
The molecular architecture of glucose-1-phosphate uridylyltransferase

JAMES B. THODEN AND HAZEL M. HOLDEN

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

(RECEIVED October 20, 2006; FINAL REVISION November 29, 2006; ACCEPTED December 4, 2006)

Abstract

Glucose-1-phosphate uridylyltransferase, also referred to as UDP-glucose pyrophosphorylase or UGPase, catalyzes the formation of UDP-glucose from glucose-1-phosphate and UTP. Not surprisingly, given the central role of UDP-glucose in glycogen synthesis and in the production of glycolipids, glycoproteins, and proteoglycans, the enzyme is ubiquitous in nature. Interestingly, however, the prokaryotic and eukaryotic forms of the enzyme are unrelated in amino acid sequence and structure. Here we describe the cloning and structural analysis to 1.9 Å resolution of the UGPase from *Escherichia coli*. The protein is a tetramer with 222 point group symmetry. Each subunit of the tetramer is dominated by an eight-stranded mixed β -sheet. There are two additional layers of β -sheet (two and three strands) and 10 α -helices. The overall fold of the molecule is remarkably similar to that observed for glucose-1-phosphate thymidylyltransferase in complex with its product, dTDP-glucose. On the basis of this similarity, a UDP-glucose moiety has been positioned into the active site of UGPase. This protein/product model predicts that the side chains of Gln 109 and Asp 137, respectively, serve to anchor the uracil ring and the ribose of UDP-glucose to the protein. The β -phosphoryl group of the product is predicted to lie within hydrogen bonding distance to the ϵ -nitrogen of Lys 202 whereas the carboxylate group of Glu 201 is predicted to bridge the 2'- and 3'-hydroxyl groups of the glucosyl moiety. Details concerning the overall structure of UGPase and a comparison with glucose-1-phosphate thymidylyltransferase are presented.

Keywords: UDP-glucose; glucose-1-phosphate uridylyltransferase; UDP-glucose pyrophosphorylase; X-ray structure; carbohydrates

The conversion of β -D-galactose to the more metabolically useful glucose-1-phosphate is accomplished in most organisms by the action of four enzymes that constitute the Leloir pathway (Holden et al. 2003). In the first step of this pathway, β -D-galactose is epimerized to α -D-galactose by galactose mutarotase. The next step involves the ATP-dependent phosphorylation of α -D-galactose by galactokinase to yield galactose-1-phosphate. Subsequently, in the third step, the enzyme galactose-1-phosphate uridylyltransferase catalyzes the

transfer of a UMP group from UDP-glucose to galactose-1-phosphate, thereby generating glucose-1-phosphate and UDP-galactose. To complete the pathway, UDP-galactose is converted to UDP-glucose by UDP-galactose 4-epimerase. In humans, mutations in the genes that encode the galactokinase, the galactose-1-phosphate uridylyltransferase, or the epimerase can result in the diseased state referred to as galactosemia with clinical manifestations including intellectual retardation, speech disorders, liver dysfunction, and cataract formation. The most common form of galactosemia arises from defects in galactose-1-phosphate uridylyltransferase and as such it has been the subject of intensive kinetic and structural investigations for many years (Holden et al. 2003). Studies have demonstrated that the reaction mechanism of this enzyme proceeds through a covalently bound

Reprint requests to: Hazel M. Holden, 433 Babcock Drive, Department of Biochemistry, University of Wisconsin, Madison, WI 53706, USA; e-mail: Hazel_Holden@biochem.wisc.edu; fax: (608) 262-1319.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.062626007>.

intermediate via a double displacement pathway (Arabshahi et al. 1986).

Glucose-1-phosphate uridylyltransferase, the focus of this investigation and hereafter referred to as UGPase, catalyzes the production of UDP-glucose from UTP and glucose-1-phosphate. As highlighted in Scheme 1, its reaction is similar to that of galactose-1-phosphate uridylyltransferase of the Leloir pathway in that a UMP moiety is transferred to a phosphate acceptor. Unlike the galactose-1-phosphate uridylyltransferase of the Leloir pathway, however, the catalytic mechanism of UGPase proceeds through a single displacement of pyrophosphate from UTP by glucose-1-phosphate (Sheu and Frey 1978). UGPase is also capable of catalyzing the conversion of UTP and galactose-1-phosphate to UDP-galactose and pyrophosphate (Knop and Hansen 1970; Turnquist et al. 1974).

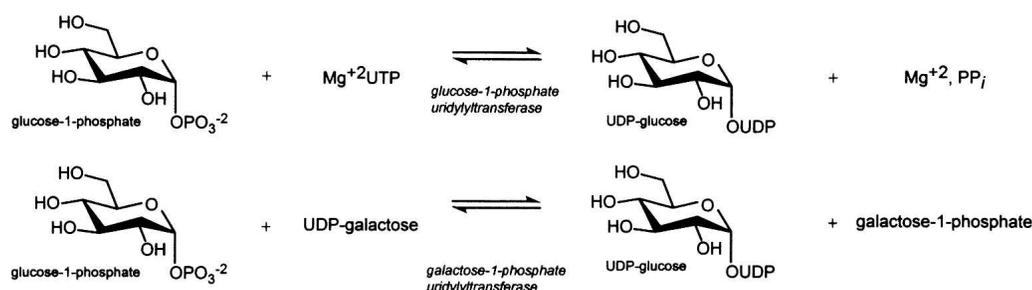
Not surprisingly, given the pivotal role of UDP-glucose in galactose utilization (Holden et al. 2003), in glycogen synthesis (Alonso et al. 1995), and in the synthesis of the carbohydrate moieties of glycolipids (Sandhoff et al. 1992), glycoproteins (Roth 1995; Verbert 1995), and proteoglycans (Silbert and Sugumaran 1995), UGPase is found in both prokaryotes and eukaryotes. Interestingly, while the prokaryotic and eukaryotic forms of UGPase catalyze the same reaction, they are completely unrelated in amino acid sequence and three-dimensional structure (Flores-Diaz et al. 1997; Mollerach et al. 1998; Mollerach and Garcia 2000). Within recent years, the bacterial UGPases have become targets for drug design given their biological roles in various Gram-negative bacteria (Genevaux et al. 1999). For example, it has been demonstrated that the enzyme is absolutely required for the production of the capsular polysaccharide, the virulence factor of *Streptococcus pneumoniae* (Bonfiglio et al. 2005).

In light of our long-standing interest in the Leloir pathway and in sugar metabolism, we recently cloned, overexpressed, and purified the UGPase from *Escherichia coli*. While crystallization reports have appeared for the UGPases from *Helicobacter pylori* and *Sphingomonas elodea* (Kim et al. 2004; Aragao et al. 2006), to date there

are no three-dimensional structures available in the public domain for either of these bacterial enzymes. Here we describe the structure of the *E. coli* UGPase determined and refined to a nominal resolution of 1.9 Å. From this study the tertiary and quaternary structures of the enzyme have been defined. As described, UGPase adopts a molecular fold similar to that of glucose-1-phosphate thymidyltransferase (Blankenfeldt et al. 2000; Barton et al. 2001; Zuccotti et al. 2001; Sivaraman et al. 2002) and UDP-*N*-acetylglucosamine pyrophosphorylase (Brown et al. 1999). These two bacterial enzymes play key roles in the synthesis of activated sugars, which ultimately serve as precursors for the synthesis of cell-surface structures. Details concerning the structure of UGPase and a comparison with the *E. coli* glucose-1-phosphate thymidyltransferase are presented.

Results and Discussion

UGPase crystallized in the space group $P2_1$ with a complete tetramer in the asymmetric unit. The model refined to an $R_{\text{factor}}/R_{\text{free}}$ of 20.2%/24.3% at 1.9 Å resolution. Shown in Figure 1 is a representative portion of the electron density for Subunit 1. Overall the electron density was well-ordered except for several residues at the N and C termini of each subunit and for the following surface loops: Glu 83–Arg 88 and Pro 233–Glu 238 in Subunit 1; Lys 84–Arg 88, Gly 110–Lys 113, and Thr 231–Gln 240 in Subunit 2; Glu 83–Arg 88 and Val 200–Pro 203 in Subunit 3; and Leu 82–Lys 87, Gly 110–Lys 113, Glu 201–Pro 209, and Thr 231–Gly 234 in Subunit 4. In each subunit, Pro 24 adopts the *cis*-conformation. Also, in each subunit, Val 37 and Asn 151 adopt dihedral angles that are outside of the allowed regions of the Ramachandran plot (near the “nucleophile elbow” region of $\varphi = \sim 60^\circ$, $\psi = \sim -100^\circ$). The electron densities for these residues are unambiguous. Val 37 is positioned in a sharp, nonclassical reverse turn found in the random coil region connecting β -strand 1 to the first α -helix of the molecule. Asn 151 is situated in a random coil area connecting β -strand 5 to α -helix 6. Other than these two



Scheme 1.

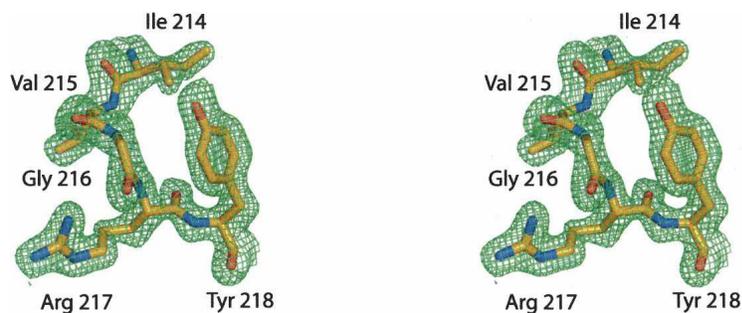


Figure 1. Representative electron density. For the most part the electron density for UGPase was well-ordered. Shown is a portion of the map for Subunit 1 that contains the β -bulge at Gly 216, which links β -strands 10 and 11. The map, contoured at 1σ , was calculated with coefficients of the form $(2F_o - F_c)$, where F_o was the native structure factor amplitude and F_c was the calculated structure factor amplitude.

amino acids, 90.6% and 9.4% of the remaining residues fall into the “most-favored” and “additionally allowed” regions of the Ramachandran plot, respectively.

A ribbon representation of the UGPase tetramer is presented in Figure 2A. The tetramer has overall dimensions of $\sim 80 \text{ \AA} \times 90 \text{ \AA} \times 90 \text{ \AA}$ and can be envisioned as a dimer of dimers. As labeled in Figure 2A, Subunits 1 and 3 and 2 and 4 form the “tight” dimers. In these tight dimers, the last two helices of the C terminus (Lys 269–Arg 282 and Gly 287–Met 298) wrap around one another (Fig. 3A). The interfaces between Subunits 1 and 2 and 3 and 4 are not as extensive (Fig. 3B). The α -carbons for the four subunits in the tetramer superimpose with a root-mean-square deviation of $\sim 0.5 \text{ \AA}$ and thus for the sake of clarity the following discussion refers only to Subunit 1 in the X-ray coordinate file.

A ribbon representation of Subunit 1 is displayed in Figure 2B. The subunit has overall dimensions of $\sim 50 \text{ \AA} \times 60 \text{ \AA} \times 45 \text{ \AA}$ and consists roughly of a single domain. There is a total 13 β -strands that form three layers of sheet. The first and largest layer is a mixed β -sheet consisting of β -strands 1, 2, 3, 4, 6, 8, 11, 12 (Lys 10–Pro 14, Thr 55–Thr 61, Thr 103–Arg 108, Val 131–Leu 135, Ser 165–Pro 171, Ser 193–Met 196, Arg 217–Leu 220, and Val 254–His 258) and flanked on either side by three α -helices. This mixed β -sheet extends across the subunit:subunit interface via β -strands 3 to form a 16-stranded sheet (Fig. 3B). As shown in Figure 2B, β -strands 5 and 13 (Val 138–Leu 140 and Ser 263–Asp 265) form a short two-stranded anti-parallel sheet that sits on one edge of the large mixed β -sheet. A third layer of anti-parallel sheet in the subunit is formed by β -strands 7, 9, and 10, (Gly 179–Asp 182, Gly 198–Glu 201, and Leu 212–Val 215). Note that β -strand 6 of the large mixed β -sheet participates in hydrogen bonding interactions with β -strand 10 of the three-stranded anti-parallel sheet (Fig. 2B). The core of the subunit is globular but is elongated as a result of the α -helices delineated by Phe 76–Glu 83 and Arg 88–Ser 96 and the final two helices

at the C terminus (Lys 269–Arg 282 and Gly 287–Met 298). These helices form the subunit:subunit interface of the “tight” dimer (Fig. 3A).

A search with DALI (Holm and Sander 1996) reveals that the closest structural relatives to UGPase are the glucose-1-phosphate thymidyltransferases (Blankenfeldt et al. 2000; Barton et al. 2001; Zuccotti et al. 2001; Sivaraman et al. 2002) and the bifunctional *N*-acetylglucosamine-1-phosphate uridylyltransferase (Brown et al. 1999). A superposition of the UGPase subunit (304 residues) onto that of the *E. coli* thymidyltransferase (293 residues) is presented in Figure 4A. These two enzymes demonstrate a 26% amino acid sequence identity and correspond with a root-mean-square deviation of 1.8 \AA for 223 structurally equivalent α -carbon positions. The two proteins begin to superimpose at Lys 9 in UGPase and Arg 4 in the thymidyltransferase and match closely until Phe 72 and Leu 68. Here, in UGPase, there is a 22-residue insertion that folds into two α -helical regions (Phe 76–Glu 83 and Arg 88–Ser 96) as indicated in Figure 4A. Following this insertion, the two proteins begin to align again at His 101 and Gly 75, for UGPase and the thymidyltransferase, respectively, and continue in close correspondence until Cys 183 and Phe 151. At this region, there is an eight-residue insertion in UGPase, relative to the thymidyltransferase. This insertion adds an additional β -strand to the mixed β -sheet in UGPase (eight versus seven strands for the thymidyltransferase). The final major difference between these two enzymes occurs at Asn 284 in UGPase and Glu 245 in the thymidyltransferase. In UGPase, there is a final α -helix, which is involved in subunit:subunit packing (Figs. 2A, 3A), whereas in the thymidyltransferase the polypeptide chain folds into three α -helices, which pack against the main body of the protein (Fig. 4A).

A close-up view of the active site for the *E. coli* thymidyltransferase in complex with its product, dTDP-glucose, is given in Figure 4B. The thymine ring is anchored to the protein via hydrogen bonds with the

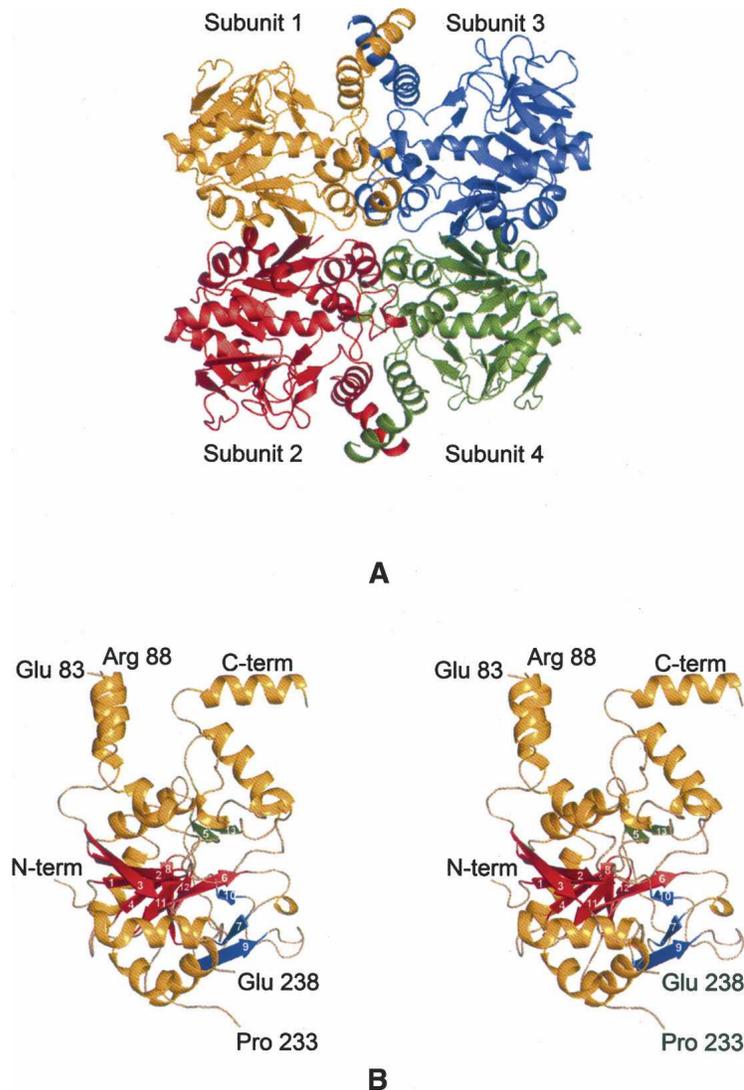


Figure 2. The molecular architecture of UGPase. (A) Ribbon representation of the UGPase tetramer. The molecule packs in the crystalline asymmetric unit with 222 symmetry and can be envisioned as a dimer of dimers. Subunits 1 and 3 and Subunits 2 and 4 form the “tight” dimer whereas Subunits 3 and 4 interact less extensively. (B) Stereo representation of Subunit 1. The subunit structure is dominated by an eight-stranded mixed β -sheet displayed in red. The two additional layers of β -sheet are colored in green and blue, respectively. There are 13 β -strands in each subunit as labeled. In Subunit 1 there are two breaks in the polypeptide chain between Glu 83 and Arg 88 and between Pro 233 and Glu 238.

backbone amide groups of Gly 11 and Gly 88 and the carboxamide group of Gln 83. The side chain of Asp 111 lies within hydrogen bonding distance to the ribose 3-hydroxyl group. Two positively charged side chains, Lys 163 and Arg 195, interact with the β -phosphoryl group of dTDP-glucose. The 2'- and 3'-hydroxyl groups of the glucosyl moiety interact with the side chain of Glu 162 whereas the 4'-hydroxyl group lies within hydrogen bonding distance to the backbone amide group of Gly 147 and the carbonyl group of Val 173.

Attempts to grow crystals of the *E. coli* UGPase in the presence of its product, UDP-glucose, or in the presence of

its substrates have not been successful. However, given the high structural homology between UGPase and the thymidylyltransferase, it was possible to build a model of UDP-glucose into the UGPase active site as presented in Figure 4C. Amino acid sequence analyses show that the thymidylyltransferases from various organisms contain the following characteristic signature sequence: G-X-G-T-R-(X₉)-K, which abuts one side of the thymine ring of the nucleotide (Fig. 4B). In UGPase, this motif begins at Gly 17 and has the following sequence: G-L-G-T-R-M-L-P-A-T-K-A-I-P-K. We postulate that the backbone amide group of Ala 16 in UGPase, which immediately precedes this

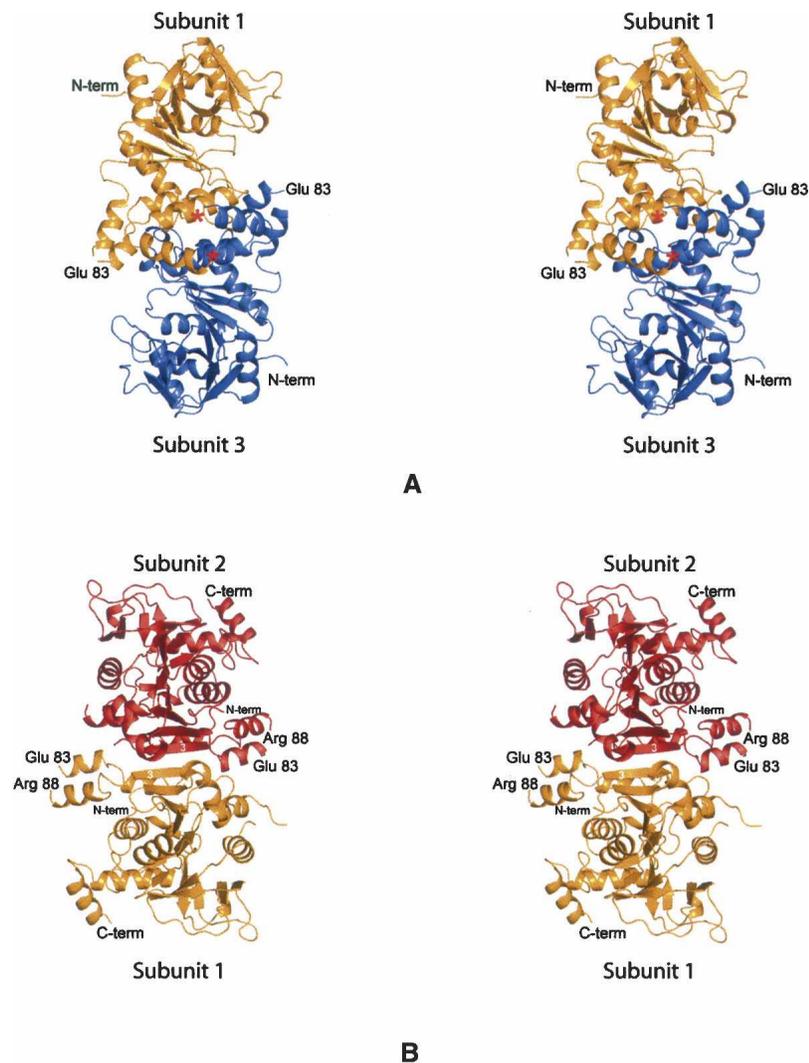


Figure 3. Close-up of the interfaces in the UGPase tetramer. (A) The “tight” dimer interface is shown. The view shown is perpendicular to the twofold rotational axis. The red asterisks indicate the positions of the C termini. (B) The interface between Subunits 1 and 2 (and by symmetry Subunits 3 and 4) is formed via β -strands 3, thereby extending the eight-stranded mixed β -sheet of one subunit into a 16-stranded β -sheet across subunit:subunit interface.

signature sequence, is involved in hydrogen bonding to the uracil ring of the product. Likewise, as in the thymidyltransferase, Gln 109 in UGPase is in the proper orientation to hydrogen bond with the uracil ring. From our model we would predict that the backbone amide group of Gly 114 hydrogen bonds with O4 of the uracil ring. In dTDP-glucose, the 2-hydroxyl group of the ribose is absent. However, by comparing the UGPase structure with that of *N*-acetylglucosamine-1-phosphate uridylyltransferase (complexed with product), we predict that the ribose 2-hydroxyl group of UDP-glucose is anchored to UGPase via the backbone amide group of Gly 17. In UGPase there is a break in the polypeptide chain between Pro 233 and Glu 238, whereas in the thymidyltransferase this region is ordered and contains Arg 95 that interacts via its guanidi-

nium group with the β -phosphoryl group of the product (Fig. 4A,B). The sequence in UGPase for this region is Gly 234, Ala 235, Gly 236, and Asp 237. Consequently, in UGPase there are no positively charged amino acids in this area that can interact with the phosphoryl group of the product (or presumably the phosphoryl group of the sugar-phosphate substrate). However, like that observed in the thymidyltransferase, Lys 202 in UGPase most likely interacts with the β -phosphoryl group of the product. From our model we predict that the 2'- and 3'-hydroxyl groups of the glucosyl moiety lie within hydrogen bonding distance to Glu 201 and that the sugar 4'-hydroxyl group is situated within hydrogen bonding distance to the backbone amide group of Gly 179 and carbonyl group of Ile 214. Note that the *cis*-proline at position 24 in UGPase is conserved in the

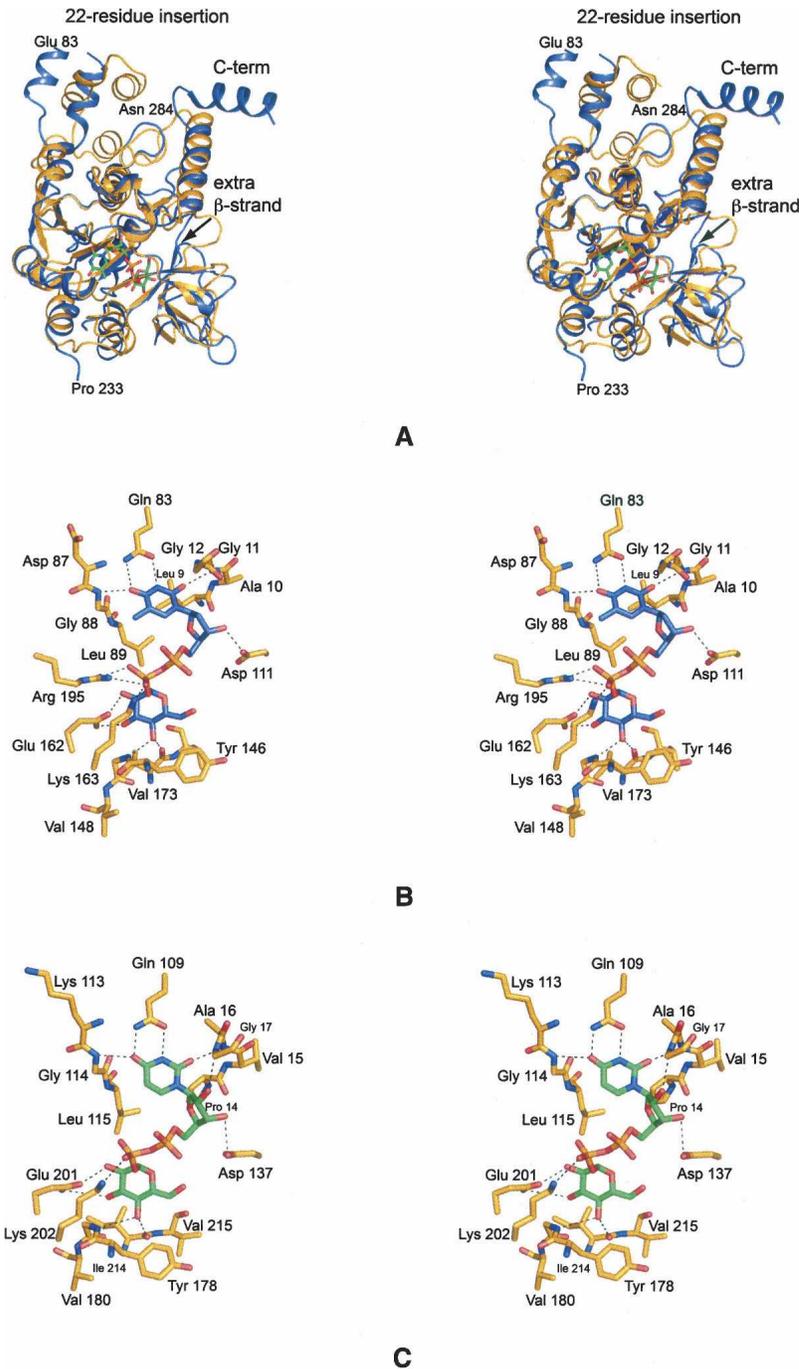


Figure 4. Comparison of UGPase with *E. coli* glucose-1-phosphate thymidyltransferase. (A) Superposition of the subunit structures for the two enzymes. UGPase is displayed in blue whereas the thymidyltransferase is depicted in gold. The bound product, dTDP-glucose in the thymidyltransferase model, is shown in a ball-and-stick representation. (B) Close-up view of the active site of the thymidyltransferase with bound product. Possible hydrogen bonds are indicated by the dashed lines. X-ray coordinates used for this figure were obtained from the Protein Data Bank (accession code 1H5T) and are based on Zuccotti et al. (2001). (C) Model of UDP-glucose built into the active site of UGPase. The dashed lines represent possible hydrogen bonding interactions.

thymidyltransferase as Pro 19. This residue is ~ 15 Å from the active site.

It should be noted that, in the thymidyltransferase/product complex models presently available, there are no

bound magnesium ions observed in the active sites (Blankenfeldt et al. 2000; Zuccotti et al. 2001). Contrastingly, in the structure of the thymidyltransferase solved in the presence of dTTP, there is a single magnesium ion

bound to the nucleotide and the carboxylate groups of Asp 108 and Asp 223 (Sivaraman et al. 2002). These residues are structurally conserved in UGPase as Asp 137 and Asp 265. Interestingly, in glucose-1-phosphate cytidylyltransferase from *Salmonella typhi*, which catalyzes a similar reaction as UGPase, one magnesium ion is observed octahedrally coordinated by side chain aspartates, waters, and phosphoryl oxygens when product is bound in the active site whereas two are observed when only CTP is present (Koropatkin and Holden 2004; Koropatkin et al. 2005). The manner in which UGPase accommodates metal ions is presently not clear. Structural and biochemical analyses to address its metal requirements are in progress.

In summary, the molecular scaffold for a bacterial UGPase has now been clearly defined. The tetrameric enzyme can be envisioned as a dimer of dimers. The active site of UGPase contains many of the amino acid residues observed in the glucose-1-phosphate thymidylyltransferases, and presumably these residues function in the same manner in UGPase. In all of the glucose-1-phosphate thymidylyltransferases studied thus far, residues within a single subunit form the active sites. This is in sharp contrast to that observed for glucose-1-phosphate cytidylyltransferase from *S. typhi*, however. While this enzyme catalyzes a similar reaction, the transfer of a CMP group from CTP to glucose-1-phosphate, its active site is formed from residues contributed by two subunits of the hexamer (Koropatkin and Holden 2004; Koropatkin et al. 2005). The question then arises as to whether the active site of UGPase is formed by residues from only one subunit or by two. As shown in Figure 5, the disordered region between Glu 83 and Arg 88 in one subunit lies within ~ 14 Å of the active site in the second subunit of the tight dimer. Recall that this is where the 22-residue insertion is located in UGPase relative to glucose-1-phosphate thymidylyltransferase. Interestingly, this disordered region is positively charged and contains Lys 84,

Arg 85, Val 86, and Lys 87. It is possible that, in the presence of its substrates or its product, this surface loop in UGPase closes down upon a neighboring subunit to form part of the active site.

Materials and methods

Molecular cloning of the UGPase gene

Genomic DNA from *E. coli* w3110 (ATCC 39936) was isolated by standard procedures. The *galU* gene was PCR-amplified using primers that introduced a 5' NdeI site and a 3' XhoI site. The purified PCR product was A-tailed and ligated into the pGEM-T (Promega) vector for screening and sequencing. A *galU*-pGEM-T vector construct of the correct sequence was then appropriately digested and ligated into a pET28b(+) vector that had been modified to provide an N-terminal TEV protease recognition site (Thoden et al. 2005).

DH5- α *E. coli* cells were transformed with the resulting plasmid and plated onto LB Agar Kan plates. Multiple colonies were picked to grow overnight in LB Kan medium from which plasmid DNA was isolated. Plasmids were tested for the presence of the *galU* gene by digestion with NdeI and XhoI.

Protein expression and purification

The *galU*-pET28JT plasmid was used to transform HMS174 (DE3) *E. coli* cells (Novagen). The culture was grown at 37°C with shaking until an optical density of 0.7 was reached at 600 nm, and then induced with 1 mM IPTG. The cells were allowed to express at 37°C for six hours after induction with IPTG.

Following growth, the cells were harvested and disrupted by sonication on ice. The lysate was cleared by centrifugation and UGPase was purified utilizing a Ni-NTA resin (Qiagen) according to the manufacturer's instructions. The N-terminal His-tag was cleaved from the protein by the addition of both 1 mM DTT and TEV protease in a 1:50 (TEV protease:UGPase) molar ratio. The cleavage reaction was carried out at 30°C for 5 h and subsequently passed through the nickel column to remove the TEV protease. The UGPase sample was dialyzed against 10 mM Tris-HCl and 200 mM NaCl at pH 8.0. Following dialysis the sample was concentrated to 15 mg/mL based on an extinction coefficient of 1.57 (mg/mL) $^{-1}$ ·cm $^{-1}$ as calculated with Protean (DNASar).

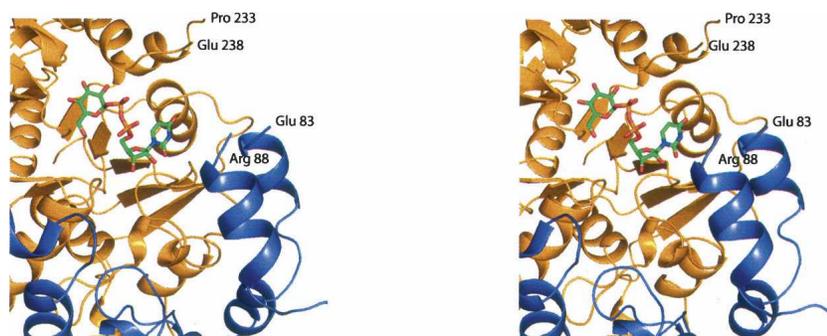


Figure 5. Close-up view of the UGPase active site. Subunits 1 and 3 are displayed in gold and blue, respectively. Strikingly, the disordered surface loop between Glu 83 and Arg 88 contains many positively charged side chains. It is tempting to speculate that this region closes down upon the active site when it is loaded with substrates or product.

Table 1. X-ray data collection statistics

Data set	Native data set no. 1	Methylmercury acetate	Native data set no. 2
Resolution limits (Å)	30–2.4 (2.51–2.4) ^a	30–2.4 (2.51–2.4)	30–1.9 (1.95–1.9)
No. independent reflections	50,498 (4,614)	45,576 (4,271)	106,815 (7,817)
Completeness (%)	90.0 (65.4)	81.2 (60.5)	98.5 (96.5)
Redundancy	1.9 (1.2)	1.3 (1.1)	6.9 (3.3)
Avg I/Avg σ (I)	10.7 (1.8)	7.5 (1.4)	13.7 (3.2)
R_{sym}^b (%)	10.5 (29.3)	8.2 (26.8)	7.6 (37.2)

^aStatistics for the highest resolution bin.

^b $R_{\text{sym}} = (\sum |I - \bar{I}| / \sum I) \times 100$.

Assessment of UGPase activity

Enzymatic activity was determined via HPLC chromatography. For each reaction, 0.1 mg of enzyme was added to a 1-mL solution containing 2 mM UTP, 4 mM MgCl₂, 2 mM sugar-1-phosphate and buffered with 25 mM HEPPS (pH 8.0). The reaction was allowed to proceed for 30 min at 25°C, after which time the enzyme was removed via filtration. The reaction mixture was loaded onto a 1-mL ResourceQ column equilibrated in 10 mM ammonium bicarbonate (pH 8.0) and the reaction products separated by a linear gradient of ammonium carbonate to a final concentration of 750 mM. Retention times for known samples of UTP, UDP, UMP, UDP-glucose, and UDP-galactose were determined to allow for comparisons with the reaction products. The *E. coli* UGPase used in this investigation was active against the following sugar 1-phosphates: glucose, galactose, xylose, mannose, glucosamine, *N*-acetylglucosamine, fucose, and galactosamine, with the latter two requiring extended reaction times of 6 to 12 h. In addition, the enzyme was able to use dTTP as well as its natural substrate UTP.

Crystallization of UGPase

Crystallization conditions were initially surveyed by the hanging drop method of vapor diffusion with a sparse matrix screen developed in the laboratory. Large crystals were subsequently grown via hanging drop against 1.8–2.0 M ammonium sulfate and 200 mM LiCl buffered at pH 6.0 with 100 mM MES. The crystals grew to maximum dimensions of 0.7 mm × 0.5 mm × 0.3 mm in ~1 wk. They belonged to the space group $P2_1$ with unit cell dimensions of $a = 65.3$ Å, $b = 110.2$ Å, $c = 101.6$ Å, and $\beta = 94.5^\circ$ and contained one tetramer in the asymmetric unit.

Structural analysis of UGPase

X-ray data from both native protein crystals and heavy-atom soaked crystals were initially measured at 4°C with a Bruker HSTAR area detector system and processed with XDS (Kabsch 1993). These data were internally scaled with XSCALIBRE (Rayment and Wesenberg, unpubl.). Prior to X-ray data collection, the crystals were stabilized by transfer into a 2.0 M ammonium sulfate solution containing 100 mM MES (pH 6.0) and 200 mM LiCl. The X-ray source was CuK α radiation from a Rigaku RU200 X-ray generator operated at 50 kV and 90 mA and equipped with Supper mirrors.

The high-resolution X-ray data set subsequently used to refine the protein model was collected at 100 K using a Bruker AXS Platinum 135 CCD detector controlled with the Proteum software

suite (Bruker AXS Inc.). In this case, the X-ray source was CuK α radiation from a Rigaku RU200 X-ray generator equipped with montel optics and operated at 50 kV and 90 mA. These X-ray data were processed with SAINT version V7.06A (Bruker AXS Inc.) and internally scaled with SADABS version 2005/1 (Bruker AXS Inc.). Crystals were prepared for low-temperature X-ray data collection by transfer into a 2.1 M ammonium sulfate solution containing 100 mM MES (pH 6.0), 500 mM LiCl, and 20% ethylene glycol. Relevant X-ray data collection statistics are presented in Table 1.

The structure was solved with one heavy atom derivative using 1 mM methylmercury acetate and a soak time of 24 h. Sixteen mercury binding sites were identified with the program SOLVE (Terwilliger and Berendzen 1999), giving an overall figure-of-merit of 0.22 to 2.5 Å resolution. Matrices relating four clusters of mercury sites were determined with RESOLVE (Terwilliger 2000), and subsequent fourfold averaging and solvent flattening with RESOLVE generated an interpretable electron density map. Once a preliminary model was constructed, it was utilized as a search model for molecular replacement with the software PHASER (Storoni et al. 2004). Subsequent least-squares refinement with TNT (Tronrud et al. 1987) reduced the R_{working} to 19.8% for all measured X-ray data from 30 to 1.9 Å resolution. Relevant refinement statistics are given in Table 2.

Table 2. Least-squares refinement statistics

Resolution limits (Å)	30–1.9
R -factor (overall) %/no. reflections ^a	20.2/106,815
R -factor (working) %/no. reflections	19.8/96,192
R -factor (free) %/no. reflections	24.3/10,623
No. protein atoms	8656 ^b
Waters	786
	Average B values (Å ²)
Protein atoms	31.0
Solvents	37.8
Weighted RMS deviations from ideality	
Bond lengths (Å)	0.014
Bond angles (deg)	2.5
Trigonal planes (Å)	0.007
General planes (Å)	0.016
Torsional angles (deg) ^c	18.3

^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.

^bThese include multiple conformations for His 101 and Val 107 in Subunit 1, and His 62, Val 107, and Arg 158 in Subunit 2.

^cThe torsional angles were not restrained during the refinement.

Acknowledgments

This research was supported by a grant from the NIH (DK47814 to H.M.H.). X-ray coordinates have been deposited in the Protein Data Bank (2E3D). We thank Prof. Ivan Rayment and Mr. Paul Cook for helpful discussions.

References

- Alonso, M.D., Lomako, J., Lomako, W.M., and Whelan, W.J. 1995. A new look at the biogenesis of glycogen. *FASEB J.* **9**: 1126–1137.
- Arabshahi, A., Brody, R.S., Smallwood, A., Tsai, T.C., and Frey, P.A. 1986. Galactose-1-phosphate uridylyltransferase. Purification of the enzyme and stereochemical course of each step of the double-displacement mechanism. *Biochemistry* **25**: 5583–5589.
- Aragao, D., Marques, A.R., Frazao, C., Enguita, F.J., Carrondo, M.A., Fialho, A.M., Sa-Correia, I., and Mitchell, E.P. 2006. Cloning, expression, purification, crystallization and preliminary structure determination of glucose-1-phosphate uridylyltransferase (UgpG) from *Sphingomonas elodea* ATCC 31461 bound to glucose-1-phosphate. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **62**: 930–934.
- Barton, W.A., Lesniak, J., Biggins, J.B., Jeffrey, P.D., Jiang, J., Rajashankar, K.R., Thorson, J.S., and Nikolov, D.B. 2001. Structure, mechanism and engineering of a nucleotidyltransferase as a first step toward glycorandomization. *Nat. Struct. Biol.* **8**: 545–551.
- Blankenfeldt, W., Asuncion, M., Lam, J.S., and Naismith, J.H. 2000. The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidyltransferase (RmlA). *EMBO J.* **19**: 6652–6663.
- Bonfiglio, L., Garcia, E., and Mollerach, M. 2005. Biochemical characterization of the pneumococcal glucose 1-phosphate uridylyltransferase (GalU) essential for capsule biosynthesis. *Curr. Microbiol.* **51**: 217–221.
- Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreux, D., Cambillau, C., and Bourne, Y. 1999. Crystal structure of the bifunctional *N*-acetylglucosamine 1-phosphate uridylyltransferase from *Escherichia coli*: A paradigm for the related pyrophosphorylase superfamily. *EMBO J.* **18**: 4096–4107.
- Flores-Diaz, M., Alape-Giron, A., Persson, B., Pollesello, P., Moos, M., von Eichel-Streiber, C., Thelestam, M., and Florin, I. 1997. Cellular UDP-glucose deficiency caused by a single point mutation in the UDP-glucose pyrophosphorylase gene. *J. Biol. Chem.* **272**: 23784–23791.
- Genevaux, P., Bauda, P., DuBow, M.S., and Oudega, B. 1999. Identification of Tn10 insertions in the *rfaG*, *rfaP*, and *galU* genes involved in lipopolysaccharide core biosynthesis that affect *Escherichia coli* adhesion. *Arch. Microbiol.* **172**: 1–8.
- Holden, H.M., Rayment, I., and Thoden, J.B. 2003. Structure and function of enzymes of the Leloir pathway for galactose metabolism. *J. Biol. Chem.* **278**: 43885–43888.
- Holm, L. and Sander, C. 1996. Mapping the protein universe. *Science* **273**: 595–603.
- Kabsch, W. 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**: 795–800.
- Kim, H., Wu, C.A., Kim, D.Y., Han, Y.H., Ha, S.C., Kim, C.S., Suh, S.W., and Kim, K.K. 2004. Crystallization and preliminary X-ray crystallographic study of UDP-glucose pyrophosphorylase (UGPase) from *Helicobacter pylori*. *Acta Crystallogr. D Biol. Crystallogr.* **60**: 1447–1449.
- Knop, J.K. and Hansen, R.G. 1970. Uridine diphosphate glucose pyrophosphorylase. IV. Crystallization and properties of the enzyme from human liver. *J. Biol. Chem.* **245**: 2499–2504.
- Koropatkin, N.M. and Holden, H.M. 2004. Molecular structure of alpha-D-glucose-1-phosphate cytidylyltransferase from *Salmonella typhi*. *J. Biol. Chem.* **279**: 44023–44029.
- Koropatkin, N.M., Cleland, W.W., and Holden, H.M. 2005. Kinetic and structural analysis of alpha-D-glucose-1-phosphate cytidylyltransferase from *Salmonella typhi*. *J. Biol. Chem.* **280**: 10774–10780.
- Mollerach, M. and Garcia, E. 2000. The galU gene of *Streptococcus pneumoniae* that codes for a UDP-glucose pyrophosphorylase is highly polymorphic and suitable for molecular typing and phylogenetic studies. *Gene* **260**: 77–86.
- Mollerach, M., Lopez, R., and Garcia, E. 1998. Characterization of the galU gene of *Streptococcus pneumoniae* encoding a uridine diphosphoglucose pyrophosphorylase: A gene essential for capsular polysaccharide biosynthesis. *J. Exp. Med.* **188**: 2047–2056.
- Roth, J. 1995. Biosynthesis 4c. *New Comprehensive Biochemistry. Glycoproteins* (eds. J. Montreuil et al.), pp. 287–312. Elsevier, Amsterdam.
- Sandhoff, K., van Echten, G., Schroder, M., Schnabel, D., and Suzuki, K. 1992. Metabolism of glycolipids: The role of glycolipid-binding proteins in the function and pathobiochemistry of lysosomes. *Biochem. Soc. Trans.* **20**: 695–699.
- Sheu, K.F. and Frey, P.A. 1978. UDP-glucose pyrophosphorylase. Stereochemical course of the reaction of glucose 1-phosphate with uridine-5′[1-thiotriphosphate]. *J. Biol. Chem.* **253**: 3378–3380.
- Silbert, J.E. and Sugumaran, G. 1995. Intracellular membranes in the synthesis, transport, and metabolism of proteoglycans. *Biochim. Biophys. Acta* **1241**: 371–384.
- Sivaraman, J., Sauve, V., Matte, A., and Cygler, M. 2002. Crystal structure of *Escherichia coli* glucose-1-phosphate thymidyltransferase (RffH) complexed with dTTP and Mg²⁺. *J. Biol. Chem.* **277**: 44214–44219.
- Storoni, L.C., McCoy, A.J., and Read, R.J. 2004. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D Biol. Crystallogr.* **D60**: 432–438.
- Terwilliger, T.C. 2000. Maximum-likelihood density modification. *Acta Crystallogr. D Biol. Crystallogr.* **56**: 965–972.
- Terwilliger, T.C. and Berendzen, J. 1999. Automated MAD and MIR structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **55**: 849–861.
- 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.
- Tronrud, D.E., Ten Eyck, L.F., and Matthews, B.W. 1987. An efficient general-purpose least-squares refinement program for macromolecular structures. *Acta Crystallogr. A* **43**: 489–501.
- Turnquist, R.L., Gillett, T.A., and Hansen, R.G. 1974. Uridine diphosphate glucose pyrophosphorylase. Crystallization and properties of the enzyme from rabbit liver and species comparisons. *J. Biol. Chem.* **249**: 7695–7700.
- Verbert, A. 1995. Biosynthesis 4c. From Glc₃ Man₉GlcNAc₂-protein to Man₅GlcNAc₂-protein: transfer “en bloc” and processing. *New Comprehensive Biochemistry. Glycoproteins* (eds. J. Montreuil et al.), pp. 145–152. Elsevier, Amsterdam.
- Zuccotti, S., Zanardi, D., Rosano, C., Sturla, L., Tonetti, M., and Bolognesi, M. 2001. Kinetic and crystallographic analyses support a sequential-ordered catalytic mechanism for *Escherichia coli* glucose-1-phosphate thymidyltransferase. *J. Mol. Biol.* **313**: 831–843.