BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 7939–7948 7939
DOI: 10.1021/bi101103s

Structural and Functional Studies of WlbA: A Dehydrogenase Involved in the Biosynthesis of 2,3-Diacetamido-2,3-dideoxy-D-mannuronic Acid[†,‡]

James B. Thoden and Hazel M. Holden*

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

Received July 10, 2010; Revised Manuscript Received August 3, 2010

ABSTRACT: 2,3-Diacetamido-2,3-dideoxy-D-mannuronic acid (ManNAc3NAcA) is an unusual dideoxy sugar first identified nearly 30 years ago in the lipopolysaccharide of *Pseudomonas aeruginosa* O:3a,d. It has since been observed in other organisms, including *Bordetella pertussis*, the causative agent of whooping cough. Five enzymes are required for the biosynthesis of UDP-ManNAc3NAcA starting from UDP-*N*-acetyl-D-glucosamine. Here we describe a structural study of WlbA, the NAD-dependent dehydrogenase that catalyzes the second step in the pathway, namely, the oxidation of the C-3′ hydroxyl group on the UDP-linked sugar to a keto moiety and the reduction of NAD^+ to NADH. This enzyme has been shown to use α-ketoglutarate as an oxidant to regenerate the oxidized dinucleotide. For this investigation, three different crystal structures were determined: the enzyme with bound NAD(H), the enzyme in a complex with NAD(H) and α-ketoglutarate, and the enzyme in a complex with NAD(H) and its substrate (UDP-*N*-acetyl-D-glucosaminuronic acid). The tetrameric enzyme assumes an unusual quaternary structure with the dinucleotides positioned quite closely to one another. Both α-ketoglutarate and the UDP-linked sugar bind in the WlbA active site with their carbon atoms (C-2 and C-3′, respectively) abutting the *re* face of the cofactor. They are positioned ∼3 Å from the nicotinamide C-4. The UDP-linked sugar substrate adopts a highly unusual curved conformation when bound in the WlbA active site cleft. Lys 101 and His 185 most likely play key roles in catalysis.

Gram-negative bacteria contain in their outer membranes complex glycoconjugates termed lipopolysaccharides (LPS). Conceptually, the LPS can be thought of in terms of three regions: lipid A, the core oligosaccharide, and the O-specific polysaccharide or O-antigen (*1*). Lipid A, which anchors the sugar components of the LPS to the bacterial cell outer membrane, has been implicated in the toxicity associated with Gram-negative bacterial infections. The O-antigen, which extends farthest from the bacterium, displays the most variation from species to species and is highly immunogenic (*2*). This region consists of repeating units, which typically contain three to five sugars, and indeed, the carbohydrate differences within the O-antigens contribute to the wide varieties of bacterial strains observed in nature. There is growing evidence that O-antigens play important biological roles, including, but not limited to, virulence, effective colonization of host tissues, protection from phagocytosis and serum-mediated killing, and resistance to antimicrobial peptides (*3*, *4*).

O-Antigens often contain some quite remarkable dideoxy sugars, including tyvelose, colitose, perosamine, legionaminic acid, and the rare di-*N*-acetylated sugar, 2,3-diacetamido-2,3-dideoxy-D-mannuronic acid, shown in Scheme 1 and hereafter termed ManNAc3NAcA. These sugars are added to the LPS via the action of glycosyltransferases, which employ nucleotide-linked derivatives such as GDP-colitose, CDP-legionaminic acid, and UDP-ManNAc3NAcA as substrates.

Although ManNAc3NAcA is a relatively rare sugar, it was first identified nearly 30 years ago in the B-band O-antigen of *Pseudomonas aeruginosa* O:3a,d (*5*). Since then, it has been observed in additional *P. aeruginosa* species and in the A-band trisaccharide of the bacterium *Bordetella pertussis*, the causative agent of whooping cough (*6*, *7*). Five enzymes are required for its biosynthesis as outlined in Scheme 1 (*7*, *8*). The first step involves the oxidation of the C-6′ hydroxyl group of UDP-*N*-acetyl-D-glucosamine by WbpO, thereby leading to the formation of UDP-*N*-acetyl-D-glucosaminuronic acid (UDP-GlcNAcA).[1] The next two enzymes in the pathway, WlbA and WlbC, apparently work in concert via an NAD(H) recycling mechanism to form UDP-2-acetamido-3-amino-2,3-dideoxy-D-glucuronic acid (UDP-GlcNAc3NA) (*9*). Specifically, it has been hypothesized that the α-ketoglutarate produced by the WlbC reaction serves as the oxidant for the regeneration of NAD^+ required for the WlbA reaction (Scheme 1). In the penultimate step of the biosynthetic pathway, UDP-GlcNAc3NA is converted to UDP-2,3-diacetamido-2,3-dideoxy-D-glucuronic acid (UDP-GlcNAc3NAcA) by WlbB, an *N*-acetyltransferase whose three-dimensional structure

[†]This research was supported in part by National Institutes of Health Grant DK47814 to H.M.H.

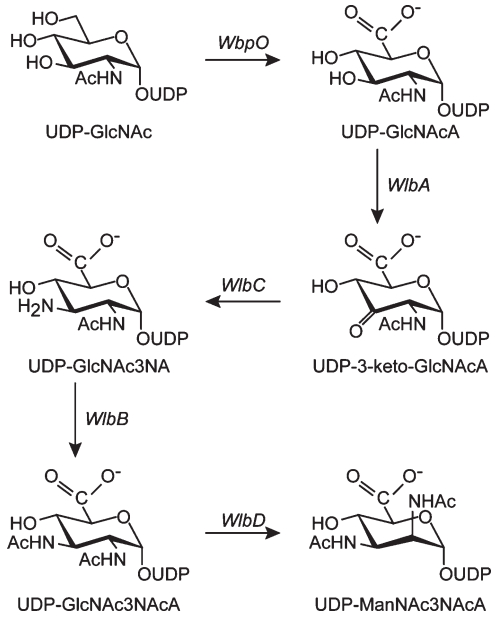
[‡]X-ray coordinates have been deposited in the Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ, as entries 3O9Z, 3OAO, and 3OA2.

*To whom correspondence should be addressed. E-mail: hazel_holden@biochem.wisc.edu. Fax: (608) 262-1319. Phone: (608) 262-4988.

[1]Abbreviations: HEPPS, 3-[4-(2-hydroxyethyl)-1-piperazinyl]propanesulfonic acid; IPTG, isopropyl β-D-thiogalactopyranoside; LB, Luria-Bertani; MOPS, 3-(*N*-morpholino)propanesulfonic acid; NAD(H), nicotinamide adenine dinucleotide; Ni-NTA, nickel-nitrilotriacetic acid; PCR, polymerase chain reaction; PIPES, piperazine-1,4-bis(2-ethanesulfonic acid); TEV, tobacco etch virus; Tris, tris(hydroxymethyl)aminomethane; UDP, uridine 5′-phosphate; UDP-GlcNAc, UDP-*N*-acetyl-D-glucosamine; UDP-GlcNAcA, UDP-*N*-acetyl-D-glucosaminuronic acid; UDP-3-keto-GlcNAcA, UDP-2-acetamido-2-deoxy-3-oxo-D-glucuronic acid; UDP-GlcNAc3NA, UDP-2-acetamido-3-amino-2,3-dideoxy-D-glucuronic acid; UDP-GlcNAc3NAcA, UDP-2,3-diacetamido-2,3-dideoxy-D-glucuronic acid; UDP-ManNAc3NAcA, UDP-2,3-diacetamido-2,3-dideoxy-D-mannuronic acid.

Published on Web 08/06/2010

Scheme 1



has recently been reported by this laboratory (*10*). The final enzyme in the pathway, WlbD, catalyzes an epimerization reaction about the C-2′ atom of the hexose, resulting in the formation of UDP-ManNAc3NAcA.

In this report, we describe a combined structural and functional analysis of WlbA, the dehydrogenase that catalyzes step 2 in UDP-ManNAc3NAcA biosynthesis. For this investigation, three high-resolution structures were determined: the enzyme with bound NAD(H), the enzyme in a complex with NAD(H) and α-ketoglutarate, and the enzyme in a complex with NAD(H) and its substrate, UDP-GlcNAcA. WlbA displays an unusual tetrameric quaternary structure with the NAD(H) moieties positioned quite close together. In addition, the UDP-GlcNAcA ligand adopts a remarkably curved conformation when bound in the active site cleft. Each subunit of the tetramer displays an overall fold similar to that reported for glucose-fructose oxidoreductase from *Zymomonas mobilis* (*11*). Details concerning the overall structure and catalytic mechanism of WlbA are presented.

MATERIALS AND METHODS

Preparation of UDP-GlcNAcA. The enzymatic synthesis of UDP-GlcNAcA was accomplished by setting up a 10 mL reaction mixture containing 50 mM HEPPS (pH 8.5), 100 mM $(NH_4)_2SO_4$, 100 mg of UDP-GlcNAc (0.154 mmol), 310 mg of NAD^+ (0.462 mmol), and 10 mg of WbpO cloned from *Bordetella petrii* (Scheme 1). The reaction was allowed to proceed at 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 by 1:4 with water. Purification was achieved by chromatography using a 6 mL Resource-Q column and a 120 mL gradient from 0 to 500 mM ammonium bicarbonate (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.

Cloning, Expression, and Purification. To produce diffraction-quality crystals of the appropriate enzyme−ligand complexes required for this investigation, two different versions of WlbA were cloned and purified, namely, one from *P. aeruginosa* PAO1-LAC (ATCC 47085) and the other from *Thermus thermophilus* (ATCC BAA-163). The necessary genomic DNA samples were obtained from the American Type Culture Collection (ATCC). Using PrimeSTAR HS DNA Polymerase (Takara BIO USA, Clontech Laboratories), the *wlba* gene from *P. aeruginosa* was PCR-amplified such that the forward primer 5′-AAACATATG*GCTAGC*AAAAATTTCGCTCTCATCGGTGCTGCCGG and the reverse primer 5′-AAA*CTCGAG*TCAACGCGCAAGCGCCGCGACGAACGGATGCCC added NdeI and XhoI cloning sites, respectively. Likewise, the *wlbA* gene from *T. thermophilus* was PCR-amplified using the forward primer 5′-AAAA*CATATG*ACCCGCTTCGCCCTCACCGGACTTGCGGG and the reverse primer 5′-AAAA*GCGGCCGC*TCACCCCAGAAAAGGATGCCGATTTTCCGGAGAAGGCTGGC. These primers added NdeI and NotI restriction sites, respectively. The purified PCR products were A-tailed and ligated into pGEM-T (Promega) vectors for screening and sequencing. The WlbA-pGEM-T vector constructs of the correct sequences were then appropriately digested and ligated into pET28JT vector plasmids for protein production with a TEV protease cleavable N-terminal His_6 tag (*12*).

The *P. aeruginosa* WlbA-pET28JT plasmid was used to transform Rosetta(DE3) *Escherichia coli* cells (Novagen). Cultures in LB medium were grown at 37 °C with shaking until the measured optical densities at 600 nm were approximately 0.8, at which time they were cooled for 15 min in an ice−water bath, induced with 1.0 mM IPTG, and allowed to express protein at 16 °C for 24 h after induction. WlbA was purified by standard procedures using Ni-NTA chromatography. Following purification, TEV protease was added to the protein at a 50:1 protein:TEV molar ratio. The digestion was allowed to proceed for 48 h at 4 °C and the sample subsequently dialyzed against 50 mM sodium phosphate (pH 8), 300 mM NaCl, and 20 mM imidazole. Cleaved enzyme was separated from the TEV protease and any undigested protein by Ni-NTA column chromatography. The cleaved protein was pooled and dialyzed against 10 mM Tris-HCl and 200 mM NaCl (pH 8.0) and then concentrated to 15.5 mg/mL using an extinction coefficient of 0.87 mg^{-1} mL cm^{-1}.

The *T. thermophilus* WlbA-pET28JT plasmid was used to transform Rosetta(DE3) *E. coli* cells (Novagen). Cultures in LB medium were grown at 37 °C with shaking until the measured optical densities at 600 nm were approximately 0.8, at which time the cells were induced with 1.0 mM IPTG and allowed to express protein at 37 °C for 7 h after induction. WlbA was purified by standard procedures using Ni-NTA chromatography. Following purification, the protein was dialyzed against 10 mM Tris-HCl and 200 mM NaCl (pH 8.0) and then concentrated to 14.5 mg/mL using an extinction coefficient of 0.82 mg^{-1} mL cm^{-1}.

Structural Analysis of WlbA. Crystallization conditions were initially surveyed via the hanging drop method of vapor diffusion and using a sparse matrix screen developed in the laboratory. Samples tested for crystallizations included the holoproteins from both organisms, and the proteins incubated with either 10 mM UDP-GlcNAcA or 20 mM α-ketoglutarate. Note that both enzymes contained bound NAD(H) upon purification, and that addition of NAD^+ to the purified protein had no effect on its activity.

For the *P. aeruginosa* enzyme, diffraction-quality crystals of the holoenzyme were obtained by mixing in a 1:1 ratio the protein and 16−22% monomethyl ether poly(ethylene glycol) 5000, 2% 1,2-ethanediol, and 100 mM PIPES (pH 6.5). The crystals belonged to space group $P2_12_12_1$ with one tetramer per asymmetric unit and the following unit cell dimensions: a = 73.2 Å, b = 129.6 Å,

Table 1: X-ray Data Collection Statistics[a]

	MAD phasing data (*T. thermophilus* enzyme)			*T. thermophilus* enzyme		*P. aeruginosa* enzyme
	peak	inflection	remote	NAD(H) α-ketoglutarate	NAD(H) UDP-GlcNAcA	NAD(H)
resolution limits	50.0−1.74 (1.77−1.74)	50−1.74 (1.77−1.74)	50.0−1.73 (1.76−1.73)	50.0−1.45 (1.48−1.45)	50.0−2.0 (2.03−2.0)	50.0−1.5 (1.53−1.5)
wavelength	0.97918	0.97953	0.97150	0.97921	0.97921	0.97921
no. of independent reflections	159786 (7515)	159970 (7521)	162213 (7289)	270515 (12317)	98909 (4534)	206888 (9562)
completeness (%)	99.5 (94.7)	99.5 (94.8)	99.0 (89.5)	96.8 (88.7)	91.9 (85.5)	93.5 (87.6)
redundancy	3.7 (2.8)	3.6 (2.8)	3.7 (2.7)	3.6 (2.3)	9.1 (6.8)	6.2 (3.4)
avg I/avg $\sigma(I)$	23.0 (2.5)	25.9 (2.9)	19.4 (1.8)	44.4 (4.2)	62.7 (17.8)	33.8 (2.3)
R_{sym}[b] (%)	7.0 (32.2)	7.4 (28.5)	6.5 (39.3)	6.1 (16.7)	7.3 (12.7)	6.3 (33.4)

[a]Statistics for the highest-resolution bin in parentheses. [b]$R_{sym} = (\sum|\Sigma I - \bar{I}|/\sum I) \times 100$.

and c = 145.2 Å. The stabilization solution required for cryo-cooling contained 26% monomethyl ether poly(ethylene glycol) 5000, 300 mM NaCl, 15% 1,2-ethanediol, and 100 mM PIPES (pH 6.5).

In the case of the *T. thermophilus* enzyme, diffraction-quality crystals were grown by mixing in a 1:1 ratio the protein incubated with 20 mM α-ketoglutarate and 15−20% monomethyl ether poly(ethylene glycol) 5000 and 100 mM MOPS (pH 7.0). The crystals belonged to space group $P2_1$ with a tetramer in the asymmetric unit and the following unit cell dimensions: a = 71.8 Å, b = 68.2 Å, c = 164.1 Å, and β = 94.7°. The stabilization solution required for cryo-cooling contained 24% monomethyl ether poly(ethylene glycol) 5000, 300 mM NaCl, 20 mM α-ketoglutarate, 20% 1,2-ethanediol, and 100 mM MOPS (pH 7.0). To obtain the ternary complex of the enzyme with its dinucleotide cofactor and UDP-GlcNAcA substrate, the crystals grown in the presence of α-ketoglutarate were first soaked for 2 days in a synthetic mother liquor [20% monomethyl ether poly(ethylene glycol) 5000, 200 mM NaCl, and 100 mM MOPS (pH 7.0)]. They were then soaked in a similar mother liquor containing 10 mM UDP-GlcNAcA and subsequently transferred for cryo-cooling to a solution containing 24% monomethyl ether poly(ethylene glycol) 5000, 300 mM NaCl, 10 mM UDP-GlcNAcA, 20% 1,2-ethanediol, and 100 mM MOPS (pH 7.0).

All X-ray data from flash-cooled crystals were collected at the Structural Biology Center Beamline 19-BM (Advanced Photon Source, Argonne National Laboratory, Argonne, IL). These data were processed and scaled with HKL-3000 (*13*). X-ray data collection statistics are listed in Table 1.

An initial structure of WlbA was determined using a selenomethionine-labeled version of the protein from *T. thermophilus*. The protein required for this was prepared via standard methods in the laboratory and purified and crystallized in the same manner as described above for the wild-type enzyme. Six of the eight possible selenium atoms were located with SOLVE (*14*), giving an overall figure of merit of 0.22 to 1.7 Å resolution. Solvent flattening and 2-fold averaging with RESOLVE (*15*) generated an interpretable electron density map with a final figure of merit of 0.66. A preliminary model was constructed using COOT (*16*). This model served as the search probe for the subsequent structural determination of the enzyme−α-ketoglutarate complex using molecular replacement as implemented in PHASER (*17*). Refinement of the model with REFMAC (*18*) led to a final overall R-factor of 19.5% for all measured X-ray data from 30 to 1.45 Å resolution. Relevant refinement statistics are given in Table 2. The structure of the ternary complex of the enzyme with NAD(H) and UDP-GlcNAcA was determined via Fourier difference techniques using the coordinates of the enzyme with bound α-ketoglutarate. Prior to refinement, all ligands and water molecules were removed. Due to some movements in the polypeptide chain backbone upon UDP-GlcNAcA binding, several stretches of the protein model had to be removed and rebuilt with iterative rounds of refinement with REFMAC and 4-fold electron density averaging with DM (*19*). The final overall R-factor was 19.2% for all measured X-ray data from 30 to 2.0 Å resolution.

The structure of the *P. aeruginosa* enzyme was determined via molecular replacement with PHASER using as a search model a single monomer of the *T. thermophilus* structure from which all ligands and solvent molecules had been removed. Iterative rounds of refinement with REFMAC, 4-fold molecular averaging with DM, and manual model building with COOT led to a final overall R-factor of 20.4% for all measured X-ray data from 30 to 1.5 Å resolution.

Determination of Kinetic Constants. The kinetic constants for WlbA were determined via a discontinuous HPLC assay using an ÄKTA HPLC system equipped with a 1 mL Resource-Q column. For the *T. thermophilus* enzyme, a total of 16 reactions were analyzed. Each 2 mL reaction mixture contained 50 μg/mL enzyme and 50 mM HEPPS (pH 8.5). The UDP-GlcNAcA concentrations were varied between 2.0, 0.3, 0.15, and 0.1 mM and the α-ketoglutarate concentrations were varied between 10.0, 3.3, 2.0, and 1.0 mM. Each reaction was conducted at 25 °C and initiated by the addition of the enzyme. For each reaction, 100 μL aliquots were taken at time zero and at 15−30 s intervals over a period of 6−10 min. The individual aliquots were immediately quenched by addition of 6 μL of 2 M HCl. Subsequently, 40 μL of CCl_4 was added to the samples, which were then vortexed and centrifuged for 1 min to remove denatured protein. An 80 μL sample of the aqueous phase was then 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 on solutions adjusted to pH 4.0 with HCl and using a gradient from 0 to 250 mM LiCl over 20 mL. The UDP-3-keto-GlcNAcA product eluted at 180 mM LiCl. The area of the HPLC peak corresponding to this product was determined and then correlated to concentration on the basis of a calibration curve created with standard samples that had been treated in the same manner as the reaction aliquots. A plot of concentration versus time was generated for each reaction, which allowed initial rates to be determined. All 16 data points were fitted to the equation $v = VAB/(K_aB + K_bA + AB)$ using

Table 2: Refinement Statistics

	T. thermophilus enzyme in a complex with NAD(H) and α-ketoglutarate	*T. thermophilus* enzyme in a complex with NAD(H) and UDP-GlcNAcA	*P. aeruginosa* enzyme in a complex with NAD(H)
resolution limits (Å)	30.0−1.45	30.0−2.0	30.0−1.5
R-factor[a] (overall) (%)/no. of reflections	19.5/270490	19.2/98901	20.4/206876
R-factor (working) (%)/no. of reflections	19.4/256862	18.9/93935	20.2/196484
R-factor (free) (%)/no. of reflections	22.6/13628	24.8/4966	24.4/10383
no. of protein atoms	9873[b]	9818[c]	9765[d]
no. of heteroatoms	1452[e]	786[f]	1194[g]
average *B* value (Å^2)			
protein atoms	24.0	35.4	28.4
ligands	18.8	51.1	13.9
solvent	33.6	36.1	32.0
weighted root-mean-square deviation from ideality			
bond lengths (Å)	0.014	0.012	0.013
bond angles (deg)	2.3	2.5	2.1
general planes (Å)	0.012	0.011	0.008

[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 Met 89, Arg 122, Ser 230, and Arg 242 in chain A, Met 89, Lys 114, and Arg 242 in chain B, Arg 242 in chain C, and Glu 53, His 185, Arg 242, and Ile 247 in chain D. [c] These include multiple conformations for Met 89, Lys 162, and Arg 242 in chain A and Arg 119 and Gln 145 in chain C. [d] These include multiple conformations for Cys 27 and Cys 78 in chain A, Cys 27 and Cys 78 in chain B, Cys 27, Cys 78, Arg 119, Arg 163, and His 286 in chain C, and Arg 19, Cys 27, Asn 35, Ser 68, and Cys 78 in chain D. [e] Heteroatoms include four NAD(H) molecules, four α-ketoglutarate molecules, three 1,2-ethanediol molecules, four chloride ions, and 1220 waters. [f] Heteroatoms include four NAD(H) molecules, four UDP-GlcNAcA ligands, four chloride ions, and 446 waters. [g] Heteroatoms include four NAD(H) molecules and 1018 waters.

Table 3: Kinetic Parameters

source of protein	V_{max} (μM/min)	K_m (mM) for α-ketoglutarate	K_m (mM) for UDP-GlcNAcA	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1}) for α-ketoglutarate	k_{cat}/K_m (M^{-1} s^{-1}) for UDP-GlcNAcA
T. thermophilus	20.6 ± 0.5	2.35 ± 0.12	0.43 ± 0.02	240 ± 10	1.0×10^5	5.6×10^5
P. aeruginosa	28.0 ± 0.8	1.35 ± 0.10	0.25 ± 0.01	5550 ± 70	4.1×10^6	2.2×10^7

SigmaPlot8, where K_a and K_b are the Michaelis constants for α-ketoglutarate and UDP-GlcNAcA, respectively, and V is the maximum velocity.

For the *P. aeruginosa* enzyme, each 2 mL reaction mixture contained 3 μg/mL enzyme and 50 mM HEPPS (pH 8.5). The UDP-GlcNAcA concentrations were varied between 2.0, 0.3, 0.15, and 0.1 mM and the α-ketoglutarate concentrations were varied between 10.0, 3.3, 2.0, and 0.5 mM. The reactions were analyzed in the same manner as described above.

RESULTS AND DISCUSSION

Functional Analysis of WlbA. A previous report on WlbA from *P. aeruginosa* suggested that it undergoes an NADH recycling mechanism using α-ketoglutarate as an oxidant (*9*). Given this, we hypothesized that catalysis by WlbA proceeds through a ping-pong reaction much like that exhibited by glucose-fructose oxidoreductase (*20*). To test this, a series of discontinuous HPLC assays were conducted as described in Materials and Methods. For our analyses, two different forms of WlbA were tested, one cloned from *T. thermophilus* and the other from *P. aeruginosa*. These enzymes contain 310 and 316 amino acid residues, respectively, per subunit, and they are 49% identical and 67% similar with respect to primary structure. Kinetic parameters were determined for both forms of the enzyme and are listed in Table 3. Double-reciprocal plots clearly revealed a series of parallel lines as presented in Figure 1. These results are suggestive of a ping-pong mechanism. Note that the enzyme from *P. aeruginosa* is more active at 25 °C than that from *T. thermophilus*. This is most likely a consequence of the fact that *T. thermophilus*, first isolated from a thermal vent, functions at an optimal growth temperature of 65 °C (*21*). In addition to the kinetic experiments, our analytical ultracentrifugation studies revealed the oligomerization state of WlbA to be tetrameric (unpublished data).

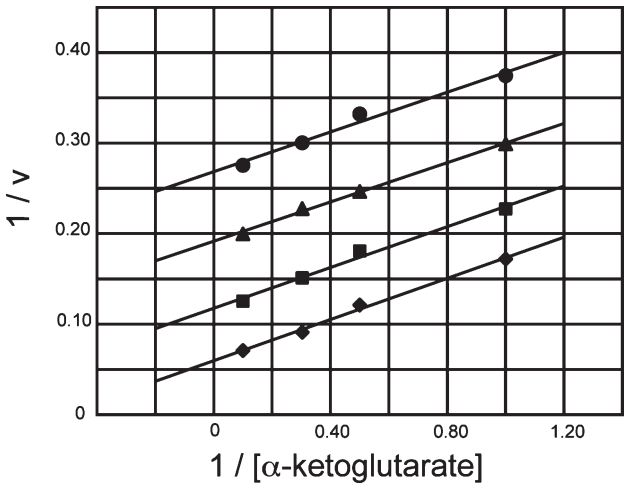


FIGURE 1: Double-reciprocal plots of initial rates of formation of UDP-3-keto-GlcNAcA using the enzyme from *T. thermophilus*. The concentration of α-ketoglutarate was varied at 10.0, 3.3, 2.0, and 1.0 mM at several fixed concentrations of UDP-GlcNAcA: 2.0 mM (◆), 0.3 mM (■), 0.15 mM (▲), and 0.1 mM (●). Initial velocities (v) are expressed in μM/min of UDP-3-keto-GlcNAcA.

Structure of WlbA with NAD(H). In all of the enzyme samples utilized in this investigation, the dinucleotide cofactor was always tightly bound to the purified protein. As a consequence,

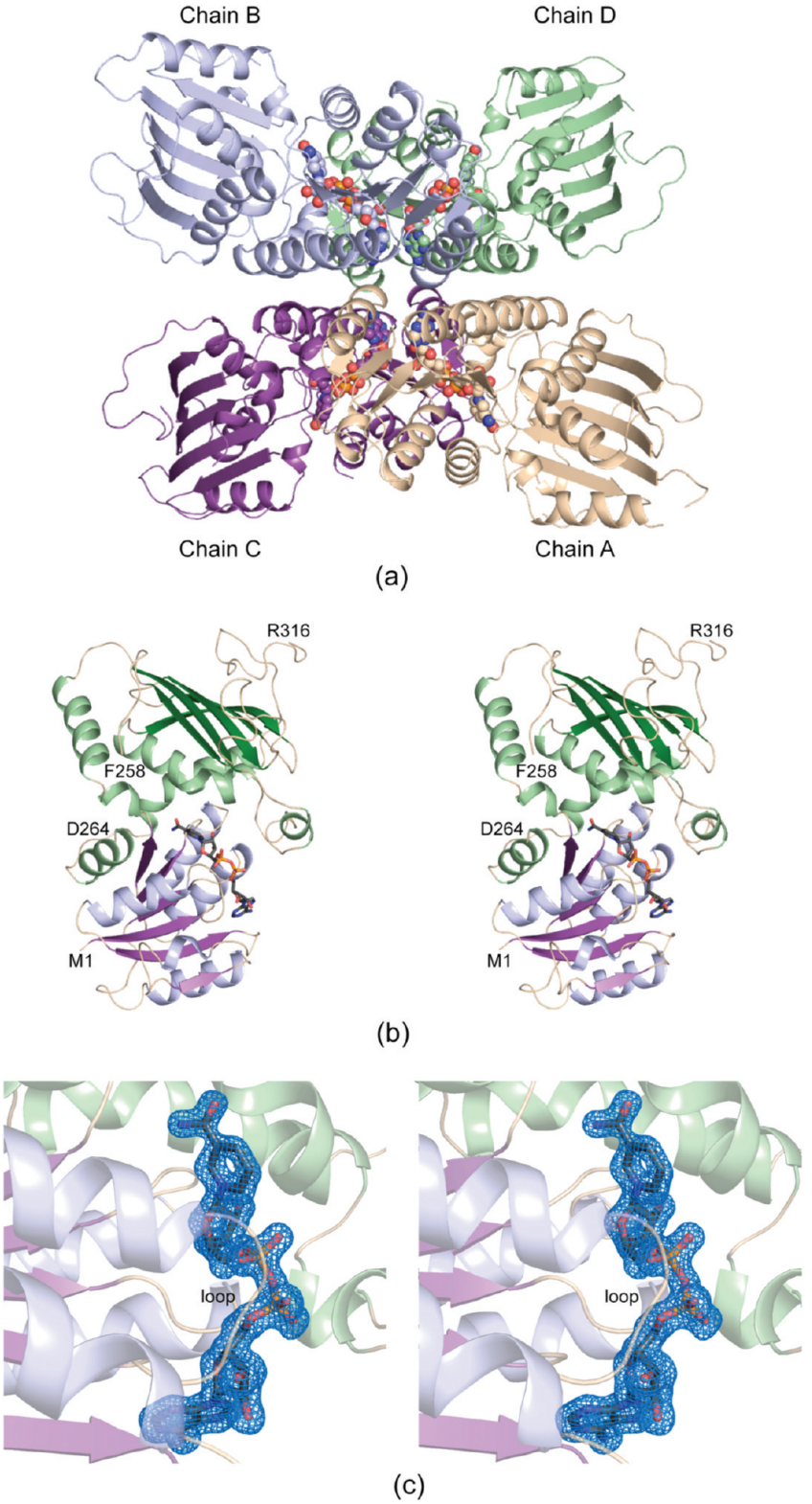


FIGURE 2: Structure of WlbA from *P. aeruginosa*. A ribbon representation of the complete tetramer is presented in panel a. The bound dinucleotides are displayed in sphere representations. A close-up stereoview of a ribbon representation for chain B is shown in panel b. The N-terminal domain is colored light blue and purple, whereas the C-terminal motif is highlighted in green. The bound NAD(H) is shown in stick representation. Electron density corresponding to the dinucleotide in chain B is displayed in panel c. The electron density map, contoured at 2.5σ, was calculated with coefficients of the form $F_o - F_c$, where F_o was the native structure factor amplitude and F_c was the calculated structure factor amplitude. The dinucleotide was not included in the coordinate file before the map was calculated. This figure and Figures 3−7 were prepared with PyMOL (*23*).

its actual oxidation state is not known, and thus, it will be referred to as NAD(H) to reflect this fact. For the structural analysis of the binary WlbA−NAD(H) complex, the enzyme from *P. aeruginosa* was crystallized and its structure determined and refined to 1.5 Å resolution. The asymmetric unit contained a tetramer, and the quality of the model was excellent, with 92.7, 7, and 0.3% of the ϕ and ψ angles lying within the core, allowed, and generously allowed regions of the Ramachandran plot, respectively (*22*).

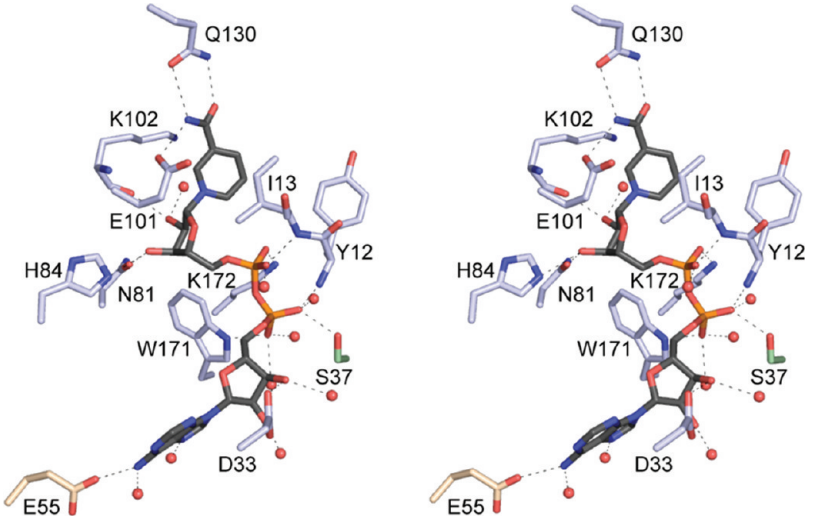


FIGURE 3: Close-up view of the NAD(H) binding pocket in *P. aeruginosa* WlbA. Those residues located within ~3.2 Å of the dinucleotide are shown. Residues belonging to chains A, B, and D are colored wheat, light blue, and green, respectively. Ordered water molecules are represented as red spheres. Possible hydrogen bonding interactions are represented by the black dashed lines.

Overall, the electron density for the polypeptide chains was very well ordered with the exception of a few breaks (Glu 147−Ser 149 and Phe 258−Asp 264 in chain A, Phe 258−Asp 264 in chain B, Pro 236−Pro 244 and Phe 258−Asp 264 in chain C, and Ala 145−Lys 152 and Met 256−Asp 264 in chain D). In each subunit, Pro 103, located in a loop abutting the nicotinamide ribose, adopted a *cis* conformation.

A ribbon representation of the WlbA tetramer is shown in Figure 2a. The molecule has overall dimensions of ~95 Å × 75 Å × 103 Å and, surprisingly, adopts an unusual quaternary structure with the dinucleotides situated closely together. For example, in chains A and C (or chains B and D), the C-2 hydroxyl groups of the adenine riboses lie within ~9 Å of each other. The surface area lost by each subunit upon formation of the tetramer is ~1800 Å^2. Shown in Figure 2b is a stereo representation of chain B. The molecule is distinctly bilobal with the N-terminal domain formed by Met 1−Ile 128 and the C-terminal domain delineated by Leu 128−Arg 316. The N-terminal motif, which harbors the binding site for NAD(H), contains a classical Rossmann fold with six parallel β-strands flanked on either side by three α-helices. One of these α-helices is contributed by the C-terminal domain as can be seen in Figure 2b. The C-terminal domain is dominated by a five-stranded mixed β-sheet covered on one side by three α-helices and on the other by random coil.

Electron density corresponding to the bound NAD(H) is presented in Figure 2c. The nicotinamide ring is planar, whereas the carboxamide group is twisted slightly out of the plane. As indicated in Figure 2c, there is a loop connecting the first β-strand to the first α-helix of the Rossmann fold that is formed by Gly 8−Ile 13. It lies against the pyrophosphoryl moiety of the cofactor and contains the characteristic signature sequence of a Rossmann fold enzyme. In *P. aeruginosa* Wlba, this sequence corresponds to Leu 6, Ile 7, Gly 8, Ala 9, Ala 10, and Gly 11. A close-up view of the NAD(H) binding site in chain B is displayed in Figure 3. Both riboses of the dinucleotide adopt the C_2'-endo pucker. The nicotinamide ring is anchored to the protein via the side chains of Glu 101 and Gln 130, and it adopts the *anti* conformation observed in A-side-specific dehydrogenases. The side chains of Asn 81 and His 84, the backbone carbonyl of Lys 102, and a water molecule lie within 3.2 Å of the nicotinamide ribose C-2 and C-3 hydroxyls. The pyrophosphoryl group of the cofactor is surrounded by the side chains of Trp 171, Lys 172, and Ser 37, the backbone amides of Tyr 12 and Ile 13, and three waters. For most NAD(H)-dependent enzymes, the binding site of the cofactor is contained within one subunit. Interestingly, Ser 37 is contributed by chain D. Four water molecules lie within hydrogen bonding distance of the adenine ring and its associated ribose. Additionally, the side chains of Asp 33 and Glu 55 serve to further anchor the adenosyl group into the active site region. Like Ser 37, Glu 55 is contributed by another subunit, in this case chain A. In total, 23 hydrogen bonds are formed between the cofactor and either the protein or solvent, and these most likely contribute to the tight binding of the dinucleotide to WlbA.

Structure of WlbA with NAD(H) and α-Ketoglutarate. Attempts to produce X-ray-quality crystals of the WlbA−NAD(H)−α-ketoglutarate complex using the enzyme cloned from *P. aeruginosa* were unsuccessful. Thus, to obtain such a complex, we cloned the enzyme from *T. thermophilus*, crystallized it, and determined its structure to a resolution of 1.45 Å. The resulting model was excellent, with 92.3 and 7.7% of the ϕ and ψ angles lying within the core and allowed regions of the Ramachandran plot, respectively. As in the case of the *P. aeruginosa* enzyme, the *T. thermophilus* protein contains one *cis* proline at position 102. Not unexpectedly, in light of the amino acid sequence identity, the structures of the *T. thermophilus* and *P. aeruginosa* enzymes are very similar, with a root-mean-square deviation (rmsd) of 1.3 Å for 279 structurally equivalent α-carbons (chain B). There are no major conformational changes that occur in the polypeptide chain backbone upon the binding of α-ketoglutarate.

The electron densities corresponding to the NAD(H) and α-ketoglutarate ligands are presented in Figure 4a. In the *T. thermophilus* enzyme, the loop connecting the first β-strand to the first α-helix of the Rossmann fold is formed by Leu 6−Gly 11 and has the signature sequence of Leu 6, Thr 7, Gly 8, Leu 9, Ala 10, and Gly 11. The dinucleotide is accommodated in the *T. thermophilus* active site in a manner similar to that described for the enzyme from *P. aeruginosa*. The carboxamide moiety attached to the nicotinamide group is also twisted out of the plane. A close-up view of the α-ketoglutarate binding site in chain B is depicted in Figure 4b. The side chain carboxylate group is bound in the active site by Tyr 156, Thr 158, Arg 160, Arg 245, and a water molecule. The α-carboxylate group interacts with the side

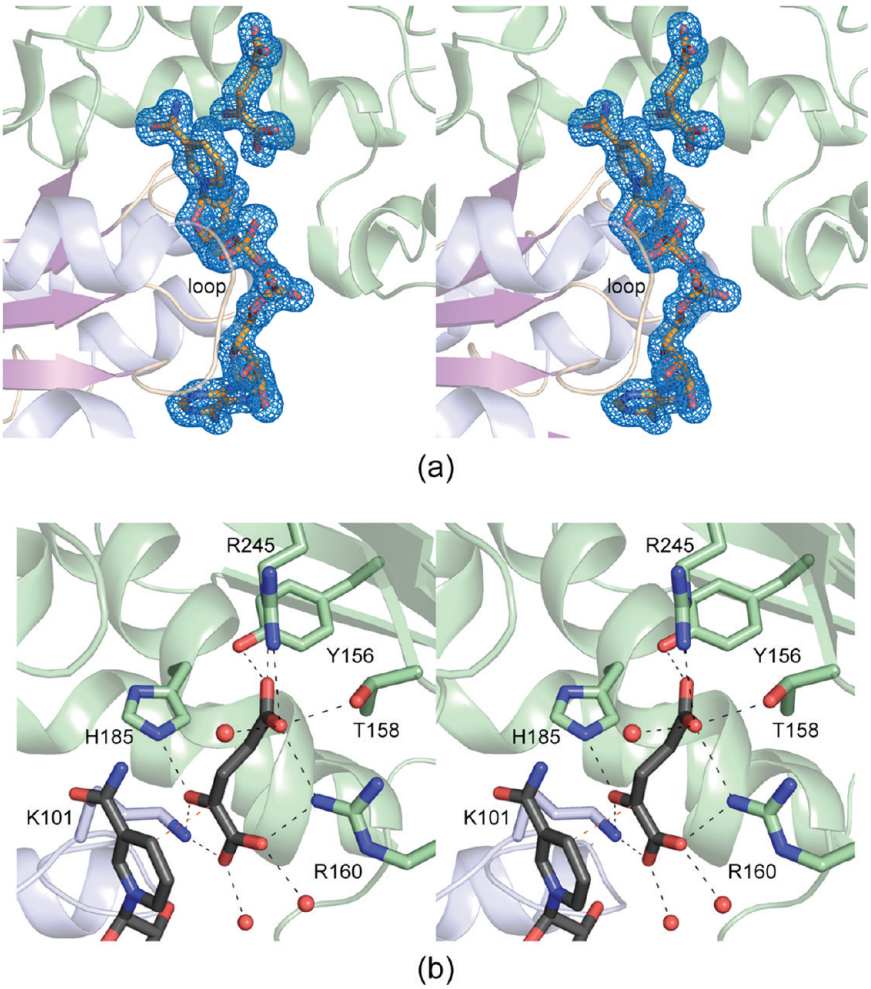


FIGURE 4: Structure of the WlbA−NAD(H)−α-ketoglutarate complex from *T. thermophilus*. Electron densities corresponding to the bound NAD(H) and α-ketoglutarate ligands in chain B are presented in panel a. The electron density map was calculated as described in the legend of Figure 2 and was contoured at 2.5σ. Those residues lying within 3.2 Å of the α-ketoglutarate are shown in panel b. Possible hydrogen bonding interactions are represented by the dashed lines, and ordered water molecules are depicted as red spheres. With the exception of Lys 101, all other residues involved in the binding of α-ketoglutarate are contributed by domain C, which is highlighted in green. The red dashed line indicates a distance of 3.1 Å between the C-2 atom of α-ketoglutarate and the nicotinamide C-4 atom.

chains of Lys 101 and Arg 160 and two waters. Finally, the α-keto group lies within 3.2 Å of Lys 101 and His 185. For α-ketoglutarate to be reduced to 2-hydroxyglutarate, a hydride must be transferred from NADH to its C-2 atom and a proton must be donated to its keto oxygen. As can be seen in Figure 4b, α-ketoglutarate and the dinucleotide cofactor adopt the proper geometry in the active site cleft for direct hydride transfer from the *re* face of the cofactor's nicotinamide moiety. The C-2 atom of α-ketoglutarate lies within 3.1 Å of the nicotinamide C-4 atom. With respect to which residue serves as the proton donor to the keto oxygen, either Lys 101 or His 185 could possibly fulfill this function. However, whereas Lys 101 is strictly conserved among members of this dehydrogenase family, His 185 is sometimes replaced with an asparagine or serine residue.

Structure of WlbA with NAD(H) and UDP-GlcNAcA. For the final structure in this investigation, crystals of the enzyme from *T. thermophilus* were transferred to a synthetic mother liquor containing its substrate, UDP-GlcNAcA. The overall structure was determined to 2.0 Å resolution with 91.2 and 8.8% of the ϕ and ψ angles lying within the core and allowed regions of the Ramachandran plot, respectively. Given that the electron density for the UDP-linked sugar was best defined in chain C, the following discussion refers to it exclusively. The models of WlbA in a complex with either α-ketoglutarate or UDP-GlcNAcA are very similar such that their polypeptide chains correspond with a root-mean-square deviation of 0.53 Å for 300 structurally equivalent α-carbons.

A stereoview of the electron densities corresponding to the bound substrate and NAD(H) is presented in Figure 5a. The electron density corresponding to the nicotinamide moiety of the dinucleotide is not as well-ordered as that observed in the WlbA−NAD(H) or WlbA−NADH(H)−α-ketoglutarate complexes. Strikingly, the UDP-linked sugar adopts a highly unusual curved conformation with the uracil ring situated ~4 Å from the hexose moiety. The hexose assumes the $^4C_1$ conformation.

As indicated in Figure 5b, the UDP-linked sugar is tightly embedded in the active site via extensive hydrogen bonding with protein side chains. Specifically, the uracil ring lies within hydrogen bonding distance of the side chain of Lys 169 and a water, whereas the ribose, which adopts the C_2'-endo pucker, hydrogen bonds to the side chain of Tyr 12 and a water. The pyrophosphoryl group of the substrate is anchored to the protein via the side chains of Arg 160, Thr 243, and Arg 245. Both Tyr 156 and Arg 245 are situated within 3.2 Å of the C-2′ *N*-acetyl group. The C-3′ hydroxyl group, which is ultimately oxidized in the reaction, hydrogen bonds to Lys 101 and His 185. The C-3′ atom

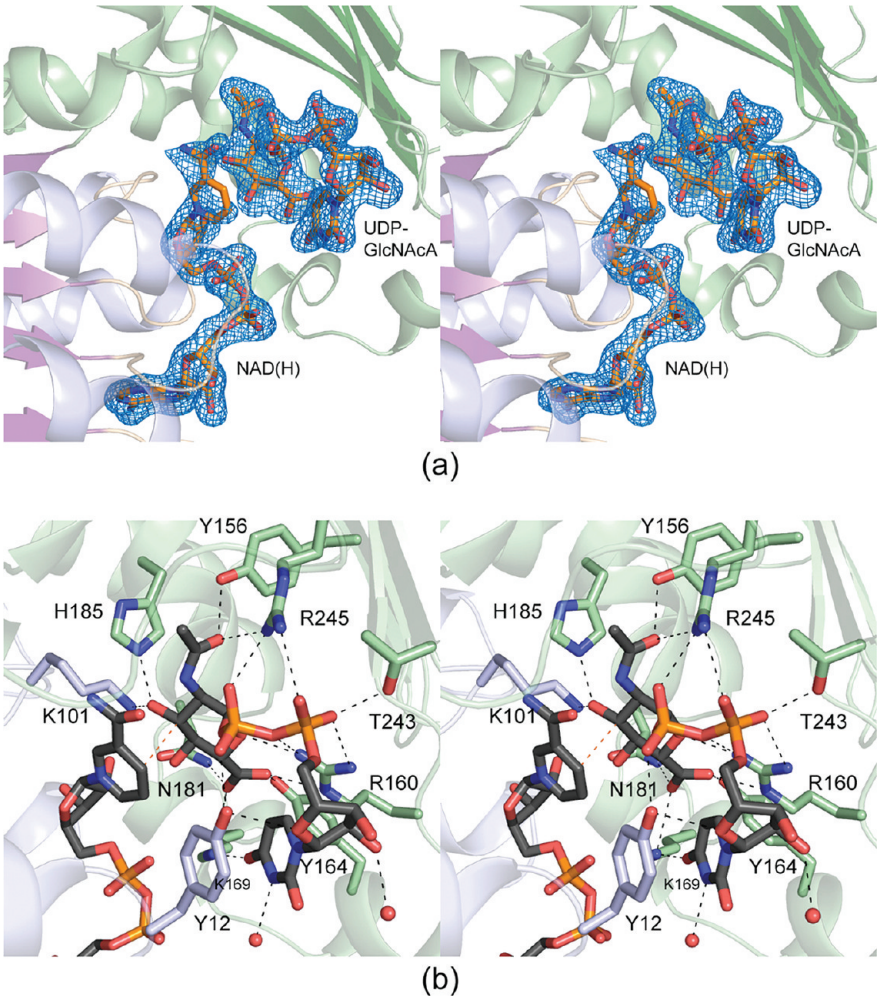


FIGURE 5: Structure of the WlbA−NAD(H)−UDP-GlcNAcA complex from *T. thermophilus*. Electron densities corresponding to the bound NAD(H) and UDP-GlcNAcA ligands in chain C are presented in panel a. The electron density map was calculated as described in the legend of Figure 2 and was contoured at 2.5σ. Those residues lying within 3.2 Å of UDP-GlcNAcA are shown in panel b. Possible hydrogen bonding interactions are represented by the dashed lines, and ordered water molecules are depicted as red spheres. With the exception of Tyr 12 and Lys 101, all other residues involved in the binding of the UDP-linked sugar are contributed by domain C, which is highlighted in green. The red dashed line indicates a distance of 3.0 Å between the C-3′ atom of the hexose and the nicotinamide C-4 atom.

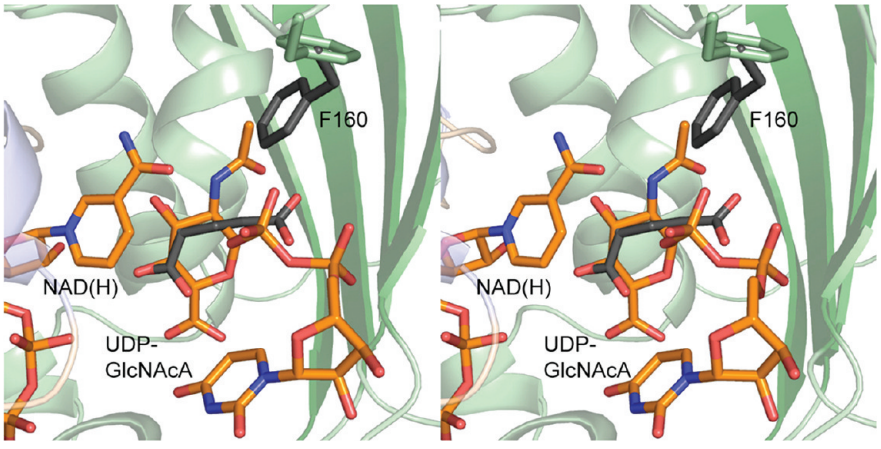


FIGURE 6: Superposition of the binding positions for α-ketoglutarate and UDP-GlcNAcA in the WlbA active site. The UDP-linked sugar and the α-ketoglutarate are depicted as gold and gray bonds, respectively. The position of Phe 160 when the active site is occupied by UDP-GlcNAcA is colored green. When α-ketoglutarate binds in the active site, Phe 160 shifts its position to that shown in gray bonds.

is located at 3.0 Å from C-4 of the nicotinamide ring and is in the correct orientation for transfer of its hydride to the dinucleotide. Finally, the side chains of Arg 160, Tyr 164, Lys 169, and Asn 181 serve to position the C-5′ carboxylate group into the active site, with Lys 169 serving as the bridge that links the uridine and hexose groups.

A comparison of the binding orientations for α-ketoglutarate and UDP-GlcNAcA is displayed in Figure 6. The α-keto and

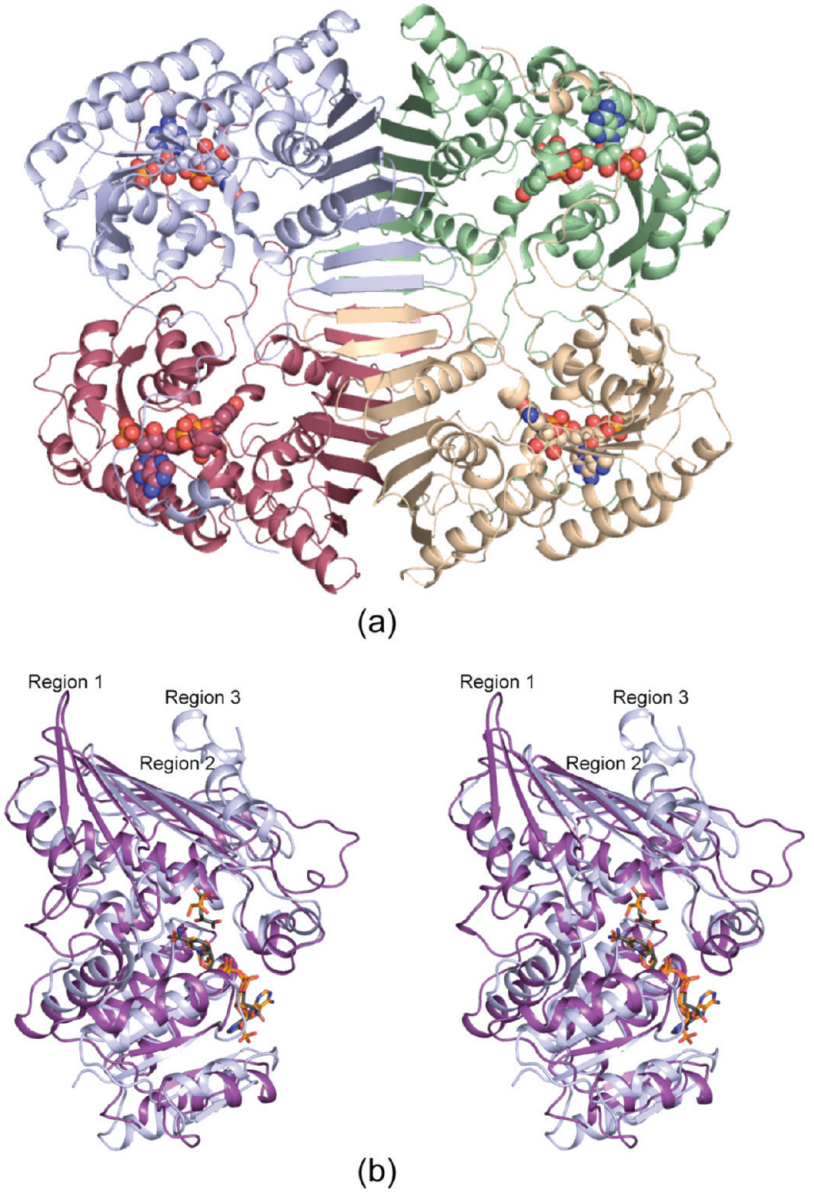


FIGURE 7: Comparison of the *T. thermus* WlbA and *Z. mobilis* GFOR structures. A ribbon representation of the GFOR tetramer is presented in panel a. Note the stunning difference between its quaternary structure compared to that of WlbA shown in Figure 2a. A superposition of a subunit of GFOR (purple) onto a subunit of WlbA (light blue) is presented in panel b. The positions of the bound ligands are displayed in stick representation. These include NAD(H) and α-ketoglutarate in WlbA (gray bonds) and NADP(H) and glycerol (gold bonds) in GFOR. There are three major conformational differences between these two enzymes that account for their widely differing quaternary structures as indicated by regions 1−3. X-ray coordinates for GFOR were obtained from Protein Data Bank entry 1H6B.

α-carboxylate groups of α-ketoglutarate occupy the same region in the active site as the C-3′, C-4′, and C-5′ atoms of UDP-GlcNAcA. Only one major change in the polypeptide chain occurs to accommodate these different substrates. Specifically, the region from Phe 256 to Leu 263 adjusts slightly depending upon the identity of the substrate occupying the active site. A significant shift occurs at Phe 260, which moves out of the active site when UDP-GlcNAcA binds. This movement is required to accommodate the bulky C-2′ *N*-acetyl group of the UDP-linked sugar.

The dinucleotide recycling system employed by WlbA, while intriguing, is not unprecedented. For example, glucose-fructose oxidoreductase (GFOR) from *Z. mobilis* catalyzes the formation of sorbitol from fructose and uses as its NADPH recycling system the conversion of glucose to gluconolactone (*20*). GFOR functions as a tetramer as depicted in Figure 7a. Clearly, the overall organization of the tetramer is completely different from that observed for WlbA (Figure 2a). The structures of the individual subunits for WlbA and GFOR are quite similar, however, as can be seen in Figure 7b. The models for WlbA and GFOR superimpose with an root-mean-square deviation of 1.9 Å for 257 structurally equivalent α-carbons. Whereas Lys 101 is conserved in GFOR, His 185 is replaced with a tyrosine residue. Importantly, Lys 101 is part of a conserved structural motif in these two enzymes, namely, Glu-Lys-Pro (*11*). In both enzymes, the prolines adopt *cis* conformations, and the glutamate residues are hydrogen bonded to the carboxamide groups of their respective cofactors.

Why are the quaternary structures so different between these two enzymes? It is a result of three major conformational differences exhibited by these proteins. One of these occurs near Phe 256 (WlbA), where in GFOR there is a 25-residue insertion that results in two additional β-strands in the C-terminal domain

(Figure 7b). Another difference arises because of a 12-residue insertion beginning at Pro 229 in WlbA that packs against the large mixed β-sheet of the C-terminal domain (Figure 7b). The final major difference arises at the C-terminus. In GFOR, the C-terminus forms an additional β-strand, whereas in WlbA, it adopts a random coil conformation and again packs against the mixed β-sheet. Taken together, these three differences result in the remarkably diverse quaternary structures exhibited by WlbA and GFOR.

In summary, the functional data presented in this investigation demonstrate that the NAD(H) recycling system displayed by WlbA is consistent with a double-displacement mechanism. Our structural data further show that the enzyme adopts a rather unique tetrameric quaternary structure with the four dinucleotides lying quite closely to one another. Contrary to what is typically observed in other NAD-dependent enzymes, the NAD(H) binding pocket in WlbA is shared among three subunits rather than contained within one. Most likely, the NAD(H) recycling mechanism is required because removal of the reduced dinucleotide would disrupt the native quaternary structure of the enzyme. We have also demonstrated in this investigation that the UDP-linked sugar substrate binds in a highly unusual conformation and is bound to the protein active site by an extensive array of hydrogen bonds. The C-3′ hydroxyl group that is oxidized to a keto moiety in the biosynthesis of UDP-ManNAc3NAcA is ideally situated in the active site for the transfer of a hydride to the nicotinamide C-4 atom of the dinucleotide. Either Lys 101 or His 185 likely serves as the catalytic base to remove the C-3′ hydroxyl proton. The accommodation of α-ketoglutarate versus UDP-GlcNAcA in WlbA is accomplished by simple flexibility in the protein region defined by Phe 256−Leu 263. The common amino acid residues utilized in binding either α-ketoglutarate or UDP-GlcNAcA in the WlbA active site include Lys 101, Tyr 156, Arg 160, His 185, and Arg 245. Experiments designed to more fully investigate the biological roles of these residues in WlbA are currently underway.

ACKNOWLEDGMENT

We gratefully acknowledge Professor W. W. Cleland for helpful discussions. 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.

REFERENCES

1. Raetz, C. R., and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu. Rev. Biochem. 71*, 635–700.
2. Liu, B., Knirel, Y. A., Feng, L., Perepelov, A. V., Senchenkova, S. N., Wang, Q., Reeves, P. R., and Wang, L. (2008) Structure and genetics of *Shigella O* antigens. *FEMS Microbiol. Rev. 32*, 627–653.
3. King, J. D., Kocincova, D., Westman, E. L., and Lam, J. S. (2009) Review: Lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa*. *Innate Immun. 15*, 261–312.
4. Wang, X., and Quinn, P. J. (2010) Lipopolysaccharide: Biosynthetic pathway and structure modification. *Prog. Lipid Res. 49*, 97–107.
5. Knirel, Y. A., Vinogradov, E. V., Shashkov, A. S., Dmitriev, B. A., and Kochetkov, N. K. (1982) 2,3-Diacetamido-2,3-dideoxy-D-mannuronic acid and its 2-imidazoline derivative: New acidic amino sugars from *Pseudomonas aeruginosa* O:3a,d lipopolysaccharide. *Carbohydr. Res. 104*, C4–C7.
6. Knirel, Y. A., Paramonov, N. A., Vinogradov, E. V., Shashkov, A. S., Dmitriev, B. A., Kochetkov, N. K., Kholodkova, E. V., and Stanislavsky, E. S. (1987) Somatic antigens of *Pseudomonas aeruginosa*. The structure of O-specific polysaccharide chains of lipopolysaccharides of *P. aeruginosa* O3 (Lanyi), O25 (Wokatsch) and Fisher immunotypes 3 and 7. *Eur. J. Biochem. 167*, 549–561.
7. 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.
8. Westman, E. L., McNally, D. J., Charchoglyan, A., Brewer, D., Field, R. A., and Lam, J. S. (2009) Characterization of WbpB, WbpE, and WbpD and reconstitution of a pathway for the biosynthesis of UDP-2,3-diacetamido-2,3-dideoxy-D-mannuronic acid in *Pseudomonas aeruginosa*. *J. Biol. Chem. 284*, 11854–11862.
9. Larkin, A., and Imperiali, B. (2009) Biosynthesis of UDP-GlcNAc-(3NAc)A by WbpB, WbpE, and WbpD: Enzymes in the Wbp pathway responsible for *O*-antigen assembly in *Pseudomonas aeruginosa* PAO1. *Biochemistry 48*, 5446–5455.
10. 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.
11. Kingston, R. L., Scopes, R. K., and Baker, E. N. (1996) The structure of glucose-fructose oxidoreductase from *Zymomonas mobilis*: An osmoprotective periplasmic enzyme containing non-dissociable NADP. *Structure 4*, 1413–1428.
12. 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.
13. 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.
14. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D55* (Part 4), 849–861.
15. Terwilliger, T. C. (2000) Maximum-likelihood density modification. *Acta Crystallogr. D56* (Part 8), 965–972.
16. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126–2132.
17. 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.
18. 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.
19. Cowtan, K. (1994) DM: An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography 31*, 34–38.
20. Hardman, M. J., and Scopes, R. K. (1988) The kinetics of glucose-fructose oxidoreductase from *Zymomonas mobilis*. *Eur. J. Biochem. 173*, 203–209.
21. Oshima, T., and Imahori, K. (1974) Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a Nonsporulating Thermophilic Bacterium from a Japanese Thermal Spa. *Int. J. Syst. Bacteriol. 24*, 102–112.
22. 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.
23. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA.