

Catalytic Mechanism of Perosamine *N*-Acetyltransferase Revealed by High-Resolution X-ray Crystallographic Studies and Kinetic Analyses

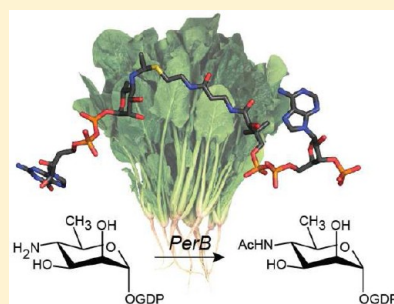
James B. Thoden,[†] Laurie A. Reinhardt,[†] Paul D. Cook,[‡] Patrick Menden,[§] W. W. Cleland,^{*,†} and Hazel M. Holden^{*,†}

[†]Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, United States

[‡]Department of Chemistry and Biochemistry, University of Mount Union, Alliance, Ohio 44601, United States

[§]McArdle Laboratory for Cancer Research, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin 53706, United States

ABSTRACT: *N*-Acetylperosamine is an unusual dideoxysugar found in the *O*-antigens of some Gram-negative bacteria, including the pathogenic *Escherichia coli* strain O157:H7. The last step in its biosynthesis is catalyzed by PerB, an *N*-acetyltransferase belonging to the left-handed β -helix superfamily of proteins. Here we describe a combined structural and functional investigation of PerB from *Caulobacter crescentus*. For this study, three structures were determined to 1.0 Å resolution or better: the enzyme in complex with CoA and GDP-perosamine, the protein with bound CoA and GDP-*N*-acetylperosamine, and the enzyme containing a tetrahedral transition state mimic bound in the active site. Each subunit of the trimeric enzyme folds into two distinct regions. The N-terminal domain is globular and dominated by a six-stranded mainly parallel β -sheet. It provides most of the interactions between the protein and GDP-perosamine. The C-terminal domain consists of a left-handed β -helix, which has nearly seven turns. This region provides the scaffold for CoA binding. On the basis of these high-resolution structures, site-directed mutant proteins were constructed to test the roles of His 141 and Asp 142 in the catalytic mechanism. Kinetic data and pH-rate profiles are indicative of His 141 serving as a general base. In addition, the backbone amide group of Gly 159 provides an oxyanion hole for stabilization of the tetrahedral transition state. The pH-rate profiles are also consistent with the GDP-linked amino sugar substrate entering the active site in its unprotonated form. Finally, for this investigation, we show that PerB can accept GDP-3-deoxyperosamine as an alternative substrate, thus representing the production of a novel trideoxysugar.



Escherichia coli O157:H7 was first identified as a pathogenic organism in 1982 during two food poisoning outbreaks in Oregon and Michigan.¹ Since then, it has been shown to be the most common disease-causing *E. coli* strain in North America.² In 2006, for example, it was responsible for the spinach contamination that led to a massive infection outbreak.

The *O*-antigen of *E. coli* O157:H7 consists of a four-sugar repeating unit with the structure [\rightarrow 2- α -D-*N*-acetylperosamine-(1 \rightarrow 3)- α -L-fucose-(1 \rightarrow 4)- β -D-glucose-(1 \rightarrow 3)- α -D-*N*-acetylgalactose \rightarrow].³ The *N*-acetylperosamine moiety is reasonably rare but has also been identified in additional bacterial species, including *Vibrio cholerae*, *Citrobacter youngae*, and *Caulobacter crescentus*, among others.^{4–6}

The biosynthesis of *N*-acetylperosamine in both *E. coli* and *C. crescentus* involves four enzymatic steps starting with mannose 1-phosphate as highlighted in Scheme 1.^{6,7} The first step involves the attachment of a GMP moiety to the sugar phosphate substrate. This reaction is catalyzed by mannose-1-phosphate guanylyltransferase. In the next step, GDP-mannose is converted to GDP-4-keto-6-deoxy-D-mannose by the action of GDP-mannose-4,6-dehydratase, an enzyme that has been well characterized both biochemically and structurally.^{8–10} The third step of the pathway is catalyzed by GDP-perosamine synthase (PerA), a pyridoxal 5'-phosphate-dependent enzyme

belonging to the aspartate aminotransferase superfamily.^{11–13} Completion of the pathway involves the acetylation of the sugar C-4''' amino group by an *N*-acetyltransferase referred to as PerB.^{6,7}

On the basis of amino acid sequence analyses, it is clear that PerB belongs to the left-handed β -helix family (L β H) of *N*-acetyltransferases. Members of this family are characterized by a left-handed β -helix, which was first observed in UDP-*N*-acetylglucosamine acyltransferase.¹⁴ In recent years, the three-dimensional structures of various *N*-acetyltransferases and *N*-acyltransferases functioning specifically on nucleotide-linked sugars have been reported.^{15–24} Some of these enzymes acetylate the amino groups attached to C-3''' of the hexose, whereas others operate on the amino groups at C-4''' as in the case of PerB. Likewise, some have catalytic mechanisms that presumably involve a histidine serving as a general base, whereas others have active sites devoid of such residues.²⁴ It is thought in these enzymes that the sulfur of CoA ultimately serves as the catalytic base.²¹ In all cases reported thus far, however, these enzymes function on either dTDP- or UDP-

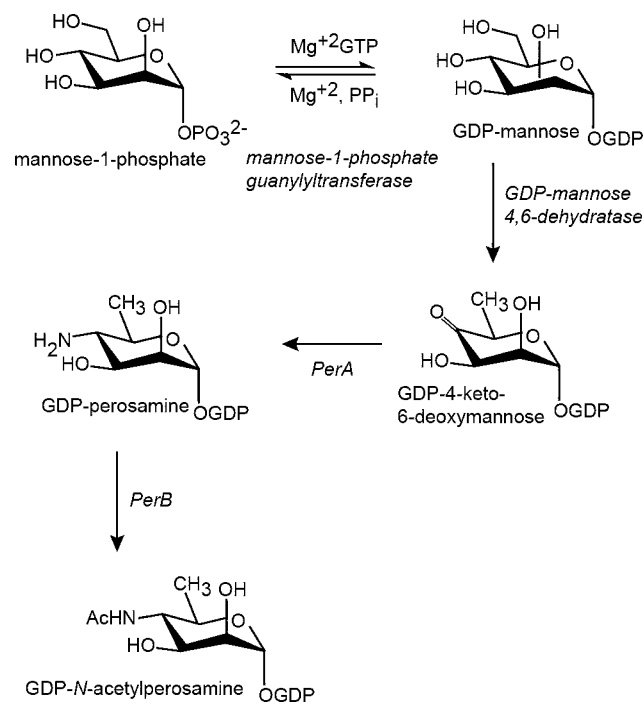
Received: February 13, 2012

Revised: March 22, 2012

Published: March 23, 2012



Scheme 1



activated hexoses as opposed to PerB that utilizes a GDP-linked sugar as its substrate.

Here we present a combined structural and functional investigation of PerB from *C. crescentus*. For this study, three crystal structures were determined to high resolution: the enzyme complexed with CoA and GDP-perosamine, the protein with bound CoA and GDP-N-acetylperosamine, and the enzyme in which a tetrahedral transition state analogue was trapped in the active site. On the basis of these structures, site-directed mutant proteins of His 141 and Asp 142 were constructed to test their roles in catalysis. Both the kinetic analyses and the pH-rate profiles are consistent with His 141 serving as a general base in the reaction mechanism. In addition, the pH-rate profiles suggest that the amino sugar enters the active site in an unprotonated form. The investigation described here represents the first structural analysis of an *N*-acetyltransferase that employs a GDP-linked hexose as its substrate. More importantly, this study provides detailed molecular snapshots along the PerB reaction mechanism from a pseudo-Michaelis complex to a tetrahedral transition state mimic to a product complex.

MATERIALS AND METHODS

Cloning, Expression, and Purification. Genomic DNA from *C. crescentus* was obtained from American Type Culture Collection. The *PerB* gene was amplified via PCR from genomic DNA such that the forward primer 5'-AAAACATATGAGCGCTTCCTCGCCATCGGGG and the reverse primer 5'-AAACTCGAGTCACGAACGGTCTCCTTTGATCTTGGCCGG 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 sequencing. A PerB-pGEM-T vector construct of the correct sequence was then appropriately digested and ligated into a pET28(b+) (Novagen) plasmid that had been previously

modified to include a TEV cleavage site for protein production with an N-terminal His₆ tag.

The PerB-pET28 plasmid was used to transform Rosetta- (DE3) *E. coli* cells (Novagen). The culture, in lysogeny broth, was grown at 37 °C with shaking until the optical density at 600 nm reached 0.7. These cultures were then cooled in an ice-water bath, induced with 1.0 mM isopropyl thiogalactoside, and transferred to a refrigerated shaker at 16 °C. The cells were allowed to express protein at 16 °C for 24 h after induction. PerB was purified by standard procedures using Ni-nitrilotriacetic acid resin. Following purification, the protein was dialyzed against 10 mM Tris and 200 mM NaCl (pH 8.0) and then concentrated to 2 mg/mL.

Structural Analysis of PerB. Crystallization conditions were initially surveyed via the hanging drop method of vapor diffusion and using a sparse matrix screen developed in the laboratory. Experiments were conducted with the apoprotein, with protein in the presence of 5 mM acetyl CoA, and with protein in the presence of 5 mM CoA and 5 mM GDP. Small crystals were obtained under a number of conditions from pH 6 to 8 with various poly(ethylene glycol)s serving as precipitants. The limited solubility of the protein, however, precluded the growth of suitable X-ray diffraction quality crystals.

Attempts to remove the N-terminal His₆ tag by TEV protease were unsuccessful. In addition, cloning PerB into expression vectors to yield an enzyme with no His₆ tag or a C-terminally tagged version yielded protein that either did not express well or was even less soluble.

To make the TEV cleavage site more accessible, the pET28 construct was modified via mutagenesis such that three alanine residues were introduced between Met 1 and Ser 2. This protein was expressed as previously described and was purified by modifications to the standard procedures at room temperature using Ni-nitrilotriacetic acid resin. Specifically, all buffers contained 500 mM NaCl and, additionally, 10% glycerol. Purified protein was dialyzed against 10 mM Tris (pH 8) with 500 mM NaCl and 10% glycerol and subsequently digested with TEV protease at a molar ratio of 1:30 at room temperature for 24 h. The TEV protease and small amounts of noncleaved protein were removed by passing the digested mixture over nitrilotriacetic acid resin. The cleaved protein was pooled and dialyzed two times against 4 L of 10 mM Tris (pH 8) with 500 mM NaCl and then concentrated to 10 mg/mL, based on a calculated extinction constant of 3.66 mg⁻¹ mL cm⁻¹.

The enzymatic activity of the modified protein was verified by its ability to convert GDP-perosamine to GDP-N-acetylperosamine with acetyl CoA. A 1 mL reaction mixture buffered at pH 8.0 with 50 mM HEPES containing 1.0 mM acetyl CoA, 1.0 mM GDP-perosamine, and 0.2 mg/mL PerB was incubated at room temperature for 2 h. The reaction products were separated from the protein by filtration through a 10 kDa ultrafiltration membrane, diluted with 2 volumes of water, and loaded onto a 1 mL Resource-Q column. Elution with a 20 mL gradient at pH 4 from 0 to 2.25 M ammonium acetate showed the elimination of the GDP-perosamine starting material (retention volume of 7.4 mL) and the generation of a new peak with a retention volume of 12.3 mL. The identity of this peak as GDP-N-acetylperosamine was verified by ESI mass spectrometry (Mass Spectrometry/Proteomic Facility at the University of Wisconsin). The GDP-perosamine required for the assay was prepared as previously described.¹³

Table 1. X-ray Data Collection Statistics^a

	PerB/CoA/GDP-perosamine	PerB/CoA/GDP-N-acetylperosamine	PerB/CoA/GDP-perosamine adduct	H141N/GDP-perosamine	H141A/GDP-perosamine
resolution limits	50.0–1.0 (1.02–1.0)	50.0–1.0 (1.02–1.0)	50.0–0.90 (0.92–0.90)	50.0–1.45 (1.55–1.45)	50.0–1.35 (1.44–1.35)
no. of independent reflections	132706 (6420)	132601 (6425)	181485 (8807)	44567 (8131)	53728 (8988)
completeness (%)	98.3 (95.4)	98.5 (95.8)	98.6 (95.9)	99.2 (97.4)	97.2 (93.4)
redundancy	7.3 (3.3)	7.2 (3.3)	8.1 (3.6)	9.3 (4.3)	6.5 (3.0)
avg <i>I</i> /avg <i>σ</i> (<i>I</i>)	44.8 (2.3)	47.1 (2.0)	50.6 (2.4)	14.5 (2.9)	16.4 (2.1)
<i>R</i> _{sym} (%) ^b	6.8 (37.4)	6.3 (39.1)	7.2 (36.0)	6.6 (36.4)	5.3 (39.1)

^aStatistics for the highest-resolution bin are given in parentheses. ^b $R_{\text{sym}} = (\sum |I - \bar{I}| / \sum I) \times 100$.

Table 2. Refinement Statistics

	PerB/CoA/GDP-perosamine	PerB/CoA/GDP-N-acetylperosamine	PerB/CoA/GDP-perosamine adduct	H141N/GDP-perosamine	H141A/GDP-perosamine
space group	<i>I</i> 23	<i>I</i> 23	<i>I</i> 23	<i>I</i> 23	<i>I</i> 23
unit cell dimensions (Å)	114.9	115.5	115.8	115.0	115.0
resolution limits (Å)	50–1.0	50.0–1.0	50.0–0.90	50.0–1.45	50.0–1.35
<i>R</i> -factor ^a (overall) (%) / no. of reflections	15.2/132551	14.4/132447	13.9/181331	18.1/43437	18.8/53728
<i>R</i> -factor (working) (%) / no. of reflections	15.1/125918	14.2/125820	13.8/172255	17.9/41237	18.7/50998
<i>R</i> -factor (free) (%) / no. of reflections	17.5/6633	16.6/6627	15.5/9076	21.0/2200	20.4/2730
no. of protein atoms	1487 ^b	1517 ^c	1495 ^d	1504 ^e	1497 ^f
no. of heteroatoms	364 ^g	421 ^h	434 ⁱ	288 ^j	317 ^k
average <i>B</i> value (Å ²)					
protein atoms	14.1	15.0	11.9	14.0	14.5
ligands	24.6	20.5	14.6	27.8	23.3
solvent atoms	34.1	38.4	33.2	26.4	27.7
weighted root-mean-square deviation from ideality					
bond lengths (Å)	0.016	0.015	0.015	0.012	0.012
bond angles (deg)	2.41	2.45	2.29	2.11	2.10
general planes (Å)	0.010	0.011	0.009	0.011	0.010

^a $R\text{-factor} = (\sum |F_o - F_c| / \sum F_o) \times 100$, where F_o is the observed structure factor amplitude and F_c is the calculated structure factor amplitude. ^bThese include multiple conformations for Ile 12, Ser 25, Arg 77, Lys 81, Ser 92, Met 116, Ile 133, Asp 140, and Arg 174. ^cThese include multiple conformations for Ile 12, Glu 31, Asp 75, Lys 81, Arg 84, Ser 92, Arg 108, Glu 165, Arg 174, Lys 209, and Lys 211. ^dThese include multiple conformations for Ile 12, Asp 75, Arg 77, Lys 81, Ser 92, Ile 133, Leu 151, Ser 155, Ser 162, Arg 174, and Lys 209. ^eThese include multiple conformations for Ser 25, Thr 32, Asp 75, Arg 77, Met 116, Trp 126, and Asp 140. ^fThese include multiple conformations for Ile 12, Lys 20, Thr 32, Asp 75, and Arg 77. ^gThese include one CoA, one GDP-perosamine, two chloride ions, and 277 waters. ^hThese include one CoA, one GDP-N-acetylperosamine, two chloride ions, and 330 waters. ⁱThese include one CoA-GDP-perosamine adduct, three chloride ions, and 343 waters. ^jThese include one CoA, one GDP-perosamine, one chloride ion, one sodium ion, and 200 waters. ^kThese include one CoA, one GDP-perosamine, one chloride ion, one sodium ion, and 229 waters.

Crystallization conditions for modified PerB were again surveyed via the hanging drop method of vapor diffusion with either the apoprotein or protein incubated with 5 mM CoA and 5 mM GDP. X-ray diffraction quality crystals were subsequently grown by mixing in a 1:1 ratio the protein incubated with CoA and GDP and 24–30% monomethylether poly(ethylene glycol) 5000 at pH 7.5. These crystals grew to maximal dimensions of 0.8 mm × 0.8 mm × 0.4 mm and belonged to the space group *I*23 with the following unit cell dimensions: *a* = *b* = *c* = 115.0 Å. The asymmetric unit contained one subunit.

Prior to X-ray data collection, all crystals were transferred to a cryoprotectant solution containing 30% monomethylether poly(ethylene glycol) 5000, 600 mM NaCl, 5 mM GDP, 5 mM CoA, and 12% ethylene glycol. The PerB structure was initially determined by single isomorphous replacement using a crystal complexed with GDP and CoA and soaked in a solution containing 1 mM methylmercury acetate for 1 day. X-ray data sets from crystals of the native protein or the native protein soaked in mercury were measured at 100 K using a Bruker AXS

Platinum 135 CCD detector controlled with the Proteum software suite (Bruker AXS Inc.). The X-ray source was Cu *Kα* radiation from a Rigaku RU200 X-ray generator equipped with Montel optics 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.). Four mercury binding sites were identified with the program SOLVE,²⁵ giving an overall figure of merit of 0.33 to 1.7 Å resolution. Solvent flattening with RESOLVE^{26,27} generated an interpretable electron density map, which allowed construction of a preliminary model using the software package COOT.²⁸ This structure, refined with REFMAC,²⁹ served as the search model for the subsequent structural analyses of the various complexes described below via molecular replacement with the software package PHASER.³⁰

All point mutations of the modified PerB-pET28 plasmid construct were created via the Stratagene QuikChange method and sequenced to verify that no other changes had been introduced into the gene. The three mutant proteins that were

studied, H141N, H141A, and D142N, were expressed and purified in the same manner as that for the wild-type enzyme. Two of the proteins, H141N and H141A, were crystallized, and X-ray data sets from these crystals were also collected using “in-house” equipment. These structures were refined with REFMAC.²⁹ X-ray data collection and relevant refinement statistics for these mutant protein structures are listed in Tables 1 and 2, respectively.

Three different ultra-high-resolution complexes were subsequently prepared for this investigation. The first was that of the enzyme with bound CoA and GDP-perosamine. For this complex, wild-type crystals were soaked overnight in a synthetic mother liquor containing 5 mM CoA and 20 mM GDP-perosamine. The second complex was that of PerB in the presence of CoA and GDP-*N*-acetylperosamine. Wild-type crystals were soaked in a synthetic mother liquor containing 10 mM acetyl CoA and 10 mM GDP-perosamine for 6 h. The enzyme turned over, and CoA and GDP-*N*-acetylperosamine were left bound in the active site. For preparation of the transition state (or intermediate) analogue, the enzyme was cocrystallized in the presence of 10 mM CoA and 10 mM GDP-perosamine. Note that all crystals were grown at room temperature, and small crystals typically appeared within 1 day. Crystal growth was generally completed within 2 weeks.

High-resolution X-ray data sets from flash-cooled crystals of these three complexes were collected at the Structural Biology Center beamline 19-ID at a wavelength of 0.667 Å (Advanced Photon Source, Argonne National Laboratory, Argonne, IL). The X-ray data sets were processed and scaled with HKL3000.³¹ X-ray data collection statistics are listed in Table 1. These structures were initially refined with REFMAC.²⁹ After major changes and most solvent molecules had been identified, the structures were then subjected to a restrained conjugate gradient least-squares process with SHELXL-97.³² All non-hydrogen atoms were refined with anisotropic thermal parameters, and hydrogen atoms were included in the final rounds of refinement in their idealized positions for all protein and ligand atoms.

Measurement of the Kinetic Constants for the Wild-Type Enzyme and the Site-Directed Mutant Proteins. The kinetic constants for the wild-type PerB and the site-directed mutant proteins were determined via a discontinuous assay using an AKTA HPLC system equipped with a 1 mL Resource-Q column. The reaction rates were determined by calculating the amount of GDP-*N*-acetylperosamine produced on the basis of the peak area in the HPLC trace. The area was correlated to concentration via a calibration curve created with standard samples that had been treated in the same manner as the reaction aliquots.

To determine the kinetic parameters for the wild-type enzyme at pH 8, nine reactions were analyzed. The GDP-perosamine concentrations were varied from 0.14 to 5.0 mM at a constant acetyl CoA concentration of 2.76 mM, whereas the acetyl CoA concentrations were varied from 0.07 to 2.76 mM at a constant GDP-perosamine concentration of 5.0 mM. All experiments were performed at 25 °C, and each reaction (550 μL) was initiated by the addition of enzyme to a final concentration of 45 μg/mL. For each reaction, 100 μL aliquots were taken at time zero and at 1–2 min intervals over a span of 8 min. The individual aliquots were immediately quenched by the addition of 6 μL of 6 M HCl. Subsequently, 40 μL of CCl₄ was added to each sample, which was then vortexed and centrifuged for 1 min to remove denatured protein. An 80 μL

aliquot of the aqueous phase was taken from each sample and diluted with 600 μL of water, and 500 μL of this solution was loaded onto the Resource-Q column for analysis. All HPLC analyses were performed using solutions at pH 4 and a gradient of 0 to 2.25 M ammonium acetate.

A plot of concentration versus time was generated for each reaction, which allowed for initial rates to be determined. All data points were fit to the equation

$$v = (VAB)/(K_aB + K_bA + K_{ia}K_b + AB)$$

using SigmaPlot8, where K_a and K_b are the Michaelis constants for GDP-perosamine and acetyl CoA, respectively, and V is the maximal velocity.

For determination of pH–rate profiles, reactions were conducted using the following buffers at concentrations of 200 mM: MES (pH 6), MOPS (pH 7), HEPPS (pH 8), CHES (pH 9), and CAPS (pH 10 and 11). The wild-type enzyme and the three mutant proteins were analyzed in a similar manner as described above. The concentrations of the substrates and enzymes were varied as needed as well as the reaction times to obtain linear plots from which we could determine the initial velocities. Note that the acetyl CoA binding constants were essentially unaffected by changes in pH (unpublished data).

The pK_a values were obtained by fitting the data to the appropriate equations using the programs BELL (eq 1) or BEL2L (eq 2).³³ The pH–rate profiles with an initial slope of 1 and a final slope of –1 were fit with eq 1. The pH–rate profiles with an initial slope of 2 and a final slope of –1 were fit with eq 2.

$$\log y = \log[C/(1 + H/K_1 + K_2/H)] \quad (1)$$

$$\log y = \log[C/(1 + H/K_1 + K_2/H + H^2/K_3)] \quad (2)$$

All kinetic data are listed in Tables 3–5.

Table 3. Kinetic Parameters Measured at pH 8.0

protein	K_m (mM) for GDP-perosamine	K_m (mM) for acetyl CoA	k_{cat}/K_m (M ^{–1} s ^{–1}) for GDP-perosamine
wild-type	0.087 ± 0.011	0.15 ± 0.02	3.5 × 10 ⁶
H141N	0.15 ± 0.01	0.094 ± 0.017	2.1 × 10 ²
H141A	0.20 ± 0.01	0.11 ± 0.01	2.0
D142N	0.11 ± 0.01	0.059 ± 0.007	3.5 × 10 ⁵

Production of GDP-*N*-acetyl-3-deoxyperosamine. GDP-3-deoxyperosamine was produced as described previously.¹³ For the production of GDP-*N*-acetyl-3-deoxyperosamine, the typical reaction mixtures contained 2 mM GDP-3-deoxyperosamine, 1 mM acetyl CoA, and 10 μM PerB in 50 mM HEPES (pH 7.5), 50 mM NaCl, and 10 mM MgCl₂. The reactions were conducted for 2–4 h at room temperature, after which they were filtered through a 10 kDa Amicon filter to remove the protein. The reaction mixtures were then analyzed via HPLC after being loaded onto a 1 mL Resource-Q (GE Healthcare) anion exchange column equilibrated with 20 mM ammonium bicarbonate. Reaction components were eluted with a linear gradient to 50% with 500 mM ammonium bicarbonate, and the fractions were analyzed by ESI mass spectrometry and ¹H NMR (in D₂O).

RESULTS AND DISCUSSION

Overall Structure of PerB in Complex with CoA and the GDP-perosamine Substrate (pseudo-Michaelis com-

Table 4. Kinetic Parameters Measured from pH 6.0 to 11.0

protein	pH 6.0	pH 7.0	pH 8.0	pH 9.0	pH 10.0	pH 11.0
wild-type ^{a,b}	0.27 ± 0.03 1.9 × 10 ⁴	0.32 ± 0.05 4.3 × 10 ⁵	0.087 ± 0.011 3.5 × 10 ⁶	0.21 ± 0.03 2.3 × 10 ⁶	0.84 ± 0.06 6.9 × 10 ⁵	9.7 ± 0.4 9.6 × 10 ⁴
H141N ^{a,b}	7.2 ± 0.3 2.0	0.27 ± 0.04 4.1 × 10	0.15 ± 0.01 2.1 × 10 ²	0.097 ± 0.006 4.3 × 10 ²	0.32 ± 0.03 2.5 × 10 ²	6.5 ± 1.7 2.5 × 10
H141A ^{a,b}	nd ^c	nd ^c	0.20 ± 0.01 2.0	nd ^c	nd ^c	nd ^c
D142N ^{a,b}	0.39 ± 0.05 8.4 × 10 ³	0.30 ± 0.03 8.0 × 10 ⁴	0.11 ± 0.01 3.5 × 10 ⁵	0.074 ± 0.004 3.9 × 10 ⁵	0.27 ± 0.04 1.6 × 10 ⁵	3.5 ± 0.3 4.2 × 10 ⁴

^a K_m (mM) for GDP-perosamine, top line. ^b k_{cat}/K_m (M⁻¹ s⁻¹), bottom line. ^cNot determined.

Table 5. pK_a Values Determined from pH–Rate Profiles

protein	pK _{a1}	pK _{a2}	pK _{a3}
wild-type ^a	7.8 ± 0.3	9.3 ± 0.2	6.5 ± 0.5
H141N ^a	8.1 ± 0.2	9.7 ± 0.2	6.0 ± 0.5
D142N ^b	7.7 ± 0.1	9.9 ± 0.1	

^aFit with BEL2L. ^bFit with BELL.

plex). The crystals used for this analysis diffracted to 1.0 Å resolution and belonged to the space group *I*23. PerB functions as a trimer, and in this crystal form, the biological unit packed

along a crystallographic 3-fold rotational axis resulting in one subunit per asymmetric unit. Each subunit of the trimer contains 215 amino acid residues.⁶ With the exception of the first two N-terminal residues, the loop between Pro 41 and Arg 43, and the final three C-terminal residues, the electron density corresponding to the polypeptide chain backbone was well ordered. The Ramachandran plot statistics for the model, as calculated with PROCHECK,³⁴ were excellent with 90.4 and 9.6% of the ϕ and ψ angles lying within the core and allowed regions, respectively.

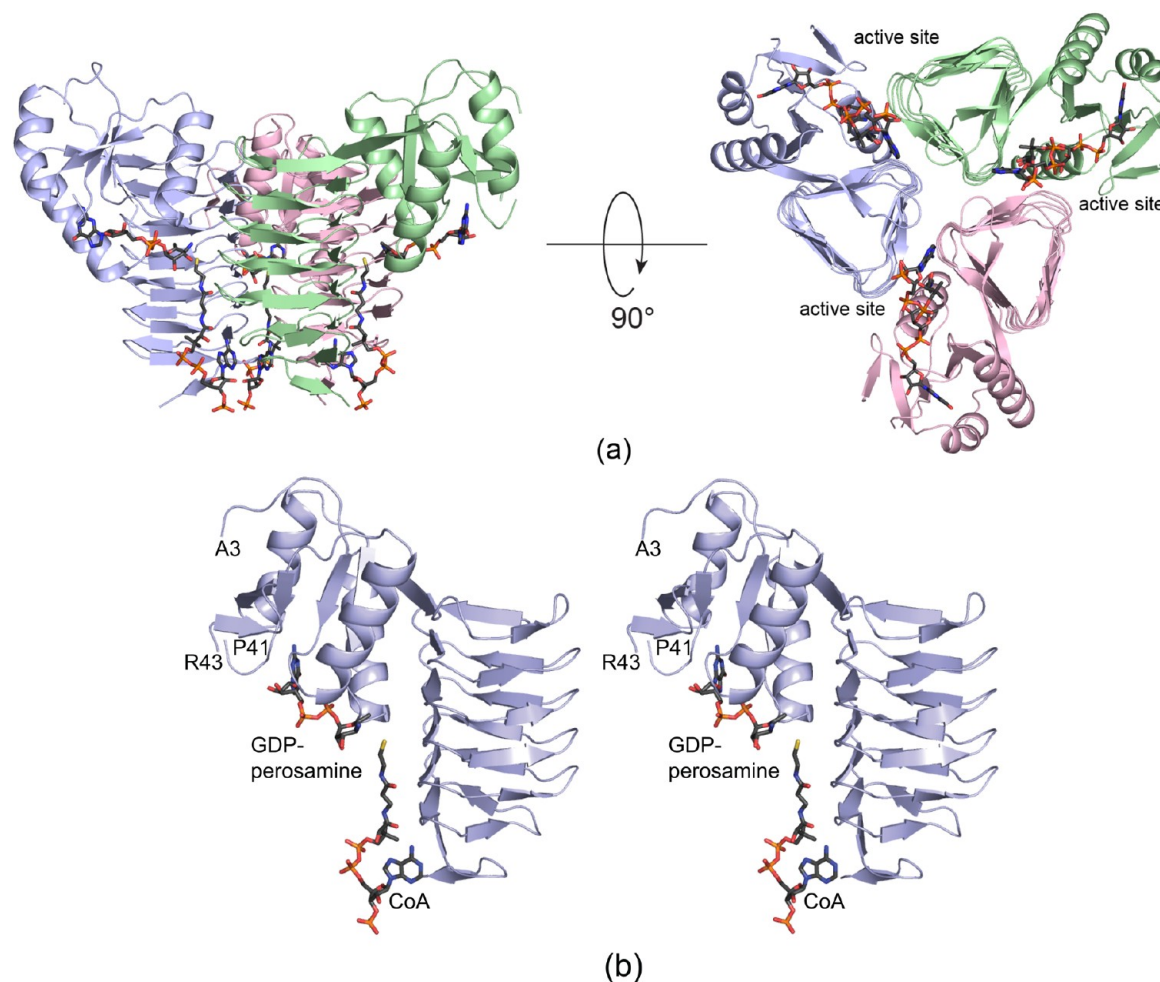


Figure 1. Ribbon representation of PerB in a complex with CoA and GDP-perosamine. The three subunits of the PerB trimer are highlighted in blue, green, and pink, respectively, in (a). The CoA and GDP-perosamine ligands are depicted in stick representations. A close-up view of one subunit of the trimer is displayed in (b). All figures were prepared with the software package PyMOL.³⁸

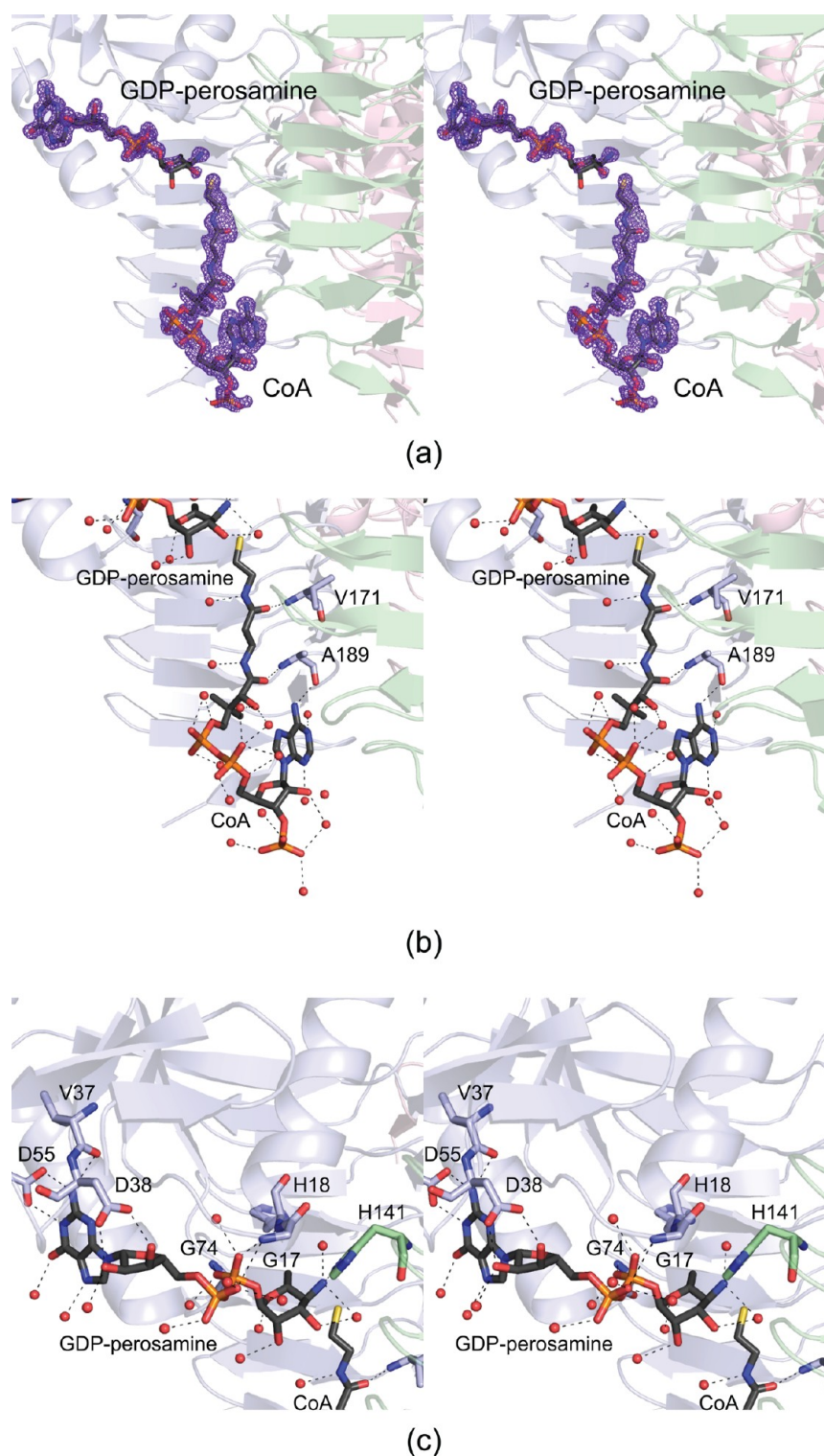


Figure 2. PerB active site with bound CoA and GDP-perosamine. The electron densities corresponding to the CoA and GDP-perosamine ligands are displayed in panel a. The map, contoured at 3σ , was calculated with coefficients of the form $F_o - F_c$, where F_o is the native structure factor amplitude and F_c the calculated structure factor amplitude. Close-up views of the CoA and GDP-perosamine binding pockets are depicted in panels b and c, respectively. Amino acid residues displayed in blue or green belong to different subunits of the trimer. Ordered water molecules are depicted as red spheres. The dashed lines indicate possible hydrogen bonding interactions.

A ribbon representation of the PerB trimer is shown in Figure 1a. It has overall dimensions of $70 \text{ \AA} \times 70 \text{ \AA} \times 70 \text{ \AA}$, and as can be seen, the active sites of the trimer are wedged between subunits. A stereoview of an individual subunit is presented in Figure 1b. The subunit can be envisioned in terms

of two domains, an N-terminal globular motif delineated by Ala 3 to Leu 93 and a β -helical region formed by Val 94 to the C-terminus. The N-terminal domain contains a six-stranded mostly parallel β -sheet flanked on one side by two α -helices and on the other by a single α -helix. The first α -helix of the subunit

is situated such that the positive end of its helix dipole moment points toward the pyrophosphoryl moiety of GDP-perosamine. The sixth β -strand of the N-terminal domain serves as a bridge to the β -helix domain, which contains nearly seven turns and displays the characteristic hexapeptide repeat that is a hallmark for this family of *N*-acetyltransferases.³⁵ Whereas the N-terminal domain provides most of the interactions between the protein and the GDP-perosamine ligand, the C-terminal domain serves to anchor the CoA moiety into the active site. Pro 207 adopts the *cis* conformation and is situated near the adenine ring of CoA.

Electron density corresponding to the bound ligands is presented in Figure 2a. The density for the CoA is very strong, as is that for the GDP portion of the substrate. The electron density for the hexose portion of GDP-perosamine is somewhat weaker, however. This most likely is a result of the short crystal soaking times employed for this structural analysis. As discussed below, if the PerB crystals were soaked in CoA and GDP-perosamine for extended periods of time, a covalent adduct formed between the two ligands.

A close-up view of the CoA binding pocket is provided in Figure 2b. The cofactor is anchored in place simply by water molecules, the backbone amide groups of Val 171 and Ala 189, and the carbonyl group of Ala 189. The sulfur of CoA is 3.6 Å from the sugar C-4' amino group. The protein region surrounding the GDP-perosamine ligand is displayed in Figure 2c. The guanine base is held in place by two water molecules, the carboxylate side chain of Asp 55, and the carbonyl group of Val 37. The carboxylate group of Asp 38 bridges the ribose C-2' and C-3' hydroxyls. In addition to numerous water molecules, the backbone amide groups of Gly 17, His 18, and Gly 74 participate in hydrogen bonding interactions with the phosphoryl oxygens. With the exception of His 141, the hexose moiety of GDP-perosamine does not interact with protein backbone atoms or side chain groups but rather simply with water molecules. His 141, which is provided by a neighboring subunit, is positioned at 2.7 Å from the hexose C-4' amino group.

Given the lack of protein side chain interactions between PerB and the hexose moiety of GDP-perosamine, we were curious as to whether we could produce a novel trideoxysugar. In a previous study, we were able to synthesize via enzymatic means GDP-3-deoxyperosamine, which is not found in nature. Using GDP-3-deoxyperosamine, we were able to, indeed, show that PerB produces an acetylated form of the sugar. Specifically, the ESI mass spectrum corresponding to a fraction from the HPLC purification containing a GDP-sugar showed a peak at 629 amu, which is the appropriate mass for GDP-*N*-acetyl-3-deoxyperosamine. The ¹H NMR spectrum for the compound isolated from the above-mentioned fraction (in D₂O) is shown in Figure 3. This spectrum corresponds closely to the ¹H NMR spectrum for GDP-3-deoxyperosamine, but with the addition of the peak near 1.8 ppm, which is the appropriate chemical shift for protons in an *N*-acetyl group.

Overall Structure of PerB in Complex with CoA and the GDP-*N*-acetylperosamine Product. The next structure determined in this study was that of PerB complexed with CoA and GDP-*N*-acetylperosamine. As in the previously described PerB/CoA/GDP-perosamine structure, the crystals employed in this analysis also belonged to space group *I*23 with one subunit in the asymmetric unit. The structure was determined to a nominal resolution of 1.0 Å, and again, 90.4 and 9.6% of the ϕ and ψ angles for the model were found to lie within the

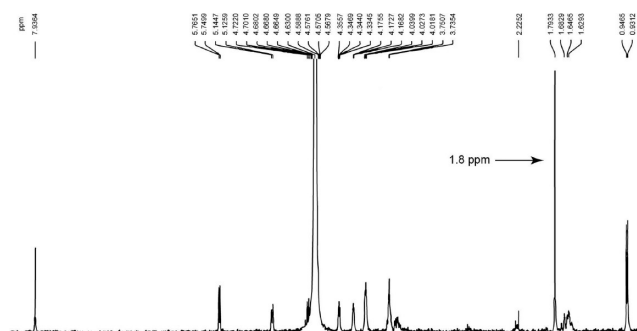


Figure 3. ¹H NMR spectrum GDP-*N*-acetyl-3-deoxyperosamine. This spectrum corresponds closely to the ¹H NMR spectrum for GDP-3-deoxyperosamine, but with the addition of the peak near 1.8 ppm, which is the appropriate chemical shift for protons in an *N*-acetyl group.

core and allowed regions, respectively, of the Ramachandran plot. There were no major conformational changes that occurred upon the binding of GDP-*N*-acetylperosamine versus GDP-perosamine. Indeed, the α -carbons for the two protein models superimpose with a root-mean-square deviation of 0.13 Å.

The electron densities corresponding to CoA and GDP-*N*-acetylperosamine are presented in panels a and b of Figure 4. As can be seen, the acetylated nucleotide-linked sugar product has clearly been trapped in the active site cleft. To the best of our knowledge, this is the first time an *N*-acetylated sugar product has been observed in the active site of any *N*-acetyltransferase that functions on nucleotide-linked sugars.

The local area surrounding the acetylated sugar is displayed in Figure 4c. His 141 shifts slightly in the active site so that it no longer interacts with the sugar amino nitrogen but rather with the C-3' hydroxyl group. The backbone amide group of Gly 159 lies within hydrogen bonding distance (3.0 Å) of the carbonyl oxygen of the acetyl moiety. Both His 141 and Gly 159 are provided by a neighboring subunit in the trimer.

Overall Structure of PerB in Complex with a Transition State Mimic. When PerB was cocrystallized in the presence of CoA and GDP-perosamine, a covalent adduct formed that bridged the sulfur of CoA and the amino nitrogen of GDP-perosamine as can be seen in the electron density map calculated to 0.9 Å resolution (Figure 5a). Given the high resolution of this structure, and the observed bond lengths, we believe this cross-linking group is CHCH₃. The hybridization about the linking carbon is clearly sp³. The question thus arises as to how this adduct was formed. The crystals were grown in the presence of poly(ethylene glycol), and commercial samples of it are notorious for being contaminated with peroxides, aldehydes, and aldehyde precursors.³⁶ A possible mechanism for the cross-linking of CoA and GDP-perosamine is shown in Scheme 2. We suggest that the C-4' amino group attacks the carbonyl carbon of acetaldehyde to form a tetrahedral intermediate. This intermediate collapses to produce a Schiff base, which is subsequently attacked by the sulfhydryl group of CoA. Clearly, PerB can accommodate acetaldehyde in the region between the CoA sulfur and the sugar nitrogen because it has to bind acetyl CoA for catalysis to ultimately occur.

A close-up view of the PerB active site with the bound CoA/GDP-perosamine adduct is displayed in Figure 5b. His 141 is positioned within 3.0 Å of the sugar nitrogen. The models of PerB with bound GDP-*N*-acetylperosamine or the covalent

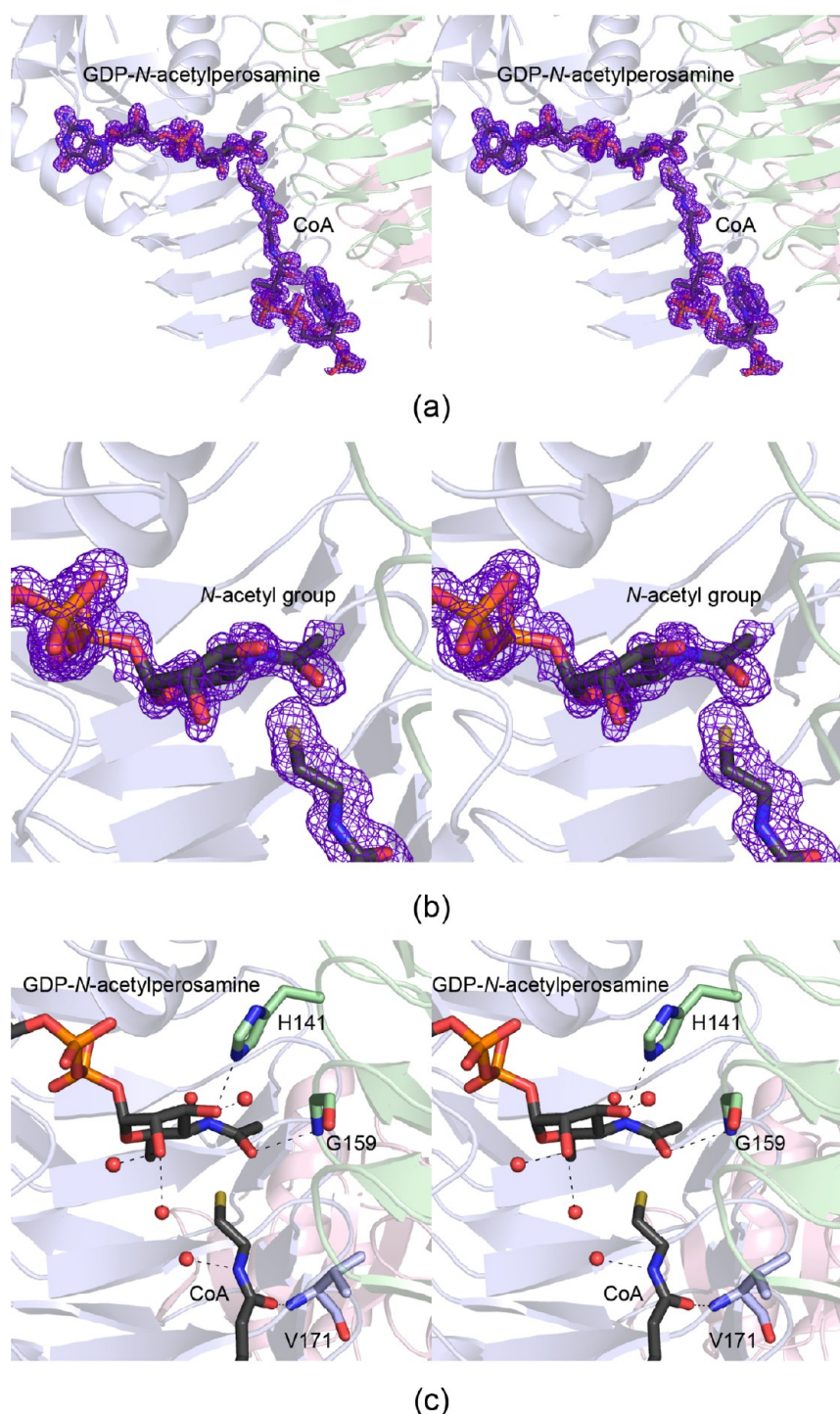


Figure 4. PerB active site with bound CoA and GDP-N-acetylperosamine. The electron densities corresponding to CoA and the nucleotide-linked sugar are presented in panel a. The map was calculated as described in the legend of Figure 2 and contoured at 3.5σ . A close-up view of the acetylated sugar product is provided in panel b. Possible hydrogen bonding interactions between the N-acetylated sugar product and the protein are depicted as dashed lines in panel c.

adduct are virtually identical such that their α -carbons superimpose with a root-mean-square deviation of 0.11 Å. The covalent adduct shown in Figure 5b serves as an excellent mimic for the tetrahedral transition state that would occur as the sugar amino nitrogen attacks the carbonyl carbon of acetyl CoA. The only difference is that the proton on the bridging carbon would be an oxygen atom in the true tetrahedral intermediate. Given that, a model for the tetrahedral

intermediate was created and is shown in Figure 5c superimposed upon the structure of the enzyme with bound GDP-N-acetylperosamine. The acetyl carbon moves by ~ 0.7 Å in the active site upon collapse of the tetrahedral intermediate. Most likely, the backbone amide of Gly 159 provides stabilization of the oxyanion that forms during the reaction mechanism.

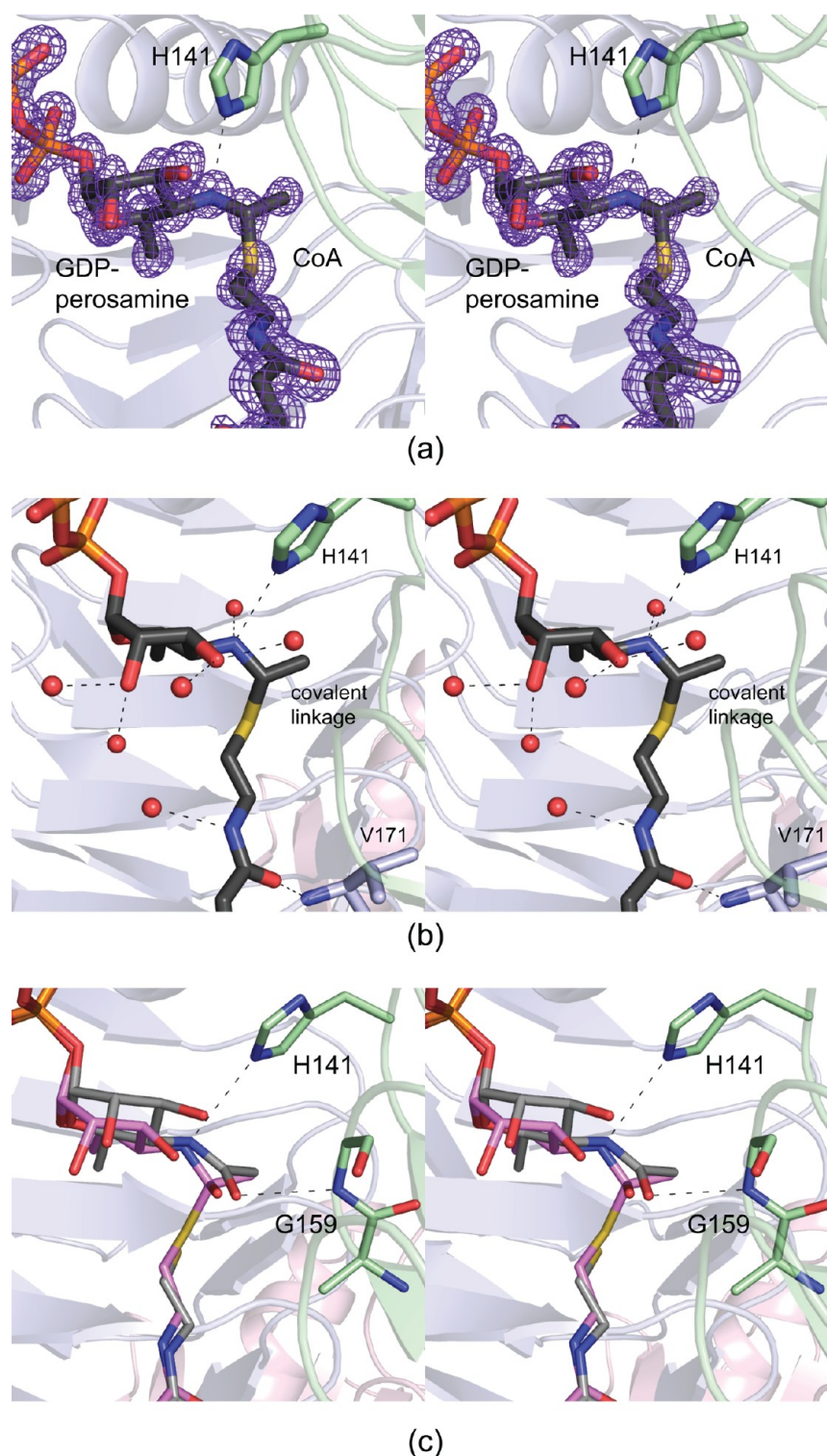
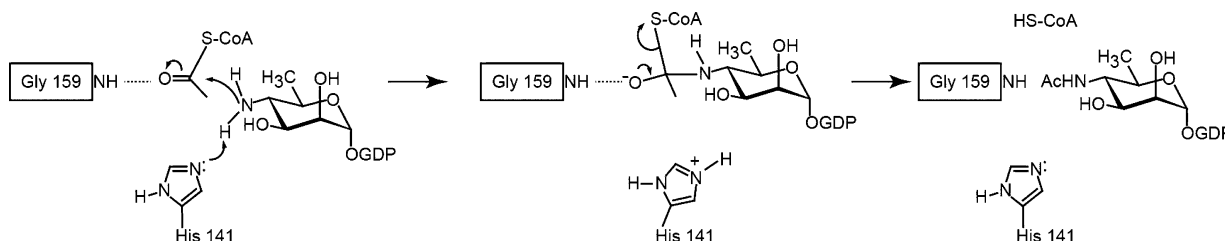


Figure 5. PerB active site with a CoA/GDP-perosamine covalent adduct. The electron density corresponding to the covalent adduct is displayed in panel a. The map was calculated as described in the legend of Figure 2 and contoured at 3.5σ . Possible hydrogen bonding interactions between the ligand and the protein are indicated by the dashed lines in panel b. Ordered water molecules are displayed as spheres. A superposition of GDP-*N*-acetylperosamine (gray filled bonds) onto the covalent adduct (pink filled bonds) is depicted in panel c.

Probing the Catalytic Mechanism of PerB. The closest structural relative to PerB is an *N*-acetyltransferase referred to as PlgD from *Campylobacter jejuni*.^{19,20} It catalyzes the last step in the biosynthesis of 2,4-diacetamido-2,4,6-trideoxy- α -D-glucose, an unusual sugar found in the glycan moieties of some eubacterial pathogens, and it employs a UDP-linked sugar

substrate. PerB and PlgD show an amino acid sequence identity of 38%, and not surprisingly, their models superimpose with a root-mean-square deviation of 0.73 Å for 157 structurally equivalent α -carbon positions. A superposition of the PerB active site onto that of PlgD is presented in Figure 6. There are two substitutions in PlgD, relative to PerB, which preclude its

Scheme 3



protons from the amino group of the substrate as it attacks acetyl CoA.

Both of these mechanisms are consistent with the pH-rate profiles, but the second allows His 141 to act in a truly catalytic manner. In the first mechanism, His 141 deprotonates the substrate only to its neutral form, but there is then no catalysis of the transacylation reaction itself. In the second mechanism, His 141 acts as a general base to catalyze the formation of the tetrahedral transition state or intermediate, and thus, it is expected that its mutation would have a major effect as observed. We thus conclude that the pH-rate profiles are reverse protonation profiles, and that the substrate enters the active site with a neutral amino group. The mechanism for PglD proposed by Rangarajan¹⁹ also assumes that the substrate binds to the enzyme in its neutral state. A possible reaction mechanism for PerB is shown in Scheme 3. The key players in the mechanism include the backbone amide of Gly 159, which provides an "oxyanion" hole, and the imidazole of His 141, which serves as the general base.

The catalytic mechanisms of PglD and PerB clearly involve a general base provided by a histidine residue. This is in marked contrast to the reaction mechanisms of WlbB, QdtC, and AntD, which are *N*-acetyl- or *N*-acyltransferases that have also been studied in our laboratory.^{21–23} Both WlbB and QdtC are involved in the biosynthesis of unusual acetylated sugars found in the *O*-antigens of certain Gram-negative bacteria, and both function on C-3'' hexose amino groups. AntD is involved in the production of D-anthrose, an important carbohydrate found in the endospores of *Bacillus anthracis*, the causative agent of anthrax. Specifically, AntD catalyzes the transfer of an acyl group from 3-hydroxy-3-methylbutyryl CoA to the C-4'' amino group of dTDP-4-amino-4,6-dideoxy- α -D-glucose. In all three enzymes, there is a decided lack of a catalytic base in their active sites. Likewise, in all three enzymes, the nucleotide-linked sugars are bound similarly, but in an orientation completely different from that observed for PerB and PglD. Clearly, the *N*-acetyltransferases (or *N*-acyltransferases) that function on nucleotide-linked sugars have evolved into two separate protein classes that differ with respect to both substrate binding orientations and reaction mechanisms.

■ ASSOCIATED CONTENT

Accession Codes

X-ray coordinates have been deposited in the Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (entries 4EA7, 4EA8, 4EA9, 4EAA, and 4EAB).

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Hazel_Holden@biochem.wisc.edu or cleland@biochem.wisc.edu. Fax: (608) 262-1319. Phone: (608) 262-4988.

Funding

This research was supported in part by National Institutes of Health Grants DK47814 (to H.M.H.) and GM18938 (to W.W.C.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

A portion of the research described in this paper was performed at Argonne National Laboratory, Structural Biology Center at the Advanced 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 and Dr. Stephan L. Ginell for assistance during the X-ray data collection at Argonne.

■ ABBREVIATIONS

CAPS, 3-(cyclohexylamino)propanesulfonic acid; CHES, 2-(cyclohexylamino)ethanesulfonic acid; CoA, coenzyme A; ESI, electrospray ionization; GDP, guanosine diphosphate; HEPPS, *N*-(2-hydroxyethyl)piperazine-*N'*-3-propanesulfonic acid; HPLC, high-performance liquid chromatography; MES, 2-(*N*-morpholino)ethanesulfonic acid; MOPS, 3-(*N*-morpholino)-propanesulfonic acid; NMR, nuclear magnetic resonance; PCR, polymerase chain reaction; TEV, tobacco etch virus; Tris, tris(hydroxymethyl)aminomethane.

■ REFERENCES

- (1) Wells, J. G., Davis, B. R., Wachsmuth, I. K., Riley, L. W., Remis, R. S., Sokolow, R., and Morris, G. K. (1983) Laboratory investigation of hemorrhagic colitis outbreaks associated with a rare *Escherichia coli* serotype. *J. Clin. Microbiol.* 18, 512–520.
- (2) Mead, P. S., and Griffin, P. M. (1998) *Escherichia coli* O157:H7. *Lancet* 352, 1207–1212.
- (3) Perry, M. B., MacLean, L., and Griffith, D. W. (1986) Structure of the *O*-chain polysaccharide of the phenol-phase soluble lipopolysaccharide of *Escherichia coli* O:157:H7. *Biochem. Cell Biol.* 64, 21–28.
- (4) Haishima, Y., Kondo, S., and Hisatsune, K. (1990) The occurrence of α (1→2) linked *N*-acetylperosamine-homopolymer in lipopolysaccharides of non-O1 *Vibrio cholerae* possessing an antigenic factor in common with O1 *V. cholerae*. *Microbiol. Immunol.* 34, 1049–1054.
- (5) Ovchinnikova, O. G., Kocharova, N. A., Katzenellenbogen, E., Zatonsky, G. V., Shashkov, A. S., Knirel, Y. A., Lipinski, T., and Gamian, A. (2004) Structures of two *O*-polysaccharides of the lipopolysaccharide of *Citrobacter youngae* PCM 1538 (serogroup O9). *Carbohydr. Res.* 339, 881–884.

- (6) Awram, P., and Smit, J. (2001) Identification of lipopolysaccharide O antigen synthesis genes required for attachment of the S-layer of *Caulobacter crescentus*. *Microbiology* 147, 1451–1460.
- (7) Albermann, C., and Beutler, H. (2008) Identification of the GDP-N-acetyl-D-perosamine producing enzymes from *Escherichia coli* O157:H7. *FEBS Lett.* 582, 479–484.
- (8) Somoza, J. R., Menon, S., Schmidt, H., Joseph-McCarthy, D., Dessen, A., Stahl, M. L., Somers, W. S., and Sullivan, F. X. (2000) Structural and kinetic analysis of *Escherichia coli* GDP-mannose 4,6 dehydratase provides insights into the enzyme's catalytic mechanism and regulation by GDP-fucose. *Struct. Folding Des.* 8, 123–135.
- (9) Mulichak, A. M., Bonin, C. P., Reiter, W. D., and Garavito, R. M. (2002) Structure of the MUR1 GDP-mannose 4,6-dehydratase from *Arabidopsis thaliana*: Implications for ligand binding and specificity. *Biochemistry* 41, 15578–15589.
- (10) Webb, N. A., Mulichak, A. M., Lam, J. S., Rocchetta, H. L., and Garavito, R. M. (2004) Crystal structure of a tetrameric GDP-D-mannose 4,6-dehydratase from a bacterial GDP-D-rhamnose biosynthetic pathway. *Protein Sci.* 13, 529–539.
- (11) Albermann, C., and Piepersberg, W. (2001) Expression and identification of the RfBE protein from *Vibrio cholerae* O1 and its use for the enzymatic synthesis of GDP-D-perosamine. *Glycobiology* 11, 655–661.
- (12) Zhao, G., Liu, J., Liu, X., Chen, M., Zhang, H., and Wang, P. G. (2007) Cloning and characterization of GDP-perosamine synthetase (Per) from *Escherichia coli* O157:H7 and synthesis of GDP-perosamine *in vitro*. *Biochem. Biophys. Res. Commun.* 363, 525–530.
- (13) Cook, P. D., and Holden, H. M. (2008) GDP-perosamine synthase: Structural analysis and production of a novel trideoxysugar. *Biochemistry* 47, 2833–2840.
- (14) 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.
- (15) Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreux, D., Cambillau, C., and Bourne, Y. (1999) Crystal structure of the bifunctional N-acetylglucosamine 1-phosphate uridylyltransferase from *Escherichia coli*: A paradigm for the related pyrophosphorylase superfamily. *EMBO J.* 18, 4096–4107.
- (16) Olsen, L. R., and Roderick, S. L. (2001) Structure of the *Escherichia coli* GlmU pyrophosphorylase and acetyltransferase active sites. *Biochemistry* 40, 1913–1921.
- (17) Kostrewa, D., D'Arcy, A., Takacs, B., and Kamber, M. (2001) Crystal structures of *Streptococcus pneumoniae* N-acetylglucosamine-1-phosphate uridylyltransferase, 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.
- (18) Sulzenbacher, G., Gal, L., Peneff, C., Fassy, F., and Bourne, Y. (2001) Crystal structure of *Streptococcus pneumoniae* N-acetylglucosamine-1-phosphate uridylyltransferase bound to acetyl-coenzyme A reveals a novel active site architecture. *J. Biol. Chem.* 276, 11844–11851.
- (19) 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.
- (20) Olivier, N. B., and Imperiali, B. (2008) Crystal structure and catalytic mechanism of PglD from *Campylobacter jejuni*. *J. Biol. Chem.* 283, 27937–27946.
- (21) 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.
- (22) Thoden, J. B., and Holden, H. M. (2010) Molecular structure of WlbB, a bacterial N-acetyltransferase involved in the biosynthesis of 2,3-diacetamido-2,3-dideoxy-D-mannuronic acid. *Biochemistry* 49, 4644–4653.
- (23) Kubiak, R. L., and Holden, H. M. (2012) Structural Studies on AntD: An N-Acyltransferase Involved in the Biosynthesis of D-Anthrose. *Biochemistry* 51, 867–878.
- (24) Holden, H. M., Cook, P. D., and Thoden, J. B. (2010) Biosynthetic enzymes of unusual microbial sugars. *Curr. Opin. Struct. Biol.* 20, 543–550.
- (25) Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D55* (Part 4), 849–861.
- (26) Terwilliger, T. C. (2000) Maximum-likelihood density modification. *Acta Crystallogr. D56* (Part 8), 965–972.
- (27) Terwilliger, T. C. (2003) Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D59*, 38–44.
- (28) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126–2132.
- (29) 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.
- (30) 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.
- (31) Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, J. (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.
- (32) Sheldrick, G. M., and Schneider, T. R. (1997) SHELXL: High-resolution refinement. *Methods Enzymol.* 277, 319–343.
- (33) Cleland, W. W. (1979) Statistical analysis of enzyme kinetic data. *Methods Enzymol.* 63, 103–138.
- (34) 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.
- (35) Vuorio, R., Harkonen, T., Tolvanen, M., and Vaara, M. (1994) The novel hexapeptide motif found in the acyltransferases LpxA and LpxD of lipid A biosynthesis is conserved in various bacteria. *FEBS Lett.* 337, 289–292.
- (36) Ray, W. J., Jr., and Puvathingal, J. M. (1985) A simple procedure for removing contaminating aldehydes and peroxides from aqueous solutions of polyethylene glycols and of nonionic detergents that are based on the polyoxyethylene linkage. *Anal. Biochem.* 146, 307–312.
- (37) McCarthy, T. J., Plog, M. A., Floy, S. A., Jansen, J. A., Soukup, J. K., and Soukup, G. A. (2005) Ligand requirements for glmS ribozyme self-cleavage. *Chem. Biol.* 12, 1221–1226.
- (38) DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.