another sialic acid. Sialyltransferases have been isolated and cloned from both

[†] 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/) as entries 2EX0 and 2EX1 for the ligand-free apo and the CMP binary complex, 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.

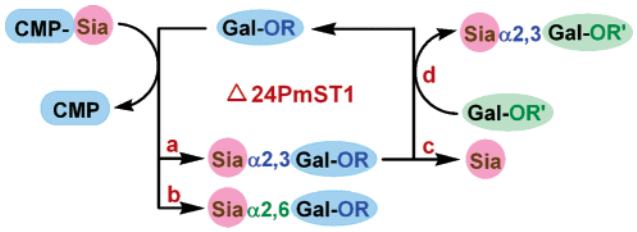
[§] Department of Chemistry, University of California.

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

[1] Abbreviations: ASU, asymmetric unit, CMP, cytidine 5′-monophosphate; DTT, dithiothreitol; GlcNAc, *N*-acetylglucosamine; GPGTF, glycogen phosporylase/glycosyltransferase superfamily; GT, glycosyltransferase; IPTG, isopropyl-1-thio-*β*-D-galactopyranoside; LOS, lipooligosaccharide; MAD, multiple wavelength anomalous dispersion; Neu5Ac; *N*-acetylneuraminic acid; rmsd, root-mean-square deviation; SeMet, Selenomethionine; SSRL, Stanford Synchrotron Radiation Laboratory; UDP, uridine 5′-diphosphate.

10.1021/bi0524013 CCC: $33.50
Published on Web 01/27/2006

Scheme 1: Multiple Functions of Δ24PmST1



eukaryotes and bacteria (*3*, *13*, *19*−*29*). It has been demonstrated that sialyltransferase activity is crucial for bacterial pathogenicity (*1*, *2*). Bacterial sialyltransferases, therefore, are potential drug targets in fighting against bacterial infection. Despite important biological roles of sialyltransferases and great interests in sialyltransferase studies, little is known about the molecular architecture and mechanism of these enzymes (*30*). The only reported sialyltransferase structure was for a recombinant bifunctional sialyltransferase, CstII, from *Campylobacter jejuni* strain OH4384, a highly prevalent food-borne pathogen (*31*).

Previously, we reported the cloning and characterization of a C-terminal His_6-tagged N-terminus truncated multifunctional sialyltransferase from *Pasteurella multocida* strain P-1059 (*32*). In this construct, the N-terminal 24 residues of the full-length protein comprising a putative membrane-spanning helix were removed. The recombinant protein was previously named tPm0188Ph for *t*runcated Pm0188 *P*rotein *h*omolog. Here, we will use Δ24PmST1 (*P. multocida* sialyltransferase 1 with an N-terminal 24 amino acid residues deletion) to represent the construct. The recombinant protein was highly active and was expressed in *E. coli* in high yields. Four types of functions have been found for the Δ24PmST1 (Scheme 1) (*32*). It is (a) an α2,3-sialyltransferase with an optimal activity at pH 7.5−9.0, (b) an α2,6-sialyltransferase with an optimal activity at pH 4.5−6.0, (c) an α2,3-sialidase with an optimal pH of 5.0−5.5, and (d) an α2,3-*trans*-sialidase with an optimal pH of 5.5−6.5. It is unusual that a single enzyme possesses so many different functions. Therefore, it is of extreme interest to elucidate the molecular structure of this protein to help understand its mechanisms of catalysis.

We report herein the crystal structures of the multifunctional sialyltransferase Δ24PmST1 in the absence and the presence of CMP at 1.65 and 2.0 Å resolutions, respectively. The overall structure of Δ24PmST1 consists of two separate Rossmann nucleotide-binding domains that forms a deep nucleotide-binding cleft between the two domains. Comparison of two structures shows a large conformational change of the enzyme upon binding to CMP. This study elucidates a new class of sialyltransferase structure that differs greatly from that of CstII.

EXPERIMENTAL PROCEDURES

Expression and Purification of Pasteurella multocida Multifunctional Sialyltransferase Δ24PmST1. These procedures were performed as previously reported (*32*). Briefly, the sialyltransferase gene was cloned excluding the predicted N-terminal membrane anchor comprising the first 24 residues. *E. coli* BL21 (DE3) harboring the recombinant plasmid in a pET23a(+) vector was grown in LB-rich medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl) containing 100 μg/mL ampicillin until the OD_{600nm} reached 0.8−1.0. The expression of the protein was induced by adding 0.1 mM IPTG (isopropyl-1-thio-β-D-galactopyranoside) followed by incubation at 37 °C for 3 h with vigorous shaking at 250 rpm in a C25KC incubator shaker (New Brunswick Scientific, Edison, NJ). The bacterial cells were harvested by centrifugation at 4 °C in a Sorvall Legend RT centrifuge with a swinging bucket rotor at 3696*g* for 60 min. The cell pellet was resuspended in 20 mL/L cell culture in lysis buffer (pH 8.0, 100 mM Tris-HCl containing 0.1% Triton X-100). Lysozyme (1 mg/L culture) and DNaseI (50 μg/L culture) were then added to the cell suspension. After the mixture was incubated at 37 °C for 50 min with vigorous shaking (250 rpm), the cell lysate was separated from inclusion bodies and other cellular debris by centrifugation (Sorvall RC-5B centrifuge with a S5-34 rotor) at 7000*g* for 30 min. The enzyme was purified directly from the lysate using a Ni^{2+}−NTA affinity column. All purification procedures were performed on ice or at 4 °C. After loading the cell lysate, the Ni^{2+} column was washed with 6 column vol of binding buffer (5 mM imidazole, 0.5 M NaCl, and 20 mM Tris-HCl, pH 7.5), followed by 8 column vol of washing buffer (20 mM imidazole, 0.5 M NaCl, and 20 mM Tris-HCl, pH 7.5). The protein was eluted with 8 column vol of elute buffer (200 mM imidazole, 0.5 M NaCl, and 20 mM Tris-HCl, pH 7.5) and detected by a UV/Vis spectrometer at 280 nm using a Smart-Spec 3000 Spectrophotometer (Bio-Rad, Hercules, CA). Fractions containing the purified enzyme were collected and stored at 4 °C.

Preparation of Selenomethionyl Δ24PmST1. Selenomethionine-containing *P. multocida* (*Pm*) sialyltransferase was prepared as reported with modifications (*33*). Briefly, recombinant *E. coli* BL21 (DE3) were grown in modified M9 minimal media (6.78 g/L Na_2HPO_4, 3.00 g/L KH_2PO_4, 0.50 g/L NaCl, 1.24 g/L $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 4 g/L glucose, trace amount of $ZnSO_4$, $MnSO_4$, H_3BO_3, $NaMoO_4$, $CoCl_2$, KI, $CuSO_4$, $FeSO_4$, and H_2SO_4) (*34*) containing 400 μg/mL ampicillin until the OD_{600nm} reached 0.6−0.8. A sterile mixture of amino acids (100 mg/L of L-lysine, L-phenylalanine, and L-threonine; 50 mg/L of L-isoleucine, L-leucine, and L-valine; and 60 mg/L of L-selenomenthionine) was then added. The expression of the protein was induced 15 min later by addition of 0.1 mM IPTG and incubation at 37 °C for 6 h with vigorous shaking at 250 rpm. The procedures for lysate preparation and protein purification were the same as described above except that 0.2 mM of dithiothreitol (DTT) was added to the lysate and all the buffers used for the purification.

Crystallization of Δ24PmST1. Purified enzyme was dialyzed in Tris-HCl buffer (20 mM, pH 7.5) and concentrated to 10 mg/mL for crystallization. Native protein was crystallized by hanging drop vapor diffusion using 3-μL drops of protein mixed with an equal volume of reservoir buffer (26% poly(ethylene glycol) 3350, 100 mM HEPES, pH = 7.5, 100 mM NaCl, and 0.4% Triton X-100). The crystals grew to full size of 0.3 mm × 0.3 mm × 0.05 mm after 2 weeks. The binary CMP crystals were grown under identical conditions except for addition of 2 mM CMP−*N*-acetylneuraminic acid (CMP−Neu5Ac). Structure determination indicated that the CMP−Neu5Ac was hydrolyzed and the *N*-acetylneuraminic acid was removed, resulting in only CMP binding to the enzyme (see below). The SeMet-derivatized

Table 1: Data Collection, Phasing, and Refinement Statistics

	Se peak	Se remote	Se inflection	native	CMP complex
x-ray source	SSRL BL 9-2	SSRL BL 9-2	SSRL BL 9-2	SSRL BL 9-2	SSRL BL 9-2
wavelength	0.97879 Å	0.91838 Å	0.97931 Å	0.9800 Å	0.97931 Å
resolution	1.80 (1.86−1.80) Å	1.80 (1.86−1.80) Å	1.80 (1.86−1.80) Å	1.65 (1.69−1.65) Å	2.00 (2.05−2.00) Å
space group	$P2_1$			$P2_1$	$P2_1$
cell parameters	$a = 52.72$ $b = 61.54$ $c = 65.05$ $\beta = 111.79°$			$a = 63.72$ $b = 61.28$ $c = 96.65$ $\beta = 101.56°$	$a = 61.25$ $b = 64.57$ $c = 64.84$ $\beta = 98.62°$
monomers/ASU	1			2	1
V_M (Å^3/Da); % solvent	2.1; 41%			2.0; 38%	2.7; 55%
no. of reflections	241355	253767	251778	280823	126172
no. unique	35255	35320	35728	86273	33403
R_{merge}[a](%)	7.2 (20.8)	6.6 (20.2)	6.4 (21.3)	5.2 (30.8)	4.5 (17.0)
mean $(I)/\sigma(I)$	24.9 (7.4)	23.5 (7.3)	23.4 (7.2)	15.1 (2.5)	17.8 (5.4)
completeness (%)	99.6 (99.9)	99.6 (100)	99.5 (99.3)	98.1 (82.1)	98.3 (97.7)
no. of Se sites	6				
figure of merit (solve/resolve)	0.66/0.74				
Refinement Statistics					
resolution (Å)				30−1.65	64.15−2.00
no. reflections ($F \geq 0$)				86273	33403
R-factor[b]				19.3	19.2
R-free[b]				21.9	22.1
rms bond length				0.005	0.006
rms bond angles				1.22	1.24
Ramachandran Plot Statistics					
residues[c]				710	343
most favorable region (%)				90.6	91.5
allowed region (%)				9.0	7.3
generously allowed region (%)				0.3	0.9
disallowed (%)				0.1	0.3
Asymmetric Unit Content					
nonhydrogen protein atoms				6365	3110
water				841	398
CMPs (atoms)					1 (21)

[a] $R_{merge} = [\sum_h \sum_i |\boldsymbol{I}_h - \boldsymbol{I}_{hi}| / \sum_h \sum_i \boldsymbol{I}_{hi}]$ where $\boldsymbol{I}_h$ is the mean of $\boldsymbol{I}_{hi}$ observations of reflection *h*. Numbers in parentheses represent highest resolution shell. [b] *R*-factor and $R_{free} = \sum||F_{obs}| - |F_{calc}||/\sum|F_{obs}| \times 100$ for 95% of recorded data (*R*-factor) or 5% of data (R_{free}). Numbers in parentheses represent highest resolution shell. [c] Number of nonproline and nonglycine residues used for calculation.

protein was crystallized under identical conditions as for the native protein. All crystals were transferred to Paratone-N and frozen in a stream of nitrogen to −173 °C for data collection.

Data Collection and Phase Determination. X-ray data from all three crystals (Native, SeMet, and CMP-bound) were collected on beam line 9-2 at Stanford Synchrotron Radiation Laboratory (SSRL). Although all three crystals were grown under identical conditions, the three crystallized in space group $P2_1$ but with different cell parameters and crystal packing (Table 1). The data were indexed and integrated with DENZO and were reduced with SCALEPACK (*35*). Complete data sets were collected to 1.65, 1.8, and 2.0 Å resolution for the native, SeMet, and CMP complex, respectively. For the native crystals, the Matthews coefficient V_M (*36*) was calculated to be 2.0 Å^3/Da assuming two monomers per asymmetric unit (ASU), which corresponds to a solvent content in the crystal of ∼38%. The SeMet crystals had one monomer per ASU with a Matthews coefficient of V_M = 2.1 Å^3/Da and a solvent content of ∼41%, while the CMP complex also had one monomer per ASU ($V_M = 2.7$ Å^3/Da, solvent content of ∼55%).

Multiwavelength anomalous dispersion (MAD) data were collected on beamline 9-2 at SSRL. Excellent 1.8 Å resolution data sets were collected at three wavelengths corresponding to the Selenium absorption edge (peak and inflection point) as well as a high-energy remote wavelength (Table 1). The three data sets were input into the program SOLVE (*37*), which was able to find 6 of the 7 selenium sites. The overall Figure of Merit (FOM) from SOLVE was 0.66. The phases were improved by solvent flattening using the program RESOLVE (*38*, *39*), which improved the FOM to 0.74 for all data to 1.8 Å resolution. The SeMet data with phases were then input into the automated protein model-building program ARP/wARP (*40*), which was able to build the entire protein structure with the exception of residues 98−107 and 145−154, with a final *R*-factor of 20.2%. This 1.8 Å resolution structure was used to solve the structure of native sialyltransferase using the 1.65 Å resolution data by molecular replacement using the program AMoRe (*41*). The SeMet structure was also used as a phasing model to solve the binary CMP structure by molecular replacement. However, initial molecular replacement attempts were not successful using the complete enzyme structure. A solution, however, was achieved by breaking the apo-enzyme structure into two separate search domains, which corresponded to the two separate Rossmann domains. The solution revealed a significant closure between the two domains upon binding to CMP (see below).

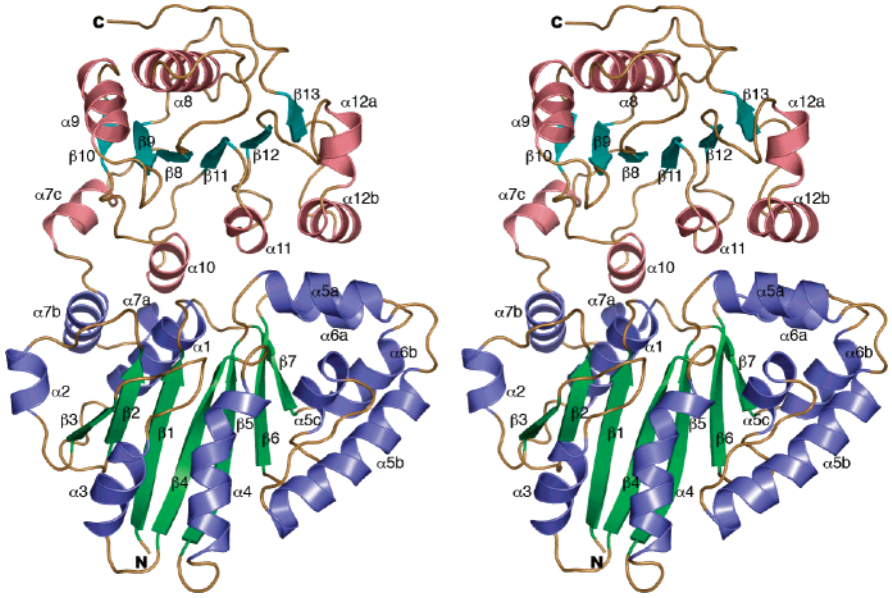


FIGURE 1: Stereoview ribbon drawing of the ligand-free Δ24PmST1 structure. The N-terminal domain is colored with blue α-helices and green β-strands. The C-terminal domain is colored with salmon helices and teal β-strands. Intervening loops are colored tan in both domains. The secondary structural elements are labeled. The α-helices are numbered after the β-strand number that precedes it, and loops between β-strands that have multiple helices are sequentially labeled with letters a, b, etc. All figures except Figure 2C were generated with the program PyMol (http://www.pymol.org).

Model Building and Refinement. Model building was carried out with the molecular graphics program ***O*** (*42*). Energy minimization-coupled coordinate and *B*-factor refinement were carried out with the program CNS (*43*) using 95% of the data to the respective maximum resolution. During each stage of refinement, the *R*-free and *R*-factor decreased. Noncrystallographic restraints for the native ligand-free structure were released during later stages of refinement resulting in a lower *R*-free. The quality of the model was checked with the program PROCHECK (*44*, *45*), and results are summarized in Table 1. Only Ala 219 falls in the disallowed region of the Ramachandran plot in all three models (two apo molecules in ASU and one binary structure). The electron density map clearly defines the same conformation in all three models ($\varphi \approx 45°$, $\psi \approx 120°$). The reason for the unfavorable main-chain angles is that Ala 219 falls in a tight turn between α6b and β7, which then proceeds into the C-terminal domain. The native and the CMP-bound atomic coordinates have been deposited in the protein data bank, accession numbers 2EX0 and 2EX1, respectively.

RESULTS AND DISCUSSION

Overall Native Structure. The sialyltransferase from *P. multocida* is likely anchored into the membrane by a predicted membrane-spanning domain consisting of an N-terminal 24 residue stretch of basic and hydrophobic residues (*32*). Other sialyltransferases are also predicted to be anchored to a membrane, through either an N-terminal or a C-terminal tail (*31*, *46*, *47*). A soluble construct was made by deleting the first 24 amino acid residues and replacing Ser 25 with the Met start codon (Δ24PmST1), which was expressed and crystallized. The monomer is made up of two separate nonidentical α/β/α domains whose structures and topologies are similar to the Rossmann nucleotide-binding domain (*48*). There is a deep cleft that forms between the two domains, which comprises the CMP binding site (see below). Thus, the PmST1 structure belongs to the glycosyltransferases B (GT-B) structural group (*46*, *49*−*53*), which has more diversity in its substrates and products compared to the glycosyltransferases A family. The PmST1 structure, therefore, differs from the only other sialyltransferase structure reported to date, the CstII sialyltransferase from *C. jejuni* (*31*). The tetrameric CstII structure belongs to glycosyltransferases-A (GT-A) or GT-A-like structural group (*46*, *49*−*54*). In the GT-A structures, each monomer consists of a single Rossmann domain and a smaller lid-like structure that folds over the active site (*31*). However, CstII lacks the DXD motif found in GT-A structures, which is known to coordinate divalent cations involved in binding to the sugar−nucleotide diphosphate moiety (*31*, *55*, *56*).

A ribbon diagram of the native ligand-free Δ24PmST1 structure is shown in Figure 1. The N-terminal Rossmann-like domain (residues 25−246) consists of a central seven-stranded parallel β-sheet (β3, β2, β1, β4, β5, β6, and β7) flanked by four α-helices on the backside (as viewed in Figure 1, α1, α2, α7a, and α7b) and seven α-helices on the front side (α3, α4, α5a, α5b, α5c, α6a, and α6b). The smaller C-terminal domain (residues 247−416) is also a α/β/α fold with a central six-stranded parallel β-sheet (β10, β9, β8, β11, β12, and β13) sandwiched by two α-helices on one side (top in Figure 1, α8 and α9) and five α-helices on the other side (α7c, α10, α11, α12a, and α12b). This domain also has a topology of the Rossmann nucleotide-binding domain. The genomic sequence of wild-type PmST1 ends at Leu 412, but the first four residues of the His-tag linker (413−416) of the recombinant Δ24PmST1 are visible in the electron density map and were modeled in extended conformation over the top of the structure opposite to the N-terminal domain.

The quaternary structure of Δ24PmST1 consists of a monomer with overall dimensions of ~73 Å × 54 Å × 45

Å. Even though the native ligand-free enzyme crystallized with two monomers in the asymmetric unit, inspection of crystal packing revealed there is little interaction between the two monomers or with other symmetry-related molecules. Size exclusion chromatography also suggested a monomeric arrangement in solution (data not shown). There is very little structural difference between the two monomers in the crystallographic asymmetric unit. The two monomers superimpose with a root-mean-squared deviation (rmsd) of 0.75 Å for 390 equivalent α-carbons. The largest structural difference is in the loop between α5b and α5c. Residues Thr 174−Lys 178 in the B monomer are pulled away from the N-terminal core slightly to make a lattice contact with a symmetry-related B monomer. The α-carbon of Gly 175 shifts 5.0 Å between the two monomers. Superposition of the individual domains independently resulted in a slightly lower rmsd of 0.64 Å for both respective domains, suggesting a minor difference in the orientation angle between the two domains and mobility between the two domains (see CMP binding below).

CMP Binary Structure. To understand the structural requirements for donor substrate binding and obtain clues for catalytic activity, Δ24PmST1 was cocrystallized with the sialyltransferase donor CMP−Neu5Ac. Inspection of the resultant electron density map, however, only revealed the CMP moiety with several well-ordered water molecules binding in the deep cleft between the two nucleotide-binding domains. It is likely that during the time required for crystallization, the enzyme hydrolyzed off the Neu5Ac moiety from CMP−Neu5Ac and only CMP was left bound to the enzyme. This phenomenon is very similar to that reported for CstII, for which crystallization of the CstII sialyltransferase in the presence of CMP−Neu5Ac also resulted in only observing CMP in the active site (*31*).

The initial CMP structure could not be solved by molecular replacement using the SeMet ligand-free structure as a search model. Nevertheless, a molecular replacement solution was found by dividing the ligand-free structure into two separate search models, each corresponding to an individual Rossmann domain. Indeed, inspection of the solution revealed a significant closure of the two domains over the CMP binding pocket in the deep cleft (Figure 2A) (see below), explaining the failure of the single-model molecular replacement search. In the final refined structure, residues Met 25−Lys 375 and Asn 385−Leu 412 are well-ordered in the electron density map. The stretch of nine disordered residues, Gln 376−Leu 384, corresponding to helix α12a, could be involved in acceptor binding (see below).

CMP makes extensive hydrogen bonds with the C-terminal nucleotide-binding domain of the enzyme, and no contacts are made with the N-terminal domain. A total of 12 hydrogen bonds develop between the enzyme C-terminal domain and atoms from the base, ribose, and phosphate of CMP (Figures 2B,C). Additionally, three ordered water molecules are in hydrogen-bonding distance to each of the terminal phosphate oxygens (Figures 2B,C). The specificity for the cytosine base comes from two hydrogen bonds between N4 of the base and the main-chain carbonyl oxygens of Gly 266 and Lys 309. In comparison, a uracil (or a thymine) base contains a deprotonated carbonyl O4, which is incapable of serving as a hydrogen donor to the main-chain carbonyl oxygens. Additionally, Nζ of Lys 309 donates a hydrogen bond to unprotonated N3 of cytosine, while in uracil (or thymine), N3 is protonated and would be unable to accept this hydrogen bond. The O2 in the cytosine ring accepts hydrogen bonds from Nζ of Lys 309 and main-chain N of Phe 337. There is insufficient space to accommodate a purine base. Additionally, unlike that in the CstII sialyltransferase structure, there is no aromatic stacking of the cytosine ring with any amino acid residues in Δ24PmST1.

Interactions between Δ24PmST1 and the ribose−phosphate moiety of CMP display some resemblance to UDP binding in UDP−GlcNAc 2-epimerase (*57*), its closest structural homologue (see below). In Δ24PmST1, the carboxyl group of Glu 338 makes a parallel bidentate hydrogen-bond pair with both hydroxyl groups (2′ and 3′) of the ribose in CMP, reminiscent of Glu 296 in UDP−GlcNAc 2-epimerase. In Δ24PmST1, the ribose ring is in a slight C2′-exo conformation, but almost planar, while in UDP−GlcNAc 2-epimerase, the sugar pucker is more pronounced and in the C2′-endo-C3′-exo conformation. The phosphate moiety of CMP is positioned at the positive dipole moment at the N-terminus of helix α11, similar to the position of the β-phosphate of UDP in UDP−GlcNAc 2-epimerase, which lies off the N-terminus of the topologically same helix. In both enzymes, the phosphate makes extensive interactions with both protein and solvent. In Δ24PmST1, each terminal phosphate oxygen hydrogen-bonds to both a solvent molecule and a hydrophilic side chain of the enzyme. CMP phosphate oxygen O1P hydrogen-bonds to Nε2 of His 311 and Wat 29. O2P hydrogen-bonds to both Oγ and N from Ser 356 and Wat 1. And finally, O3P hydrogen-bonds to Oγ of Ser 355 and Wat 11 (Figure 2B,C). His 311 and Ser 355 are analogous to His 213 and Ser 290 in UDP−GlcNAc 2-epimerase, which make similar hydrogen bonds to the UDP β-phosphate. These residues, His 311, Ser 355, and Ser 356 in Δ24PmST1, could play catalytic roles in one or more of the measured activities of this enzyme. In particular, His 311, which stabilizes the phosphate, may serve as a general acid and protonate the departing phosphate oxygen of CMP. This histidine is comparable to His 213 in UDP−GlcNAc 2-epimerase, which hydrogen-bonds to a β-phosphate oxygen of UDP and is hypothesized to serve as a general acid (*57*).

Apo and Binary Structure Comparison. Comparison of the ligand-free and the binary CMP−enzyme structures reveals a large N-terminal domain movement upon the enzyme binding to CMP. However, it is possible that the domain shift occurred during or after CMP−Neu5Ac binding or hydrolysis. This domain closure is a result of substrate or CMP binding and not of crystal packing, because comparison of the apo SeMet crystal structure to the apo native enzyme (natural amino acids) structure revealed little structural differences and an rmsd of only 0.486 and 0.345 Å (A and B subunits, respectively) between the two apo forms, which pack in diverse crystal arrangements.

Since the CMP moiety only contacts the C-terminal domain, the C-terminal domain was thus used as a fixed point for superposition and comparison of the CMP-bound and the native apo structures. When the two structures are overlaid at their C-terminal domain, there is a substantial rotation of the N-terminal domain upon binding CMP (Figure 3A). The N-terminal domain rotates ∼23° up toward the CMP binding site to close up the cleft between the two Rossmann domains. This results in a movement of >18 Å

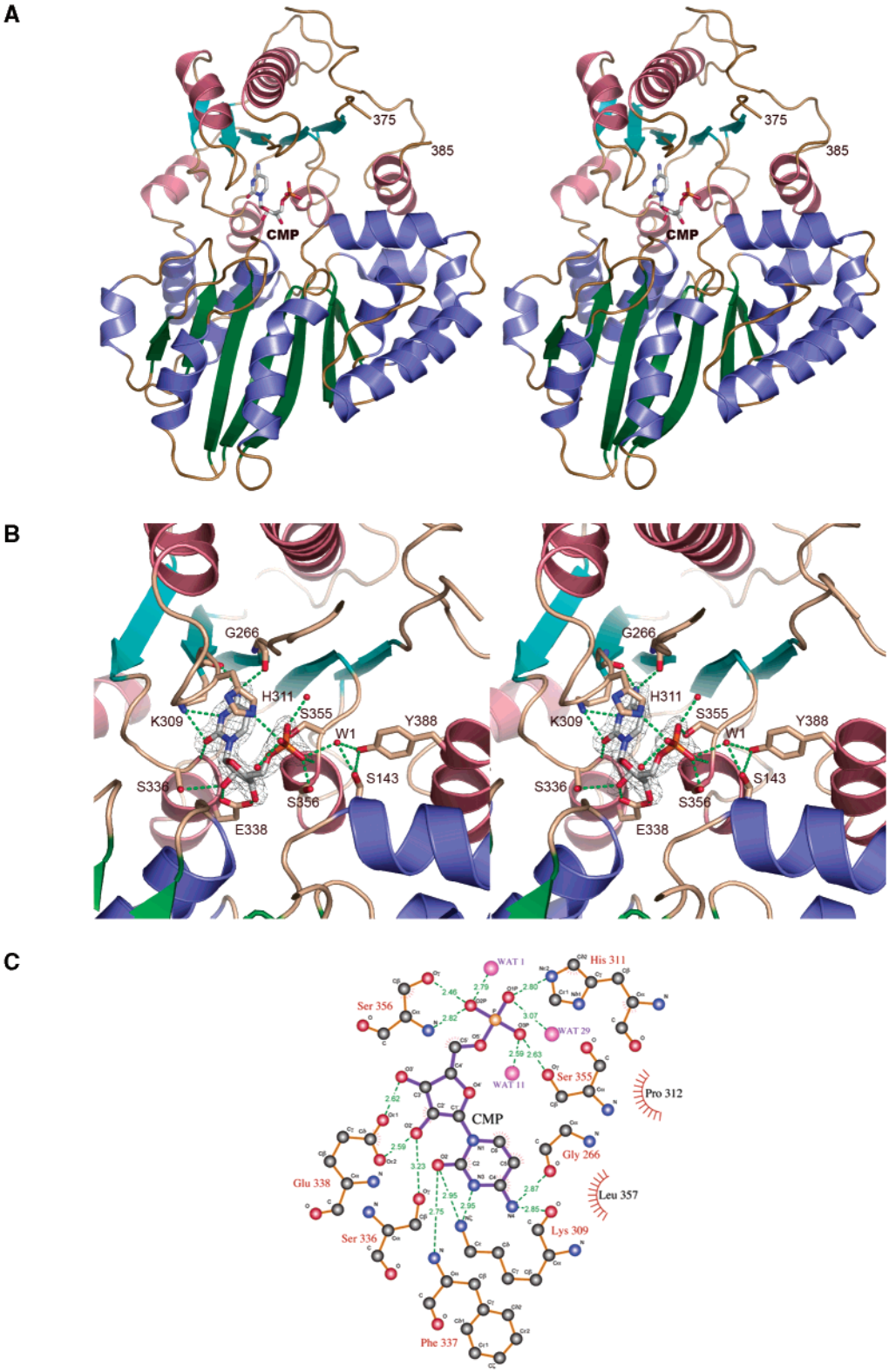


FIGURE 2: Binary CMP-bound Δ24PmST1 structure. (A) Stereo-ribbon diagram revealing the CMP-bound (stick model with white-colored carbons) in the deep cleft between the two Rossmann domains. Coloring of the protein is the same as in Figure 1. Residues 375 and 385 mark the region of helix α12a, which becomes disordered upon binding CMP. (B) Close-up stereoview of the CMP binding site. The CMP is drawn as a stick model with white-colored carbon atoms (nitrogen, oxygen, and phosphorus atoms are colored blue, red, and orange, respectively). The amino acid side chains of residues that interact with CMP are also drawn as a stick model with tan-colored carbon atoms. Three water molecules that hydrogen-bond to terminal phosphate oxygens in CMP are represented as small red spheres. Many other well-ordered solvent molecules in the active site are not shown for clarity. Potential hydrogen bonds are drawn as green dashes. Electron density corresponding to a simulated-annealing omit map is drawn in gray mesh around the CMP moiety contoured at 5σ. (C) Flatten two-dimensional representation showing the interactions between CMP and protein or solvent atoms. Covalent bonds are drawn as solid purple lines for CMP and tan lines for protein residues. Hydrogen bonds are drawn as green dashed lines with numbers corresponding to the distances in angstroms. van der Waals contacts are represented by hashed red lines around an atom, which point to the direction of the corresponding atom (also with red hashed lines) with which it interacts. Panel C was generated with the program LIGPLOT (*68*).

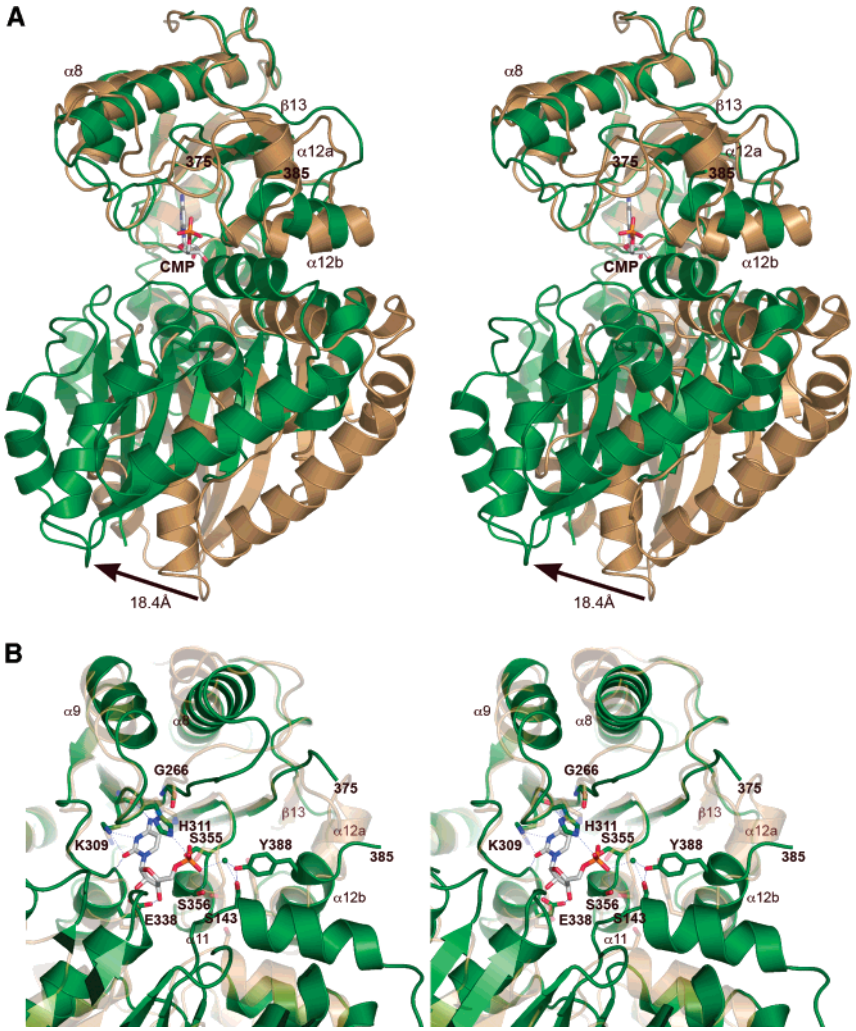


FIGURE 3: Conformational changes in Δ24PmST1 upon binding to CMP. (A) Stereo ribbon drawing of the ligand-free apo structure drawn in tan and the CMP-bound binary structure drawn in green with CMP drawn as stick model with white-colored carbons. The structures were superimposed at the C-terminal CMP binding domain to illustrate the large conformational change upon binding to CMP. The C-terminal domains superimpose with an rmsd value of 2.41 Å for 156 equivalent α-carbons. The N-terminal domain rotates ~23° resulting in a movement of >18 Å for some loops in a direction shown by the arrow. (B) Close-up stereoview of the active site showing the movement of Δ24PmST1 upon binding to CMP. The CMP binary structure is drawn in green and the ligand-free apo structure is drawn semitransparent in tan color. Tyr 388 in helix α12b moves up and flips around to hydrogen-bond to Ser 143 and Wat 1 (green sphere). Upon binding to CMP, helix α11 moves slightly away from Tyr 388 to accommodate this large movement. Ser 143 in the N-terminal domain moves 5.2 Å to hydrogen-bond to both Tyr 388 and Wat 1. This interaction could stabilize the conformational change seen in the CMP-bound structure. Residues 375−385 correspond to the disordered helix α12a in the CMP-bound state, which could be involved in sugar binding that would likely occur to the right of the CMP phosphate as shown.

for some residues at the distal end of the N-terminal domain (Figure 3A). The pivot point is Asp 249 in the loop linking the two domains.

Comparison of the individual N- or C-terminal domains reveals that significant structural changes have also occurred within the C-terminal domain upon binding CMP, but little difference is seen in the N-terminal domain. The rmsd values for overlaying the C-terminal domains between the CMP-bound and the apo structures are 2.41 and 2.48 Å for A and B apo subunits, respectively (156 equivalent α-carbons). In contrast, the rmsd values for the superposition of the N-terminal domains alone are only 0.55 and 0.72 Å for the A and B apo subunits, respectively (223 equivalent α-carbons).

The significant differences observed when comparing the two C-terminal domain structures center around helix α12a in the C-terminal domain. This helix, which is well-defined in the ligand-free structure, becomes completely disordered in the CMP-bound structure, indicating its mobility in the latter (Figure 3B). The conformational flexibility of this helix suggests that it may swing down and interact with sugar substrates. The α12a helix contains many hydrophilic residues (including three lysines, an aspartate, a glutamate, a serine, an asparagine, and a glutamine), which may be involved in binding to the sialic acid residue in donor CMP−sialic acid and/or the acceptor sugar and possibly contribute residues for catalysis. This is reminiscent of the CstII sialyltransferase structures in which a small loop was disordered in the CMP complex but became ordered upon binding to a CMP−sialic acid analogue (*31*).

In addition to the disorder of helix α12a, helix α12b rotates up away from the N-terminal domain (Figure 3B). This rotation is due to the flipping of Tyr 388 in helix α12b. In the CMP structure, the phenolic oxygen (Oη) of Tyr 388 hydrogen-bonds to ordered water molecule Wat 1, which in turn hydrogen-bonds to the terminal phosphate oxygen of

CMP. In the apo structure, there would be insufficient room for this flipping because of the position of helix α11. When CMP binds to the enzyme, the phosphate in CMP binds to Ser 355 and Ser 356 at the N-terminal end of helix α11 and pulls it slightly away from helix α12b, thus, creating room for Tyr 388 to flip. Other significant C-terminal domain structural differences include the movement of β13, which can no longer hydrogen-bond with β12 because of the rotation of helix α12b. Additionally, helix α8 also rotates down toward the bound CMP. This movement is likely facilitated by the hydrogen bonding of main-chain carbonyl oxygen of Gly 266 (in the β8−α8 loop) to the N4 of the CMP cytosine base.

One must investigate the nature of the closing of the two Rossmann domains in Δ24PmST1 upon its binding to CMP, since it does not interact with the N-terminal domain (residues 25−246). Therefore, what induces the N-terminal domain to move? Inspection of the CMP-bound structure reveals that Ser 143 of helix α5a moves up 5.2 Å in the CMP-bound structure where the Oγ hydrogen-bonds to both Water 1 and the Oη of Tyr 388. Water 1 in turn hydrogen-bonds to the terminal phosphate oxygen in the CMP moiety. The side chain of Tyr 388 flips ∼180°, and the Cα moves 5.5 Å causing helix α12b to shift up upon CMP binding (Figure 3B). Therefore, it is likely that this water molecule may occupy the same position of a glycosidic oxygen in CMP−Neu5Ac. A similar conformational change, thus, may be induced in Δ24PmST1 upon its binding to CMP−Neu5Ac.

Comparison to Other GT-B Structures. A search of the PDB database for other structures with similar folds revealed that the closest match was bacterial UDP-*N*-acetylglucosamine 2-epimerase (*57*) (PDB ID, 1f6d). The DALI search resulted in a *Z* score of 10.3 (*58*). The rmsd between the two enzymes is 3.8 Å based on the structural alignment of 248 α-carbons. Despite the structural similarity, these two enzymes only share 9% sequence identity. Both structures belong to the glycosyltransferase group B (GT-B) structural family and have very similar topologies. Both have an N-terminal Rossmann nucleotide-binding domain consisting of a central seven-stranded parallel β-sheet sandwiched between α-helices and a C-terminal six-stranded nucleotide-binding domain. However, the Δ24PmST1 structure contains additional helices inserted between some of the β-strands. While the UDP−GlcNAc 2-epimerase structure only has a single helix between each β-strand except for the interdomain linkage, the Δ24PmST1 has three helices between β5 and β6 (α5a, α5b, and α5c), two helices between β6 and β7 (α6a and α6b), and two helices between β12 and β13 (α12a and α12b). Another major divergence in topology between these two structures is that the bacterial epimerase contains a C-terminal helix that extends out from the C-terminal nucleotide-binding domain and interacts extensively with the N-terminal domain (*57*). This interaction, which is uniquely absent in the Δ24PmST1 structure, appears to be common to all other members of the GT-B family whose structures are known (*49*, *59*). In Δ24PmST1, the absence of this C-terminal helix tethering to the N-terminal domain may allow for greater flexibility of the substrate binding pocket between the two Rossmann domains to accommodate a wider variety of substrates and catalytic activities as measured for this enzyme (*32*).

The second highest structural similarity DALI score (*Z* score = 9.3) corresponds to the structure of the UDP−glucosyltransferase GtfB that modifies the heptapeptide aglycone in biosynthesis of vancomycin group antibiotics (*60*) (PDB ID, 1iir). The rmsd between this structure and that of Δ24PmST1 is 4.9 Å for 260 equivalent α-carbons, for which it shares only 6% identity in sequence. GtfB also contains two domains with similar topologies, but GtfB binds UDP−glucose. However, since this structure was determined in the absence of nucleotide, it is difficult to compare the active sites. Given that these enzymes have similar structures and bind sugar nucleotides, an evolutionary link is suggested between this family of sialyltransferases and the bacterial epimerase and glucosyltransferase.

The DALI search did not identify any significant correlation between the *P. multocida* sialyltransferase reported here and the only other sialyltransferase whose structure is known, CstII from *C. jejuni* (*31*). This was expected, since the CstII structure belongs to the GT-A structural family, which consists of a single Rossmann nucleotide-binding domain with a small lid inserted in the center of the Rossmann domain (between β4 and β5) involved in substrate binding (*31*). It was a surprise, though, that even DALI searches using each individual Rossmann domain of the Δ24PmST1 structure did not result in any homology with a *Z* score greater than 2.0 to the CstII structure.

In conclusion, the crystal structure of Δ24PmST1 represents the first GT-B structure type sialyltransferase and a new member of the GT-B glycosyltransferase family. Despite their significant structure similarity, GT-B-type glycosyltransferases have great sequence diversity and dramatic substrate differentiation (*49*). For example, T4 phage β-glucosyltransferase (BGT) (*50*, *61*) transfers glucose from UDP−glucose (donor substrate) to hydroxymethylcytosine on duplex DNA (acceptor). *E. coli* MurG, another GT-B glycosyltransferase, transfers GlcNAc from UDP−GlcNAc (donor) to the C4 position of a lipid-linked *N*-acetylmuramyl peptide (acceptor) (*46*, *62*). Additionally, GtfB transfers glucose from UDP−glucose to heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics (*60*). Thus, the donor substrates of the GT-B glycosyltransferases diverge from UDP−glucose, UDP−GlcNAc, to CMP−sialic acid. The acceptor substrates vary dramatically from duplex DNA, lipid-linked glycopeptide, peptide, to oligosaccharide. A common feature shared among these GT-B-type glycosyltransferases is that they belong to microbial or bacteriophage inverting type glycosyltransferases (*63*). Because of their structural similarity to GT-B glycosyltransferases, both UDP−GlcNAc 2-epimerase (*57*) and glycogen phosphorylase (*64*−*66*) have been categorized together with GT-B glycosyltransferases into a superfamily named GPGTF (glycogen phosphorylase/glycosyltransferase) superfamily (*67*). There is excellent structural conservation between protein members in this superfamily, particularly in the C-terminal nucleotide-binding domain. The GPGTF topology, with a large cleft between the two Rossmann domains, appears to be well-adapted to accommodate substrate molecules of varied sizes and diverse structures.

REFERENCES

1. Angata, T., and Varki, A. (2002) Chemical diversity in the sialic acids and related α-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. 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.
4. Smith, H., Parsons, N. J., and Cole, J. A. (1995) Sialylation of neisserial lipopolysaccharide: a major influence on pathogenicity, *Microb. Pathog. 19*, 365−377.
5. Weisgerber, C., Hansen, A., and Frosch, M. (1991) Complete nucleotide and deduced protein sequence of CMP−NeuAc: poly-alpha-2,8 sialosyl sialyltransferase of *Escherichia coli* K1, *Glycobiology 1*, 357−365.
6. Frosch, M., Edwards, U., Bousset, K., Krausse, B., and Weisgerber, C. (1991) Evidence for a common molecular origin of the capsule gene loci in gram-negative bacteria expressing group II capsular polysaccharides, *Mol. Microbiol. 5*, 1251−1263.
7. Vimr, E., Lichtensteiger, C., and Steenbergen, S. (2000) Sialic acid metabolism's dual function in *Haemophilus influenzae*, *Mol. Microbiol. 36*, 1113−1123.
8. Vimr, E., and Lichtensteiger, C. (2002) To sialylate, or not to sialylate: that is the question, *Trends Microbiol. 10*, 254−257.
9. Schilling, B., Goon, S., Samuels, N. M., Gaucher, S. P., Leary, J. A., Bertozzi, C. R., and Gibson, B. W. (2001) Biosynthesis of sialylated lipooligosaccharides in *Haemophilus ducreyi* is dependent on exogenous sialic acid and not mannosamine. Incorporation studies using *N*-acylmannosamine analogues, *N*-glycolylneuraminic acid, and ^{13}C-labeled *N*-acetylneuraminic acid, *Biochemistry 40*, 12666−12677.
10. Mansson, M., Bauer, S. H., Hood, D. W., Richards, J. C., Moxon, E. R., and Schweda, E. K. (2001) A new structural type for *Haemophilus influenzae* lipopolysaccharide. Structural analysis of the lipopolysaccharide from nontypeable *Haemophilus influenzae* strain 486, *Eur. J. Biochem. 268*, 2148−2159.
11. 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.
12. 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.
13. 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.
14. 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.
15. Mandrell, R. E., Griffiss, J. M., and Macher, B. A. (1988) Lipooligosaccharides (LOS) of *Neisseria gonorrhoeae* and *Neisseria meningitidis* have components that are immunochemically similar to precursors of human blood group antigens. Carbohydrate sequence specificity of the mouse monoclonal antibodies that recognize crossreacting antigens on LOS and human erythrocytes, *J. Exp. Med. 168*, 107−126.
16. Campagnari, A. A., Spinola, S. M., Lesse, A. J., Kwaik, Y. A., Mandrell, R. E., and Apicella, M. A. (1990) Lipooligosaccharide epitopes shared among gram-negative non-enteric mucosal pathogens, *Microb. Pathog. 8*, 353−362.
17. Mandrell, R. E., and Apicella, M. A. (1993) Lipo-oligosaccharides (LOS) of mucosal pathogens: molecular mimicry and host-modification of LOS, *Immunobiology 187*, 382−402.
18. Harvey, H. A., Porat, N., Campbell, C. A., Jennings, M., Gibson, B. W., Phillips, N. J., Apicella, M. A., and Blake, M. S. (2000) Gonococcal lipooligosaccharide is a ligand for the asialoglycoprotein receptor on human sperm, *Mol. Microbiol. 36*, 1059−1070.
19. Tsuji, S., Datta, A. K., and Paulson, J. C. (1996) Systematic nomenclature for sialyltransferases, *Glycobiology 6*, v−vii.
20. Tsuji, S. (1996) Molecular cloning and functional analysis of sialyltransferases, *J. Biochem. (Tokyo) 120*, 1−13.
21. Harduin-Lepers, A., Recchi, M. A., and Delannoy, P. (1995) 1994, the year of sialyltransferases, *Glycobiology 5*, 741−758.
22. Wakarchuk, W. W., Watson, D., St Michael, F., Li, J., Wu, Y., Brisson, J. R., Young, N. M., and Gilbert, M. (2001) Dependence of the bi-functional nature of a sialyltransferase from *Neisseria meningitidis* on a single amino acid substitution, *J. Biol. Chem. 276*, 12785−12790.
23. Steenbergen, S. M., Wrona, T. J., and Vimr, E. R. (1992) Functional analysis of the sialyltransferase complexes in *Escherichia coli* K1 and K92, *J. Bacteriol. 174*, 1099−1108.
24. Yamamoto, T., Nakashizuka, M., Kodama, H., Kajihara, Y., and Terada, I. (1996) Purification and characterization of a marine bacterial beta-galactoside alpha 2,6-sialyltransferase from *Photobacterium damsela* JT0160, *J. Biochem. (Tokyo) 120*, 104−110.
25. Gilbert, M., Watson, D. C., Cunningham, A. M., Jennings, M. P., Young, N. M., and Wakarchuk, W. W. (1996) Cloning of the lipooligosaccharide alpha-2,3-sialyltransferase from the bacterial pathogens *Neisseria meningitidis* and *Neisseria gonorrhoeae*, *J. Biol. Chem. 271*, 28271−28276.
26. Bozue, J. A., Tullius, M. V., Wang, J., Gibson, B. W., and Munson, R. S., Jr. (1999) *Haemophilus ducreyi* produces a novel sialyltransferase. Identification of the sialyltransferase gene and construction of mutants deficient in the production of the sialic acid-containing glycoform of the lipooligosaccharide, *J. Biol. Chem. 274*, 4106−4114.
27. Harduin-Lepers, A., Vallejo-Ruiz, V., Krzewinski-Recchi, M. A., Samyn-Petit, B., Julien, S., and Delannoy, P. (2001) The human sialyltransferase family, *Biochimie 83*, 727−737.
28. Datta, A. K., and Paulson, J. C. (1997) Sialylmotifs of sialyltransferases, *Indian J. Biochem. Biophys. 34*, 157−165.
29. Nakayama, J., Angata, K., Ong, E., Katsuyama, T., and Fukuda, M. (1998) Polysialic acid, a unique glycan that is developmentally regulated by two polysialyltransferases, PST and STX, in the central nervous system: from biosynthesis to function, *Pathol. Int. 48*, 665−677.
30. Williams, S. J., and Davies, G. J. (2001) Protein−carbohydrate interactions: learning lessons from nature, *Trends Biotechnol. 19*, 356−362.
31. 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.
32. Yu, H., Chokhawala, H., Karpel, R., Yu, H., 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.
33. Doublie, S. (1997) Preparation of selenomethionyl proteins for phase determination, in *Methods in Enzymology*, (Carter, C. W. J., and Sweet, R. M., Eds.) Vol. 276, pp 523−530, Academic Press, New York.
34. Sambrook, J. R., and Russell, D. W. (2001) *Molecular cloning: A laboratory manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
35. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, in *Methods in Enzymology*, (Carter, C. W. J., and Sweet, R. M., Eds.) Vol. 276, pp 307−326, Academic Press, New York.
36. Matthews, B. W. (1968) Solvent content of protein crystals, *J. Mol. Biol. 33*, 491−497.
37. Terwilliger, T. C., and Berendzen, J. (1999) Automated structure solution for MIR and MAD, *Acta Crystallogr. D55*, 849−861.
38. Terwilliger, T. C. (1999) Reciprocal-space solvent flattening, *Acta Crystallogr. D55*, 1863−1871.
39. Terwilliger, T. C. (2000) Maximum likelihood density modification, *Acta Crystallogr. D56*, 965−972.
40. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol. 6*, 458−463.
41. Navaza, J. (1993) On the computation of the fast rotation function, *Acta Crystallogr. D49*, 588−591.
42. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. (1991) Improved methods for building protein models in electron density

maps and the location of errors in these models, *Acta Crystallogr. A47* (Pt. 2), 110−119.
43. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D 54* (Pt. 5), 905−921.
44. Ramakrishnan, C., and Ramachandran, G. N. (1965) Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units, *Biophys. J. 5*, 909−933.
45. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) Procheck−a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
46. 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.
47. Hodson, N., Griffiths, G., Cook, N., Pourhossein, M., Gottfridson, E., Lind, T., Lidholt, K., and Roberts, I. S. (2000) Identification that KfiA, a protein essential for the biosynthesis of the *Escherichia coli* K5 capsular polysaccharide, is an alpha−UDP-GlcNAc glycosyltransferase. The formation of a membrane-associated K5 biosynthetic complex requires KfiA, KfiB, and KfiC, *J. Biol. Chem. 275*, 27311−27315.
48. Rossmann, M. G., Moras, D., and Olsen, K. W. (1974) Chemical and biological evolution of nucleotide-binding protein, *Nature 250*, 194−199.
49. Hu, Y., and Walker, S. (2002) Remarkable structural similarities between diverse glycosyltransferases, *Chem. Biol. 9*, 1287−1296.
50. 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.
51. 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.
52. Bourne, Y., and Henrissat, B. (2001) Glycoside hydrolases and glycosyltransferases: families and functional modules, *Curr. Opin. Struct. Biol. 11*, 593−600.
53. Unligil, U. M., and Rini, J. M. (2000) Glycosyltransferase structure and mechanism, *Curr. Opin. Struct. Biol. 10*, 510−517.
54. Breton, C., Snajdrova, L., Jeanneau, C., Koea, J., and Imberty, A. (2005) Structures and mechanisms of glycosyltransferases superfamily, *Glycobiology*, Epub ahead of print, PMID 16037492.
55. Hagen, F. K., Hazes, B., Raffo, R., deSa, D., and Tabak, L. A. (1999) Structure−function analysis of the UDP-*N*-acetyl-D-galactosamine: polypeptide *N*-acetylgalactosaminyltransferase. Essential residues lie in a predicted active site cleft resembling a lactose repressor fold, *J. Biol. Chem. 274*, 6797−6803.
56. Wiggins, C. A., and Munro, S. (1998) Activity of the yeast MNN1 alpha-1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases, *Proc. Natl. Acad. Sci. U.S.A. 95*, 7945−7950.
57. Campbell, R. E., Mosimann, S. C., Tanner, M. E., and Strynadka, N. C. (2000) The structure of UDP-*N*-acetylglucosamine 2-epimerase reveals homology to phosphoglycosyl transferases, *Biochemistry 39*, 14993−15001.
58. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices, *J. Mol. Biol. 233*, 123−138.
59. Unligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., and Rini, J. M. (2000) X-ray crystal structure of rabbit *N*-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily, *EMBO J. 19*, 5269−5280.
60. Mulichak, A. M., Losey, H. C., Walsh, C. T., and Garavito, R. M. (2001) Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics, *Structure (London) 9*, 547−557.
61. Morera, S., Imberty, A., Aschke-Sonnenborn, U., Ruger, W., and Freemont, P. S. (1999) T4 phage beta-glucosyltransferase: substrate binding and proposed catalytic mechanism, *J. Mol. Biol. 292*, 717−730.
62. Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., and Walker, S. (2003) Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases, *Proc. Natl. Acad. Sci. U.S.A. 100*, 845−849.
63. Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2005) Substrate-induced conformational changes in glycosyltransferases, *Trends Biochem. Sci. 30*, 53−62.
64. Mitchell, E. P., Withers, S. G., Ermert, P., Vasella, A. T., Garman, E. F., Oikonomakos, N. G., and Johnson, L. N. (1996) Ternary complex crystal structures of glycogen phosphorylase with the transition state analogue nojirimycin tetrazole and phosphate in the T and R states, *Biochemistry 35*, 7341−7355.
65. Artymiuk, P. J., Rice, D. W., Poirrette, A. R., and Willett, P. (1995) beta-Glucosyltransferase and phosphorylase reveal their common theme, *Nat. Struct. Biol. 2*, 117−120.
66. Holm, L., and Sander, C. (1995) Evolutionary link between glycogen phosphorylase and a DNA modifying enzyme, *EMBO J. 14*, 1287−1293.
67. Wrabl, J. O., and Grishin, N. V. (2001) Homology between *O*-linked GlcNAc transferases and proteins of the glycogen phosphorylase superfamily, *J. Mol. Biol. 314*, 365−374.
68. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein−ligand interactions, *Protein Eng. 8*, 127−134.

BI0524013