Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of CDP-D-glucose 4,6-dehydratase from *Salmonella typhi* complexed with CDP-D-xylose

Nicole M. Koropatkin and Hazel M. Holden*

Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706-1544, USA

Correspondence e-mail: hazel_holden@biochem.wisc.edu

Received 22 November 2004
Accepted 20 December 2004

PDB Reference: CDP-D-glucose 4,6-dehydratase, 1wvg, r1wvgsf.

Tyvelose is a unique 3,6-dideoxyhexose found in the O antigens of some pathogenic species of *Yersinia* and *Salmonella*. It is produced *via* a complex biochemical pathway that employs CDP-D-glucose as the starting ligand. CDP-D-glucose 4,6-dehydratase catalyzes the first irreversible step in the synthesis of this 3,6-dideoxysugar by converting CDP-D-glucose to CDP-4-keto-6-deoxyglucose *via* an NAD^+-dependent intramolecular oxidation–reduction reaction. Here, the cloning, protein purification and X-ray crystallographic analysis of CDP-D-glucose 4,6-dehydratase from *Salmonella typhi* complexed with the substrate analog CDP-D-xylose are described. Each subunit of the tetrameric enzyme folds into two domains. The N-terminal region contains a Rossmann fold and provides the platform for NAD(H) binding. The C-terminal motif is primarily composed of α-helices and houses the binding pocket for the CDP portion of the CDP-D-xylose ligand. The xylose moiety extends into the active-site cleft that is located between the two domains. Key residues involved in anchoring the sugar group to the protein include Ser134, Tyr159, Asn197 and Arg208. Strikingly, Ser134 O^{γ} and Tyr159 O^{η} sit within 2.9 Å of the 4′-hydroxyl group of xylose. Additionally, the side chains of Asp135 and Lys136 are located at 3.5 and 3.2 Å, respectively, from C-5 of xylose. In the structurally related dTDP-D-glucose 4,6-dehydratase, the Asp/Lys pair is replaced with an Asp/Glu couple. On the basis of this investigation, it can be speculated that Tyr159 serves as the catalytic base to abstract the 4′-hydroxyl proton from the sugar and that Asp135 and Lys136 play critical roles in the subsequent dehydration step that leads to the final product.

1. Introduction

The 3,6-dideoxyhexoses, including abequose, paratose and tyvelose, are located primarily in the O antigens of surface lipopolysaccharides of some Gram-negative bacteria including *Yersinia pseudotuberculosis* and *Salmonella typhi* (Luderitz *et al.*, 1966; Ashwell & Hackman, 1971; Bishop & Jennings, 1982; Lindberg, 1990). These unusual carbohydrates have been characterized as the main antigenic determinants of these bacteria and thus contribute to their serological specificities (Raetz, 1990). The first committed step in the synthesis of tyvelose is catalyzed by CDP-D-glucose 4,6-dehydratase, the focus of this investigation (Liu & Thorson, 1994; He *et al.*, 1996). On the basis of amino-acid sequence homology, CDP-D-glucose 4,6-dehydratase is known to be a member of the short-chain dehydrogenase/reductase (SDR) superfamily (Oppermann *et al.*, 2003). These NAD(P)-dependent enzymes are widely distributed in nature, where they are involved in a number of important physiological processes including normal

and metastatic growth, fertility and hypertension (Duax *et al.*, 2000). In spite of limited amino-acid sequence identities of between 15 and 30%, members of this superfamily share highly conserved structural features (Duax *et al.*, 2003). While their quaternary structures vary, in all cases the individual subunits are distinctly bilobal. The N-terminal regions contain a Rossmann fold, which is responsible for anchoring the NAD(H) or NADP(H) dinucleotide into the active-site cleft. In addition to similar overall molecular folds, members of the SDR superfamily contain two characteristic signature sequences. The first of these is a Y*XXX*K motif in which the conserved tyrosine plays a key role in catalysis. The second of the signature sequences is a G*XXX*G*X*G motif, which is involved in dinucleotide binding.

As indicated in Fig. 1, CDP-D-glucose 4,6-dehydratase catalyzes the conversion of CDP-D-glucose to its 4-keto-6-deoxyhexose derivative by oxidizing the 4′-hydroxyl group and removing the 6′-hydroxyl group of the substrate (He *et al.*, 1996). Enzymes catalyzing similar reactions, such as dTDP-D-glucose 4,6-dehydratase and GDP-D-mannose 4,6-dehydratase, have been extensively studied both kinetically and structurally (Somoza *et al.*, 2000; Allard *et al.*, 2001, 2002, 2004; Gross *et al.*, 2001; Hegeman *et al.*, 2001, 2002; Mulichak *et al.*, 2002; Webb *et al.*, 2004). According to all presently available data, it is believed that in the first step of the reaction mechanism for the dTDP-D-glucose 4,6-dehydratases the phenolate group of the conserved tyrosine in the Y*XXX*K motif acts as the general base to abstract a proton from the 4′-hydroxyl group of the sugar, while the hydride on C-4 is transferred to NAD^+. This 4-dehydrogenation of the substrate serves as the activating step for the subsequent β-elimination of water between C-5 and C-6 by lowering the pK_a of the C-5 proton into the range 18–19 (March, 1985). In the second step of the reaction mechanism, a conserved glutamate removes the proton from C-5 of the sugar. As proton removal occurs at C-5, the C-6 hydroxyl is eliminated as water by donation of a proton from the side chain of a conserved aspartate, thereby forming a dTDP-4-ketoglucose-5,6-ene intermediate. Subsequent proton transfer from the conserved glutamate to C-5 and hydride transfer from NADH to C-6 yields the final product.

Recently, the crystal structure of CDP-D-glucose 4,6-dehydratase from *Y. pseudotuberculosis* complexed with NAD^+ was solved to 1.8 Å resolution (Vogan *et al.*, 2004). While this investigation represented the first structure of a CDP-dependent 4,6-dehydratase to be reported and revealed the overall tetrameric nature of the enzyme, details concerning the manner in which the substrate CDP-D-glucose is accommodated in the active site were limited to model building from the known molecular architecture of UDP-galactose 4-epimerase (Thoden *et al.*, 1996). Here, we describe the three-dimensional structure of CDP-D-glucose 4,6-dehydratase from *S. typhi* complexed with CDP-D-xylose. From this investigation, the manner in which the enzyme accommodates the nucleotide-linked sugar has been determined and the conformational changes that occur upon ligand binding defined.

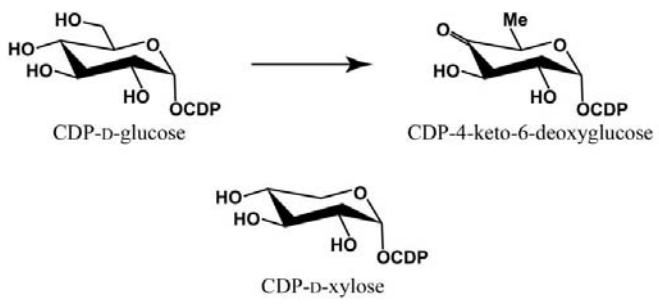

Figure 1
Reaction catalyzed by CDP-D-glucose 4,6-dehydratase.

2. Materials and methods

2.1. Cloning of the CDP-D-glucose 4,6-dehydratase gene

The sequence of the gene encoding CDP-D-glucose 4,6-dehydratase, *ddhB*, has previously been reported and allowed the design of primers for gene amplification (Hobbs & Reeves, 1995). Genomic DNA for *S. typhi* isolates CDC Nos. 87-2059 and 88-2009 was obtained as a generous gift from the laboratory of Dr Stanley Maloy at the University of Illinois, Urbana-Champaign. The *ddhB* gene was PCR amplified from genomic DNA such that the forward primer 5′-CTAGC-TAGCATTGATAAAAATTTTTGGCAAGGT-3′ included an *Nhe*I site, while the reverse primer 5′-CCCAAGCTTT-TACACCACCACCACCACCACACGAGTAGTTGCAGA-CATATA-3′ included a *Hin*dIII site for cloning into the expression vector pET-21a (Novagen). The reverse primer also encoded a His_6 tag added to the C-terminus of the protein. To create an *Nhe*I-recognition site in the forward primer, the codons GCT (Ala) and AGC (Ser) were inserted after the starting methionine of the dehydratase. The *ddhB* gene was PCR amplified with *Pfx* Platinum DNA polymerase (Invitrogen) according to the manufacturer's instructions and using standard cycling conditions. The PCR product was purified with the QIAquick PCR purification kit (Qiagen) and digested for 16 h at 310 K with both *Nhe*I and *Hin*dIII. The resulting gene fragment was gel purified with a QIAquick gel-purification kit (Qiagen). The purified fragment was subsequently ligated into *Escherichia coli* expression plasmid pET-21a (Novagen) that had been previously cut with the same restriction enzymes. The ligation mixture was used to transform *E. coli* DH5α cells, which were then plated onto LB media supplemented with 100 µg ml^{-1} ampicillin. Individual colonies were selected, cultured overnight and the plasmid DNA extracted with a Qiaprep spin miniprep kit (Qiagen Inc.). Plasmids were tested for incorporation of the *ddhB* gene by digestion with both *Nhe*I and *Hin*dIII. Positive clones were sequenced with the ABI prism Big Dye Primer Cycle sequencing kit (Applied Biosystems, Inc.) to check for any introduced mutations.

2.2. Protein expression

E. coli Rosetta(DE3) pLysS cells were transformed with the pET21a-*ddhB* plasmid. Transformants were plated onto LB

media supplemented with 30 μg ml^{-1} chloramphenicol and 100 μg ml^{-1} ampicillin. After ∼16 h of growth at 310 K, the plates were scraped and the cells resuspended in LB media for use in the inoculation of 8 × 2 l baffled flasks containing 500 ml TB media supplemented with the previously mentioned antibiotics. Cultures were grown at 310 K to an A_{600} of ∼0.7 and then induced with 0.5 m*M* isopropyl-1-thio-β-D-galactopyranoside (IPTG). The cells were allowed to grow for an additional 2 h at 310 K before harvesting by centrifugation at 6000*g* for 8 min. The cell paste was frozen in liquid nitrogen and stored at 193 K.

2.3. Expression of the selenomethionine-labeled protein

E. coli Rosetta(DE3) pLysS cells were transformed with the pET21a-*ddhB* plasmid and plated onto LB media supplemented with ampicillin as described above. The transformants were grown overnight (∼16 h) at 310 K. Several colonies were subsequently selected and used to inoculate 100 ml M9 minimal media supplemented with 30 μg ml^{-1} chloramphenicol and 100 μg ml^{-1} ampicillin for growth overnight at 310 K. Following this growth, 10 ml culture was used to inoculate 6 × 2 l baffled flasks each containing 500 ml M9 minimal media supplemented with 5 μg ml^{-1} thiamine, 30 μg ml^{-1} chloramphenicol and 100 μg ml^{-1} ampicillin. Cultures were incubated at 310 K to an A_{600} of ∼0.45 before adjusting the temperature to 303 K for the remainder of the growth period. At this time, the flasks were supplemented with 50 mg each of L-lysine, L-threonine, L-phenylalanine and 25 mg each of L-leucine, L-isoleucine, L-valine and L-selenomethionine (Van Duyne *et al.*, 1993). After an additional 20 min of growth, the cells were induced with 0.5 m*M* IPTG and allowed to grow for 7 h. The cultures were harvested by centrifugation at 6000*g* for 8 min and the cell paste was frozen in liquid nitrogen.

2.4. Protein purification

All protein-purification steps were carried out at 277 K unless otherwise noted. The purification scheme was the same for both the native and selenomethionine-substituted proteins. Approximately 13 g of cell paste was thawed in 60 ml Ni–NTA lysis buffer containing 50 m*M* NaH_2PO_4, 300 m*M* NaCl and 10 m*M* imidazole pH 8.0. Cells were lysed on ice by four cycles of sonication (30 s) separated by 3 min cooling. The lysate was centrifuged at 277 K for 40 min at 20 000*g* to remove cellular debris. The clarified lysate was loaded onto a 10 ml Ni–NTA agarose column (Qiagen, Inc.) previously equilibrated with Ni–NTA lysis buffer. After loading, the column was washed with Ni–NTA wash buffer (50 m*M* NaH_2PO_4, 300 m*M* NaCl and 20 m*M* imidazole pH 8.0) until the A_{280} of the flowthrough decreased to a stable baseline. The protein was eluted from the column with a gradient of 20–300 m*M* imidazole in Ni–NTA lysis buffer. Protein-containing fractions were pooled on the basis of SDS–PAGE and dialyzed into 25 m*M* HEPPS pH 8.0. The protein was concentrated to 20 mg ml^{-1} based on an extinction coefficient of 1 ml mg^{-1} cm^{-1} as calculated with the program *PROTEAN* (DNAstar). Aliquots of the protein were quickly frozen in liquid nitrogen and stored at 193 K prior to crystallization trials.

2.5. Molecular-weight determination

Analytical ultracentrifugation experiments were performed at the University of Wisconsin Biochemistry Instrumentation Facility. Three samples of the wild-type dehydratase were tested at 0.25, 0.5 and 1.0 mg ml^{-1} in 20 m*M* HEPPS pH 8.0. The experiments were performed with a Beckman Optima XL-A centrifuge and an An60-Ti rotor at 293 K with 12 mm double-sector charcoal-filled Epson centerpieces. Analysis of the radial position of the sample (by UV absorption at 280 nm) at various rotor speeds showed CDP-D-glucose 4,6-dehydratase to be a homotetramer.

2.6. Enzymatic assay

The activity of CDP-D-glucose 4,6-dehydratase was measured *via* an assay which detects the presence of CDP-4-keto-6-deoxyglucose in 0.1 *M* NaOH by its characteristic absorbance at 320 nm (Matsuhashi *et al.*, 1966; Yu *et al.*, 1992). The assay was performed in triplicate. Each 100 μl reaction contained 25 m*M* HEPES pH 8.0, 1 m*M* CDP-D-glucose and 6 μg CDP-D-glucose 4,6-dehydratase. The addition of NAD^+ to a final concentration of 1 m*M* did not affect the outcome of the assay. The assay mixture was incubated for 15 min at 310 K followed by the addition of 400 μl sterile water and 500 μl 0.2 *M* NaOH. The assay mixture was then incubated a second time at 310 K for 20 min. The spectrophotometer was zeroed with 0.1 *M* NaOH. A negative control was included in which no CDP-D-glucose was added to the reaction mixture. The extinction coefficient of CDP-4-keto-6-deoxyglucose at 320 nm under the assay conditions described has been reported as 6500 M^{-1} cm^{-1} (Matsuhashi *et al.*, 1966).

2.7. Enzymatic synthesis of CDP-D-xylose

CDP-D-xylose was synthesized in an enzymatic reaction utilizing *S. typhi* glucose-1-phosphate cytidylyltransferase. The expression and purification of this enzyme has been previously reported (Koropatkin & Holden, 2004). All other reagents and enzymes were purchased from Sigma–Aldrich. The enzymatic reaction, in a total volume of 30 ml, contained 10 m*M* D-xylose-1-phosphate, 10 m*M* CTP, 10 m*M* $MgCl_2$, 200 U inorganic pyrophosphatase and 5 mg cytidylyltransferase in 25 m*M* HEPPS pH 8.0. The reaction mixture was incubated at 310 K for 1.5 h followed by 16 h at room temperature. The mixture was filtered through a Centriprep YM-30 to remove the enzymes and then diluted 1:1 with 20 m*M* ammonium carbonate pH 8.5. The CDP-D-xylose product was purified in two chromatographic steps on an Äkta HPLC (Amersham-Pharmacia) equipped with a 6 ml Resource Q column. CDP-D-xylose was purified from CDP and CTP using a linear gradient of 20–500 m*M* ammonium carbonate pH 8.5. The CDP-D-xylose-containing fractions were identified by comparing the retention time to that of authentic samples of CDP, CTP and CDP-D-glucose. The CDP-D-xylose fractions were lyophilized and resuspended in

Table 1
X-ray data-collection statistics.

Values in parentheses are for the highest resolution bin.

	Peak	Inflection	Remote	Native
Wavelength (Å)	0.97920	0.97932	0.96401	0.96411
Resolution (Å)	50–2.2 (2.28–2.2)	50–2.2 (2.28–2.2)	50–2.2 (2.28–2.2)	50–1.8 (1.86–1.8)
No. of independent reflections	39920 (3894)	39910 (3904)	39944 (3927)	68364 (6117)
Completeness (%)	99.7 (99.0)	99.6 (99.0)	99.6 (99.6)	94.3 (85.7)
Redundancy	8.8 (6.2)	9.1 (6.1)	9.4 (7.9)	5.2 (2.5)
$\langle I\rangle/\langle\sigma(I)\rangle$	41.5 (6.3)	40.4 (5.5)	45.3 (7.7)	37.1 (3.1)
$R_{\rm sym}$† (%)	7.8 (20.2)	7.0 (18.8)	6.5 (16.7)	5.5 (24.8)

† $R_{\rm sym} = (\sum|I-\bar{I}|/\sum I) \times 100$.

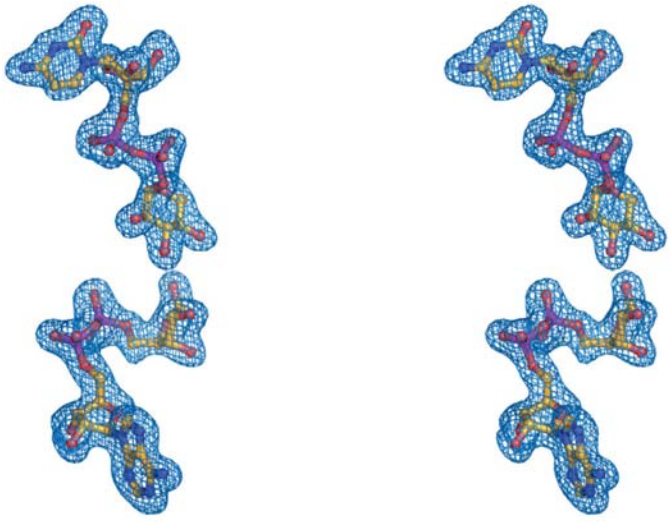

Figure 2
Electron density corresponding to the ADP-ribose and CDP-D-xylose ligands in subunit I. The map was calculated with coefficients of the form $(F_o - F_c)$, where F_o is the native structure-factor amplitude and F_c is the calculated structure-factor amplitude from the model lacking the coordinates for the ligands. The map was contoured at 3σ.

20 m*M* ammonium acetate pH 4.0. In the second chromatographic step, CDP-D-xylose was purified from any remaining D-xylose-1-phosphate *via* a linear gradient of 20–500 m*M* ammonium acetate pH 4.0. The CDP-D-xylose-containing fractions were pooled and lyophilized. Electrospray ionization mass spectrometry in the negative-ion mode was performed at the University of Wisconsin Biotechnology Center. The mass of the parent ion from a sample of purified CDP-D-xylose was 534.6, which is consistent with the mass of CDP-D-xylose.

2.8. Crystallization of selenomethionine-labeled CDP-D-glucose 4,6-dehydratase

Crystallization trials of the purified selenomethionine-labeled dehydratase were carried out with both an in-house sparse-matrix screen and Crystal Screen I from Hampton Research. The hanging-drop method of vapor diffusion was utilized for crystallization trials. Small crystals were observed at room temperature from a solution containing 1.6 *M* sodium/potassium phosphate, 100 m*M* HEPES pH 7.5. Refinement of the crystallization conditions led to large single crystals growing by batch methods at room temperature. Specifically, the protein at 20 mg ml^{-1} was mixed in a 1:1 ratio with a precipitant solution containing 100 m*M* HOMOPIPES and 0.9–1.1 *M* sodium/potassium phosphate pH 4.2. The diamond-shaped crystals achieved typical dimensions of ∼0.5 × 0. 7 × 0.4 mm in two to four weeks and belonged to space group $P3_121$, with unit-cell parameters $a = b = 94.0$, $c = 151.8$ Å and two monomers per asymmetric unit.

2.9. Preparation of the wild-type enzyme–CDP-D-xylose complex

The substrate analog CDP-D-xylose was soaked into crystals of the wild-type enzyme grown under similar conditions as described for the selenomethionine-substituted protein. Specifically, the crystals were transferred to a synthetic mother liquor containing 0.5 *M* sodium/potassium phosphate, 100 m*M* HOMOPIPES pH 4.2, 10 m*M* NAD^+ and 25 m*M* CDP-D-xylose and soaked for 24 h. The unit-cell parameters were the same within experimental error.

2.10. High-resolution X-ray data collection

Both wild-type and selenomethionine-labeled protein crystals were treated in similar manners. Crystals were harvested from batch experiments and soaked for 24 h in a synthetic mother liquor consisting of 0.6 *M* sodium/potassium phosphate, 100 m*M* HOMOPIPES pH 4.2. In the case of the wild-type protein, the soaking solution also contained 10 m*M* NAD^+ and 25 m*M* CDP-D-xylose. The crystals were then gradually transferred first to a final synthetic mother liquor containing 1.5 *M* sodium/potassium phosphate, 100 m*M* HOMOPIPES pH 4.2 and next to a cryoprotectant solution containing 2 *M* sodium/potassium phosphate, 100 m*M* HOMOPIPES pH 4.2 and 15% ethylene glycol. The crystals were flash-cooled to 123 K in a stream of nitrogen vapor and stored under liquid nitrogen until synchrotron beam time became available. X-ray data sets from the selenomethionine-labeled crystals were collected on a MAR detector at the COM-CAT 32-ID beamline (Advanced Photon Source, Argonne National Laboratory, Argonne, IL, USA). X-ray data from the wild-type protein–CDP-D-xylose crystals were collected on a 3 × 3 tiled SBC3 CCD detector at the Structural Biology Center 19-ID beamline (Advanced Photon Source, Argonne National Laboratory, Argonne, IL, USA). All X-ray data were processed with *HKL*2000 and scaled with *SCALEPACK* (Otwinowski & Minor, 1997). Relevant X-ray data-collection statistics are presented in Table 1.

2.11. X-ray structural analysis

The structure of CDP-D-glucose 4,6-dehydratase from *S. typhi* was solved *via* multiwavelength anomalous dispersion phasing with X-ray data obtained from the selenomethionine-substituted protein crystals. The software package *SOLVE*

was utilized to determine and refine the positions of the Se atoms (Terwilliger, 2000). A readily interpretable electron-density map to 2.3 Å was obtained directly from *SOLVE* (figure of merit of 0.7) and allowed complete manual tracing of the dimer in the asymmetric unit. The model was subjected to ten cycles of least-squares refinement with the software package *TNT*, which reduced the *R* factor to 23.9% (Tronrud *et al.*, 1987). Non-crystallographic symmetry restraints were not imposed during any of the least-squares refinement procedures. The structure of the wild-type enzyme–CDP-D-xylose complex was solved *via* molecular replacement with the software package *AMoRe* using this partially refined structure as a search model (Navaza, 1987). Alternate cycles of manual model building with the graphics program *TURBO* and least-squares refinement with *TNT* reduced the *R* factor to 20.3% for all measured X-ray data from 50 to 1.8 Å (Roussel & Cambillau, 1998). The model includes ADP-ribose in both subunits and CDP-D-xylose in one subunit. Relevant least-squares refinement statistics are presented in Table 2. The electron density corresponding to ADP-ribose and CDP-D-xylose in subunit I is shown in Fig. 2. A Ramachandran plot analysis indicates that 89.9, 9.1 and 1.0% of the residues are in the most favored, additional allowed and generously allowed regions, respectively. The most significant outlier is Thr133 in each subunit ($\varphi = -104.5$, $\psi = -111.9^\circ$ and $\varphi = -109.2$, $\psi = -106.4^\circ$ in subunits I and II, respectively). Thr133 resides at the end of the fifth β-strand and the electron density is unambiguous in this region. Subunit I extends from Ser3 to Ala358, with one break in the polypeptide chain between Asp294 and Pro299. Subunit II extends from Ser3 to Ser357, with one break in the polypeptide chain between His300 and His303.

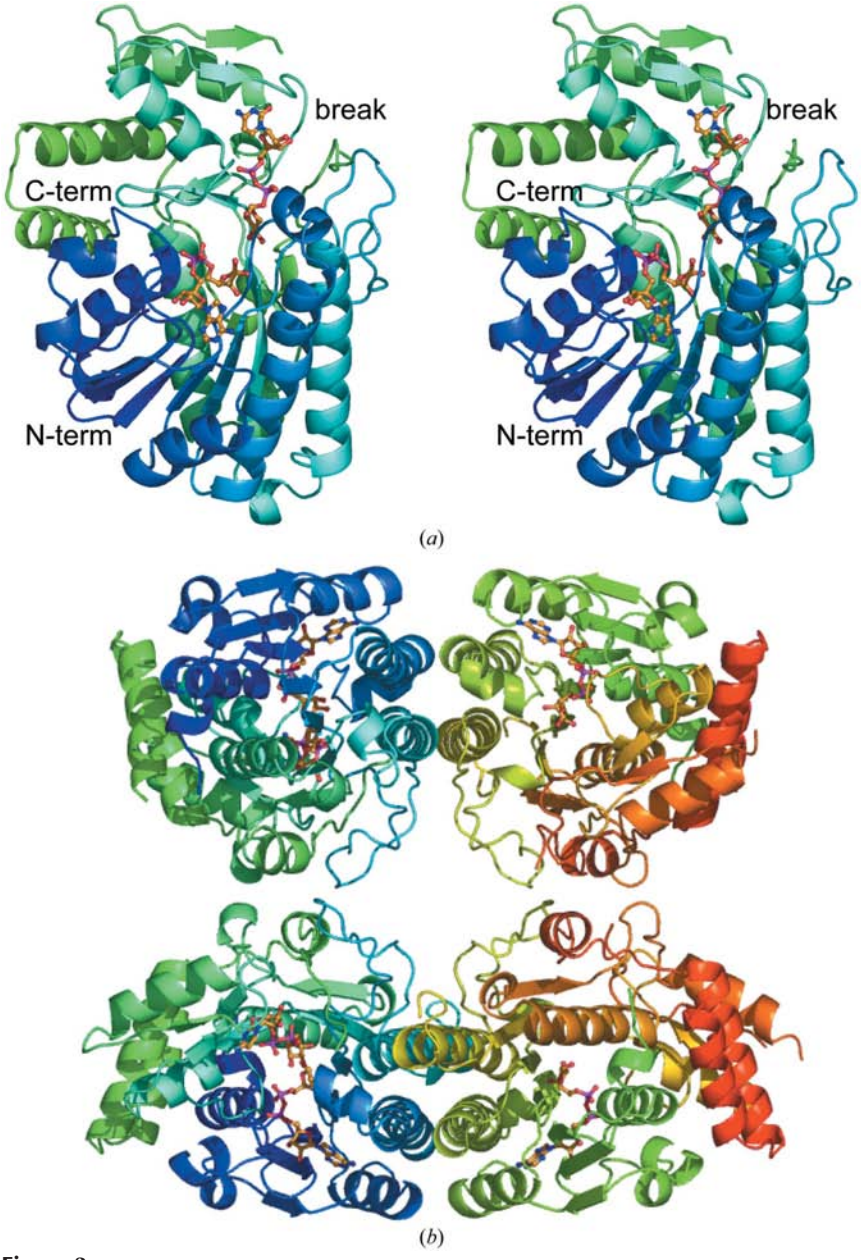

Figure 3
Ribbon representation of CDP-D-glucose 4,6-dehydratase. A stereoview of one subunit, color-ramped (blue to green) from the N-terminus to the C-terminus, is shown in (*a*). The active site is wedged between the N- and C-terminal domains, with the ADP-ribose and CDP-D-xylose displayed in a ball-and-stick representation. The quaternary structure of the enzyme is shown in (*b*).

3. Results and discussion

As expected from amino-acid sequence homology, the subunit architecture of the *S. typhi* CDP-D-glucose 4,6-dehydratase is that typically observed for other members of the SDR superfamily, with the N-terminal domain dominated by a seven-stranded parallel β-sheet (formed by Arg13–Thr17, Ile37–Ala42, Glu62–Ile65, Ile85–His88, Ala128–Thr133, Gly189–Ile199 and Glu260–Phe264) and seven helical regions (Fig. 3*a*). The C-terminal domain is composed mostly of α-helix, with a two-stranded parallel β-sheet formed by Val223–Ile225 and Ser290–Asp294 and a second one formed by Ile231–Gln235 and Tyr304–Asp308.

Crystals employed for this investigation belonged to space group $P3_121$ and contained two subunits per asymmetric unit. While NAD^+ was included in the crystallization trials, in each subunit the nicotinamide ring of the dinucleotide was absent. A previous enzymatic characterization of the enzyme indicated that the purified enzyme contained a high proportion of NADH (He *et al.*, 1996). The presence of NADH may be a consequence of the *in vivo* formation of naturally occurring abortive complexes of the enzyme, as has been observed for the closely related enzyme UDP-galactose 4-epimerase. In enzymatic assays, however,

Table 2
Least-squares refinement statistics.

Resolution limits (Å)	50 – 1.8
R factor† (overall) (%)/No. of reflections	20.3/68364
R factor (working) (%)/No. of reflections	20.1/61460
R factor (free) (%)/No. of reflections	27.0/6904
No. of protein atoms‡	5666
No. of heteroatoms§	432
Average *B* values (Å^2)	
Protein atoms	37.7
CDP-D-xylose	27.2
ADP-ribose	37.7
Solvents	41.8
Weighted r.m.s. deviations from ideality	
Bond lengths (Å)	0.011
Bond angles (°)	2.25
Trigonal planes (Å)	0.007
General planes (Å)	0.011
Torsional angles¶ (°)	17.9

† 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. ‡ This value includes multiple conformations for Asp70 in subunit I and Ser25, Val237 and Glu269 in subunit II. § These include 333 water molecules. ¶ The torsional angles were not restrained during the refinement.

exogenous NAD$^+$ was not required for full catalytic activity. It is most likely the nicotinamide ring of NADH was hydrolyzed from the dinucleotide over the period of time required for crystal growth at low pH (4.2) (Lowry *et al.*, 1961). The two subunits in the asymmetric unit are related by a local twofold rotational axis and superimpose with a root-mean-square deviation of 0.45 Å for 285 C$^\alpha$ atoms. Subunit–subunit interactions are mediated primarily by residues positioned in the α-helices delineated by Val95–Val123 and Pro158–Phe177. The total buried surface area at this interface is ∼2300 Å^2 according to the algorithm of Lee & Richards (1971) using a probe sphere radius of 1.4 Å. Ultracentrifugation experiments indicate that the quaternary structure of the *S. typhi* CDP-D-glucose 4,6-dehydratase is tetrameric. Examination of the crystalline packing arrangement reveals that the homotetramer is positioned in the lattice with one of its dyads coincident to the crystallographic twofold along the z axis. A ribbon representation of the complete homotetramer, which displays 222 symmetry, is given in Fig. 3(*b*). The total buried surface area at the second subunit–subunit interface is less extensive at ∼1400 Å^2 and thus the quaternary structure of the dehydratase can be aptly described as a dimer of dimers.

Subunit I contains bound CDP-D-xylose, while subunit II does not. Fig. 4 shows the conformational changes that occur upon CDP–sugar binding. The most significant differences occur in the region defined by Ala205–Pro233 and Glu269–His303. In subunit II, the β-strands formed by Val223–Ile225 and Ser290–Asp294 are shifted away from the active site by approximately 2.5 Å. Strikingly, the polypeptide chains for both subunits I and II contain breaks in this region suggesting conformational flexibility. Specifically, subunit I contains a break between Asp294 and Pro299, while the break in subunit II occurs between His300 and His303.

Despite the absence of the nicotinamide rings, the ADP-ribose moieties observed in both subunits I and II are bound in virtually identical positions to those observed for the NAD$^+$ cofactors in the dTDP-D-glucose 4,6-dehydratases and the *Yersinia* CDP-D-glucose 4,6-dehydratase (Thoden *et al.*, 1996; Allard *et al.*, 2002, 2004; Vogan *et al.*, 2004). As subunit I represents a ternary complex of the enzyme, the rest of the discussion will refer to it only.

A close-up view of the active site of the *S. typhi* CDP-D-glucose 4,6-dehydratase is shown in Fig. 5(*a*). There are eight well ordered water molecules lying within 3.2 Å of the dinucleotide. The adenine ribose adopts the C$_{2'}$-*endo* conformation, while the adenine base is anchored to the protein *via* hydrogen-bonding interactions with the carboxylate side chain of Asp67 and the backbone amide of Ile68. The ribose-ring O atom is located at 3.2 Å from the backbone amide N atom of Ala91, while its 2′- and 3′-hydroxyl groups are bridged by Thr20 O$^\gamma$. The 2′-hydroxyl group also lies within 3.1 and 2.6 Å of the backbone amide N atom of Leu43 and an ordered water molecule, respectively. The backbone amide N atom of Phe22 sits at 2.7 Å from one of the phosphoryl O atoms (Fig. 5*a*). Both Thr20 and Phe22 belong to the highly conserved sequence GHTGFKG found in the CDP-D-glucose 4,6-dehydratases. In the *S. typhi* enzyme, this conserved sequence begins at Gly18. As can be seen in Fig. 5(*a*), the guanidinium group of Arg208 projects towards the second phosphoryl group of the dinucleotide. The terminal ribose group, which technically is the nicotinamide ribose in NAD(H), adopts the C$_{2'}$-*endo* conformation and is hydrogen bonded through its 2′- and 3′-hydroxyl groups to Lys163 N$^\zeta$. This residue is the lysine of the Y*XXX*K signature sequence that is characteristic to all members of the

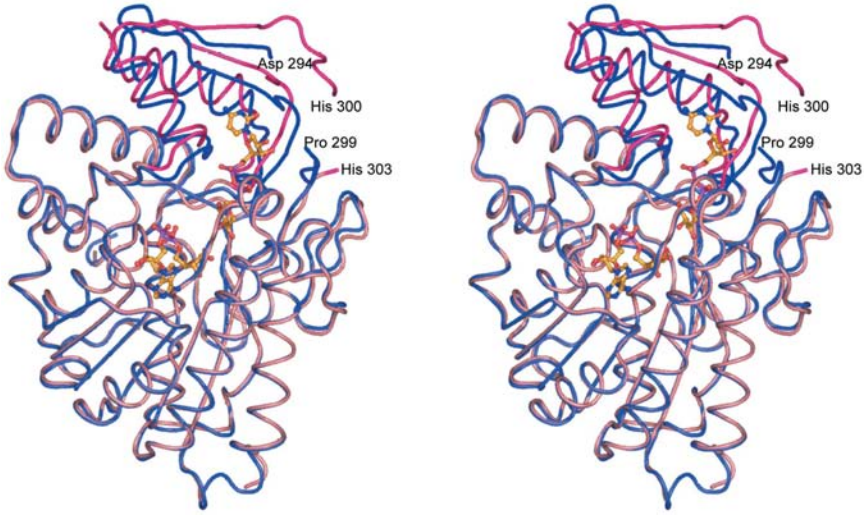

Figure 4
Overlay of subunits I and II in the asymmetric unit. Subunit I (with bound CDP-D-xylose) is depicted in blue, whereas subunit II is displayed in pink. The areas of the polypeptide chain that shift significantly upon substrate binding are highlighted in darker colors. The CDP-D-xylose and ADP-ribose ligands are displayed in a ball-and-stick representation.

SDR superfamily. Two biochemical roles have been postulated for this lysine. One hypothesis suggests that this basic residue serves to activate the catalytic base, which is the conserved tyrosine in the Y*XXX*K motif. In the structure presented here, Lys163 N^{ζ} is 4.1 Å from the catalytic tyrosine, a distance comparable to that observed in the crystal structures of UDP-galactose 4-epimerase, CDP-tyvelose 2-epimerase and dTDP-D-glucose 4,6-dehydratase, among others (Thoden *et al.*, 1996; Koropatkin *et al.*, 2003; Allard *et al.*, 2004). The second hypothesis is that the primary role for this conserved lysine is simply for proper positioning of the dinucleotide within the active site. In addition to its interaction with Lys163, the 2′-hydroxyl group of the terminal ribose is located within 2.9 Å of Tyr159 O^{η} (the catalytic base).

CDP-D-xylose is bound primarily by residues in the C-terminal domain, with the sugar moiety extending into the active-site cleft. In particular, the backbone amides of Arg226 and Asn227, the carbonyl O atom of Ile224 and the carboxylate group of Asp212 all provide hydrogen-bonding interactions with various atoms of the cytosine ring. Interestingly, the amino group of the cytosine base lies at 2.8 Å from both the carboxylate side chain of Asp212 and the backbone carbonyl O atom of Ile224. The specificity of this enzyme for CDP sugars can be explained by the presence of Asp212. Both thymine and uracil have a carbonyl rather than an amino moiety in the same position. Indeed, in the dTDP-D-glucose 4,6-dehydratase from *Streptomyces venezuelae*, for example, the structurally comparable residue is Leu194 (Allard *et al.*, 2004). The aliphatic portion of Arg226 and the side chain of Leu209 abut opposite sides of the cytosine base. Both Asn227 and Glu301 shift upon substrate binding and are located within ∼2.8 Å of the 2′-hydroxyl group of the ribose (which also adopts the $C_{2'}$-*endo* conformation). The phosphoryl O atoms of CDP-D-xylose interact with the protein *via* hydrogen-bonding and electrostatic interactions with the backbone amide of Leu209 and the side-chain atoms of Lys136, Asn197 and Arg232. There are seven well ordered water molecules located within 3.2 Å of CDP-D-xylose, several of which are hydrogen bonded to the phosphoryl O atoms.

The 4′-hydroxyl group of the xylose moiety is positioned at 2.9 Å from both Ser134 O^{γ} and Tyr159 O^{η} with ideal hydrogen-bonding geometry. The distance and geometry displayed by Tyr159 with respect to the 4′-hydroxyl group of xylose is in keeping with what has been observed in the similarly related dTDP-D-glucose 4,6-dehydratases. The 2′-hydroxyl group of xylose interacts with the guanidinium side chain of Arg208 (2.9 Å) and a water molecule (2.5 Å), while the 3′-hydroxyl group sits within 2.7 and 3.1 Å of the carbonyl O atom of Pro93 and Tyr159 O^{η}, respectively. Finally, the ring O atom of the ribose sits near the side chain of Asn197 (3.1 Å). Strikingly, Asp135 and Lys136 are located at 3.5 and 3.2 Å, respectively, from C-5 of xylose. In the crystal structures of several dTDP-D-glucose 4,6-dehydratases, an aspartate/glutamate pair occupy equivalent positions to Asp135 and Lys136 (Allard *et al.*, 2002, 2004).

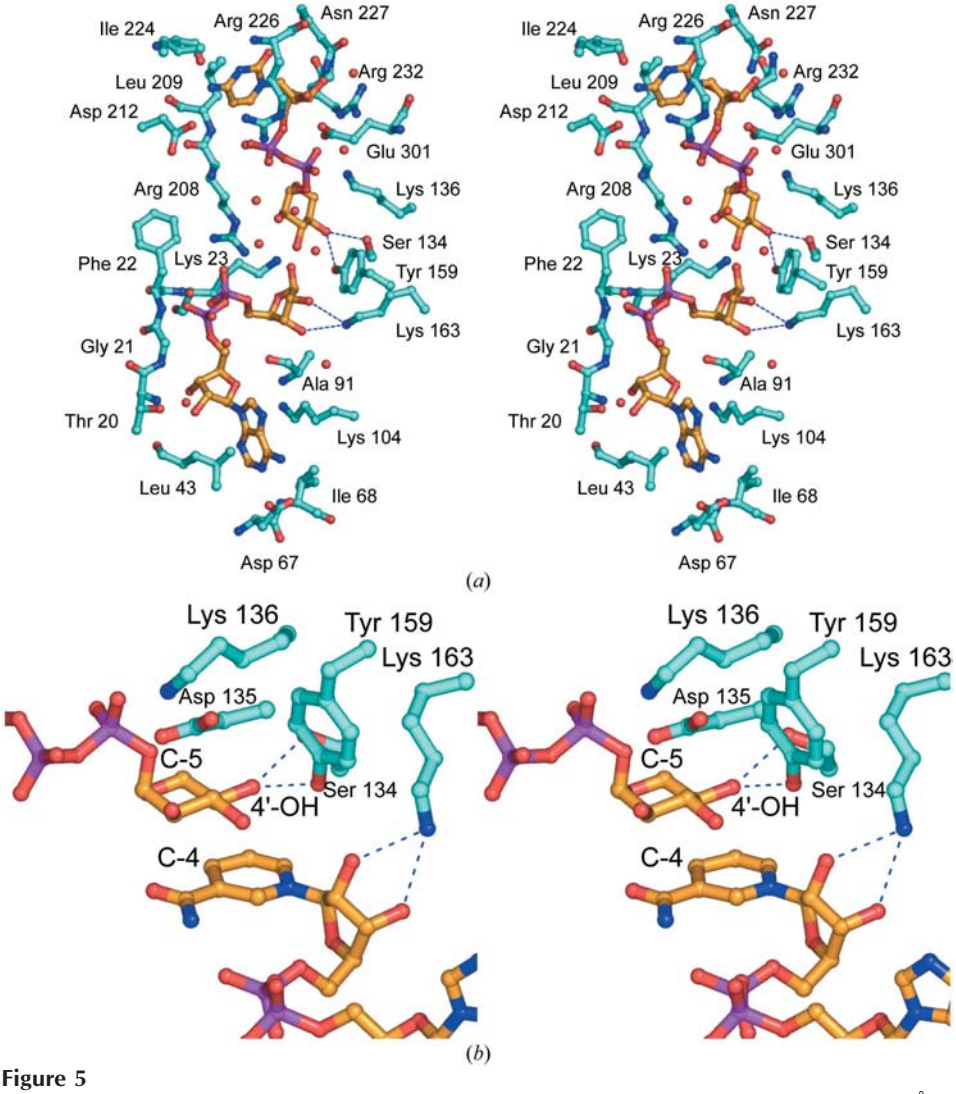

Figure 5
Close-up view of the CDP-D-glucose 4,6-dehydratase active site. Those residues located within 3.5 Å of the ADP-ribose and CDP-D-xylose ligands are shown in (*a*). Ordered water molecules are displayed as red spheres. Important hydrogen-bonding interactions are indicated by the blue dashed lines. For the sake of clarity, Gly18, Met89, Pro93, Asn197, Asp207, Ser230 and Asp135 have been omitted from the figure. A model of the active site with bound NAD^{+} and CDP-D-xylose is depicted in (*b*). The ligands are displayed in gold, while the protein is shown in cyan. Only residues thought to play key roles in the reaction mechanism are displayed.

Previous kinetic analyses of CDP-D-glucose 4,6-dehydratase have suggested that it differs from the dTDP-D-glucose 4,6-dehydratases and GDP-D-mannose 4,6-dehydratases in its requirement for exogenous NAD^+ for full catalytic activity (Matsuhashi *et al.*, 1966; Yu *et al.*, 1992; He *et al.*, 1996). Indeed, it has been speculated that the relatively low affinity of CDP-D-glucose 4,6-dehydratase for the NAD(H) dinucleotide might be attributed to an extended NAD(H)-binding motif relative to that found in the dTDP-D-glucose 4,6-dehydratases (He *et al.*, 1996). In order to test this theory, site-directed mutant proteins of the *Yersinia* CDP-D-glucose 4,6-dehydratase were created. Specifically, His17 was changed to a glycine and Lys21 was mutated to an isoleucine in order to mimic the binding region observed in the dTDP-D-glucose dehydratases (He *et al.*, 1996). Most notably, the H17G mutant protein demonstrated an approximate threefold increase in catalytic efficiency and a greater than tenfold decrease in the K_d for NAD^+. In contrast, the K21I mutant protein did not display a significant change in catalytic efficiency, but still showed enhanced affinity for NAD^+. On a structural level, however, these observations are somewhat difficult to reconcile. Superposition of the X-ray coordinates for the *S. typhi* CDP-D-glucose 4,6-dehydratase presented here with those of dTDP-D-glucose 4,6-dehydratase from *S. venezuelae* demonstrates that the polypeptide chains adopt nearly identical conformations in the area surrounding the NAD^+ cofactor (Allard *et al.*, 2004). The backbone atoms of the signature motif form similar hydrogen-bonding contacts with the phosphoryl O atoms of the NAD^+ cofactor in both of these structures. In fact, the φ, ψ angles of the histidine in the *S. typhi* enzyme are nearly identical to those of Gly9, which is found in the equivalent position in the *S. venezuelae* TDP-D-glucose 4,6-dehydratase. It should also be noted that the mechanistic and structural homolog GDP-D-mannose 4,6-dehydratase, which utilizes NADP(H) rather than NAD(H), contains an extended signature motif G(I/V)TGQDG and retains the cofactor throughout purification (Mulichak *et al.*, 2002; Webb *et al.*, 2004). Additionally, both ADP-ribose and CDP were observed in the X-ray crystal structure of the selenomethionine-substituted CDP-D-glucose 4,6-dehydratase, despite the fact that the dinucleotide was not added at any point during the purification or crystallization experiments. These results suggest that, as observed in other dehydratases, the NAD(H) cofactor is tightly bound to the protein and is not released during the catalytic cycle. Further studies will be required in order to clarify this issue.

In order to visualize the manner in which the cofactor and substrate are aligned for catalysis, the nicotinamide ring was modeled into the active site on the basis of that observed in the structure of the dTDP-D-glucose 4,6-dehydratase from *S. venezuelae*. This model is presented in Fig. 5(*b*). The nicotinamide ring was built in the *syn* conformation as observed for other B-side specific enzymes (Thoden *et al.*, 1996; Thoden & Holden, 2000; Allard *et al.*, 2002, 2004; Mulichak *et al.*, 2002). In this model, none of the atoms of the ADP-ribose had to be moved to accommodate the nicotinamide ring. However, in order to prevent steric clashing between the side chain of Arg208 and the nicotinamide ring, this residue was rotated away from the active site. In the *Yersinia* enzyme structure this arginine also points away from the active site (Vogan *et al.*, 2004).

In the model presented in Fig. 5(*b*), C-4 of the nicotinamide ring is approximately 2.9 Å from C-4 of xylose. The most striking difference between the active sites of CDP-D-glucose 4,6-dehydratase and dTDP-D-glucose 4,6-dehydratase is the substitution of a conserved glutamate, Glu129 in the *S. venezuelae* dTDP-D-glucose 4,6-dehydratase, for Lys136 in the CDP-D-glucose 4,6-dehydratase. In the structure of CDP-D-glucose 4,6-dehydratase, Lys N^{ζ} is positioned at 3.2 Å from C-5 of the xylose moiety. On the basis of both kinetic isotope-exchange experiments and X-ray crystallographic analyses, the glutamate of dTDP-D-glucose 4,6-dehydratase is thought to act as a catalytic base, abstracting a proton from the activated sugar C-5 in the dehydration step (Allard *et al.*, 2002, 2004; Hegeman *et al.*, 2002). It is most likely that Lys136 in CDP-D-glucose 4,6-dehydratase serves a similar role assuming that is it deprotonated, which is likely given that its ε-amino group is 2.8 Å

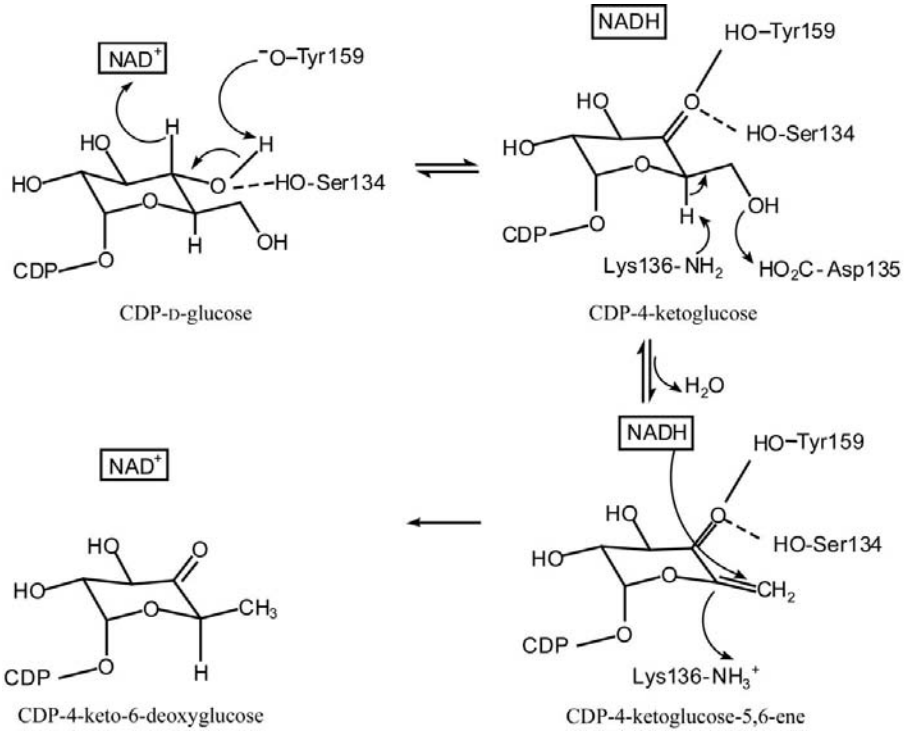

Figure 6
Proposed mechanism for CDP-D-glucose 4,6-dehydratase.

from a negatively charged phosphoryl O atom of the nucleotide. The other step of the dehydration reaction presumably occurs *via* a catalytic acid, which donates a proton to the C-6 hydroxyl leading to the elimination of water across the C-5 and C-6 bond. The catalytic acid in the dTDP-D-glucose 4,6-dehydratase from *S. venezuelae* has been identified as Asp128. In CDP-D-glucose 4,6-dehydratase, Asp135 occupies a similar position. The carboxylate side chain of Asp135 is located at 3.4 Å from C-5 of xylose. It is possible that this residue shifts somewhat towards the 6′-hydroxyl group of the sugar upon the binding of the natural substrate CDP-D-glucose.

From the current analysis of the *S. typhi* CDP-D-glucose 4,6-dehydratase, it can be speculated that the mechanism of this enzyme is similar to that proposed for the dTDP-D-glucose 4,6-dehydratases. Accordingly, as outlined in Fig. 6, Tyr157 abstracts a proton from C-4 of the glucose moiety and a hydride is transferred to NAD^+, resulting in the formation of a CDP-4-ketohexose intermediate and NADH. In the dehydration step, Lys136 abstracts a proton from C-5 of the intermediate, while Asp135 donates a proton to the 6′-hydroxyl, resulting in a CDP-4-ketoglucose-5,6-ene intermediate. In the final step, the hydride from NADH is transferred back to C-6 of the intermediate and a proton from Lys136 is donated back to C-5, yielding the final product, CDP-4-keto-6-deoxyglucose.

In conclusion, the structure of CDP-D-glucose 4,6-dehydratase from *S. typhi* complexed with CDP-D-xylose has been solved to high resolution and has allowed a more complete description of its active-site geometry. Despite catalyzing nearly identical reactions, the NDP-sugar 4,6-dehydratases have evolved with differing quaternary structures and nucleotide specificities. In addition, it appears that CDP-D-glucose 4,6-dehydratase utilizes a lysine rather than a glutamate as the catalytic base in the dehydration step of the reaction.

This research was supported in part by a grant from the NIH (DK47814 to HMH).

References

Allard, S. T., Beis, K., Giraud, M. F., Hegeman, A. D., Gross, J. W., Wilmouth, R. C., Whitfield, C., Graninger, M., Messner, P., Allen, A. G., Maskell, D. J. & Naismith, J. H. (2002). *Structure*, **10**, 81–92.

Allard, S. T., Cleland, W. W. & Holden, H. M. (2004). *J. Biol. Chem.* **279**, 2211–2220.

Allard, S. T., Giraud, M. F., Whitfield, C., Graninger, M., Messner, P. & Naismith, J. H. (2001). *J. Mol. Biol.* **307**, 283–295.

Ashwell, G. & Hackman, J. (1971). *Microbial Toxins*, edited by G. Weinbaum, S. Kodis & S. A. Aji, pp. 235–266. New York: Academic Press.

Bishop, C. T. & Jennings, H. J. (1982). *The Polysaccharides*, edited by G. O. Aspinall, pp. 291–330. New York: Academic Press.

Duax, W. L., Ghosh, D. & Pletnev, V. (2000). *Vitam. Horm.* **58**, 121–148.

Duax, W. L., Pletnev, V., Addlagatta, A., Bruenn, J. & Weeks, C. M. (2003). *Proteins*, **53**, 931–943.

Gross, J. W., Hegeman, A. D., Gerratana, B. & Frey, P. A. (2001). *Biochemistry*, **40**, 12497–12504.

He, X., Thorson, J. S. & Liu, H. W. (1996). *Biochemistry*, **35**, 4721–4731.

Hegeman, A. D., Gross, J. W. & Frey, P. A. (2001). *Biochemistry*, **40**, 6598–6610.

Hegeman, A. D., Gross, J. W. & Frey, P. A. (2002). *Biochemistry*, **41**, 2797–2804.

Hobbs, M. & Reeves, P. R. (1995). *Biochim. Biophys. Acta*, **1245**, 273–277.

Koropatkin, N. M. & Holden, H. M. (2004). *J. Biol. Chem.* **279**, 44023–44029.

Koropatkin, N. M., Liu, H. W. & Holden, H. M. (2003). *J. Biol. Chem.* **278**, 20874–20881.

Lee, B. & Richards, F. M. (1971). *J. Mol. Biol.* **55**, 379–400.

Lindberg, B. (1990). *Adv. Carbohydr. Chem. Biochem.* **48**, 279–318.

Liu, H. W. & Thorson, J. S. (1994). *Annu. Rev. Microbiol.* **48**, 223–256.

Lowry, O. H., Passonneau, J. V. & Rock, M. K. (1961). *J. Biol. Chem.* **236**, 2756–2759.

Luderitz, O., Staub, A. M. & Westphal, O. (1966). *Bacteriol. Rev.* **30**, 192–255.

March, J. (1985). *Advanced Organic Chemistry*, pp. 218–236. New York: Wiley.

Matsuhashi, S., Matsuhashi, M., Brown, J. G. & Strominger, J. L. (1966). *J. Biol. Chem.* **241**, 4283–4287.

Mulichak, A. M., Bonin, C. P., Reiter, W. D. & Garavito, R. M. (2002). *Biochemistry*, **41**, 15578–15589.

Navaza, J. (1987). *Acta Cryst.* A**43**, 645–653.

Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X., Lindh, M., Shafqat, J., Nordling, E., Kallberg, Y., Persson, B. & Jörnvall, H. (2003). *Chem.-Biol. Interact.* **143–144**, 247–253.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Raetz, C. R. (1990). *Annu. Rev. Biochem.* **59**, 129–170.

Roussel, A. & Cambillau, C. (1998). *Silicon Graphics Partners Directory*, pp. 77–78. Mountain View, CA, USA: Silicon Graphics.

Somoza, J. R., Menon, S., Schmidt, H., Joseph-McCarthy, D., Dessen, A., Stahl, M. L., Somers, W. S. & Sullivan, F. X. (2000). *Structure*, **8**, 123–135.

Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.

Thoden, J. B., Frey, P. A. & Holden, H. M. (1996). *Biochemistry*, **35**, 5137–5144.

Thoden, J. B. & Holden, H. M. (2000). *Trans. Am. Crystallogr. Assoc.* **35**, 103–114.

Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). *Acta Cryst.* A**43**, 489–501.

Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993). *J. Mol. Biol.* **229**, 105–124.

Vogan, E. M., Bellamacina, C., He, X., Liu, H. W., Ringe, D. & Petsko, G. A. (2004). *Biochemistry*, **43**, 3057–3067.

Webb, N. A., Mulichak, A. M., Lam, J. S., Rocchetta, H. L. & Garavito, R. M. (2004). *Protein Sci.* **13**, 529–539.

Yu, Y., Russell, R. N., Thorson, J. S., Liu, L. D. & Liu, H. W. (1992). *J. Biol. Chem.* **267**, 5868–5875.