# Molecular Structure of WlbB, a Bacterial *N*-Acetyltransferase Involved in the Biosynthesis of 2,3-Diacetamido-2,3-dideoxy-D-mannuronic Acid<sup>†,‡</sup>

James B. Thoden and Hazel M. Holden\*

Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706

Received April 14, 2010; Revised Manuscript Received April 29, 2010

ABSTRACT: The pathogenic bacteria Pseudomonas aeruginosa and Bordetella pertussis contain in their outer membranes the rare sugar 2,3-diacetamido-2,3-dideoxy-D-mannuronic acid. Five enzymes are required for the biosynthesis of this sugar starting from UDP-*N*-acetylglucosamine. One of these, referred to as WlbB, is an N-acetyltransferase that converts UDP-2-acetamido-3-amino-2,3-dideoxy-D-glucuronic acid (UDP-GlcNAc3NA) to UDP-2,3-diacetamido-2,3-dideoxy-D-glucuronic acid (UDP-GlcNAc3NAcA). Here we report the three-dimensional structure of WlbB from *Bordetella petrii*. For this analysis, two ternary structures were determined to 1.43 Å resolution: one in which the protein was complexed with acetyl-CoA and UDP and the second in which the protein contained bound CoA and UDP-GlcNAc3NA. WlbB adopts a trimeric quaternary structure and belongs to the L $\beta$ H superfamily of N-acyltransferases. Each subunit contains 27  $\beta$ -strands, 23 of which form the canonical left-handed  $\beta$ -helix. There are only two hydrogen bonds that occur between the protein and the GlcNAc3NA moiety, one between  $O^{\delta 1}$  of Asn 84 and the sugar C-3' amino group and the second between the backbone amide group of Arg 94 and the sugar C-5' carboxylate. The sugar C-3' amino group is ideally positioned in the active site to attack the si face of acetyl-CoA. Given that there are no protein side chains that can function as general bases within the GlcNAc3NA binding pocket, a reaction mechanism is proposed for WlbB whereby the sulfur of CoA ultimately functions as the proton acceptor required for catalysis.

*N*-Acetyltransferases catalyze the transfer of acetyl groups from acetyl-CoA<sup>1</sup> to primary amine acceptors. They are distributed throughout nature where they play key roles in biological processes such as the sleep—wake cycles of vertebrates, the reversible modifications of histone proteins, and the inactivations of aminoglycoside antibiotics (1, 2). Substrates for these enzymes range from simple small molecules to large proteins. Thus far, two major classes of *N*-acyltransferases have been identified and are referred to as the GNAT and L $\beta$ H superfamilies. Members of the GNAT superfamily adopt an overall fold built around a central mixed  $\beta$ -sheet, which is typically composed of six  $\beta$ -strands (1, 2). As the name implies, proteins that belong to the L $\beta$ H superfamily have a structure dominated by a stunning parallel  $\beta$ -helix with repeating isoleucine-rich hexapeptide motifs and rare left-handed crossover connections (3).

Within the last two years, the molecular architectures of two intriguing *N*-acetyltransferases that function on nucleotide-linked sugars have been reported. The first of these was that of PglD, from *Campylobacter jejuni*, which specifically catalyzes the *N*-acetylation

of UDP-2-acetamido-4-amino-2,4,6-trideoxy-D-glucose to yield UDP-2,4-diacetamido-2,4,6-D-trideoxyglucose or UDP-Qui-NAc4NAc as indicated in Scheme 1 (4-6). C. jejuni, itself, is a highly unusual Gram-negative bacterium that contains N-glycosylated proteins, and QuiNAc4NAc serves as the bridge between the asparagine residue of a given glycosylated protein and the attached N-glycan moiety (4, 7). PglD functions as a trimer with each subunit folding into two separate domains, an N-terminal region containing a three-stranded parallel  $\beta$ -sheet flanked by two  $\alpha$ -helices and a C-terminal domain dominated by a lefthanded  $\beta$ -helix motif (5). From detailed structural and kinetic studies, it has been proposed that His 125 functions as an active site base to remove a proton from the amino group of the UDPlinked sugar as it attacks the carbonyl carbon of acetyl-CoA(5, 6). The actual role of His 125 is still open to debate, however, given the recent biochemical data suggesting that while it is important for catalysis, it is not absolutely essential (8).

The second *N*-acetyltransferase structure recently reported was that of QdtC from *Thermoanaerobacterium thermosaccharolyticum* E207-71 (9). As indicated in Scheme 1, QdtC catalyzes the last step in the biosynthesis of dTDP-3-acetamido-3,6dideoxy-D-glucose or Quip3NAc, an unusual dideoxysugar found in the *O*-antigens of some Gram-negative bacteria (10) and in the *S*-layer glycoprotein glycans of some Gram-positive bacteria (11). QdtC, like PglD, functions as a trimer. Each subunit of QdtC is dominated by 32  $\beta$ -strands that form the canonical L $\beta$ H motif. QdtC lacks, however, the N-terminal domain observed in PglD. Also, as opposed to PglD, there are no potential catalytic bases located within the active site cleft of QdtC. Indeed, PglD and QdtC accommodate their nucleotidelinked sugars in completely different orientations as shown in Figure 1. Given the lack of a potential catalytic base in QdtC and

<sup>&</sup>lt;sup>†</sup>This research was supported in part by National Institutes of Health Grant DK47814 to H.M.H.

<sup>&</sup>lt;sup>\*</sup>X-ray coordinates have been deposited in the Protein Data Bank of the Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (entries 3MQG and 3MQH). \*To whom correspondence should be addressed. E-mail:

<sup>\*</sup>To whom correspondence should be addressed. E-mail: Hazel\_Holden@biochem.wisc.edu. Fax: (608) 262-1319. Phone: (608) 262-4988.

<sup>&</sup>lt;sup>1</sup>Abbreviations: CoA, coenzyme A; HEPES, 4-(2-hydroxyethyl)-1piperazineethanesulfonic acid; HEPPS, 3-[4-(2-hydroxyethyl)]-1-piperazinepropanesulfonic acid; IPTG, isopropyl β-D-thiogalactopyranoside; LB, Luria-Bertani; MOPS, 3-(*N*-morpholino)propanesulfonic acid; NiNTA, nickel-nitrilotriacetic acid; PCR, polymerase chain reaction; rms, root-mean-square; TB, Terrific broth; UDP, uridine diphosphate; UMP, uridine monophosphate; Tris, tris(hydroxymethyl)aminomethane.

Scheme 1



in light of subsequent site-directed mutagenesis experiments and enzymatic activity assays, a novel catalytic mechanism for it was proposed (9). Specifically, the sulfur of acetyl-CoA ultimately serves as the catalytic base by accepting a proton from the sugar amino group as it attacks the carbonyl carbon of acetyl-CoA.

Herein we report the molecular architecture of WlbB, an Nacetyltransferase from Bordetella petrii. WlbB catalyzes the penultimate step in the biosynthesis of UDP-2,3-diacetamido-2,3-dideoxy-D-mannuronic acid or UDP-ManNAc3NAcA (Scheme 2). Specifically, as shown in Schemes 1 and 2, WlbB catalyzes the N-acetylation of UDP-2-acetamido-3-amino-2.3-dideoxy-D-glucuronic acid or UDP-GlcNAc3NA to yield UDP-GlcNAc3NAcA (12). Although a reasonably rare di-Nacetylated sugar, ManNAc3NAcA is found not only in B. petrii but also in the B-band O-antigen of the opportunistic bacterium Pseudomonas aeruginosa, a major source of nosocomial infections, and in the A-band trisaccharide of the bacterium Bordetella pertussis, the causative agent of whopping cough (13). Interestingly, mutant strains of B. pertussis that lack the A-band trisaccharide have been shown to be defective in the colonization of the respiratory tracts of BALB/c mice (14). In addition, the O-antigen in P. aeruginosa is thought to protect the bacterium from phagocytosis (15) and from serum-mediated killing (16).

For the investigation described here, WlbB was crystallized in the presence of either acetyl-CoA and UDP or CoA and its substrate UDP-GlcNAc3NA. The structures were determined and refined to a nominal resolution of 1.43 Å. The binding mode for the UDP-GlcNAc3NA ligand is similar to that observed for QdtC. WlbB is considerably smaller than QdtC, however, and thus, its L $\beta$ H motif is shorter by nine  $\beta$ -strands. The pyranosyl moiety of UDP-GlcNAc3NA is accommodated within the active site region of WlbB via only two specific interactions: one between  $O^{\delta 1}$  of Asn 84 and the sugar C-3' amino group and the second between the backbone amide group of Arg 94 and the sugar C-5' carboxylate. Importantly, the sugar C-3' amino nitrogen is in the proper position to attack the si face of acetyl-CoA. Details concerning the overall active site architecture of WlbB are presented, and implications for its catalytic mechanism are discussed.

### MATERIALS AND METHODS

*Cloning, Expression, and Purification.* Genomic DNA from *B. petrii* (ATCC BAA-461) was isolated by standard procedures. The *wlbB* gene was PCR-amplified from genomic DNA such that the forward primer 5'-AAACATATGGCCAC-TATCCATCCTACCGCCATCG and the reverse primer 5'-AAAACTCGAGTCAGGCCAGTCGGCATACGCCGTCAG added NdeI and XhoI cloning sites, respectively. The purified PCR product was A-tailed and ligated into a pGEM-T (Promega) vector for screening and DNA sequencing. A WlbB-pGEM-T vector construct of the correct sequence was then appropriately digested and ligated into a pET28b(+) (Novagen) plasmid that had been previously modified for the production of a protein with a TEV protease cleavable N-terminal hexahistidine tag (*17*).

The WlbB-pET28jt plasmid was used to transform Rosetta-(DE3) Escherichia coli cells (Novagen). The culture in LB medium was grown at 37 °C with shaking until the absorbance at 600 nm reached 0.7. The cultures were subsequently cooled in an ice bath and induced with 1.0 mM IPTG, and the cells were allowed to express protein at 16 °C for 24 h after induction. WlbB was purified by a modification to the standard procedures for Ni-NTA chromatography. Specifically, all buffers contained 500 mM NaCl; the lysis buffer contained 10% glycerol, and all steps, other than the initial sonication of the cell pellet at 4 °C, were performed at room temperature. Following purification via Ni-NTA chromatography, TEV protease was added to the pooled protein (at a 1:50 TEV protease:WlbB molar ratio), and this mixture was dialyzed against 50 mM sodium phosphate, 500 mM NaCl, 40 mM imidazole, and 1 mM DTT (pH 8.0) for 24 h. The cleaved and intact versions of WlbB, as well as the TEV protease itself, were separated by passage through a column of Ni-NTA resin. The cleaved protein was dialyzed against 10 mM Tris and 400 mM NaCl at pH 8.0 and then concentrated to 12.5 mg/mL using an extinction coefficient of  $0.83 \text{ mL mg}^{-1} \text{ cm}^{-1} \text{ at } 280 \text{ nm}.$ 

Production of UDP-2-acetamido-3-amino-2,3-dideoxy*p*-glucuronic Acid (UDP-GlcNAc3NA). UDP-GlcNAc3NA was synthesized from UDP-GlcNAc. To aid in the purification process, the synthesis was divided into two steps. The first involved the oxidation of UDP-GlcNAc to UDP-GlcNAcA (Scheme 2). A typical 10 mL reaction mixture contained 50 mM HEPPS (pH 8.5), 100 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mg of UDP-GlcNAc (0.154 mmol), 310 mg of NAD<sup>+</sup> (0.462 mmol), and 10 mg of WpbO (from *B. petrii*). The reaction was allowed to proceed at



FIGURE 1: Differences in the orientations of the nucleotide-linked sugar substrates when bound to PglD or QdtC. The ribbon representation in the background corresponds to that for QdtC. The substrates for PglD and QdtC are highlighted with blue and black bonds, respectively. All figures were prepared with PyMOL (*31*).

Scheme 2



37 °C for 4 h. The enzyme was removed via filtration with a 10 kDa cutoff Centriprep concentrator, and the enzyme-free reaction products were diluted with water in a 1:4 ratio. Purification was achieved by chromatography using a 6 mL Resource-Q column and a 120 mL gradient from 0 to 500 mM ammonium bicarbonate at pH 8.5. The desired product was identified by mass spectrometry (m/z 620.1). Fractions containing UDP-GlcNAcA were pooled and lyophilized until all traces of buffer were removed. The second step involved the oxidation and subsequent amination of UDP-GlcNAcA to yield UDP-GlcNAc3NA. A typical 400 mL reaction mixture contained 0.8 mM UDP-GlcNAcA, 50 mM

glutamate, 50 mM HEPPS (pH 8), 7  $\mu$ M WlbA (from *Thermus thermophilus*), and 6  $\mu$ M WlbC (from *B. petrii*). The reaction was allowed to proceed for 12 h at 30 °C. The enzymes were removed via filtration with a 10 kDa cutoff Centriprep concentrator, and the enzyme-free reaction products were diluted with water in a 1:3 ratio. Purification was achieved by chromatography using a 6 mL Resource-Q column and a 150 mL gradient from 0 to 450 mM ammonium bicarbonate (pH 8.5). The desired product was identified by mass spectrometry (*m*/*z* 619.0). Fractions containing the oxidized sugar product were pooled and lyophilized until all traces of buffer were removed.

The enzymatic activity of *B. petrii* WlbB was verified by its incubation (10  $\mu$ M) with 0.5 mM UDP-GlcNAc3NA and 0.5 mM acetyl-CoA at 30 °C for 2 h. Removal of the enzyme and purification of the resulting reaction products on a 1 mL Resource-Q column (20 mL gradient, 0 to 600 mM ammonium bicarbonate, pH 8.5) showed the disappearance of the peaks corresponding to UDP-GlcNAc3NA and acetyl-CoA and the formation of peaks corresponding to CoA and a new uridine-containing product. Mass spectrometry of the uridine-containing product gave an ion at m/z 661.1, which corresponds to UDP-GlcNAc3NAcA.

Structural Analysis of WlbB. Crystallization conditions were initially surveyed via the hanging drop method of vapor diffusion and using a sparse matrix screen developed in the laboratory. Trials were conducted with both the apoprotein and the protein incubated with 10 mM acetyl-CoA. Diffraction quality crystals were subsequently grown via the hanging drop method by mixing in a 1:1 ratio the protein incubated with acetyl-CoA and a precipitant solution containing 15–20% poly-(ethylene glycol) 8000 [100 mM MOPS (pH 7.5)]. The crystals belonged to the space group C222<sub>1</sub> with two trimers per asymmetric unit and the following unit cell dimensions: a = 85.8 Å, b = 158.4 Å, and c = 118.4 Å.

An initial X-ray data set was measured from this orthorhombic crystal form at 4 °C using a Bruker AXS HiSTAR area detector. The X-ray source was Cu K $\alpha$  radiation from a Rigaku RU200 generator equipped with Supper mirrors and operated at 50 kV and 90 mA. These X-ray data were processed with SAINT version 7.06A (Bruker AXS Inc.) and internally scaled with SADABS version 2005/1 (Bruker AXS Inc.).

The structure was solved via single isomorphous replacement using a crystal soaked in 1 mM methylmercury(I) chloride for

#### Table 1: X-ray Data Collection Statistics

	enzyme complexed with acetyl-CoA and UDP	enzyme complexed with CoA and UDP-GlcNAc3NA
resolution limits	$50.0-1.43(1.48-1.43)^{b}$	$50.0-1.43(1.48-1.43)^b$
no. of independent reflections	$226783(20999)^{b}$	$213775(17996)^{b}$
completeness (%)	95.6 (88.7) <sup>b</sup>	90.0 $(76.1)^{b}$
redundancy	$3.9(2.7)^{b}$	$2.9(2.1)^{b}$
avg $I/avg \sigma(I)$	$38.2(8.9)^b$	$28.3 (6.6)^b$
$R_{\rm sym}$ (%) <sup>a</sup>	$7.0(14.1)^b$	$7.7(12.5)^b$
${}^{a}R_{\text{sym}} = (\sum  I - \overline{I}  / \sum I) \times 100.$	Statistics for the highest-resolution bin in parentheses.	

### Table 2: Model Refinement Statistics

	enzyme complexed with acetyl-CoA and UDP	enzyme complexed with CoA and UDP-GlcNAc3NA
resolution limits (Å)	30.0-1.43	30.0-1.43
<i>R</i> -factor <sup><i>a</i></sup> (overall) $(\%)/no.$ of independent reflections	19.2/226763	17.6/213775
<i>R</i> -factor (working) $(\%)/no.$ of independent reflections	19.1/215388	17.4/203089
<i>R</i> -factor (free) $(\%)/no.$ of independent reflections	21.7/11375	20.6/10686
no. of protein atoms	$8749^{b}$	8796 <sup>c</sup>
no. of heteroatoms	$1945^{d}$	$2310^{e}$
average <i>B</i> value ( $Å^2$ )		
protein atoms	12.2	10.7
ligands	17.0	9.7
solvent	26.0	25.9
weighted rms deviation from ideality		
bond lengths (Å)	0.012	0.012
bond angles (deg)	2.18	2.12
general planes (Å)	0.009	0.007

 ${}^{a}R$ -factor =  $(\sum |F_o - F_c|/\sum |F_o|) \times 100$ , where  $F_o$  is the observed structure factor amplitude and  $F_c$  is the calculated structure factor amplitude.  ${}^{b}$ These include multiple conformations for Lys 57 and Arg 166 in subunit A, Glu 36, Lys 57, Glu 159, and Glu 170 in subunit B, Arg 21, Ile 121, and Glu 159 in subunit C, Ile 16 in subunit D, and Lys 57 and Glu 159 in subunit E. These include multiple conformations for Arg 49 and Arg 166 in subunit A, Glu 33, Glu 36, Arg 55, Glu 93, and Glu 159 in subunit B, Ile 22, Ile 121, Arg 149, and Glu 159 in subunit C, Ile 16 and Ile 22 in subunit D, Glu 33, Glu 93, and Glu 159 in subunit E. These include multiple conformations for Arg 49 and Arg 166 in subunit A, Glu 33, Glu 36, Arg 55, Glu 93, and Glu 159 in subunit B, Ile 22, Ile 121, Arg 149, and Glu 159 in subunit C, Ile 16 and Ile 22 in subunit D, Glu 33, Glu 93, and Glu 159 in subunit E, and Glu 159 in subunit F. <sup>d</sup>Heteroatoms include six acetyl-CoA molecules, three UDP molecules, three UMP molecules, 1144 waters, two phosphate ions, four ethylene glycol molecules, six UDP-GlcNAc3NA molecules, 180 waters, two phosphate ions, 12 ethylene glycol molecules, two sodium ions, and one 2-{2-[2-(2-{2-(2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-(2-{2-(2-(2-{2-(2-{2-(2-{2-(2-{2-(2-(2-{2-(2-{2-(2-{2-(2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-{2-(2-(2-{2-(2-(2-{2-(2-{2-(2-(2-{2-(2-(2-{2-(2-(2-{2-(2-(2-{2-(2-(2-{2-(2-(2-(2-(2-{2-(2-(2-(

24 h. Six mercury binding sites were identified with SOLVE (18), giving an overall figure of merit of 0.23 to 3.0 Å resolution. Solvent flattening and 6-fold averaging with RESOLVE (19, 20) generated an interpretable electron density map, which allowed a preliminary model to be constructed using COOT (21).

A second crystal form was obtained when 10 mM UDP was included with the acetyl-CoA in the crystallization experiments. These crystals were grown at pH 8.0 in the presence of 100 mM HEPPS, 16-20% poly(ethylene glycol) 3400, and 210 mM tetramethylammonium chloride. They belonged to the space group  $P2_1$  with two trimers per asymmetric unit and the following unit cell dimensions: a = 68.5 Å, b = 107.9 Å, c = 91.3 Å, and  $\beta = 102.6^{\circ}$ . High-resolution X-ray data from these crystals were collected at beamline 19-BM (Structural Biology Center, Argonne National Laboratory, Argonne, IL), and the data were processed and scaled with HKL-3000 (22). Prior to X-ray data collection, the crystals were transferred to a stabilization solution containing 22% poly(ethylene glycol) 3400, 300 mM NaCl, 10 mM UDP, 210 mM tetramethylammonium chloride, 25 mM acetyl-CoA, 100 mM HEPPS (pH 8.0), and 18% ethylene glycol. For preparation of the ternary complex of WlbB with CoA and UDP-GlcNAc3NA, the crystals were soaked for 6 h in a solution containing 17% poly(ethylene glycol) 3400, 300 mM NaCl, 100 mM HEPPS (pH 8.0), 210 mM tetramethylammonium chloride, 25 mM acetyl-CoA, and 25 mM UDP-GlcNAc3NA prior to being transferred to the cryoprotectant. Although acetylCoA was included in the soaking solution, only CoA was observed in the electron density map. Most likely, the UDP-sugar reacted with acetyl-CoA in the enzyme's active site, and the UDPsugar product dissociated, leaving bound CoA. A second UDPsugar substrate entered the active site, resulting in the trapping of an "abortive complex". Relevant X-ray data collection statistics are presented in Table 1.

The original structure solved from crystals belonging to the space group  $C222_1$  served as the search model for the subsequent three-dimensional analyses of the WlbB-acetyl-CoA-UDP and WlbB-CoA-UDP-GlcNAc3NA ternary complexes via molecular replacement with PHASER (23). Each model was refined with Refmac using maximum likelihood (24). Relevant refinement statistics are listed in Table 2.

Measurement of Enzymatic Activity. N-Acetyltransferase activity was monitored spectrophotometrically by following the increase in absorbance at 412 nm due to the reaction of the sulfhydryl group of the CoASH product with 5,5'-dithiobis-(2-nitrobenzoic acid) resulting in a disulfide interchange. This interchange leads to the formation of 5-thio-2-nitrobenzoic acid, which has a characteristic absorbance at 412 nm and an extinction coefficient of 14150  $M^{-1}$  cm<sup>-1</sup>. The use of this compound for quantification of CoASH was first reported by Tomkins et al. (25), and our assay method was similar to that described in ref 26. Reactions were monitored continuously with a Beckman DU 640B spectrophotometer, and enzyme activities were



FIGURE 2: The structure of WlbB complexed with acetyl-CoA and UDP. A ribbon representation of the WlbB trimer is shown in panel a with the three subunits displayed in red, wheat, and light blue and the bound ligands shown as sticks. The overall three-dimensional fold of one subunit is presented in stereo in panel b. The loop defined by Asn 84–Asp 100 interrupts the regularity of the L $\beta$ H motif.

calculated from the initial rates. Assay reaction mixtures were 100  $\mu$ L in volume and contained, in addition to enzyme and substrates, 100 mM HEPES (pH 7.5) and 5 mM 5,5'-dithiobis-(2-nitrobenzoic acid). The reactions were initiated by the addition of enzyme and monitored at 25 °C. Kinetic parameters were determined by having one substrate at a saturating concentration (>10K<sub>m</sub>) while varying the concentration of the second. Specifically,  $K_m$  for acetyl-CoA was determined by varying its concentration from 0.01 to 5.0 mM using 5.0 mM UDP-GlcNAc3-NA and the  $K_m$  for UDP-GlcNAc3NA concentration from 0.01 to 5.0 mM. Individual substrate kinetic data were fitted to eq 1 using Sigma Plot 8.

$$v = VA/(K_{\rm a} + A) \tag{1}$$

Calculated  $K_{\rm m}$  values were 83  $\pm$  5  $\mu$ M for UDP-GlcNAc3NA and 175  $\pm$  10  $\mu$ M for acetyl-CoA.

## **RESULTS AND DISCUSSION**

Structure of WlbB Complexed with Acetyl-CoA and UDP. The crystals utilized in this investigation belonged to the space group  $P2_1$ , contained two trimers per asymmetric unit, and

diffracted to 1.43 Å resolution. These trimers were not related by a local 2-fold axis, but rather by a translation. The quality of the models for the six polypeptide chains in the asymmetric unit was excellent, with 87, 12.4, and 0.5% of the  $\phi$  and  $\psi$  values lying within the core, allowed, and generously allowed regions of the Ramachandran plot, respectively (27). All six subunits demonstrated exceedingly similar architectures such that their  $\alpha$ -carbons superimposed with root-mean-square deviations between 0.08 and 0.19 Å. In light of this, the following discussion refers only to the first trimer in the X-ray coordinate file and specifically to the subunits designated "A" and "B".

The WlbB trimer, as displayed in Figure 2a, has dimensions of  $\sim 70 \text{ Å} \times 66 \text{ Å} \times 67 \text{ Å}$ , and each active site is wedged between two subunits. The surface area lost upon trimerization is  $\sim 2200 \text{ Å}^2$ . Each subunit contains all 190 amino acid residues of the crystallized protein, and Pro 147 adopts a *cis*-peptide conformation. A close-up view of one subunit is displayed in Figure 2b. There are 27  $\beta$ -strands, 23 of which form the  $\beta$ -helix. The regularity of the  $\beta$ -helix is broken between  $\beta$ -strands 14 and 15 by a loop formed between Asn 84 and Asp 100. This loop reaches from one subunit into the next of the trimer and provides major binding interactions for both the acetyl-CoA and UDP-linked sugar ligands as



(a)



FIGURE 3: Active site architecture of the WlbB-acetyl-CoA-UDP complex model. Shown in panel a is the initial electron density corresponding to the bound ligands (before they were included in the model). The map was contoured at  $\sim 1.5\sigma$  and calculated with coefficients of the form  $2F_{o} - F_{c}$ , where  $F_{o}$  is the native structure factor amplitude and  $F_{c}$  the calculated structure factor amplitude. Those amino acid residues lying within  $\sim 3.2$  Å of acetyl-CoA are shown in panel b. The active site is formed by residues contributed from both subunits in the trimer as indicated by the wheat and blue bonds and ribbon representations. Water molecules are depicted as red spheres, and potential hydrogen bonds are represented by the dashed lines.

described below. Following the last  $\beta$ -strand of the L $\beta$ H motif, the polypeptide chain folds into a four-stranded antiparallel  $\beta$ -barrel, which splays away from the main body of the trimer (Figure 2a).

For the initial structural analysis of WlbB, the protein was crystallized in the presence of acetyl-CoA and UDP. The electron densities corresponding to these bound ligands are shown in Figure 3a. As can be seen, the electron density for acetyl-CoA is extremely well-ordered whereas that for UDP is somewhat less so. A close-up view of the binding pocket for acetyl-CoA is displayed in Figure 3b. The CoA adopts a curved conformation such that N6 of the adenine ring lies within 3.1 Å of one of the carbonyl oxygens of the pantothenate unit. The adenine ring is positioned within hydrogen bonding distance of the carbonyl oxygen of Ala 129 from subunit B and two waters, whereas the ribose C-2 hydroxyl group is 2.9 Å from Asn 134 in subunit A.

The ribose adopts the C<sub>2</sub>'-endo pucker. A salt bridge occurs between the C-3 phosphoryl group on the ribose and the guanidinium group of Arg 149 of subunit A. The pyrophosphoryl group of the acetyl-CoA ligand interacts with seven waters, O<sup> $\eta$ </sup> of Tyr 86 from subunit A, and N<sup> $\varepsilon$ 2</sup> of Gln 160 from subunit B. Note that Tyr 86 lies in the loop that interrupts the regularity of the L $\beta$ H motif. The pantothenate group adopts an extended structure much like a strand of  $\beta$ -sheet where it hydrogen bonds with the backbone carbonyl groups of Val 85 and Tyr 86 from subunit A and the backbone amide groups of Ala 111 and Ala 129 from subunit B. Finally, the carbonyl oxygen of the cofactor's acetyl group interacts with the backbone amide of Asn 84 from subunit A, which initiates the extended loop (Figure 2b).

Structure of WlbB Complexed with CoA and UDP-GlcNAc3NA. The binding of a nucleotide-linked sugar to WlbB resulted in very little overall structural perturbations. Indeed, the



(a)



FIGURE 4: Active site architecture of the WlbB–CoA–UDP-GlcNAc3NA complex model. Shown in panel a is the initial electron density corresponding to the bound ligands (before they were included in the model). The map was contoured at  $\sim 1.5\sigma$  and calculated with coefficients of the form  $2F_o - F_c$ , where  $F_o$  is the native structure factor amplitude and  $F_c$  the calculated structure factor amplitude. Those amino acid residues lying within  $\sim 3.2$  Å of UDP-GlcNAc3NA are shown in panel b. Water molecules are depicted as red spheres, and potential hydrogen bonds are represented by the dashed lines. Note that the water molecule located within hydrogen bonding distance of the sugar C-3' amino group is displaced when acetyl-CoA is bound in the active site.

 $\alpha$ -carbons for the structures, with or without the UDP-linked sugar, superimpose with a rms deviation of ~0.8 Å, and the side chain conformations within the active site regions of these two complexes are virtually identical. The only notable difference is that in the WlbB-acetyl-CoA-UDP complex, water molecules occupy the position assumed by the sugar in the WlbB-CoA-UDP-sugar complex.

For the structural analysis of this ternary complex, crystals were prepared in the presence of acetyl-CoA and UDP-GlcNAc3NA, but only CoA and UDP-GlcNAc3NA were visible in the electron density map (Figure 4a). This also occurred in our previous structural analysis of QdtC where crystals were grown in the presence of acetyl-CoA and a nucleotide-linked sugar, but only CoA was observed (9). As a consequence, the structure of the WlbB ternary complex reported in this section represents that of an abortive complex. A close-up view of the active site

surrounding the UDP-linked sugar is presented in Figure 4b. The uridine ring lies within 3.2 Å of three water molecules. The ribose, which adopts the  $C_3'$ -endo pucker, hydrogen bonds to two waters and the carboxamide side chain of Asn 60 in subunit B. The carboxamide side chain of Gln 59 in subunit B and numerous solvent molecules surround the pyrophosphoryl group of the UDP-linked sugar. The pyranose of UDP-GlcNAc3NA assumes the <sup>4</sup>C<sub>1</sub> chair conformation. Only two specific interactions occur between the protein (subunit A) and the GlcNAc3NA moiety: the backbone amide group of Arg 94 interacts with the sugar carboxylate group at C-5', and  $O^{\delta 1}$  of Asn 84 lies within 2.8 Å of the sugar C-3' amino group that is ultimately acetylated during the reaction. As indicated by the red dashed line in Figure 4b, the sugar C-3' amino group is positioned within 3.2 Å of the CoA sulfur atom. The lack of specific interactions between the pyranosyl moiety and the protein suggests that WlbB



FIGURE 5: Superposition of WlbB and QdtC. Ribbon representations for the WlbB and QdtC subunits are highlighted in light blue and white, respectively. The CoA and nucleotide-linked sugars, when bound to WlbB and QdtC, are displayed with black and white bonds, respectively.



FIGURE 6: Close-up of the WlbB active site. For this figure, the two structures described in this report were superimposed, and hydrogens were added to the C-3' amino group of the sugar substrate. The green dashed line represents the hypothetical approach of the lone pair of electrons on the amino nitrogen toward the *si* face of acetyl-CoA. The black dashed lines do not specifically imply hydrogen bonds but merely indicate distances of 2.2-2.9 Å between the hydrogens on the C-3' nitrogen and either O<sup>\delta</sup>1 of Asn 84 or the sulfur of CoA.

may accept alternative substrates. Experiments to test this are presently underway.

Comparison of WlbB to QdtC. As discussed previously, QdtC is an N-acetyltransferase that catalyzes the acetylation of the C-3' amino group of dTDP-Quip3N (Scheme 1). The overall level of amino acid sequence identity between QdtC and WlbB is 36%. A superposition of the two structures is presented in Figure 5. Whereas WlbB is considerably smaller than QdtC (190 and 265 amino acid residues, respectively), both display nearly identical binding orientations of the nucleotide-linked sugars. The two proteins initially align at Ala 2 in WlbB and Ala 77 in OdtC. A five-residue insertion occurs after Pro 88 in WlbB such that the two proteins begin to align again at Glu 97 in WlbB and Glu 167 in QdtC. Their close structural correspondence continues until Met 154 in WlbB and Ile 224 in QdtC. Indeed, the  $\alpha$ -carbons for 123 target pairs between WlbB and QdtC superimpose with a rms deviation of 0.71 Å. The specific interactions between WlbB or QdtC and their respective nucleotide-linked sugars are quite different, however. In QdtC, three amino acid side chains are located within hydrogen bonding distance of the sugar's amino and hydroxyl groups, namely, Glu 141, Asn 159, and Asp 160 (9). These residues were tested for their roles in catalysis by preparing the following site-directed mutant proteins: E141O, E141A,

N159D, N159A, D160N, and D160A (9). All of these proteins retained catalytic activity, albeit the activity was seriously reduced for the D160A but not the D160N mutant enzyme. In WlbB, these residues correspond to Asp 66, Asn 84, and Val 85. The only conserved interaction between the nucleotide-linked sugars and these two proteins is the hydrogen bond between the side chain of Asn 84 in WlbB (Asn 159 in QdtC) and the C-3' amino group. As in the case of QdtC, there are no potential catalytic bases within the active site of WlbB to deprotonate the C-3' amino group for its attack at the carbonyl carbon of acetyl-CoA.

Implications for the Catalytic Mechanism of WlbB. The two complex structures of WlbB described here show the binding conformations for acetyl-CoA and CoA and thus represent the "pre" and "post" reaction states of the cofactor, respectively. A superposition of these two "states" is presented in Figure 6. Note that the sulfur atom of the cofactor moves in the active site by ~2 Å upon departure of the acetyl group. In Figure 6, these two conformations of the cofactor are displayed as well as the binding position of the UDP-GlcNAc3NA substrate. Assuming that the amino group of the UDP-GlcNAc3NA substrate enters the WlbB active site in an unprotonated state, the hypothetical positions for the two hydrogens attached to nitrogen are also included in Figure 6. A potential catalytic mechanism can thus be envisioned as follows. The C-3' amino nitrogen of the substrate attacks the si face of the acetyl moiety of acetyl-CoA to produce a tetrahedral oxyanion intermediate. This intermediate is most likely stabilized by interaction of the oxyanion with the backbone amide group of Asn 84 (Figure 3b). In the model shown in Figure 6, the amino nitrogen is located within 2.2 A of the carbonyl carbon of acetyl-CoA and is oriented such that its lone pair of electrons is in the proper location for nucleophilic attack. The two hydrogens attached to the amino nitrogen lie within hydrogen bonding distance of Asn 84, which most likely plays a key role in proper substrate positioning. As the oxyanion intermediate collapses, the bond between the carbon and the sulfur of acetyl-CoA breaks and the sulfur shifts its position in the active site as indicated in Figure 6 and accepts a proton from the C-3' amino group.

We previously suggested a similar reaction mechanism for QdtC on the basis of X-ray crystallographic analyses, sitedirected mutagenesis, and kinetic experiments (9). At the time, we assumed that the catalytic mechanism of QdtC might be an "outlier" among sugar-modifying N-acetyltransferases in light of the structure of PglD, which contains a putative catalytic base. However, from this study, we now know that QdtC and WlbB accommodate their sugar substrates in nearly identical fashions and in marked contrast to that observed for PglD (Figure 1). In addition, both QdtC and WlbB contain asparagine residues that form hydrogen bonds to the substrate amino groups. Interestingly, these enzymes operate on amino groups attached to the C-3' atoms of the sugar substrates (Scheme 1). PglD, on the other hand, acetylates an amino group attached to C-4'. On the basis of these observations, we suggest that there are two subfamilies of  $L\beta H$  N-acetyltransferases that operate on nucleotide-linked sugars: those that belong to the "PglD" group, which function on amino groups attached to C-4' atoms, and those that belong to the "QdtC" family, which operate on substrates containing C-3' amino groups.

As shown in Scheme 1, GDP-perosamine N-acetyltransferase (PerB) catalyzes the acetylation of GDP-perosamine to yield GDP-N-acetylperosamine, a 6-deoxyhexose found on the lipopolysaccharide of E. coli O157:H7, the causative agent of enterohemorrhagic diarrhea first identified in contaminated hamburgers and subsequently in other foods, including spinach (28, 29). We suggest that its structure will belong to the PglD subfamily. Another N-acetyltransferase, FdtC from Aneurinibacillus thermoaerophilus L420-91<sup>T</sup>, catalyzes the acetylation of dTDP-3-amino-3,6-dideoxygalactose or dTDP-Fucp3N to yield dTDP-Fucp3NAc (11). This unusual dideoxy sugar is found in the S-layer of many Gram-positive and Gram-negative bacteria (30) and in the O-antigens of some Gram-negative bacteria (10). Given that FdtC functions on a C-3' amino group, we would predict that it belongs to the QdtC subfamily. X-ray crystallographic analyses of these two enzymes are presently underway.

#### ACKNOWLEDGMENT

A portion of the research described in this paper was performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source (U.S. Department of Energy, Office of Biological and Environmental Research, under Contract DE-AC02-06CH11357). We gratefully acknowledge Dr. Norma E. C. Duke for assistance during the X-ray data collection at Argonne and Professor W. W. Cleland for helpful discussions.

#### REFERENCES

- Dyda, F., Klein, D. C., and Hickman, A. B. (2000) GCN5-related N-acetyltransferases: A structural overview. Annu. Rev. Biophys. Biomol. Struct. 29, 81–103.
- Vetting, M. W., S de Carvalho, L. P., Yu, M., Hegde, S. S., Magnet, S., Roderick, S. L., and Blanchard, J. S. (2005) Structure and functions of the GNAT superfamily of acetyltransferases. *Arch. Biochem. Biophys.* 433, 212–226.
- Raetz, C. R., and Roderick, S. L. (1995) A left-handed parallel β helix in the structure of UDP-*N*-acetylglucosamine acyltransferase. *Science* 270, 997–1000.
- Szymanski, C. M., and Wren, B. W. (2005) Protein glycosylation in bacterial mucosal pathogens. *Nat. Rev. Microbiol.* 3, 225–237.
- Rangarajan, E. S., Ruane, K. M., Sulea, T., Watson, D. C., Proteau, A., Leclerc, S., Cygler, M., Matte, A., and Young, N. M. (2008) Structure and active site residues of PglD, an *N*-acetyltransferase from the bacillosamine synthetic pathway required for *N*-glycan synthesis in *Campylobacter jejuni. Biochemistry* 47, 1827–1836.
- Olivier, N. B., and Imperiali, B. (2008) Crystal structure and catalytic mechanism of PglD from *Campylobacter jejuni. J. Biol. Chem. 283*, 27937–27946.
- Guerry, P., and Szymanski, C. M. (2008) Campylobacter sugars sticking out. Trends Microbiol. 16, 428–435.
- Demendi, M., and Creuzenet, C. (2009) Cj1123c (PglD), a multifaceted acetyltransferase from *Campylobacter jejuni*. *Biochem. Cell Biol.* 87, 469–483.
- Thoden, J. B., Cook, P. D., Schaffer, C., Messner, P., and Holden, H. M. (2009) Structural and functional studies of QdtC: An *N*acetyltransferase required for the biosynthesis of dTDP-3-acetamido-3,6-dideoxy-α-D-glucose. *Biochemistry* 48, 2699–2709.
- Feng, L., Wang, W., Tao, J., Guo, H., Krause, G., Beutin, L., and Wang, L. (2004) Identification of *Escherichia coli* O114 *O*-antigen gene cluster and development of an O114 serogroup-specific PCR assay. *J. Clin. Microbiol.* 42, 3799–3804.
- Pfoestl, A., Hofinger, A., Kosma, P., and Messner, P. (2003) Biosynthesis of dTDP-3-acetamido-3,6-dideoxy-α-D-galactose in *Aneurinibacillus thermoaerophilus* L420-91T. J. Biol. Chem. 278, 26410–26417.
- Westman, E. L., Preston, A., Field, R. A., and Lam, J. S. (2008) Biosynthesis of a rare di-*N*-acetylated sugar in the lipopolysaccharides of both *Pseudomonas aeruginosa* and *Bordetella pertussis* occurs via an identical scheme despite different gene clusters. *J. Bacteriol. 190*, 6060–6069.
- Kochetkov, N. K., and Knirel, Y. A. (1994) Structure of lipopolysaccharides from Gram-negative bacteria. III. Structure of O-specific polysaccharides. *Biochemistry (Moscow, Russ. Fed.)* 59, 1325–1383.
- Harvill, E. T., Preston, A., Cotter, P. A., Allen, A. G., Maskell, D. J., and Miller, J. F. (2000) Multiple roles for *Bordetella* lipopolysaccharide molecules during respiratory tract infection. *Infect. Immun.* 68, 6720–6728.
- Engels, W., Endert, J., Kamps, M. A., and van Boven, C. P. (1985) Role of lipopolysaccharide in opsonization and phagocytosis of *Pseudomonas aeruginosa. Infect. Immun.* 49, 182–189.
- Dasgupta, T., de Kievit, T. R., Masoud, H., Altman, E., Richards, J. C., Sadovskaya, I., Speert, D. P., and Lam, J. S. (1994) Characterization of lipopolysaccharide-deficient mutants of *Pseudomonas aeruginosa* derived from serotypes O3, O5, and O6. *Infect. Immun. 62*, 809–817.
- Thoden, J. B., Timson, D. J., Reece, R. J., and Holden, H. M. (2005) Molecular structure of human galactokinase: Implications for Type II galactosemia. J. Biol. Chem. 280, 9662–9670.
- Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D55* (Part 4), 849–861.
- Terwilliger, T. C. (2000) Maximum-likelihood density modification. Acta Crystallogr. D56 (Part 8), 965–972.
- Terwilliger, T. C. (2003) Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D59*, 38–44.
- Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126–2132.
- Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: The integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr. D62*, 859–866.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674.

- 24. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D53*, 240–255.
- Alpers, D. H., Appel, S. H., and Tomkins, G. M. (1965) A spectrophotometric assay for thiogalactoside transacetylase. *J. Biol. Chem.* 240, 10–13.
- Magalhaes, M. L., and Blanchard, J. S. (2005) The kinetic mechanism of AAC3-IV aminoglycoside acetyltransferase from *Escherichia coli*. *Biochemistry* 44, 16275–16283.
- 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.
- Riley, L. W., Remis, R. S., Helgerson, S. D., McGee, H. B., Wells, J. G., Davis, B. R., Hebert, R. J., Olcott, E. S., Johnson, L. M., Hargrett, N. T., Blake, P. A., and Cohen, M. L. (1983) Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N. Engl. J. Med.* 308, 681–685.
- Albermann, C., and Beuttler, H. (2008) Identification of the GDP-Nacetyl-D-perosamine producing enzymes from *Escherichia coli* O157: H7. *FEBS Lett.* 582, 479–484.
- Pfostl, A., Zayni, S., Hofinger, A., Kosma, P., Schaffer, C., and Messner, P. (2008) Biosynthesis of dTDP-3-acetamido-3,6-dideoxyα-D-glucose. *Biochem. J.* 410, 187–194.
- DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA.