BIOCHEMISTRY

 Volume 47, Number 7 February 19, 2008

Articles

Structure and Active Site Residues of PglD, an *N*-Acetyltransferase from the Bacillosamine Synthetic Pathway Required for *N*-Glycan Synthesis in *Campylobacter jejuni*[†,‡]

Erumbi S. Rangarajan,[§,||] Karen M. Ruane,[⊥] Traian Sulea,[@] David C. Watson,[+] Ariane Proteau,[§] Sonia Leclerc,[+] Miroslaw Cygler,[§,@] Allan Matte,[*,@] and N. Martin Young[*,+]

Department of Biochemistry, McGill University, Montreal, QC, H3G 1Y6 Canada, York Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5YW, U.K., Biotechnology Research Institute, National Research Council of Canada, Montreal, QC, H4P 2R2 Canada, and Institute for Biological Sciences, National Research Council of Canada, Ottawa, ON, K1A 0R6 Canada

Received October 10, 2007; Revised Manuscript Received December 7, 2007

ABSTRACT: *Campylobacter jejuni* is highly unusual among bacteria in forming *N*-linked glycoproteins. The heptasaccharide produced by its *pgl* system is attached to protein Asn through its terminal 2,4-diacetamido-2,4,6-trideoxy-D-Glc (QuiNAc4NAc or *N*,*N*′-diacetylbacillosamine) moiety. The crucial, last part of this sugar's synthesis is the acetylation of UDP-2-acetamido-4-amino-2,4,6-trideoxy-D-Glc by the enzyme PglD, with acetyl-CoA as a cosubstrate. We have determined the crystal structures of PglD in CoA-bound and unbound forms, refined to 1.8 and 1.75 Å resolution, respectively. PglD is a trimer of subunits each comprised of two domains, an N-terminal α/β-domain and a C-terminal left-handed β-helix. Few structural differences accompany CoA binding, except in the C-terminal region following the β-helix (residues 189−195), which adopts an extended structure in the unbound form and folds to extend the β-helix upon binding CoA. Computational molecular docking suggests a different mode of nucleotide−sugar binding with respect to the acetyl-CoA donor, with the molecules arranged in an "L-shape", compared with the "in-line" orientation in related enzymes. Modeling indicates that the oxyanion intermediate would be stabilized by the NH group of Gly143′, with His125′ the most likely residue to function as a general base, removing H^+ from the amino group prior to nucleophilic attack at the carbonyl carbon of acetyl-CoA. Site-specific mutations of active site residues confirmed the importance of His125′, Glu124′, and Asn118. We conclude that Asn118 exerts its function by stabilizing the intricate hydrogen bonding network within the active site and that Glu124′ may function to increase the pK_a of the putative general base, His125′.

The biosynthetic pathways that produce the complex hexoses present in bacterial glycans and antibiotics frequently introduce new amino groups into common precursors such as UDP-GlcNAc. These amino sugars are then further modified by acylation, often by transfer of an acetyl group from acetyl-CoA. The *N*-acetyltransferases that carry out these reactions belong to more than one structural family, and a prototypic structure for one family, the GNAT

[†] This work was supported in part by the NRC's Genomics and Health Initiative and by a grant from the Canadian Institutes of Health Research (MOP-48370) to M.C.; K.M.R. was funded by the Biotechnology and Biological Sciences Research Council (BBSRC), in the group of Professor Gideon Davies.

[‡] Coordinates for the structures described in this paper have been deposited in the Protein Data Bank as entries 2VHE and 3BFP.

* To whom correspondence should be addressed. N.M.Y.: phone, (613) 990-0855; fax, (613) 941-1327; e-mail, Martin.Young@nrc-cnrc.gc.ca. A.M.: phone, (514) 496-2557; fax, (514) 496-5143; e-mail, Allan.Matte@nrc-cnrc.gc.ca.

[§] McGill University.

[||] Current address: Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN 38105.

[⊥] University of York.

[@] Biotechnology Research Institute, National Research Council of Canada.

[+] Institute for Biological Sciences, National Research Council of Canada.

10.1021/bi702032r CCC: $40.75
Published on Web 01/17/2008

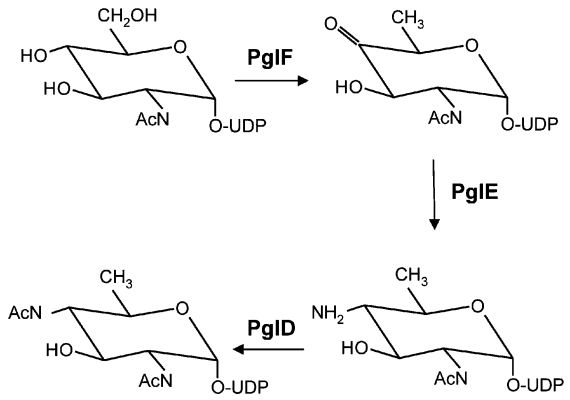

FIGURE 1: Bacillosamine biosynthetic pathway from UDP-GlcNAc to UDP-QuiNAc4NAc.

acetyltransferases, has recently been determined (*1*). This is an *N*-acetyltransferase that modifies 4-amino-L-fucose, a component of the enterobacterial common antigen. The sequences of many other acetyltransferases indicate that they belong to a second family, which display left-handed parallel β-helix (LβH)[1] structures. This type of structure is found in many enzymes that employ CoA-linked donors, including a sugar *O*-acetyltransferase from *Escherichia coli*, GAT (*2*), which modifies some galactose species to prevent them triggering the lac operon (*3*). Two other members of the LβH family *N*- or *O*-acylate simple amino sugars but are more complex in their structures and reactions. For example, the enzyme GlmU converts GlcN-1-PO_4 into UDP-GlcNAc by a two-stage process (*4*), and in the lipid A biosynthetic pathway, an *O*-acyltransferase, LpxA, modifies UDP-GlcNAc with fatty acids from an acyl carrier protein (*5*, *6*).

A complex diacetamidohexose is central to the *N*-glycosylation system of *Campylobacter jejuni*. This Gram-negative organism is highly unusual among bacteria in forming glycoproteins with glycans attached to Asn residues in the manner of eukaryotes (recently reviewed in ref *7*). The *N*-glycan heptasaccharide produced by the *pgl* (protein glycosylation) system is attached to the protein through its terminal 2,4-diacetamido-2,4,6-trideoxy-D-glucose (*N*,*N*′-diacetylbacillosamine or QuiNAc4NAc) residue (*8*). The biosynthesis of this key sugar proceeds from UDP-GlcNAc (Figure 1) via the formation of a 4-keto-6-deoxy compound by PglF (Cj1120c), which is converted to QuiNAc4N by PglE (Cj1121c). This amino sugar is the substrate for the acetyltransferase PglD (Cj1123c) along with acetyl-CoA (*9*). Mutational analysis has shown this enzyme is essential for *N*-glycosylation in *C. jejuni* (*10*). A 2-acetamido group on the sugar substrate is a requirement of the oligosaccharyltransferase, PglB, but the 4-acetamido group is not (*11*).

PglD is a highly active protein, with a k_{cat}/K_m of 1.18×10^6 min^{-1} mM^{-1} (*9*). It is smaller in size than many of the acetyltransferases described previously, and from a mechanistic point of view, it carries out a simpler and more facile reaction, acetylating an amino group rather than a hydroxyl group. The enzyme's sequence strongly suggests it is a member of the LβH family. To improve our understanding of the structure and function of this second family of complex sugar *N*-acetyltransferases, we have determined its crystal structure with and without CoA and probed the active site region using a combination of computational molecular docking and site-directed mutagenesis.

MATERIALS AND METHODS

Cloning, Overexpression, and Purification of PglD. For the determination of the native structure, a construct consisting of the complete open reading frame of Cj1123c with an N-terminal His_6 sequence was cloned into expression plasmid pCWori$^+$ and maintained in *E. coli* AD202. Cells were grown at 37 °C in 1 L of 2YT medium supplemented with 150 μg/mL ampicillin to an A_{600} of 0.5, followed by induction with 1 mM isopropyl thiogalactopyranoside, and growth continued for 6 h. Cells were lysed in 10 mM Hepes buffer (pH 7.5) containing protease inhibitors (Roche) by mechanical disruption. After centrifugation, the soluble fraction was adjusted to 500 mM NaCl and 50 mM imidazole and loaded onto a 5 mL HisTrap column (GE Healthcare). PglD was eluted with a linear gradient from 50 to 500 mM imidazole, and the PglD-containing fractions were dialyzed against 10 mM Hepes (pH 7.5) and 250 mM NaCl. Further purification on a Hi-Load Superdex 200 16/60 column at 4 °C calibrated with molecular mass standards (Sigma Chemical Co.) showed the protein was a trimer with an apparent mass of 62.7 kDa. The mass of monomeric PglD was determined to be 22 025 Da by ESI-MS using an Agilent 1100 Series LC-MSD instrument (Agilent Technologies, Mississauga, ON) and analyzed using Agilent Chemstation version A.09.01.

The determination of the structure with bound CoA utilized a *pglD* gene expression construct kindly provided by B. Wren and S. Hinchcliffe (London School of Hygiene & Tropical Medicine, London, U.K.). The pCRT7 TOPO (Invitrogen) plasmid containing the *pglD* gene was transformed into BL21 (RIPL) *E. coli* cells and cultured in 0.5 L of autoinduction medium (*12*) supplemented with 50 mg/L ampicillin, 34 mg/L chloramphenicol, and 50 mg/L streptomycin at 37 °C, for 8 h. Protein expression was induced overnight at 25 °C, and the enzyme was purified by immobilized metal affinity chromatography in a manner similar to that described above.

Crystallization and Data Collection. To produce native crystals, initial crystallization screening was performed by sitting drop vapor diffusion using sparse matrix screens from Hampton Research (Aliso Viejo, CA). Crystallization drops were set using a HydraII crystallization robot (Matrix Technologies, Manchester, NH). Crystals suitable for structure determination were obtained in 1 day at 22 °C by hanging drop vapor diffusion by mixing 1 μL of protein (9 mg/mL) in buffer [10 mM Hepes (pH 7) and 100 mM NaCl] with 1 μL of reservoir solution [0.1 M sodium acetate (pH 4.5) and 1.9 M ammonium sulfate]. Crystals were cryoprotected in reservoir solution supplemented with 25% (v/v) glycerol prior to being flash-cooled in a N_2 cold stream at 100 K (Oxford Cryosystems, Oxford, U.K.). Superior crystals were later obtained with a reservoir solution containing 0.1 M Bis-Tris (pH 6.5), 0.15 M ammonium acetate, 26% (v/v) PEG 400 [or 0.1 M Bis-tris (pH 6.5), 0.2 M potassium acetate, and 21% (v/v) PEG 400], and 10 mM CoA. Despite the addition of CoA, no density for this ligand was observed in the final electron density maps. To obtain single crystals, an additional step of streak seeding was necessary with the resulting crystals then used as macro seeds to produce larger

[1] Abbreviations: Bac, bacillosamine or 2,4-diamino-2,4,6-trideoxy-D-Glc; LβH, left-handed β-helix; TI, tetrahedral intermediate.

Table 1: X-ray Data Collection and Refinement Statistics

Data Collection			
	PglD−iodide	apo-PglD	PglD−CoA
space group	$P6_3$	$P6_3$	$H3$
cell dimensions			
a (Å)	86.60	86.46	99.25
c (Å)	65.54	65.55	143.03
Z	12	12	18
resolution range (Å)	50−2.1	25−1.75	15−1.8
wavelength (Å)	1.54	1.1	0.9756
total no. of reflections	218625	110357	156459
no. of unique reflections	16208	28128	48692
average redundancy	13.5	3.9	3.2
completeness (%)	98.5	99.7	100
[1] R_{sym}	0.134	0.048	0.074
$I/\alpha(I)$	11.5	19.5	11.4

Refinement		
	apo-PglD	PglD−CoA
resolution range (Å)	32−1.75	15−1.8
[2] R_{work}, *#hkl*	0.198 (26695)	0.190 (46151)
R_{free}, *#hkl*	0.225 (1414)	0.225 (2457)
no. of atoms (*B*-values, Å^2)		
protein	1398 (22.43)	2928 (23.57)
CoA	−	96 (39.16)
solvent	116 (36.59)	263 (40.28)
ions (SO_4)	−	10 (35.87)
ions (citrate)	13 (29.32)	−
Ramachandran plot		
allowed	100	99.4
generously allowed	0.0	0.6
disallowed	0.0	0.0
root-mean-square deviation		
bonds (Å)	0.01	0.01
angles (deg)	1.19	1.60
PDB entry	3BFP	2VHE

crystals, for which no additional cryoprotectant was needed. The crystal structure of apo-PglD was obtained by soaking a native PglD crystal in reservoir solution containing 500 mM NaI for 30 s, mounting the sample in a nylon loop, and flash-cooling it in the nitrogen cold stream at 100 K. Diffraction data were collected using a MAR imaging plate detector mounted on a Rigaku FRD generator (Boehringer-Ingelheim, Laval, QC). Subsequently, a 1.75 Å data set was collected from a native PglD crystal at beamline X8C, National Synchrotron Light Source, using a ADSC Quantum-4 CCD detector and processed using HKL2000 (*13*).

For the CoA complex, the protein was exchanged into a buffer consisting of 10 mM Tris (pH 8.0) and 5 mM tris-(2-carboxyethyl)phosphine hydrochloride (TCEP) and concentrated to 25 mg/mL. The PglD crystals were grown by vapor phase diffusion by the hanging drop method. Crystallization was conducted in 0.1 M sodium citrate (pH 5.6), 0.16 M sodium tartrate, and 2.0 M ammonium sulfate. Crystals appeared overnight and were cryoprotected in mother liquor solution with the addition of 25% (v/v) glycerol before being flash frozen in liquid nitrogen. Data were collected from a single crystal on ID14-1 at the European Synchrotron Radiation Facility (ESRF) (Grenoble, France), to 1.8 Å resolution (Table 1). As described herein, these crystals serendipitously contained the CoA moiety of the acetyl-CoA donor.

Structure Determination and Refinement. A sodium iodide-soaked crystal was collected at the Cu Kα wavelength and the structure determined by SAD using the anomalous signal from iodine. Iodine positions and phases were computed using data to 2.7 Å using auto-SHARP (*14*) and these phases improved by density modification using DM (*15*). An initial model was autobuilt using ARP/wARP (*16*) and partially refined to 2.3 Å resolution using Refmac5 (*17*), alternating with manual refitting using COOT (*18*). A subsequent data set was collected to 1.75 Å resolution at beamline X8C, NSLS, BNL, and used for final refinement of the PglD model. Residues 12 and 13 and the loop containing residues 36−44 are disordered in the apo structure. No σ cutoff was applied to the data during refinement. Final data collection and refinement statistics are presented in Table 1.

Data for the CoA complex were processed with Mosfilm, and all subsequent computing was conducted with the CCP4 suite of programs (*19*). The structure of the PglD−CoA complex was determined by molecular replacement using AMoRe (*20*) with the coordinates of the native structure. There are two molecules in the asymmetric unit. The structure was refined using iterative cycles of Refmac5 and model building and solvent addition with COOT. The electron density map clearly revealed density both for coenzyme A and for the loop region disordered in the apo structure. Relevant data collection and refinement statistics are summarized in Table 1.

Site-Directed Mutagenesis of PglD. The *pglD* gene was mutagenized using a two-stage mutagenesis protocol. The 5′ and 3′ PCR primers both containing the desired mutation were used in separate PCRs to generate two overlapping gene fragments. Both PCR products were then used as template with the primer pair used to obtain the gene originally in a third PCR to generate full-length mutagenized PglD. The PCR products were subsequently cloned and verified by sequence analysis; the mutant proteins were expressed and purified as described above. The N118A, E124A, H125A, and H134A mutations were constructed and expressed in this way.

Enzyme activity was determined by a capillary electrophoresis assay with a UDP-amino sugar substrate prepared by means of the previous enzymes in the pathway, e.g., PglE and PglF. The 100 μL test mixtures contained 1.5 mM UDP-amino sugar substrate, 3.5 mM AcCoA, 6 mM HEPES buffer (pH 7.5), and 5 μg of BSA, in addition to the appropriate amounts of enzyme. The CE separation was carried out on a Beckman P/ACE 5510 instrument at 18 kV, with a running buffer of 25 mM sodium tetraborate (pH 9.4), and the products were detected at 260 nm.

Computational Docking and Molecular Modeling. The tetrahedral intermediate (TI) species formed by the attack of the 4-amino group of the Bac4N-UDP substrate at the acetyl carbonyl carbon of acetyl-CoA was built in complex with the PglD trimer starting from the CoA-bound structure. Phosphate ions and solvent molecules were removed from the crystal structure, with the exception of six buried water molecules, two of which are located in the putative catalytic site. The protonation state of the protein at neutral pH was assigned according to the acid ionization constants (pK_a values) calculated for all ionizable side chains in the native PglD trimer with H++ (*21*, *22*) using the Poisson−Boltzmann solvation model at a salt concentration of 0.15 M and solute and solvent dielectric constants of 20 and 80, respectively, and computing the partition sum over all ionization states via the Monte Carlo algorithm. The TI

molecule was assigned a protonation state corresponding to a net charge of −7*e*, i.e., with the oxyanion and neutral amine as two of the TI center substituents.

Given the flexibility of the TI molecule, docking of the QuiNAc4N-UDP portion was accomplished with a "ligand grow" stepwise protocol (*23*), including three growing steps guided either by knowledge from structural data for related enzymes or by flexible docking and energy minimization. In the first step, CoA-bound PglD was modified at the S atom to generate the chiral TI center, i.e., S−H → S−C(CH_3)-(O^-)(NH_2), which was built on the basis of the mutual three-dimensional arrangement between the attacked thioester group of the succinyl-CoA and the attacking amino group of the L-2-aminopimelate substrate in the crystal structures of the tetrahydrodipicolinate *N*-succinyltransferase from *Mycobacterium bovis* (PDB entries 1KGQ and 1KGT) (*24*). The PglD-bound tetrahedral TI center, which results in the *S* absolute configuration, was subsequently refined by energy minimization. In the second step, the QuiNAc4N-1-phosphate fragment, built from standard covalent bond and angle geometries with Sybyl 6.9 (Tripos, Inc., St. Louis, MO), was linked to the refined tetrahedral TI center by replacing its N atom with the N^4 atom of QuiNAc4N-1-phosphate. The amino group of the TI center being located in a sterically confined region close to PglD severely limits the conformational freedom of the QuiNAc4N-1-phosphate fragment to one dominant, sterically allowed conformation, which was refined by energy minimization. In the third step, the UMP fragment, built in a canonical RNA conformation using Sybyl 6.9, was added to the TI refined in the previous step, by establishing the appropriate connection between the 1-phosphate constituent of QuiNAc4N and the 5′-phosphate of UMP. The binding mode of the linked UMP was explored by flexible docking using the Monte Carlo minimization (MCM) conformational sampling procedure (*25*) adapted to flexible ligand docking (*1*, *26*−*28*). The acyclic rotatable bonds in the UDP region were sampled for a total of 5000 MCM cycles.

All energy evaluations and conjugate gradient minimizations, including those performed during the MCM search, were carried out with the AMBER all-atom molecular mechanics force field energy (*29*), using a distance-dependent dielectric constant ($4R_{ij}$), an 8 Å nonbonded cutoff, and MMFF94 partial atomic charges for the TI molecule. Protein residues and explicit water molecules surrounding the energy-minimized regions of the ligand were allowed to relax, while the ligand was anchored by rigidifying the adenine ring of CoA, located remotely from the putative catalytic center and the QuiNAc4N-UDP binding site.

RESULTS AND DISCUSSION

Structure of the PglD Monomer. The crystal structure of PglD has been determined by SAD phasing using a crystal soaked in sodium iodide and refined to an R_w of 0.198 and an R_f of 0.225 at 1.75 Å resolution. In the native PglD model, residues Ala12, Ser13, and Asp36−Ser44 are disordered. The model shows good geometry, with 90.4% of residues in the most favored region of the Ramachandran plot [PROCHECK (*30*)]. Each monomer consists of two domains, an N-terminal α/β-domain (residues 2−72) and the C-terminal Lβ*H* domain consisting of six turns of parallel β-helix (residues 75−187) (Figure 2A). The α/β-domain consists of a central three-stranded parallel β-sheet flanked by two α-helices that are nearly parallel to the β-strands and lie along the edges of the β-sheet. The overall topology of the α/β-domain is β1-α1-β2-β3-α2. A short segment of polypeptide (residues 72−75) acts as a linker between the two domains such that the α/β-domain sits near the N-terminal end of the Lβ*H*.

Each coil of the Lβ*H* shows the expected features of this fold, including three β-strands, S1, S2, and S3, each joined by three, two-residue connectors, L1, L2, and L3 (*31*, *32*). The interior core of the Lβ*H* is dominated by stacks of small, hydrophobic residues (Leu, Ile, and Val) located at the middle of each β-strand, with the exterior showing many polar and charged residues. Glycine residues are found to dominate near the turns, as previously observed (*31*). Unlike most structures having this fold, no loop insertions are found in any of the β-helix coils. The first coil is composed of 20 residues; the next five are composed of 18 residues, and the sixth has only 16 residues. The C-terminus in apo-PglD (residues 189−195) assumes an extended conformation pointing away from the body of the β-helix.

Trimer Structure. The molecules of PglD are arranged into trimers related by 3-fold crystallographic symmetry. Each monomer within the PglD trimer contacts one other monomer through the side chains of the S2 β-strands of the Lβ*H* and another monomer through the side chains of the L1 turn (Figure 2B). The N-terminal α/β-domain of each monomer contacts only one neighboring monomer. The extended C-terminal segment of each apo-PglD monomer interacts with the base of the Lβ*H* from an adjacent subunit, acting as a "clamp" to assist in stabilizing the trimer arrangement. The contacts between the trimer subunits include van der Waals contacts and hydrogen bonding interactions. The charged residues are present predominantly outside of the intersubunit interfaces, on β-strands not involved in trimer formation.

Related Structures. The Lβ*H* fold is common to a number of enzymes that act on sugar or nucleotide−sugar substrates (*33*), many of which utilize acetyl-CoA as a cosubstrate, including ADP-glucose pyrophosphorylase (*34*), maltose acetyltransferase (*35*), galactoside acetyltransferase (*2*), UDP-glucose pyrophosphorylase (*36*, *37*), UDP-GlcNAc 3-*O*-acyltransferase, LpxA (*6*, *38*), and the acetyl-transferase domain of the bifunctional enzyme *N*-acetylglucosamine-1-phosphate uridyltransferase, GlmU (*4*, *39*−*42*). Many, but not all, of these enzymes, including PglD, have a trimer as the biological unit. Like PglD, in most of these proteins the Lβ*H* domain follows an α-helical or α/β-domain. However, other arrangements have also been found, e.g., in virginiamycin acetyltransferase, VatB (*43*), and a xenobiotic acetyltransferase from *Pseudomonas aeruginosa* (*44*), where the Lβ*H* domain precedes the α-helical domain, and tetrahydrodipicolinate *N*-succinyltransferase, which has domains both N- and C-terminal to the Lβ*H* (*45*).

Another PDB entry has been deposited for apo-PglD from *C. jejuni* (2NPO). This structure exhibits a root-mean-square deviation (rmsd) of 0.36 Å for 362 common backbone atoms with our apo-PglD model. Both our apo-PglD model and 2NPO show the same extended conformation of the C-terminus.

CoA-Binding Site. Crystals were obtained from a PglD sample that copurified with CoA, by using ammonium sulfate

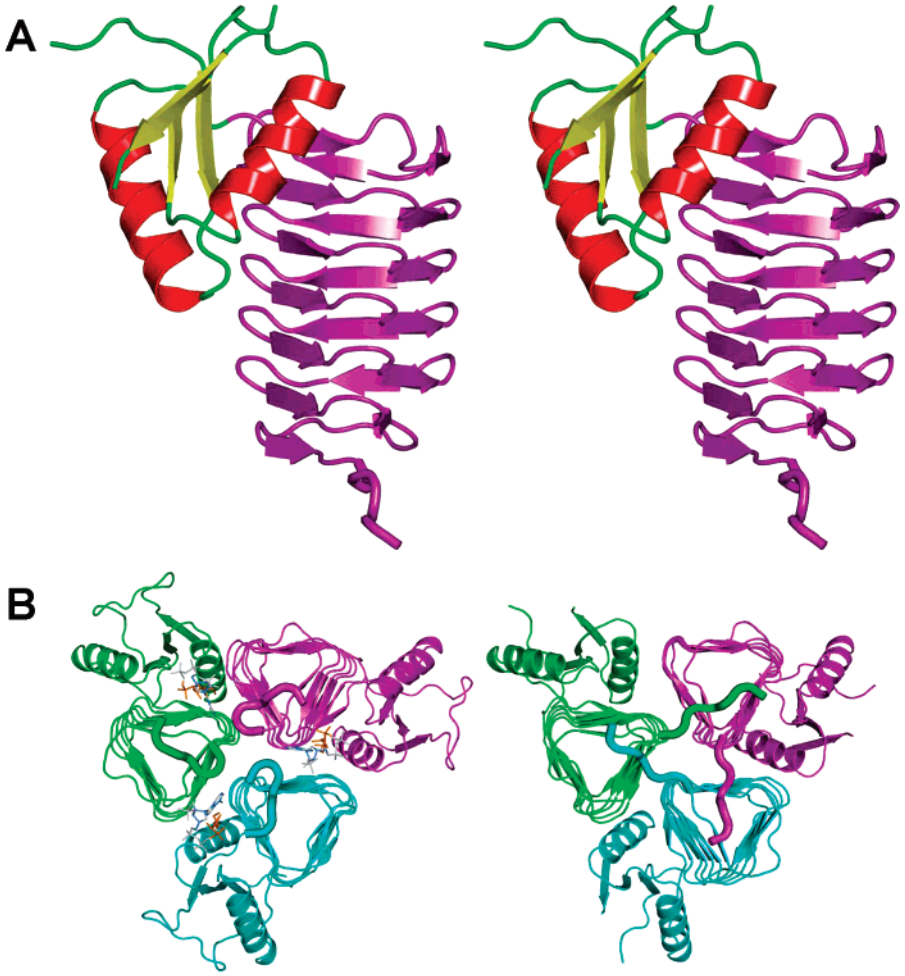

FIGURE 2: Structure of PglD. (A) Monomer, with the N-terminal α/β-domain colored red (helices) and yellow (β-strands) and the LβH domain colored magenta. (B) Structure of the PglD trimer, colored by subunit: (left) PglD in complex with CoA and (right) apo-PglD. The C-termini, shown as thicker tubes, have different conformations in the two structures. These and subsequent figures were prepared using PyMOL (http://pymol.sourceforge.net).

rather than PEG400 as the precipitant. They diffracted to 1.8 Å resolution and showed good density for the cofactor (Figure 3A). Superposition of the structures of CoA-bound and apo-PglD subunits gives a rmsd of 0.5 Å for 356 common main chain atoms. Small structural differences occur within the loop of residues Gly11−His15 connecting β1 and α1 in the N-terminal domain and the loop of residues Asp36−Glu43, which is disordered in the apo structure but is ordered well in the PglD−CoA complex. The Gly189−Met195 segment at the C-terminus, which in apo-PglD is extended and interacts with an adjacent subunit, folds back and interacts with the last coil of the LβH in the complex, following the same path taken by the extended C-terminus from the adjacent subunit in the apo form (Figure 2B). Thus, swapping of this segment in apo-PglD is not essential for trimer formation.

Coenzyme A binds in a shallow cleft formed between adjoining LβH domains from two neighboring subunits, near the C-terminal end of the LβH, with residues from S1 of one subunit and S2 of an adjacent subunit contributing to binding (Figure 3B). There are three CoA molecules bound to the trimer, consistent with the fact that trimer formation plays a critical role in the acetyl-transferase activity of another protein with the same fold, GlmU (*46*). The cofactor binds to PglD in such a way that the 3′-phosphoryl group of the phosphoadenosine and the diphosphate moiety of the cofactor are largely solvent-exposed. A hydrogen bond is formed between the carbonyl of Gly173 and N6 of adenine. The adenine ring is sandwiched by several hydrophobic side chains, including Val188, Met195, and Val178′, Pro191′, and Val190′ from an adjacent subunit. The pantetheine moiety of CoA is also predominantly surrounded by apolar residues, including the side chains of Phe152, Ile155, and Val137 from one subunit and Ala142′, Leu160′, and Pro161′ from an adjacent subunit. The *gem*-methyl groups of the pantetheine moiety are anchored by Ile170, Val186, and Val188. Altogether, hydrophobic residues appear to dominate binding of the cofactor to the enzyme (Figure 3B). The amide carbonyls of the pantetheine moiety form hydrogen bonds with the main chain NH groups of residues Ile155 and Gly173.

Nucleotide−Sugar Binding and Active Site Region. Co-crystallization experiments with PglD and its UDP-sugar substrate in the presence or absence of CoA did not lead to interpretable electron density for the substrate. The putative binding site for the UDP-sugar substrate is within a continuation of the intersubunit, CoA-binding cavity, proximal to the sulfur atom of CoA. This cavity is formed in part at the interface of the LβH domains of each subunit, and also involves the α1 helix from the α/β-domain from one of the subunits. On the basis of their chemical properties and the proximity to the sulfur atom of CoA, Asn118, His134,

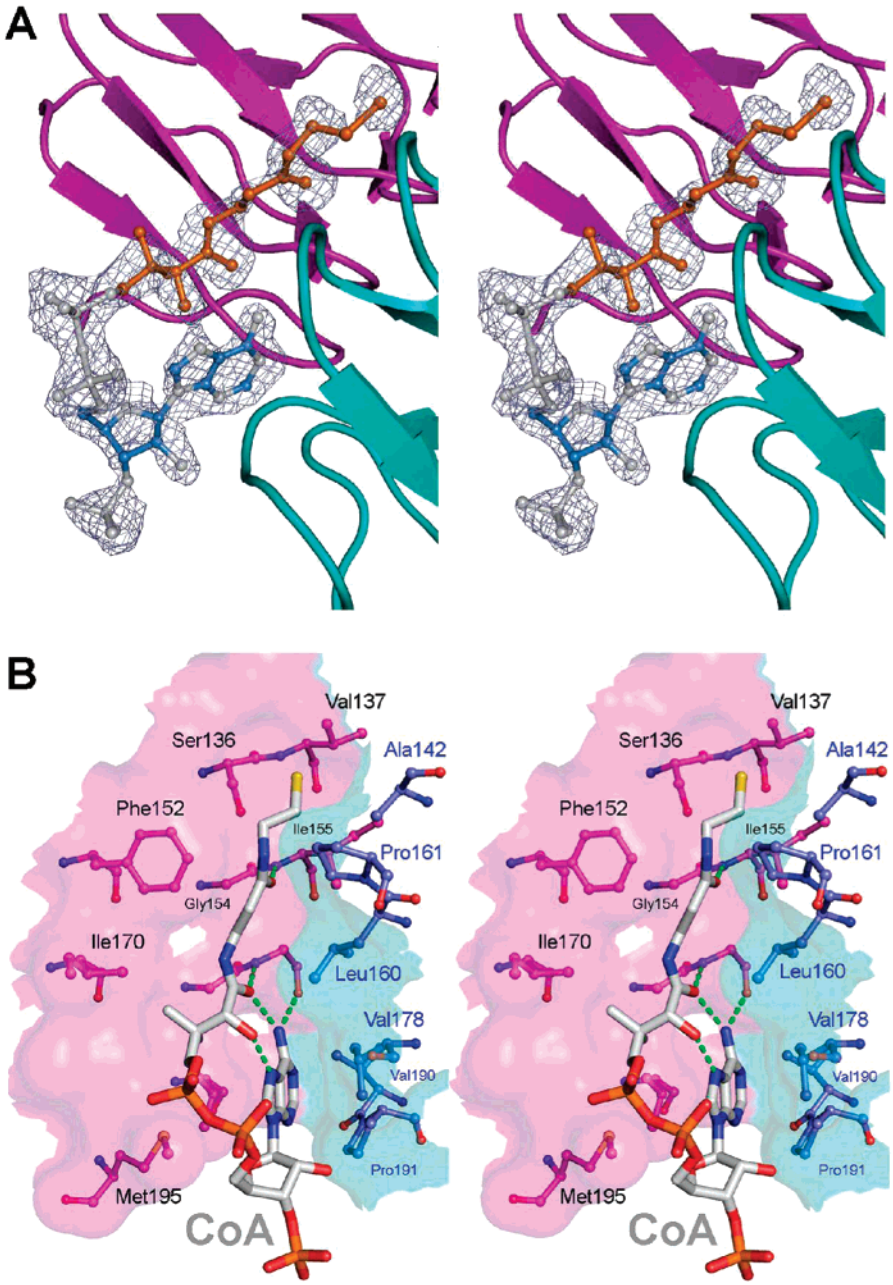

FIGURE 3: CoA-binding site. (A) Stereo $2F_o - F_c$ (omit) electron density for CoA contoured at a level of 2σ from a PglD crystal obtained from an enzyme sample copurified with CoA. (B) Stereoview of the CoA-binding site, dominated by apolar residues. The semitransparent surface of the binding site at the subunit interface is colored cyan and magenta as in panel A. CoA and neighboring residues are colored according to a CPK scheme.

Glu124′, and His125′ were identified as potential active site residues and were selected for site-directed mutagenesis to alanine. Other residues within this region, such as His15, also line this cleft and may have a potential role within the active site. Of the four residues selected for mutagenesis, in preliminary enzyme kinetics experiments H125A showed the greatest loss of activity, its specific activity being only 39 milliunits/mg compared to 5000 units/mg for the wild type. Residues Glu124′ and Asn118 also appear to play key roles, as mutating these to Ala resulted in reductions in activity to 7.5 and 4.5 units/mg, respectively. Mutation of His134 to Ala reduced activity to only 300 units/mg, suggesting a less important role within the active site. In the crystal structure, His134 forms a hydrogen bond with Asn118 and may help to orient this later residue in the active site.

Model of the Tetrahedral Intermediate Complex. The available mechanistic data for this class of enzyme point to formation of an enzyme−substrate ternary complex with a chemical mechanism in which the acceptor substrate performs a nucleophilic attack at the acetyl group of the cofactor (*47*). We built a model of the PglD trimer in complex with a tetrahedral intermediate species (TI) that would be formed by the attack of the 4-amino group of the UDP-QuiNAc4N substrate at the carbonyl C atom of acetyl-CoA. The overall geometry of the PglD-bound AcCoA-QuiNAc4N-UDP covalent intermediate is L-shaped with an ~90° bend at the TI center (Figure 4A). This indicates that the substrate and cofactor molecules are displayed for reaction on the surface of the enzyme in a mutual perpendicular arrangement. The L-shaped approach of the substrate is dictated by the

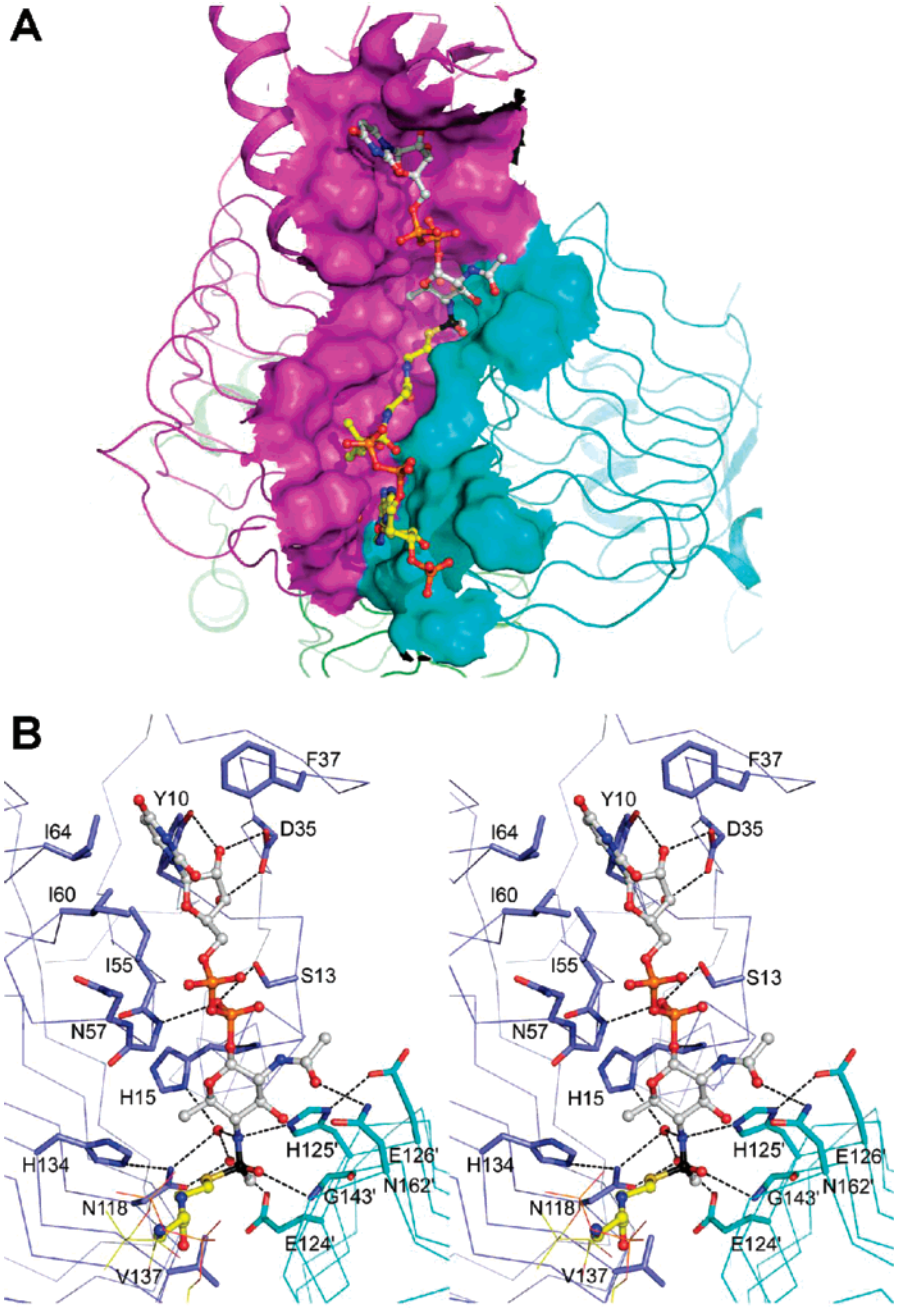

FIGURE 4: Model of PglD in complex with the tetrahedral intermediate (TI) formed between acetyl-CoA and the 4-amino UDP-sugar. (A) Overall structure showing the L shape of the TI bound at the subunit interface of PglD. (B) Stereoview of interactions of PglD at the tetrahedral TI center and with the UDP-sugar fragment. In panel A, the PglD molecular surface forming the putative substrate cofactor-binding site is colored by subunit as in Figure 3. Different regions of the TI ligand are color-coded by C atoms: yellow for the CoA fragment, black for the TI center C atom, and white for the UDP-sugar and the methyl of the transferred acetyl of acetyl-CoA. Hydrogen bonds are shown as black dashed lines.

N-terminal α/β-domain of PglD, which creates a wall in the direction defined by the extended conformation of the pantetheine moiety of the cofactor. This topology differs from those in related enzymes, where additional domains, differently shaped and positioned relative to the LβH domains, allow for an in-line display of the reacting cofactor and substrate (*2*, *24*). The binding site for the Bac4N sugar portion is formed jointly by two subunits in the PglD trimer, whereas binding of the UDP fragment is predicted to reside entirely in one monomer, and specifically in the N-terminal α/β-domain.

The chiral TI center adopts the *S* absolute configuration, which was inferred from the observed spatial arrangement of the substrate amine nucleophile and the cofactor thioester group in the structurally related tetrahydrodipicolinate *N*-succinyltransferase, DapD (*24*). In the refined model of the PglD−TI complex, the substituents at the tetrahedral center establish direct and water-mediated interactions with the enzyme (Figure 4B). The oxyanion is stabilized by a hydrogen bond with the main chain NH group of Gly143′. The methyl group is buried and forms nonpolar interactions with the Val137 and Glu124′ side chains. The amino group forms a direct hydrogen bond with the imidazole ring (NE atom) of His125′, as well as hydrogen bonds mediated by a structured water molecule with His15, from the α/β-domain, and with Asn118. The sulfur atom contacts the side chains

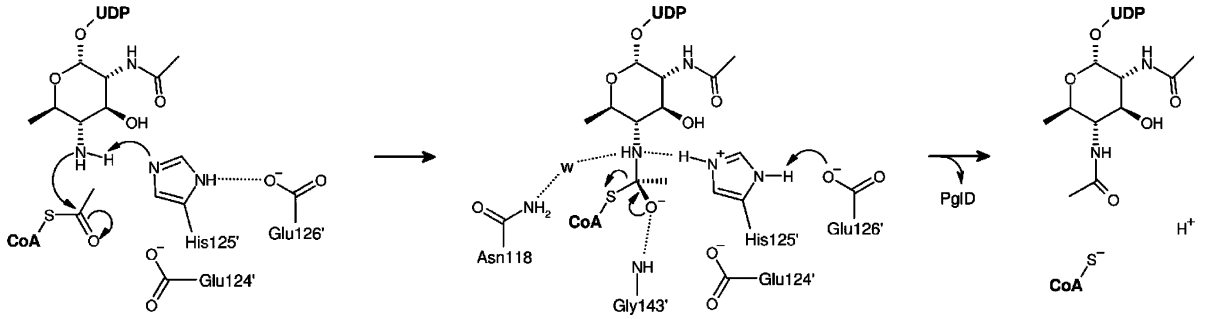

FIGURE 5: Summary of the proposed catalytic mechanism for PglD based on modeling of the tetrahedral intermediate and activity analysis of site-directed mutants. The putative general base, His125′, activated by the adjacent Glu124′ and Glu126′ residues, abstracts a proton from the 4-amino position of the substrate sugar, facilitating the nucleophilic attack at the carbonyl carbon of acetyl-CoA and stabilizing the tetrahedral intermediate with the cofactor. Gly143′ contributes to oxyanion stabilization. Asn118 and His15 (not shown) may also participate in tetrahedral intermediate stabilization, possibly via a water molecule (W).

of Asn118, His134, and Val137 and deviates slightly from its position in the PglD−CoA complex. The positions of the O and N substituents at the TS center match closely those of two water molecules from the PglD−CoA crystal structure.

The QuiNAc4N sugar appears conformationally restricted for rotations around the C^4-N^4 and N^4-TI center bonds. This is due in part to the location of its N^4 amine nucleophile close to the PglD surface, as well as to the substitution pattern on the chair pyranose ring with axial 1-phosphate, and equatorial 2-acetamide, 3-hydroxyl, 4-amino, and 5-methyl substituents. In the refined TI model, these substituents establish complementary interactions with several residues of the enzyme, including Gly14, Gly56, His134, His125′, and Asn162′ (Figure 4B). The UDP moiety is predicted to bind exclusively to a single subunit of the PglD trimer (Figure 4A), specifically in a groove shaped by loops $\beta1-\alpha1$, $\beta2-\beta3$, and $\beta3-\alpha2$ and the first half of helix $\alpha2$, all from the N-terminal α/β-domain. Parts of these loops are relatively flexible as revealed by PglD crystal structures, but their conformations may be stabilized upon substrate binding. The nucleotide α-phosphate is largely solvent exposed, whereas the β-phosphate faces the protein toward the main chain of Gly14, His15, and Gly56 and the side chain of Ser13 (Figure 4B). The ribose 2′- and 3′-hydroxyls are anchored by hydrogen bonds to the Asp35 carboxylate. The uracil ring is projected further into a narrow hydrophobic groove, sandwiched between the Phe37 aromatic ring (perpendicular stacking) and Ile60 side chain.

Roles of Specific Residues in the Catalytic Mechanism. The structural model of the TI, which positions the amino group of QuiNAc4N in direct hydrogen bond contact with the NE atom of His125′, suggests that His125′ may function as the general base to abstract a proton from the QuiNAc4N amino group (Figure 5). This appears to be supported by the significantly reduced activity of the H125A mutant versus the wild-type enzyme. It is possible that the proximity of Glu124′ and Glu126′ to the His125′ side chain may contribute to a slight basic pK_a shift of the latter. It is difficult to see how the active site could abstract the two protons required if the substrate was bound with its amino group in the charged form. A spatially similar residue, His363, has been proposed to function as a general base, activating the 2-amino group of GlcN-1-PO_4 in *E. coli* GlmU (*42*). Histidinyl residues have also been implicated in having a similar functional role in a number of CoA-dependent, *O*-acetyltransferases with the LβH fold (*2*, *43*, *47*−*50*). The residue proposed to function as the general base in PglD (His125′), DapD (Asp141), and GlmU (His363) originates at the surface of the core LβH, not from an extended loop protruding from the LβH, as in some *O*-acetyl-transferases (*2*).

In addition to the putative catalytic base (His125′), the TI model also suggests that the nucleophilic N^4 amino group of the substrate is stabilized by hydrogen bonds to the His15 and Asn118 side chains through a water molecule, and to the Glu124′ side chain via two water molecules. The residues equivalent to Asn118 of PglD in related enzymes have been suggested to fulfill important catalytic roles, e.g., Asn85 of a *lac* operon galactoside acetyltransferase implicated in oxyanion stabilization (*2*, *3*), Asp141 of *M. bovis* DapD acting as a general base and possibly as a general acid (*24*), or His363 of *E. coli* GlmU, the proposed general base (*42*). Our mutagenesis experiments with PglD indicate that mutation of either Asn118 or Glu124′ compromises the enzymatic activity (>1000-fold). In addition to disrupting the hydrogen bond network that may be necessary for a productive nucleophilic attack, mutations of Asn118 (to Ala and His) and Glu124′ (to Ala) may impair the biologically active trimeric state of PglD, as both residues are significantly buried and located within the intersubunit contact range. Glu124′ may also contribute to a possible basic pK_a shift of the putative general base His125′.

The proposed stabilization of the oxyanion by the main chain NH group of Gly143′ in PglD is structurally conserved in DapD, where the Gly166B amide has been proposed to fulfill this role (*24*). Main chain amide groups of small residues have also been indicated as oxyanion-stabilizing groups in other related enzymes, e.g., Gly143 of *E. coli* UDP-GlcNAc acyltransferase (LpxA) (*50*) or Ala380 of *E. coli* GlmU (*42*).

There is no obvious candidate side chain to act as a general acid in the breakdown of the TI by protonation of the sulfur atom of the leaving CoA product. Although His134 is in contact with the sulfur atom in the TI model, it is essentially required in neutral form to function as a hydrogen bond acceptor for the NH_2 side chain group of Asn118. This requirement, corroborated by the modest (up to 10-fold) effect on activity upon its mutation to nontitratable residues, argues against His134 fulfilling a role as a general acid. Interestingly, His134 is also present in DapD (His157) where it also does not appear to be implicated directly in the

catalytic mechanism (*24*, *44*). One residue proposed as a potential general acid in DapD, Glu189B (*24*), is equivalent to Cys158 in PglD and therefore is not conserved. It is possible that the role of general acid in the catalytic mechanism of PglD is instead fulfilled by a solvent molecule or that the TI breaks down into the thiolate form of CoA, which may be kinetically favorable for product release.

ACKNOWLEDGMENT

We thank Dr. Warren Wakarchuk for invaluable advice, Dr. Ian Schoenhofen for a sample of the amino sugar substrate, Drs. Brendan Wren and Stewart Hinchcliffe (London School of Hygiene & Tropical Medicine) for provision of the *pglD* expression construct to the York laboratory, and Wendy Offen for invaluable technical assistance. Karen Ruane is a Ph.D. student, in the group of Professor Gideon Davies, who is thanked for his assistance with this work. We also thank Rene Coloumbe, Boehringer-Ingelheim Canada, for access to their X-ray diffraction facilities. We thank Leon Flaks for assistance in synchrotron X-ray data collection. Data for this study were measured at beamline X8C of the National Synchrotron Light Source. Financial support comes principally from the Office of Biological and Environmental Research and of Basic Energy Sciences of the U.S. Department of Energy, and the National Center for Research Resources of the National Institutes of Health. This is publication NRC 49533 from the National Research Council of Canada.

SUPPORTING INFORMATION AVAILABLE

Atomic coordinates for the modeled AcCoA-QuiNAc4N-UDP tetrahedral intermediate bound to PglD. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hung, M. N., Rangarajan, E., Munger, C., Nadeau, G., Sulea, T., and Matte, A. (2006) Crystal structure of TDP-fucosamine acetyltransferase (WecD) from *Escherichia coli*, an enzyme required for enterobacterial common antigen synthesis, *J. Bacteriol. 188*, 5606−5617.
2. Wang, X.-G., Olsen, L. R., and Roderick, S. L. (2002) Structure of the *lac* operon galactoside acetyltransferase, *Structure 10*, 581−588.
3. Roderick, S. L. (2005) The *lac* operon galactoside acetyltransferase, *C. R. Biol. 328*, 568−575.
4. Olsen, L. R., and Roderick, S. L. (2001) Structure of the *Escherichia coli* GlmU pyrophosphorylase and acetyltransferase active sites, *Biochemistry 40*, 1913−1921.
5. Raetz, C. R. H., and Roderick, S. L. (1995) A left-handed parallel β helix in the structure of UDP-*N*-acetylglucosamine acyltransferase, *Science 270*, 997−1000.
6. Williams, A. H., Immoromino, R. M., Gewirth, D. T., and Raetz, C. R. H. (2006) Structure of UDP-N-acetylglucosamine acyltransferase with a bound antibacterial pentadecapeptide, *Proc. Natl. Acad. Sci. U.S.A. 103*, 10877−10882.
7. Szymanski, C. M., and Wren, B. W. (2005) Protein glycosylation in bacterial mucosal pathogens, *Nat. Rev. Microbiol. 3*, 225−237.
8. Young, N. M., Brisson, J. R., Kelly, J., Watson, D. C., Tessier, L., Lanthier, P. H., Cadotte, N., St. Michael, F., Aberg, E., and Szymanski, C. M. (2002) Structure of the N-linked glycan present on multiple glycoproteins in the gram-negative bacterium, *Campylobacter jejuni*, *J. Biol. Chem. 277*, 42530−42539.
9. Olivier, N. B., Chen, M. M., Behr, J. R., and Imperiali, B. (2006) In vitro biosynthesis of UDP-N,N′-diacetylbacillosamine by enzymes of the *Campylobacter jejuni* general protein glycosylation system, *Biochemistry 45*, 13659−13669.
10. Kelly, J., Jarrell, H., Millar, L., Tessier, L., Fiori, L. M., Lau, P. C., Allan, B., and Szymanski, C. M. (2006) Biosynthesis of the N-linked glycans in *Campylobacter jejuni* and addition onto protein through block transfer, *J. Bacteriol. 188*, 2427−2434.
11. Wacker, M., Feldman, M. F., Callewaert, N., Kowarik, M., Clarke, B. R., Pohl, N. L., Hernandez, M., Vines, E. D., Valvano, M. A., Whitfield, C., and Aebi, M. (2006) Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems, *Proc. Natl. Acad. Sci. U.S.A. 103*, 7088−7093.
12. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures, *Protein Expression Purif. 41*, 207−234.
13. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
14. Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2006) Automated structure solution with autoSHARP, *Methods Mol. Biol. 364*, 215−230.
15. Cowtan, K. D., and Main, P. (1993) Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints, *Acta Crystallogr. D49*, 148−157.
16. Morris, R. J., Perrakis, A., and Lamzin, V. S. (2003) ARP/wARP and automatic interpretation of protein electron density maps, *Methods Enzymol. 374*, 229−244.
17. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT, *Acta Crystallogr. D55*, 247−255.
18. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics, *Acta Crystallogr. D60*, 2126−2132.
19. Collaborative Computational Project, Number 4 (1994) The *CCP*4 suite: Programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
20. Navaza, J. (2001) Implementation of molecular replacement in AMoRe, *Acta Crystallogr. D57*, 1367−1372.
21. Gordon, J. C., Myers, J. B., Folta, T., Shoja, V., Heath, L. S., and Onufriev, A. (2005) H++: A server for estimating pKas and adding missing hydrogens to macromolecules, *Nucleic Acids Res. 33*, W368−W371.
22. Myers, J., Grothaus, G., Narayanan, S., and Onufriev, A. (2006) A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules, *Proteins 63*, 928−938.
23. Kohls, D., Sulea, T., Purisima, E. O., MacKenzie, R. E., and Vrielink, A. (2000) The crystal structure of the formiminotransferase domain of formiminotransferase-cyclodeaminase: Implications for substrate channeling in a bifunctional enzyme, *Structure 8*, 35−46.
24. Beaman, T. W., Vogel, K. W., Drueckhammer, D. G., Blanchard, J. S., and Roderick, S. L. (2002) Acyl group specificity at the active site of tetrahydridipicolinate N-succinyltransferase, *Protein Sci. 11*, 974−979.
25. Li, Z., and Scheraga, H. A. (1987) Monte Carlo-minimization approach to the multiple-minima problem in protein folding, *Proc. Natl. Acad. Sci. U.S.A. 84*, 6611−6615.
26. Lin, L. Y., Sulea, T., Szittner, R., Vassilyev, V., Purisima, E. O., and Meighen, E. A. (2001) Modeling of the bacterial luciferase-flavin mononucleotide complex combining flexible docking with structure-activity data, *Protein Sci. 10*, 1563−1571.
27. Chowdhury, S. F., Sivaraman, J., Wang, J., Devanathan, G., Lachance, P., Qi, H., Menard, R., Lefebvre, J., Konishi, Y., Cygler, M., Sulea, T., and Purisima, E. O. (2002) Design of noncovalent inhibitors of human cathepsin L. From the 96-residue proregion to optimized tripeptides, *J. Med. Chem. 45*, 5321−5329.
28. Sivaraman, J., Myers, R. S., Boju, L., Sulea, T., Cygler, M., Davisson, V. J., and Schrag, J. D. (2005) Crystal structure of *Methanobacterium thermoautotrophicum* phosphoribosyl-AMP cyclohydrolase HisI, *Biochemistry 44*, 10071−10080.
29. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc. 117*, 5179−5197.
30. 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.
31. Lengar, P., Joshi, N. V., and Balaram, P. (2006) Conformational and sequence signatures in β-helix proteins, *Structure 14*, 529−542.

32. Jenkins, J., and Pickersgill, R. (2001) The architecture of parallel β-helices and related folds, *Prog. Biophys. Mol. Biol. 77*, 111−175.
33. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) SCOP: A structural classification of proteins database for the investigation of sequences and structures, *J. Mol. Biol. 247*, 536−540.
34. Jin, X., Ballicora, M. A., Preiss, J., and Geiger, J. H. (2005) Crystal structure of potato tuber ADP-glucose pyrophosphorylase, *EMBO J. 24*, 694−704.
35. Leggio, L. L., Degan, F. D., Poulsen, P., Andersen, S. M., and Larsen, S. (2003) The structure and specificity of *Escherichia coli* maltose acetyltransferase give new insight into the LacA family of acyltransferases, *Biochemistry 42*, 5225−5235.
36. Roeben, A., Plitzko, J. M., Körner, R., Böttcher, U. M., Siegers, K., Hayer-Hartl, M., and Bracher, A. (2006) Structural basis for subunit assembly in UDP-glucose pyrophosphorylase from *Saccharomyces cerevisae*, *J. Mol. Biol. 364*, 551−560.
37. Steiner, T., Lamerz, A.-C., Hess, P., Breithaupt, C., Krapp, S., Bourenkov, G., Huber, R., Gerardy-Schahn, R., and Jacob, W. (2007) Open and closed structures of the UDP-glucose pyrophosphorylase from *Leishmania major*, *J. Biol. Chem. 282*, 13003−13010.
38. Lee, H. I., and Suh, S. W. (2003) Crystal structure of UDP-N-acetylglucosamine acyltransferase from *Helicobacter pylori*, *Proteins: Struct., Funct., Genet. 53*, 772−774.
39. Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreuix, D., Cambillau, C., and Bourne, Y. (1999) Crystal structure of the bifunctional N-acetylglucosamine 1-phosphate uridyltransferase from *Escherichia coli*: A paradigm for the related pyrophosphorylase superfamily, *EMBO J. 18*, 4096−4107.
40. Kostrewa, D., D'Arcy, A., Takacs, B., and Kamber, M. (2001) Crystal structures of *Streptococcus pneumoniae* N-acetylglucosamine-1-phosphate uridyltransferase, GlmU, in apo form at 2.33 Å resolution and in complex with UDP-N-acetylglucosamine and Mg^{2+} at 1.96 Å resolution, *J. Mol. Biol. 305*, 279−289.
41. Sulzenbacher, G., Gal, L., Peneff, C., Fassy, F., and Bourne, Y. (2001) Crystal structure of *Streptococcus pneumoniae* N-acetylglucosamine-1-phosphate uridyltransferase bound to acetyl-coenzyme A reveals a novel active site architecture, *J. Biol. Chem. 276*, 11844−11851.
42. Olsen, L. R., Vetting, M. W., and Roderick, S. L. (2007) Structure of the *E. coli* bifunctional GlmU acetyltransferase active site with substrates and products, *Protein Sci. 16*, 1230−1235.
43. Kehoe, L. E., Snidwongse, J., Courvalin, P., Rafferty, J. B., and Murray, I. A. (2003) Structural basis of Synercid (Quinupristin-Dalfopristin) resistance in gram-positive bacterial pathogens, *J. Biol. Chem. 278*, 29963−29970.
44. Beaman, T. W., Sugantino, M., and Roderick, S. L. (1998) Structure of the hexapeptide xenobiotic acetyltransferase from *Pseudomonas aeruginosa*, *Biochemistry 37*, 6689−6696.
45. Beaman, T. W., Binder, D. A., Blanchard, J. S., and Roderick, S. L. (1997) Three-dimensional structure of tetrahydrodipicolinate *N*-succinyltransferase, *Biochemistry 36*, 489−494.
46. Pompeo, F., Bourne, Y., van Heijenoort, J., Fassy, F., and Mengin-Lecreulx, D. (2001) Dissection of the bifunctional *Escherichia coli N*-acetylglucosamine-1-phosphate uridyltransferase enzyme into autonomously functional domains and evidence that trimerization is absolutely required for glucosamine-1-phosphate acetyltransferase activity and cell growth, *J. Biol. Chem. 276*, 3833−3839.
47. Johnson, C. M., Roderick, S. L., and Cook, P. F. (2005) The serine acetyltransferase reaction: Acetyl transfer from an acylpantothenyl donor to an alcohol, *Arch. Biochem. Biophys. 433*, 85−95.
48. Beaman, T. W., Blanchard, J. S., and Roderick, S. L. (1998) The conformational change and active site structure of tetrahydrodipicolinate *N*-succinyltransferase, *Biochemistry 37*, 10363−10369.
49. Pye, V. E., Tingey, A. P., Robson, R. L., and Moody, P. C. E. (2004) The structure and mechanism of serine acetyltransferase from *Escherichia coli*, *J. Biol. Chem. 279*, 40729−40736.
50. Williams, A. H., and Raetz, C. R. H. (2007) Structural basis for the acyl chain selectivity and mechanism of UDP-N-acetylglucosamine acyltransferase, *Proc. Natl. Acad. Sci. U.S.A. 104*, 13543−13550.

BI702032R