THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 23, Issue of June 6, pp. 20874–20881, 2003
© 2003 by The American Society for Biochemistry and Molecular Biology, Inc.
Printed in U.S.A.

High Resolution X-ray Structure of Tyvelose Epimerase from *Salmonella typhi**

Received for publication, February 24, 2003, and in revised form, March 14, 2003
Published, JBC Papers in Press, March 17, 2003, DOI 10.1074/jbc.M301948200

Nicole M. Koropatkin‡, Hung-wen Liu§, and Hazel M. Holden‡¶

From the ‡Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706 and the §Division of Medicinal Chemistry, College of Pharmacy, and Department of Chemistry and Biochemistry, University of Texas, Austin, Texas 78712

Tyvelose epimerase catalyzes the last step in the biosynthesis of tyvelose by converting CDP-D-paratose to CDP-D-tyvelose. This unusual 3,6-dideoxyhexose occurs in the *O*-antigens of some types of Gram-negative bacteria. Here we describe the cloning, protein purification, and high-resolution x-ray crystallographic analysis of tyvelose epimerase from *Salmonella typhi* complexed with CDP. The enzyme from *S. typhi* is a homotetramer with each subunit containing 339 amino acid residues and a tightly bound NAD^+ cofactor. The quaternary structure of the enzyme displays 222 symmetry and can be aptly described as a dimer of dimers. Each subunit folds into two distinct lobes: the N-terminal motif responsible for NAD^+ binding and the C-terminal region that harbors the binding site for CDP. The analysis described here demonstrates that tyvelose epimerase belongs to the short-chain dehydrogenase/reductase superfamily of enzymes. Indeed, its active site is reminiscent to that observed for UDP-galactose 4-epimerase, an enzyme that plays a key role in galactose metabolism. Unlike UDP-galactose 4-epimerase where the conversion of configuration occurs about C-4 of the UDP-glucose or UDP-galactose substrates, in the reaction catalyzed by tyvelose epimerase, the inversion of stereochemistry occurs at C-2. On the basis of the observed binding mode for CDP, it is possible to predict the manner in which the substrate, CDP-paratose, and the product, CDP-tyvelose, might be accommodated within the active site of tyvelose epimerase.

Carbohydrate epimerases are found in all living organisms and play crucial roles in numerous biochemical pathways as recently reviewed (1). The ability of an organism to invert the stereochemistry at a specific chiral center of a carbohydrate can determine its potential to utilize a particular carbon source, form a functional cell wall, and/or act pathogenically toward a host (2). Because of the diverse nature of both the chemical composition of carbohydrates and their various metabolic roles, several different epimerization mechanisms have evolved. These mechanisms can be broadly classified into two categories: those that function on activated carbon centers lying adjacent to carbonyl, carboxylic acid, or ester functional groups, and those that epimerize unactivated carbon centers.

Perhaps the best characterized member of the latter group, both kinetically and crystallographically, is UDP-galactose 4-epimerase (3, 4). This enzyme plays a key role in the Leloir pathway for galactose metabolism by interconverting UDP-galactose and UDP-glucose. The proposed catalytic mechanism of this epimerase involves several key features: 1) transfer of the hydride from C-4 of the sugar to NAD^+ to form NADH and abstraction of the 4′-hydroxyl hydrogen of the sugar by an enzymatic base to yield a 4′-ketopyranose intermediate; 2) rotation of this intermediate by ~180° about the phosphoryl-oxygen bond connecting the UDP and sugar moieties; and 3) transfer of the hydride from NADH back to the opposite face at C-4 of the keto-sugar intermediate and protonation of the resulting hydroxyl group. The enzymatic base in UDP-galactose 4-epimerase is thought to be a conserved tyrosine residue (4).

UDP-galactose 4-epimerase belongs to a subclass of NAD(P)-dependent enzymes referred to as the short-chain dehydrogenases/reductases (5). Members of this family contain a characteristic Tyr-*XXX*-Lys motif in which the conserved tyrosine is thought to play a key role in catalysis. Other proteins belonging to this subfamily include dTDP-D-glucose 4,6-dehydratase (6), GDP-mannose 4,6-dehydratase (7, 8), and ADP-L-glycero-D-*manno*-heptose 6-epimerase (9), among others. Alignments of the amino acid sequences for these enzymes with the Swiss-Prot data bank indicate that they share at least 20 and 40% sequence identity and similarity, respectively. On the basis of such amino acid sequence alignments another member of the superfamily has recently been identified, namely CDP-tyvelose 2-epimerase (10).

CDP-tyvelose 2-epimerase, hereafter referred to simply as tyvelose epimerase, is involved in the synthesis of tyvelose, a 3,6-dideoxyhexose that occurs in the *O*-antigens of some types of Gram-negative bacteria (11). Tyvelose is produced via a complex biochemical pathway that employs CDP-D-glucose as the starting ligand. The last step in the pathway, the conversion of CDP-D-paratose to CDP-D-tyvelose, is catalyzed by tyvelose epimerase as outlined in Scheme 1. Interestingly, tyvelose is found in the *O*-antigens of some pathogens such as *Salmonella typhi* and *Yersinia pseudotuberculosis* IVA. In that tyvelose is the major immunological determinant in these bacteria, screening for the tyvelose epimerase gene via PCR amplification is being considered for earlier identification and treatment of *S. typhi* infections (12).

Here we describe the cloning, purification, and high-resolution x-ray crystallographic analysis of tyvelose epimerase from *S. typhi*. Crystals amenable for a high-resolution x-ray investigation of the enzyme were grown in the presence of CDP. The enzyme is a homotetramer displaying 222 symmetry. Each subunit contains 339 amino acid residues and one tightly

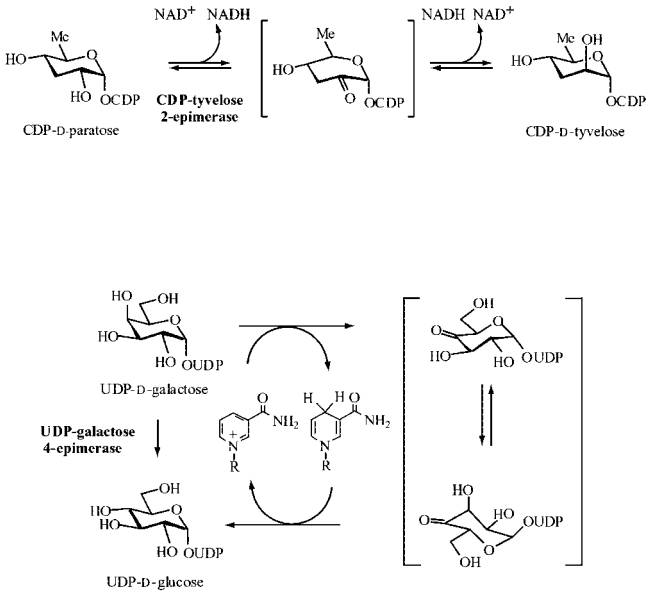
* This work was supported in part by National Institutes of Health Grants DK47814 (to H. M. H.) and GM35906 (to H.-w. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1ORR) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

¶ To whom correspondence should be addressed. Tel.: 608-262-4988; Fax: 608-262-1319; E-mail: Hazel_Holden@biochem.wisc.edu.

This paper is available on line at http://www.jbc.org

Downloaded from http://www.jbc.org/ at University of Wisconsin-Madison on September 16, 2020



SCHEME 1

bound NAD^+ moiety. The subunit is distinctly bilobal with an N-terminal region responsible for dinucleotide binding, a C-terminal motif important for CDP positioning, and an active site wedged between these two lobes. Side chain functional groups directly involved in NAD^+ binding include Asp-32, Asp-58, Ser-123, and Lys-169 whereas Ser-194, Asn-126, Lys-127, Gln-235, Arg-237, Arg-299, and Asp-302 function in anchoring the CDP moiety into the active site. The x-ray analysis described here defines both the quaternary structure of tyvelose epimerase and its active site geometry, and confirms that the enzyme belongs to the short-chain dehydrogenase/reductase superfamily.

EXPERIMENTAL PROCEDURES

Cloning of the Tyvelose Epimerase Gene—The sequence of the tyvelose epimerase gene, *tyv*, from *S. typhi* has been reported and allowed for the design of primers for gene amplification from genomic DNA. Genomic DNA samples from *S. typhi* clinical isolates CDC numbers 87-2059 and CDC 88-2009 were generous gifts from Dr. Stanley Maloy at the University of Illinois, Urbana-Champaign. The tyvelose epimerase gene was PCR amplified from genomic DNA such that the forward primer 5′-CATGCCATGGCCAAGCTTTTAATTACCGGTGGA-3′ and reverse primer 5′-CCGCTCGAGTATAGAACTAGTCCATCATACATTTC-3′ added *Nco*I and *Xho*I cloning sites, respectively. To preserve the *Nco*I recognition sequence of CCATGG, the codon GCC (encoding alanine) was inserted before the second residue of the gene. The gene was PCR amplified with Platinum *Pfx* DNA polymerase (Invitrogen) according to the manufacturer's instructions and standard cycling conditions. The PCR product was purified with the QIAquick PCR purification kit (Qiagen Inc.), followed by digestion with both *Nco*I and *Xho*I at 37 °C overnight. The gene was separated from digestion by-products on a 1.0% agarose gel, excised from the gel, and purified with QIAquick gel purification kit (Qiagen Inc.). The purified tyvelose epimerase gene was then ligated into the expression vector pET-28a (Novagen) that was previously cut with the same restriction enzymes. *Escherichia coli* DH5α cells were transformed with the ligation mixture and then plated onto LB media supplemented with 30 μg/ml kanamycin. Individual colonies were selected, cultured overnight, and plasmid DNA was extracted with the QIAprep spin miniprep kit (Qiagen Inc.). Plasmids were tested for insertion of the *tyv* gene by digesting with *Nco*I and *Xho*I. Positive clones were sequenced with the ABI Prism™ Big Dye Primer Cycle sequencing kit (Applied Biosystems, Inc.) to confirm that no mutations were introduced during PCR amplification.

Protein Expression—For protein expression, *E. coli* Rosetta-(DE3)pLysS (Novagen) cells were transformed with the pET28a-*tyv* plasmid and plated onto LB media supplemented with 30 μg/ml kanamycin. After ~16 h, the plates were scraped and the cells were resuspended in LB media to use for the inoculation of 6× 2-liter baffled flasks containing 500 ml of LB plus 30 μg/ml kanamycin and 30 μg/ml chloramphenicol. The cells were grown at 37 °C with aeration to an A_{600} of ~0.4, at which time they were transferred to a 30 °C shaker for 30 min before inducing with 1 mM isopropyl-1-thio-β-D-galactopyranoside. The cells were allowed to grow for an additional 6 h at 30 °C before harvesting by centrifugation at 6000 × *g* for 8 min. The cell paste was frozen in liquid nitrogen and stored at −80 °C.

Expression of the Selenomethionine-labeled Protein—*E. coli* Rosetta-(DE3)pLysS (Novagen) cells were transformed with the pET28a-*tyv* plasmid and plated onto LB media supplemented with kanamycin. Approximately 24 h later, several large colonies were selected to inoculate 500 ml of M9 minimal media supplemented with 30 μg/ml kanamycin and 30 μg/ml chloramphenicol for growth overnight at 37 °C. Subsequently 10 ml of the overnight culture was used to inoculate 3× 2-liter baffled flasks containing 500 ml of M9 minimal media with 5 μg/ml thiamine, 30 μg/ml kanamycin, and 30 μg/ml chloramphenicol. Cultures were grown at 37 °C to an A_{600} of ~0.2. At that point, the temperature of the incubator was adjusted to 30 °C for the remainder of the growth. Cultures were grown to an A_{600} of ~0.75, before each flask was supplemented with 50 mg each of L-lysine, L-threonine, and L-phenylalanine, and 25 mg each of L-leucine, L-isoleucine, L-valine, and L-selenomethionine (13). After 20 additional minutes of growth, the cells were induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside and allowed to grow for 6 h. Cultures were harvested by centrifugation at 6000 × *g* for 8 min, and the cell paste was frozen in liquid nitrogen for storage at −80 °C.

Protein Purification—All protein purification steps were carried out on ice or at 4 °C. The cell paste was thawed in 100 ml of cold Ni-NTA[1] lysis buffer containing 50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole at pH 8.0. Cells were lysed on ice by four cycles of sonication (30 s) separated by 3 min of cooling. Cellular debris was removed by centrifugation at 4 °C for 25 min at 20,000 × *g*. The clarified lysate was loaded onto a 10-ml Ni-NTA-agarose (Qiagen Inc.) column pre-equilibrated with Ni-NTA lysis buffer. After loading, the column was washed with about 60 ml of Ni-NTA wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0), followed by gradient elution of the protein from 20 to 300 mM imidazole in Ni-NTA lysis buffer. Protein-containing fractions were pooled based on SDS-PAGE, and dialyzed against 20 mM HEPPS (pH 8.0) with 200 mM NaCl. The dialyzed protein was concentrated to ~23 mg/ml, based on an extinction coefficient of ~1.3 ml/(mg cm) as calculated by the program Protean (DNAstar, Inc., Madison, WI).

Molecular Weight Determination—Analytical ultracentrifugation was performed by the University of Wisconsin, Madison Biochemistry Instrumentation Facility. Three samples of the purified native tyvelose epimerase at 0.25, 0.5, and 1.0 mg/ml in 10 mM HEPPS (pH 8.0) and 200 mM NaCl were provided for analysis. The experiments were performed with a Beckman Optima XL-A centrifuge and an An 60-Ti rotor at 20 °C using 12-mm double sector charcoal-filled Epson centerpieces. Rotor speeds of 3,000, 7,600, 11,000, and 13,500 rpm were utilized while monitoring the radial position of the sample by absorption of light at 280 nm. The molecular weight from a global fit of 9 data sets containing 1246 points was 154,000 indicating that the quaternary structure of tyvelose epimerase is tetrameric.

Crystallization of Native and Selenomethionine-labeled Tyvelose Epimerase—Prior to crystallization, the protein was diluted to 13 mg/ml with 20 mM HEPPS (pH 8.0), and 200 mM NaCl and CDP were added to a final concentration of 5 mM. Potential crystallization conditions were examined with a sparse matrix screen composed of 144 conditions at both room temperature and 4 °C via the hanging drop method of vapor diffusion. Single crystals were observed at room temperature at pH 5–6 with 10% poly(ethylene glycol) 8000. Refinement of the crystallization conditions led to large single crystals grown at room temperature with precipitant solutions of 4–6% poly(ethylene glycol) 8000, 250 mM tetramethylammonium chloride, and 100 mM succinate (pH 5.5). Crystals achieved typical dimensions of ~0.5 × 0.3 × 0.15 mm in 2 to 3 weeks and belonged to the space group $P2_1$ with unit cell dimensions of a = 47.9 Å, b = 167.3 Å, c = 89.4 Å, and β = 105.5°. The solvent content of the crystals was ~50%, with one homotetramer in the asymmetric unit.

High Resolution X-ray Data Collection—Both the native and selenomethionine-substituted protein crystals were flash-frozen in the same manner. Briefly, crystals were harvested from the hanging drop experiments and soaked for several hours to several days in a synthetic mother liquor composed of 10–12% poly(ethylene glycol) 8000, 200 mM

[1] The abbreviations used are: Ni-NTA, nickel-nitrilotriacetic acid; HEPPS, *N*-(2-hydroxyethyl)piperazine-*N*′-3-propanesulfonic acid.

Downloaded from http://www.jbc.org/ at University of Wisconsin-Madison on September 16, 2020

TABLE I
X-ray data collection statistics

	Wavelength	Resolution	No. independent reflections	Completeness	Redundancy	Avg I/Avg $\sigma(I)$	R_{sym}[a]
	Å			%			%
Peak	0.979371	50.0—2.30	59,298	99.0	3.78	36.3	6.5
		2.38—2.30[b]	5,918	98.9	3.76	24.6	9.2
Inflection	0.97960	50.0—2.30	59,576	99.1	3.81	35.8	6.7
		2.38—2.30	5,936	99.1	3.69	18.5	10.9
Remote	0.96411	50.0—2.30	59,535	99.4	3.89	33.7	5.8
		2.38—2.30	5,944	99.4	3.71	12.5	12.0
Native	0.97940	50.0—1.46	225,494	95.2	3.80	40.8	6.1
		1.51—1.46	21,022	89.2	2.80	3.9	21.2

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

NaCl, 300 mM tetramethylammonium chloride, 5 mM CDP, and 100 mM succinate (pH 5.5). These crystals were then transferred to 4% ethylene glycol, 14% poly(ethylene glycol) 8000, 325 mM NaCl, 300 mM tetramethylammonium chloride, 5 mM CDP, and 100 mM succinate (pH 5.5) for 30 s, followed by a final transfer to 15% ethylene glycol, 20% poly(ethylene glycol) 8000, 400 mM NaCl, 300 mM tetramethylammonium chloride, 5 mM CDP, and 100 mM succinate (pH 5.5). Subsequently, the crystals were flash-cooled to −150 °C in a stream of nitrogen gas. Frozen crystals were stored under liquid nitrogen until synchrotron beam time became available. X-ray data sets from both the native protein and the selenomethionine-substituted enzyme crystals were collected on a 3 × 3 tiled "SBC3" CCD detector at the Structural Biology Center 19-BM beamline (Advanced Photon Source, Argonne National Laboratory, Argonne, IL). The x-ray data were processed with HKL2000 and scaled with SCALEPACK (14). Relevant x-ray data collection statistics are presented in Table I.

X-ray Structural Analyses—The structure of tyvelose epimerase was solved via multiwavelength anomalous dispersion phasing with x-ray data collected from the selenomethionine-substituted protein crystals. The software package SOLVE was utilized to determine and refine the positions of the selenium atoms (15). Visual inspection of the selenium sites with the graphics program TURBO revealed four clusters of atoms related by non-crystallographic 222 symmetry (16). Subsequent non-crystallographic symmetry averaging and solvent flattening with the software package DM resulted in a readily interpretable electron density map calculated to 2.6-Å resolution (17). The averaged map was used to build an initial model of one subunit of tyvelose epimerase, after which this subunit was rotated back into the unit cell to create the full tetramer. This tetramer served as a search model for molecular replacement with the program AMORE against the native x-ray data (18). Alternate cycles of least-squares refinement with the software package TNT (19) and manual model building reduced the R-factor to 17.9% for all measured x-ray data from 30- to 1.5-Å resolution. Relevant least-squares refinement statistics are summarized in Table II. A Ramachandran plot for all nonglycinyl main chain ϕ,ψ values is displayed in Fig. 1*a*. In each monomer there are two residues that lie outside of the allowed regions: Ser-163 and Thr-269. The electron density for these residues is unambiguous. Ser-163 is located at the beginning of the sixth α-helix in the N-terminal region whereas Thr-269 lies within a Type III turn delineated by Thr-269 to Asn-272. Neither one of these residues is directly involved in ligand binding. The electron density corresponding to the bound ligands, CDP and NAD^+ in Subunit I is shown in Fig. 1*b*.

TABLE II
Least squares refinement statistics

Resolution limits (Å)	30.0–1.50
R-factor (overall) %[a]/No. reflections	17.9/22,5493
R-factor (working) %/No. reflections	17.8/20,3000
R-factor (free) %/No. reflections	22.9/22,493
No. protein atoms[b]	10,711
No. Hetero-atoms	1355
Average B values (Å^2)	
Protein atoms	25.8
NAD^+	17.4
CDP	31.1
Solvents	37.4
Weighted root mean square deviations from ideality	
Bond lengths (Å)	0.012
Bond angles (deg)	2.23
Trigonal planes (Å)	0.005
General planes (Å)	0.010
Torsional angles[c] (deg)	16.2

[a] R-factor = $(\Sigma|F_o - F_c|/\Sigma|F_o|) \times 100$, where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude.
[b] This value includes multiple conformations for His-44, Ser-76, Val-84, Asn-126, Thr-145, Gln-158, Ile-182, Asn-186, Ser-247, Thr-314, Asn-315 in Subunit I; Ser-19, Val-271, Thr-294, Ser-337 in Subunit II; Lys-3, Thr-125, Gln-158, Thr-251, Leu-276, Thr-314 in Subunit III; and Ser-19, Lys-62, Gln-158, Thr-314, Lys-322 in Subunit IV.
[c] The torsional angles were not restrained during the refinement.

RESULTS AND DISCUSSION

Quaternary Structure of Tyvelose Epimerase—On the basis of ultracentrifugation experiments and analysis of the crystalline packing arrangement, tyvelose epimerase is a homotetramer as shown in Fig. 2*a*. The enzyme has overall dimensions of ~60 × 100 × 110 Å and displays 222 symmetry. It can be envisioned as a dimer of dimers with the A/B and C/D pairs mimicking the quaternary structural interactions observed in UDP-galactose 4-epimerase (20). In both UDP-galactose 4-epimerase and the A/B or C/D dimers of tyvelose epimerase, there are two parallel α-helices contributed by each monomer that provide the structural framework for the subunit:subunit interface. In tyvelose epimerase, these α-helices are formed by Pro-93 to Gln-113 and Pro-164 to Phe-183. The buried surface area in this subunit: subunit interface is ~2470 Å^2, as calculated according to the method of Lee and Richards (21) with a probe sphere of 1.4 Å. There are two additional regions per subunit (Asn-33 to Thr-40 and Glu-53 to Arg-67) that participate in monomer:monomer interactions along the A/C and B/D pairs. The polypeptide chain lying between Glu-53 to Arg-67 contains the third β-strand and the third α-helix of the Rossmann fold. The interfaces between the A/C and B/D pairs are not as extensive with buried surface areas of ~ 1330 Å^2. In the A/B and C/D subunit pairs, the adenine rings of the dinucleotides are separated by ~27 Å whereas in the A/C and B/D pairs, these rings are positioned at ~11 Å.

Tertiary Structure of the Individual Subunit—The crystals employed in this investigation contained a complete tetramer in the asymmetric unit. The α-carbon traces for all of the monomers are virtually identical, however, such that they superimpose with root mean square deviations of 0.4 Å or less. For the sake of simplicity the following discussion will only refer to Subunit I of the coordinate file. As depicted in Fig. 2*b*, the overall molecular architecture of the tyvelose epimerase subunit can be envisioned as two lobes: the N-terminal region formed by Ala-2 to Gln-200 and Ala-242 to Gly-268 and the C-terminal motif delineated by Phe-201 to His-241 and Thr-269 to Ile-339. The N-terminal motif is dominated by a seven stranded parallel β-sheet defined by Ala-2 to Thr-7, Asp-27 to Asp-32, Glu-53 to Gly-57, Ser-76 to Leu-80, Asn-119 to Ser-123, Asn-186 to Tyr-196, and Ala-263 to Ile-266. There are two

Downloaded from http://www.jbc.org/ at University of Wisconsin-Madison on September 16, 2020

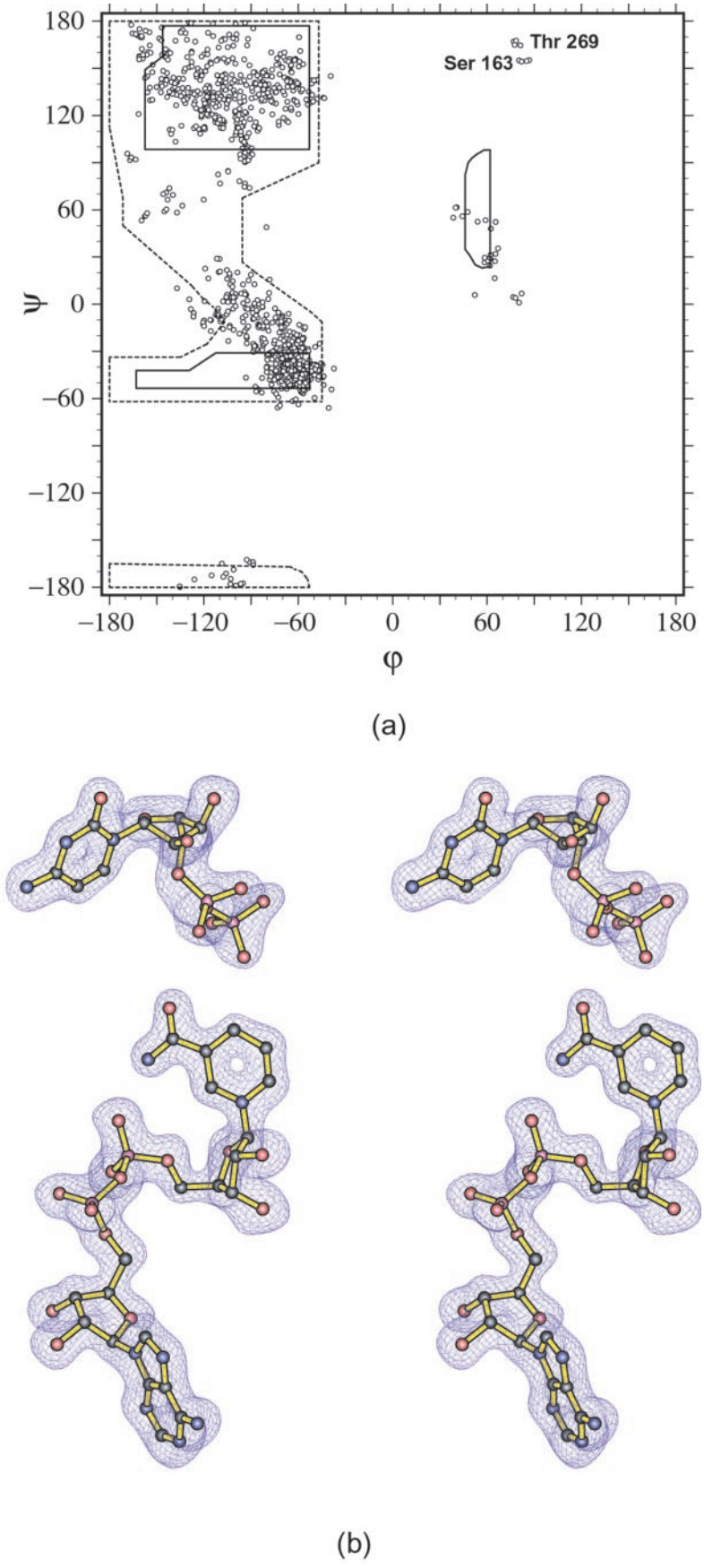


FIG. 1. **Quality of the x-ray model.** A plot of the ϕ,ψ angles for all non-glycinyl residues is shown in *a*. Those regions in the Ramachandran plot that are fully or partially allowed are indicated by the *solid* and *dashed lines*, respectively. Electron density corresponding to the bound CDP and NAD^+ in Subunit I is displayed in *b*. The map 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 from the model lacking the coordinates for the ligand. The map was contoured at 3 σ.

additional anti-parallel β-strands formed by Lys-136 to Glu-139 and Tyr-144 to Cys-146 and connected by a Type I turn (Thr-140 to Arg-143). These secondary structural elements of the N-terminal region are connected by a total of seven α-helices (Phe-12 to Ala-21, Thr-40 to Leu-49, Lys-62 to Tyr-72, Met-86 to Asp-91, Pro-93 to Gln-113, Pro-164 to Phe-183, and

Downloaded from http://www.jbc.org/ at University of Wisconsin-Madison on September 16, 2020

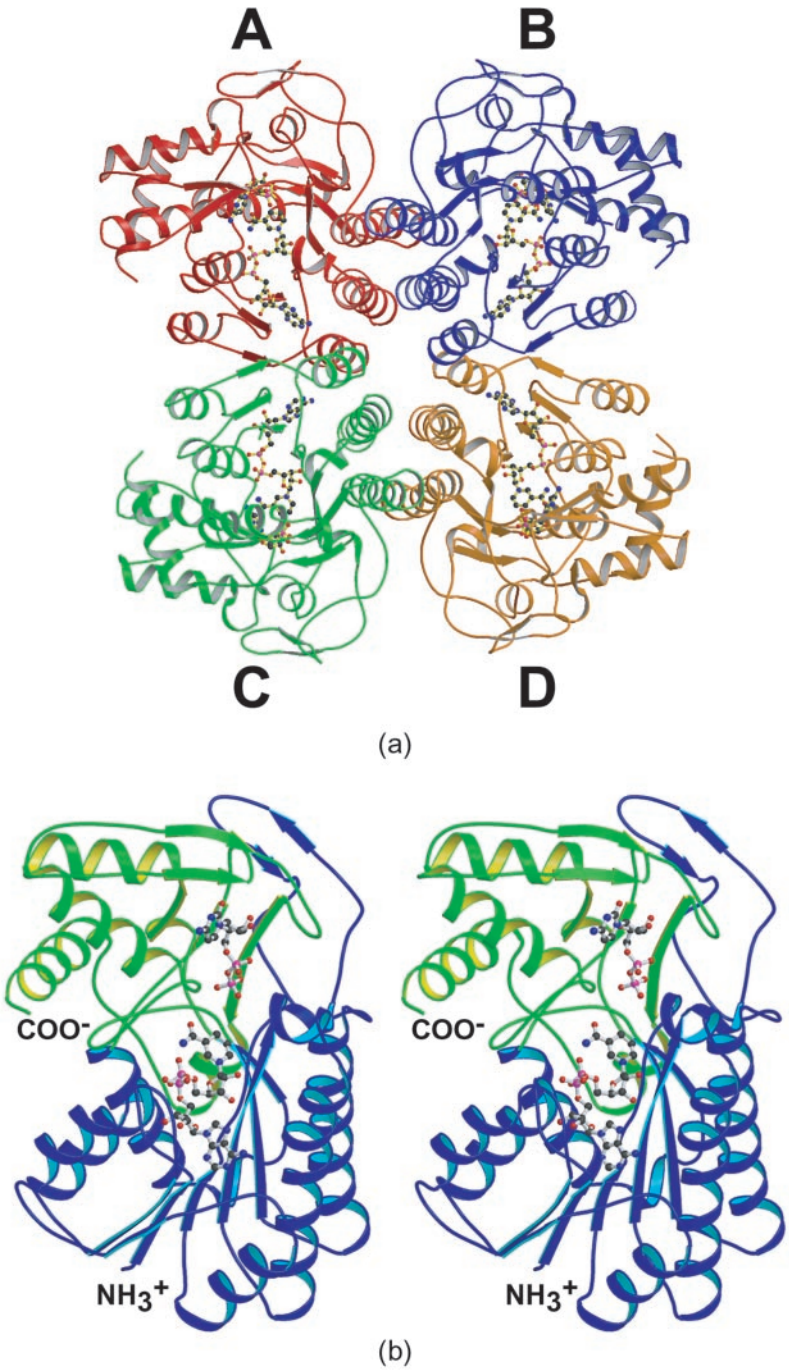


FIG. 2. **Ribbon representation of tyvelose epimerase.** The tetrameric structure of the enzyme is shown in *a*. The A/B and C/D pairs of dimers are similar to the dimers observed in both the human and bacterial forms of UDP-galactose 4-epimerase. Bound ligands are depicted in ball-and-stick representations. A stereo view of one subunit of the tetramer is shown in *b*. The molecular architecture of the subunit can be envisioned as two lobes, as indicated in *blue* and *green*. The active site is wedged between these two lobes.

Ala-242 to Ala-254) and numerous reverse turns. The C-terminal motif is less complicated with three strands of mixed β-sheet (Val-236 to Leu-240, Ser-273 to Ser-275, and Arg-304 to Phe-306), two strands of parallel β-sheet (Phe-227 to Gly-231 and Thr-294 to Leu-296), and four major α-helices (Trp-208 to Glu-218, Leu-276 to Tyr-286, Lys-311 to Ala-316, and Ala-325 to Ile-339). The overall dimensions of the subunit are ~50 × 50 × 50 Å.

Binding Sites for NAD$^+$ and CDP—Shown in Fig. 3 is a close-up view of the active site for the tyvelose epimerase subunit. There are 15 well ordered water molecules located within 3.2 Å of the CDP or NAD$^+$ moieties. The cytosine ring of CDP is hydrogen bonded to the protein via the carbonyl oxygen of Tyr-204 and the peptidic NH of Ser-230. An ordered water molecule is located at 2.9 Å from the amino group of the pyrimidine ring. A strikingly aromatic patch of residues, composed of Tyr-204, Trp-208, Trp-211, and Phe-212, surrounds one side of the cytosine ring. The CDP ribose adopts the C_2'-*endo* conformation and its 3-hydroxyl group is anchored to the protein via the side chains of Asp-302 and Gln-235. Both Arg-237 and Arg-299, through their side chain guanidinium groups, are involved in electrostatic interactions with the α-phosphoryl oxygens. Additionally, N^ϵ of Arg-237 lies within 3 Å of one of the β-phosphoryl oxygens. Asn-126 and Ser-194 form hydrogen bonding interactions with the other two β-phosphoryl oxygens. The β-phosphoryl group of CDP is situated at ~5 Å from the nicotinamide ring of NAD$^+$.

The nicotinamide ring of the dinucleotide adopts the *syn*-conformation as typically observed for *B*-side specific dehydrogenases. Both riboses of the dinucleotide adopt C_2'-*endo* conformations. The adenine ring of NAD$^+$ forms hydrogen bonding interactions with the carboxylate group of Asp-58 and the back-

Downloaded from http://www.jbc.org/ at University of Wisconsin-Madison on September 16, 2020

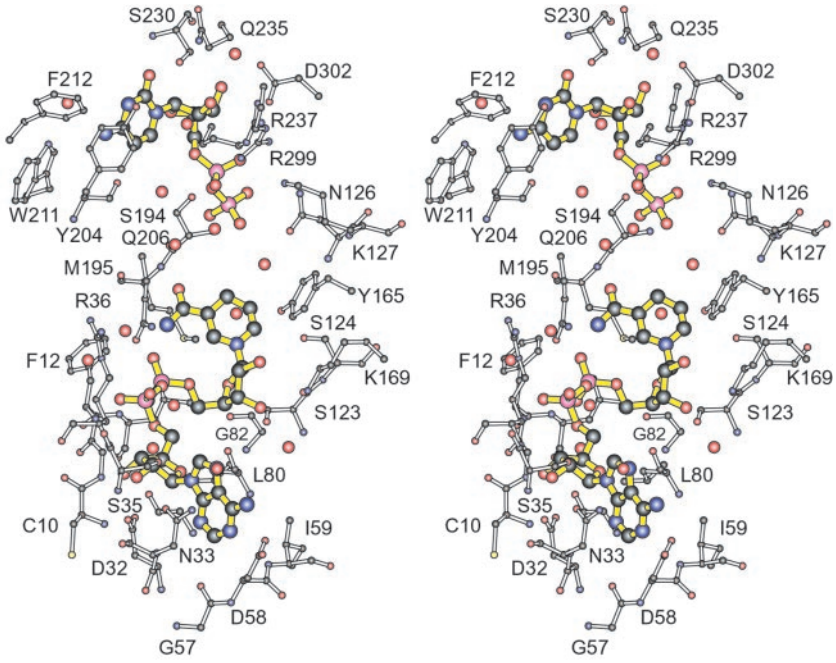


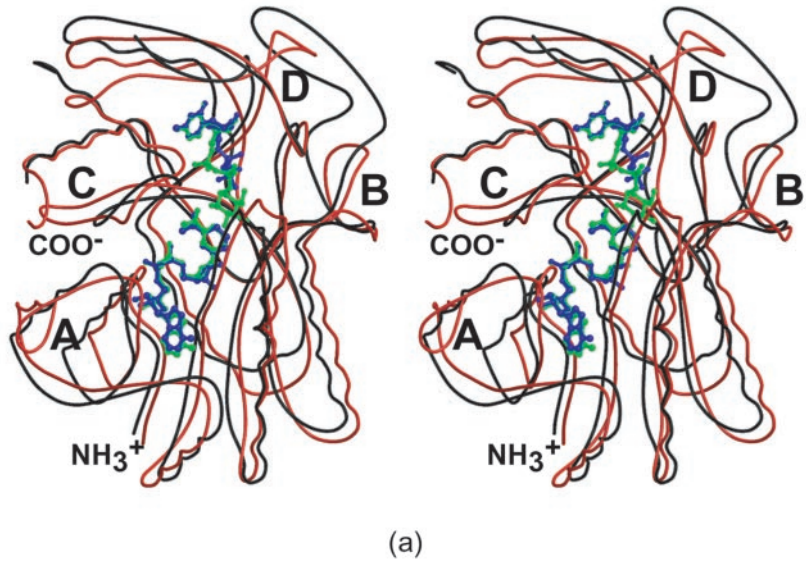
FIG. 3. **Close-up view of the tyvelose epimerase active site.** Those amino acid residues that are located within ~3.2 Å of the NAD^+ and CDP ligands are shown. The ligands are *highlighted* in *yellow* bonds. Ordered water molecules are indicated by the *red spheres*. For the sake of clarity, Val-84 and Trp-208 were omitted from the figure.

bone peptidic NH group of Ile-59. Asp-32 bridges the 2- and 3-hydroxyl groups of the adenine ribose. The ribose 2-hydroxyl group also lies within hydrogen bonding distance to the peptidic NH group of Arg-36. There are four water molecules located within 3.2 Å of the phosphoryl oxygens of the NAD^+ ligand. Additionally, the peptidic NH groups of Phe-12 and Leu-13 form hydrogen bonds with OP2 and OP5 of the NAD^+, respectively. These residues reside in the first α-helix of the Rossmann fold. The 3-hydroxyl group of the nicotinamide ribose lies with 3.2 Å of the backbone carbonyl oxygen of Leu-80, a water molecule, and N^{ζ} of Lys-169. This interaction between Lys-169 and the dinucleotide is conserved among the members of the short-chain dehydrogenase/reductase family. The 2-hydroxyl group of the nicotinamide ribose is positioned near a water molecule that, in turn, lies within 3 Å of the nicotinamide ring nitrogen. In the short-chain dehydrogenase/reductase family, it has been postulated that Lys-169 activates Tyr-165 that is the ultimate base in the reaction mechanism. The distance between Lys-169 and Tyr-165 in tyvelose epimerase is 4.9 Å, a value that is comparable with that observed in other members of the superfamily. Interestingly, the distance between the nicotinamide ring nitrogen, which carries the positive charge in the oxidized form of the cofactor, and O^{η} of Tyr-165 is 4.3 Å. It is possible that the chemical role of Lys-169 is in proper dinucleotide positioning and not in activation of the conserved tyrosine residue. Regarding the carboxamide moiety of the nicotinamide ring, there is an intramolecular hydrogen bond between its nitrogen and OP4 of the phosphate backbone (2.9 Å). The carbonyl oxygen of the carboxamide group hydrogen bonds with a water molecule and the backbone peptidic NH group of Met-195.

Comparison of Tyvelose Epimerase with UDP-galactose 4-Epimerase—Tyvelose epimerase is highly homologous to both UDP-galactose 4-epimerase from *E. coli* (27% identity and 43% similarity), and human UDP-galactose 4-epimerase (23% identity and 41% similarity). In contrast to tyvelose epimerase, however, both the human and bacterial forms of UDP-galactose 4-epimerase are homodimers. Shown in Fig. 4*a* is a superposition of the polypeptide chain backbones for the human UDP-galactose 4-epimerase and the tyvelose epimerase subunits. There are only four major differences in conformations between these two enzymes as indicated by the letters *A–D* in Fig. 4*a*. For example, there is a 7-residue deletion in tyvelose epimerase, relative to UDP-galactose 4-epimerase, starting at Arg-36 and extending to His-56. This particular region is involved in the A/C and B/D interfaces in tyvelose epimerase and thus may be one factor for the differing quaternary states between these two proteins. There is a 16-residue insertion in tyvelose epimerase, defined by Leu-132 to Tyr-153, which forms two antiparallel β-strands that are connected by a Type I turn (Thr-140 to Arg-143). Interestingly, this region in not involved in subunit:subunit contacts nor located near the NAD^+ binding site. The functional significance of this insertion in tyvelose epimerase is unknown. The third region of difference occurs as a 6-residue deletion in tyvelose epimerase relative to UDP-galactose 4-epimerase from Gly-197 to Arg-199. This area is not located within the immediate vicinity of the NAD^+ or CDP ligands. Finally, there is a 6-residue deletion in tyvelose epimerase from Gly-231 to Val-236, which corresponds to the region defined by Gly-227 to Val-238 in UDP galactose 4-epimerase. In both of these enzymes, this area of polypeptide chain is involved in binding the nucleotide portion of the substrate, either CDP or UDP, in tyvelose epimerase or UDP-galactose 4-epimerase, respectively. Indeed, while the positions of the NAD^+ ligands in both enzymes are nearly identical, the locations of the CDP *versus* UDP moieties are somewhat different as can be seen in Fig. 4*a*. The α- and β-phosphorus atoms of the CDP *versus* UDP moieties are shifted by ~2 Å. This shift may be a function of the need for the substrate, CDP-paratose, to position differently in the active site of tyvelose epimerase as compared with UDP-glucose (or UDP-galactose) in the active site of UDP-galactose 4-epimerase. Recall that in the reaction mechanism of UDP-galactose 4-epimerase, the inversion of stereochemistry occurs about C-4 of the hexose rather than C-2 as in tyvelose epimerase. Also, paratose is a 3,6-dideoxysugar. Strikingly, in the model of the human UDP-galactose 4-epimerase, the 2′-hydroxyl group of UDP-glucose is hydrogen bonded to Asn-207 whereas the 6′-hydroxyl group lies within 2.6 Å of $O^{\delta 1}$ of Asn-187. In tyvelose epimerase, these residues have been replaced by Gly-207 and Ser-194, respectively.

Previous detailed mechanistic investigations of tyvelose epimerase from *Y. pseudotuberculosis* IVA have suggested that its catalytic mechanism may be reminiscent to that proposed for UDP-galactose 4-epimerase (22). The reaction catalyzed by

Downloaded from http://www.jbc.org/ at University of Wisconsin-Madison on September 16, 2020



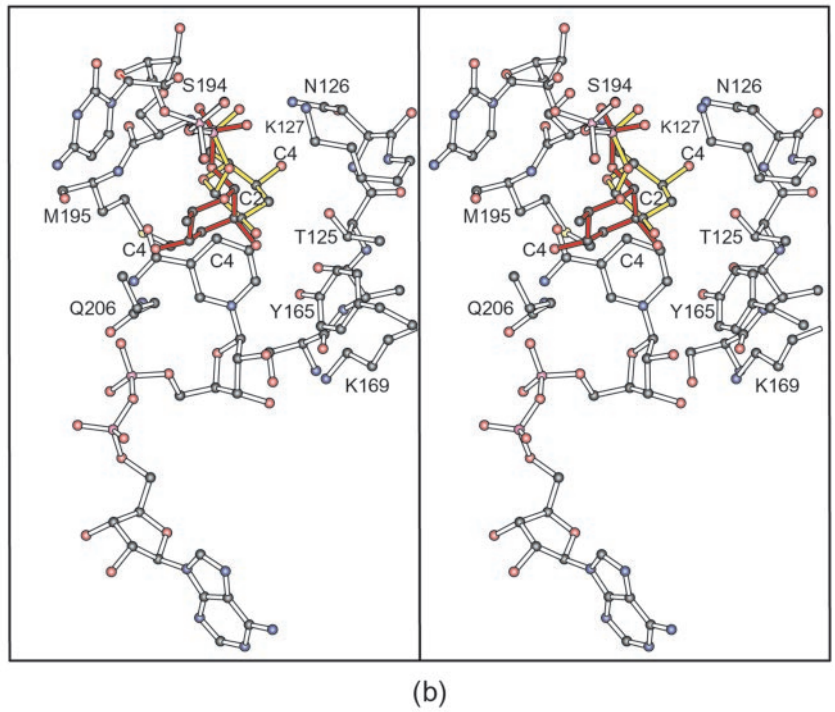


FIG. 4. **Comparison of tyvelose epimerase with UDP-galactose 4-epimerase.** Shown in *a* is a superposition of the α-carbons for these two enzymes. The bound ligands for tyvelose epimerase (NAD^+ and CDP) and UDP-galactose 4-epimerase (NADH and UDP-glucose) are displayed in *blue* and *green*, respectively. On the basis of the observed binding for CDP to tyvelose epimerase, models of CDP-paratose and CDP-tyvelose were built into the active site as shown in *b*. The CDP-paratose moiety is highlighted in *yellow-filled* bonds, whereas the CDP-tyvelose ligand is depicted in *red-filled* bonds.

UDP-galactose epimerase is an epimerization about C-4 of UDP-glucose (or UDP-galactose) as indicated in Scheme 1. Extensive structural investigations in this laboratory have concluded that in the human UDP-galactose 4-epimerase, Tyr-157 functions to initiate the reaction by abstracting the 4′-hydroxyl hydrogen from UDP-glucose (or UDP-galactose) with the concomitant transfer of the hydride at C-4 to the nicotinamide ring of NAD^+. Additionally, Ser-132 in the human epimerase is thought to facilitate proton transfer from the sugar 4′-hydroxyl group to O^{η} of Tyr-157 possibly via a low barrier hydrogen bond (4). Tyr-157 in human UDP-galactose 4-epimerase is conserved in tyvelose epimerase as Tyr-165 whereas Ser-132 in the human enzyme has been replaced with Thr-125. Both Tyr-165 and Thr-125 occupy similar positions within the active site of tyvelose epimerase as observed for Tyr-157 and Ser-132 in the UDP-galactose 4-epimerase active site region. Assuming that the reaction mechanisms are similar for these two epimerases as suggested by the mechanistic data, it is possible to speculate on the manner in which CDP-paratose or CDP-tyvelose might bind in the active site of the enzyme. Previous studies on the

Downloaded from http://www.jbc.org/ at University of Wisconsin-Madison on September 16, 2020

E. coli UDP-galactose 4-epimerase have demonstrated that the binding of the UDP moiety in the protein·NADH·UDP complex is virtually identical to that observed in the protein·NADH·UDP-glucose complex (20, 23). In light of this, a model of CDP-paratose was positioned into the tyvelose epimerase active site as indicated in Fig. 4*b*. This model was built by following the position of the CDP moiety as experimentally determined in this investigation and rotating the sugar portion to orient its C-2 near C-4 of the nicotinamide ring. In this model, C-2 of paratose is located at ~3.7 Å of C-4 of NAD^+ and its 2′-hydroxyl group is situated within ~3.6 Å of O^{η} of Tyr-165. In the model presented in Fig. 4*b*, the 4′-hydroxyl group of CDP-paratose lies within hydrogen bonding distance to Thr-125 and possibly to Asn-126. For the model depicted in Fig. 4*b*, Asn-126 had to be rotated in the active site to avoid overlap with the paratose 4′-hydroxyl group. Indeed, in support of this idea of rotation, Asn-126 demonstrates multiple conformations in the high resolution structure of the protein·NAD^+·CDP complex determined here.

For the model of CDP-tyvelose presented in Fig. 4*b*, a series of rotations was made about the dihedral angles starting at the β-phosphorus atom. Accordingly, this model places C-2 of the tyvelose moiety within 3.4 Å of C-4 of the nicotinamide ring and its hydroxyl group within 2.8 Å of O^{η} of Tyr-165. The tyvelose 4′-hydroxyl group is now located within hydrogen bonding distance to Gln-206. In the models presented in Fig. 4*b* for the two sugar ligands, the positions of their β-phosphorus atoms differ by ~1.0 Å. Support for this type of rotation in the active site is demonstrated by the high resolution structural analysis of the *E. coli* UDP-galactose 4-epimerase complexed with either UDP-glucose or UDP-galactose (24). In that study, significant changes were observed between the two bound substrates beginning at the β-phosphorus atoms where their positions differed relative to one another by 0.51 Å.

The putative binding mode depicted in Fig. 4*b* can be tested by preparing and crystallizing abortive complexes of tyvelose epimerase with NADH and either CDP-paratose or CDP-tyvelose. These studies are presently underway. Regardless of the outcome of these investigations, it can be concluded that rotation of a 2′-ketopyranose intermediate in the active site of tyvelose epimerase must be different from that proposed for UDP-galactose 4-epimerase because of the difference in its regioselectivity. In UDP-galactose 4-epimerase, the putative 4′-ketopyranose intermediate displays a pseudo-C_2 axis, which is not present in a 2-ketose intermediate. As such, a more complicated set of rotations about the phosphoryl backbone of the CDP-2′-ketopyranose intermediate must occur to present the opposite side of the oxidized sugar back to the reduced NADH for subsequent hydride transfer.

Acknowledgments—We thank Professors W. W. Cleland and Frank M. Raushel for carefully reading this manuscript and Dr. James B. Thoden for helpful advice throughout the course of this investigation.

REFERENCES

1. Allard, S. T., Giraud, M. F., and Naismith, J. H. (2001) *Cell. Mol. Life Sci.* **58,** 1650–1665
2. He, X. M., and Liu, H.-w. (2002) *Annu. Rev. Biochem.* **71,** 701–754
3. Frey, P. A. (1996) *FASEB J.* **10,** 461–470
4. Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L., and Holden, H. M. (2000) *Biochemistry* **39,** 5691–5701
5. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) *Protein Sci.* **5,** 2149–2161
6. Allard, S. T., Giraud, M. F., Whitfield, C., Graninger, M., Messner, P., and Naismith, J. H. (2001) *J. Mol. Biol.* **307,** 283–295
7. Somoza, J. R., Menon, S., Schmidt, H., Joseph-McCarthy, D., Dessen, A., Stahl, M. L., Somers, W S., and Sullivan, F. X. (2000) *Structure* **8,** 123–135
8. Mulichak, A. M., Bonin, C. P., Reiter, W.-D., and Garavito, R. M. (2002) *Biochemistry* **41,** 15578–15589
9. Deacon, A. M., Ni, Y. S., Coleman, W. G., Jr., and Ealick, S. E. (2000) *Structure* **8,** 453–462
10. Hallis, T. M., and Liu, H.-w. (1999) *Acc. Chem. Res.* **32,** 579–588
11. Madigan, M. T., Martinko, J. M., and Parker, J. (1996) in *Biology of Microorganisms*, Prentice Hall, Upper Saddle River, NJ
12. Hirose, K., Itoh, K., Nakajima, H., Kurazono, T., Yamaguchi, M., Moriya, K., Ezaki, T., Kawamura, Y., Tamura, K., and Watanabe, H. (2002) *J. Clin. Microbiol.* **40,** 633–636
13. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) *J. Mol. Biol.* **229,** 105–124
14. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
15. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. Sect. D* **55**, 849–861
16. Roussel, A., Fontecilla-Camps, J. C., and Cambillau, C. (1990) *Acta Crystallogr. Sect. A* **4**, C66–C67
17. Cowtan, K., and Main, P. (1998) *Acta Crystallogr. Sect. D* **54**, 487–493
18. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50**, 157–163
19. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. Sect. A* **43,** 489–501
20. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) *Biochemistry* **35,** 2557–2566
21. Lee, B., and Richards, F. M. (1971) *J. Mol. Biol.* **55,** 379–400
22. Hallis, T. M., Zhao, Z., and Liu, H.-W. (2000) *J. Am. Chem. Soc.* **122,** 10493–10503
23. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) *Biochemistry* **35,** 5137–5144
24. Thoden, J. B., and Holden, H. M. (1998) *Biochemistry* **37,** 11469–11477

Downloaded from http://www.jbc.org/ at University of Wisconsin-Madison on September 16, 2020

High Resolution X-ray Structure of Tyvelose Epimerase from *Salmonella typhi*

Nicole M. Koropatkin, Hung-wen Liu and Hazel M. Holden

J. Biol. Chem. 2003, 278:20874-20881.
doi: 10.1074/jbc.M301948200 originally published online March 17, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M301948200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 23 references, 1 of which can be accessed free at
http://www.jbc.org/content/278/23/20874.full.html#ref-list-1

Downloaded from http://www.jbc.org/ at University of Wisconsin-Madison on September 16, 2020