Three-Dimensional Structure of Galactose-1-phosphate Uridylyltransferase from *Escherichia coli* at 1.8 Å Resolution†,‡

Joseph E. Wedekind, Perry A. Frey,* and Ivan Rayment*

Institute for Enzyme Research, Graduate School, and Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin−Madison, Madison, Wisconsin 53705

Received May 16, 1995; Revised Manuscript Received June 19, 1995⊗

ABSTRACT: Galactose-1-phosphate uridylyltransferase catalyzes the reversible transfer of the uridine 5′-monophosphoryl moiety of UDP-glucose to the phosphate group of galactose 1-phosphate to form UDP-galactose. This enzyme participates in the Leloir pathway of galactose metabolism, and its absence is the primary cause of the potentially lethal disease galactosemia. The three-dimensional structure of the dimeric enzyme from *Escherichia coli* complexed with uridine 5′-diphosphate is reported here. The structure was solved by multiple isomorphous replacement and electron density modification techniques and has been refined to 1.8 Å resolution. Enzyme subunits consist of a single domain with the topology of a "half-barrel". The barrel staves are formed by nine strands of antiparallel β-sheet. The barrel axis is approximately parallel to the local dyad that relates each subunit. Two amphipathic helices fill the half-barrel sequestering its hydrophobic interior. An iron atom resides on the outside of the barrel, centered in the subunit interface. Intrasubunit coordination to iron resembles a distorted square pyramid formed by the equatorial ligation of two histidines and a bidentate carboxylate group and a single axial histidine. The subunit interface is stabilized by this coordination and is further characterized by the formation of two intermolecular "mini-sheets" distinct from the strands of the half-barrel. Loops that connect the mini-sheet strands contribute to the formation of the active site, which resides on the external surface of the barrel rim. Loops of the barrel strands are tethered together by a structural zinc atom that orients the local fold in a manner essential for catalysis. In one of the latter loops, Sγ of a cysteine is modified by β-mercaptoethanol, which prevents the α-phosphorus of the nucleotide from access to the nucleophile His^{166}. This conformation does not appear to perturb the interactions to the uracil and ribose moieties as mediated through the side chains of Leu^{54}, Phe^{75}, Asn^{77}, Asp^{78}, Phe^{79}, and Val^{108}. Several of the latter residues have been implicated in human galactosemia. The present structure explains the deleterious effects of many of those mutations.

Galactose-1-phosphate uridylyltransferase (hexose-1-phosphate uridylyltransferase, EC 2.7.7.12) catalyzes the nucleotide exchange between UDP-hexoses and hexose 1-phosphates. This provides an essential balance of UDP-glc[1] and UDP-gal for the cell. Such activated hexose sugars are consumed in the synthesis of disaccharides, glycoproteins, glycolipids, and glycogen (Leloir, 1971). These sugars are vital in the synthesis of cell walls in bacteria, fungi, and plants (Delmer, 1983, 1987). Attenuation or deficiency of uridylyltransferase activity in mammals results in the potentially lethal disease galactosemia (Kalckar, 1960). This condition occurs as an autosomal recessive trait afflicting approximately 1 in every 40 000−60 000 live-born infants (Tedesco, 1972; Levy & Hammerson, 1978; Fridovich-Keil & Jinks-Robertson, 1993). Numerous clinical studies have identified mutations and polymorphisms common among human galactosemics (Elsas et al., 1994). However, the absence of structural information has made these studies difficult to interpret.

Kinetic analysis of this enzyme from various sources has demonstrated that the reaction displays ping-pong kinetics and a double displacement mechanism (Wong & Frey, 1974a,b; Wu et al., 1974, Markus et al., 1977, Hester & Raushel, 1987). Catalysis proceeds through the formation of a covalent enzyme−substrate intermediate according to eqs 1 and 2 in which the nucleophilic histidine at residue

$$\text{UDP-glc} + \text{E}-\text{His}^{166} \rightleftarrows \text{Glc-1-P} + \text{E}-\text{His}^{166}-\text{UMP} \quad (1)$$

$$\text{Gal-1-P} + \text{E}-\text{His}^{166}-\text{UMP} \rightleftarrows \text{UDP-gal} + \text{E}-\text{His}^{166} \quad (2)$$

166 in the enzyme from *Escherichia coli* is transiently nucleotidylated (Yang & Frey, 1979; Field et al., 1989; Kim et al., 1990).

The amino acid sequence of the uridylyltransferase is conserved. The enzymes from *Saccharomyces cerevisiae*

† Supported by Grant AR35186 from the National Institute of Arthritis, Musculoskeletal and Skin Diseases (I.R.), by Grant GM30480 from the National Institute of General Medical Sciences (P.A.F.), and by an NRSA fellowship (J.E.W.) from NIH Training Grant GM08293.

‡ The X-ray coordinates have been deposited in the Brookhaven Protein Data Bank (file name 1HXP).

* Address correspondence to these authors at the Institute for Enzyme Research, University of Wisconsin, 1710 University Ave., Madison, WI 53705 [phone, 608-262-0437 (I.R.) and 608-262-0055 (P.A.F.); Fax, 608-265-2904].

⊗ Abstract published in *Advance ACS Abstracts,* August 1, 1995.

[1] Abbreviations: UDP-glc, uridine 5′-diphosphate glucose; UDP-gal, uridine 5′-diphosphate galactose; UMP, uridine 5′-monophosphate; UDP, uridine 5′-diphosphate; glc-1-P, glucose 1-phosphate; gal-1-P, galactose 1-phosphate; phenyl-UDP, P^1-5′-uridine-P^2-phenyl diphosphate; PEG, poly(ethylene glycol); HEPES, *N*-2-(hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; rms, root mean square; MIR, multiple isomorphous replacement; NDP kinase, nucleoside diphosphate kinase.

0006-2960/95/0434-11049$09.00/0

and humans share approximately 39% and 46% sequence identity with the enzyme from *E. coli*; these enzymes exist as α_2 dimers (Saito et al., 1967; Segawa & Fukasawa, 1979; Dale & Popjak, 1976). The dimeric enzyme from *E. coli* has a molecular mass of 80 kDa, with 348 amino acids per subunit (Lemaire & Mueller-Hill, 1986; Cornwell et al., 1987). Recently, the uridylyltransferase from *E. coli* has been found to contain approximately 0.67 ± 0.14 mol of iron and 1.21 ± 0.09 mol of zinc per mole of enzyme subunit (Ruzicka et al., 1995). From the latter study it appears that at least 1 equiv of divalent metal is required to support catalysis. Cations such as Co^{2+}, Cd^{2+}, and Mn^{2+} can be substituted for iron and zinc to provide partial restoration of enzymatic activity. Fully active enzymes have been prepared that contain either 2 equiv of iron or 2 equiv of zinc per subunit (Ruzicka et al., 1995). The variety of metal ions that can be accommodated by the enzyme suggests these cations participate in a noncatalytic role, although this has not been proven.

In order to understand the catalytic mechanism and the metal ion requirements of gal-1-P uridylyltransferase, a structural study of the nucleotide-bound enzyme from *E. coli* was initiated. The resulting crystal structure at 1.8 Å resolution reveals a novel single domain β-sheet architecture fortified by the tight association of two metal ions. These metals appear to stabilize the structure; zinc contributes to the formation of the active site fold, whereas iron secures the subunit interface. In addition, the three-dimensional structure has provided a means to interpret the clinical work obtained through studies of human galactosemic patients. These studies, in conjunction with existing kinetic and chemical information, delineate the fundamental structural requirements necessary for catalysis by this enzyme.

EXPERIMENTAL PROCEDURES

Chemicals. PEG 10 000 and mercaptosuccinic acid were obtained from Aldrich Chemical Co. (Milwaukee, WI). CMDNP [2-(chloromercuri)-4,6-dinitrophenol] was purchased from the Alfred Bader Library of Rare Chemicals (a division of Aldrich Chemical Co., Milwaukee, WI). Ammonium bromoosmate [$(NH_4)_2OsBr_6$] was purchased from Johnson Matthey (a division of Alpha Products, Danvers, MA). Uranyl acetate [$UO_2(C_2H_3O_2)_2$] was purchased from Ted Pella Inc. (Redding, CA). Sources for other supplies are described in Wedekind et al. (1994).

Crystallization and Heavy Atom Derivatives. Protein purification and crystallization were by the method of Wedekind et al. (1994). Crystals were grown from approximately 14.5% (w/v) PEG 10 000 at pH 5.9 and belonged to the space group $P2_12_12$ with unit cell dimensions a = 58.6 Å, b = 217.2 Å, and c = 69.6 Å. One dimer was found to be present in the asymmetric unit. In order to prepare heavy atom derivatives, heavy atom compounds were dissolved into 0.5 mL reservoirs of synthetic mother liquor and allowed to sit at 4.0 °C for 24 h prior to crystal transfer. The synthetic mother liquor consisted of 20% (w/v) PEG 10 000, 0.40 M Li_2SO_4, 0.20 M NaCl, 0.10 M sodium succinate at pH 6.0, 0.001 M NaN_3, and 4 mM phenyl-UDP. Precession photographs (10°) of the *hk*0 and 0*kl* zones of crystals soaked in solutions of heavy atoms were compared to the respective photographs of native crystals for the purpose of screening. Suitable intensity changes prompted X-ray data collection.

X-ray Data Collection. Crystals were mounted in thin-walled quartz capillary tubes. Crystal slippage was prevented by coating the crystals with a thin film of Formvar (Rayment et al., 1977). Precession photographs were recorded at a crystal-to-film distance of 10 cm for 26 h at 4 °C. Ni-filtered Cu Kα radiation was from a Rigaku RU-200 rotating anode X-ray generator with a 200 μm focal spot, operated at 50 kV × 50 mA. X-ray intensities were collected at 4 °C on a Siemens X-1000D multiwire area detector at a crystal-to-detector distance of 24 cm. Complete X-ray data sets required between four and six crystals. Friedel pairs were recorded for the uranyl acetate derivative. Data were processed with the XDS data reduction software of Kabsch (1988a,b) and scaled by the program XSCALIBRE written by G. Wesenberg and I. Rayment (University of Wisconsin). Relevant intensity statistics are provided in Table 1a.

Synchrotron Radiation. Native X-ray intensities were recorded to 1.8 Å resolution at beam line 7-1 of the Stanford Synchrotron Radiation Laboratory (Palo Alto, CA). Prior to data collection, crystals were soaked for 1 day in the synthetic mother liquor described. X-ray data were collected at 4 °C. The 15 cm radius of the MAR image plate was utilized at a crystal-to-detector distance of 21.5 cm. Exposure times were approximately 12−17 s per 1° rotation. Three or four rotations were possible per crystal volume before radiation damage compromised the quality of the data. Owing to their length (approximately 1.0 mm), crystals were translated along the spindle axis c^*, such that three or four volumes per crystal could be exposed. One degree of overlap was allowed between successive crystals. This provided 151 rotations from 17 capillary mounted crystals. Data were processed by use of the DENZO data reduction software (Otwinowski, 1993) and scaled by the versions of ROTAVATA and AGROVATA available through the program suite assembled by the Collaborative Computational Project Number 4 (CCP4, 1979). Intensity statistics are provided in Table 1b.

Structural Determination. The positions of the heavy atom sites for three derivatives were located by inspection of difference Patterson maps calculated with all X-ray data between 30 and 5.0 Å resolution. These sites were placed on a common origin by difference Fourier maps. The positions and occupancies of the heavy atoms were refined by the origin-removed-Patterson-correlation method of Rossmann (1960) and Terwilliger and Eisenberg (1983). Refined heavy atom parameters are provided in Table 2. MIR phases were calculated by use of the program HEAVY (Terwilliger & Eisenberg, 1983). Relevant phase calculation statistics are provided in Table 3. The correct hand of the heavy atom constellation was determined from an anomalous difference Fourier map calculated for the uranyl acetate derivative using all X-ray data to 3.0 Å resolution.

The high- and low-resolution X-ray data sets for the mercury and osmium derivatives were collected at different times spanning several months. Due to crystal sensitivity to the heavy atom compounds and radiation damage, the resulting high- and low-resolution data sets did not scale internally in a satisfactory manner. Therefore, the high- (2.5−4.78 Å) and low- (4.78−100 Å) resolution data sets were treated as separate derivatives (Tables 1a and 2).

Electron Density Modification. The initial electron density map calculated with MIR phases to 5.0 Å resolution was stacked on Plexiglas sheets in order to define the contents

Table 1: Intensity Statistics for Native and Derivative X-ray Data Sets

(a) Native and Derivative X-ray Data Sets

	total[c]	shell (Å) 100−4.78	−3.80	−3.32	−3.01	−2.80	−2.63	−2.50
native l = 1.54 Å								
observations	67184	17377	17668	12956	7225	6401	4855	702
independent reflections	24152	4594	4509	4186	3895	3643	2860	465
completeness (%)	76	96	100	93	85	84	62	11
intensity (av)	6624	7723	10000	6451	3586	2359	1746	1559
σ (av)	629	499	729	699	618	599	599	603
R-factor(%)[a]	5.26	4.2	4.7	6.1	7.8	10.5	13.0	14.1
uranyl acetate								
observations	121374	33451	32468	25763	13582	9056	5264	1790
independent reflections	23290	4643	4549	4263	3673	3089	2071	1002
intensity (av)	6940	8439	10000	6105	3509	2213	1561	1177
σ (av)	473	409	554	504	455	427	394	369
R-factor(%)[a]	5.85	5.4	5.2	6.4	8.0	10.1	12.4	13.9
CMDNP[b]								
observations	48670	10020	11005	6959	6286	5637	4984	3779
independent reflections	29672	4241	6522	4252	4104	3935	3679	2939
intensity (av)	5424	10000	10000	6698	3820	2539	1843	1450
σ (av)	503	741	550	532	481	465	462	464
R-factor (%)[a]	5.61	6.6	4.7	5.4	6.7	8.8	10.7	13.2
bromoosmate[b]								
observations	38992	6903	16356	6915	5702	3116	0	0
independent reflections	22866	4503	8392	4092	3677	2202	0	0
intensity (av)	7125	10000	10000	5906	3182	1952	0	0
σ (av)	826	598	790	887	847	837	0	0
R-factor (%)[a]	7.11	3.10	6.5	7.2	10.8	16.1	0	0

(b) Native Data X-ray Data Set

	total	shell (Å) 50−4.74	−3.36	−2.75	−2.38	−2.13	−1.94	−1.80
native l = 1.08 Å								
observations	169713	10566	19214	23501	25828	28076	30990	31548
independent reflections	68553	4376	8024	10011	11131	11455	11729	11827
completeness (%)	82	90	96	94	89	81	75	70
intensity (av)d	17544	48648	60688	23765	11174	7453	4403	2949
σ (av)	1269	2314	2949	850	426	332	315	788
R-factor (%)[a]	2.2	1.8	1.7	1.8	2.2	2.8	4.5	9.6

[a] $R\text{-factor} = \Sigma|I_j - \bar{I}|/\Sigma|I_j| \times 100$. [b] Low- (100−4.78 Å) and high- (4.78−2.5 Å) resolution data sets were scaled separately. [c] The overall statistics for the number of observations and independent reflections combine both the high- and low-resolution data for CMDNP and bromoosmate, whereas the overall average intensity, σ, and merging *R*-factor are for the high-resolution data only. [d] Intensities are on a different scale compared to (a).

Table 2: Refined Heavy Atom Parameters

derivative	soak time (h), concn (mM)	site no.	relative occupancy	x[a]	y	z	B (Å^2)[b]	R_{iso}[c]	location
uranyl acetate	39, 2.5	1	4.2	0.363	0.494	0.791	20.0	13.1	Glu 215 and Asp 340 subunit II
		2	2.1	0.129	0.357	0.788	20.0		Asp 49 subunit II
		3	2.0	0.060	0.618	0.660	20.0		Glu 188 subunit II
		4	2.6	0.743	0.123	0.753	20.0		Asp 49 subunit I
CMDNP									
low[d]	12−14, 0.5	1	6.3	0.185	0.822	0.541	20.0	23.6	Cys 272 Subunit I
high		1	8.2	0.191	0.815	0.545	15.0	28.9	Cys 99 subunit II
low		2	3.8	0.230	0.727	0.691	20.0		
high		2	4.7	0.247	0.712	0.694	15.0		
NH_4 bromoosmate									
low	36, 5	1	5.5	0.060	0.872	0.677	20.0	15.9	His 166 subunit I
high		1	6.5	0.118	0.862	0.691	20.0	15.5	His 166 subunit II
low		2	4.6	0.643	0.638	0.766	20.0		
high		2	6.5	0.621	0.659	0.758	20.0		

[a] Fractional coordinates (x, y, z). [b] Temperature factors were not refined. [c] $R_{iso} = \Sigma(|F_N| - |F_H|)/\Sigma|F_N| \times 100$. [d] A 1 mM stock solution of CMDNP was diluted to 0.5 mM by addition of synthetic mother liquor made 2 mM in mercaptosuccinic acid.

of the asymmetric unit. This confirmed that the asymmetric unit was composed of a single molecular dimer. The asymmetry of the heavy atom constellation (Table 2) did not reveal a clear relationship between the noncrystallographically related subunits by inspection. Therefore, the iron sites from each subunit were identified in a native anomalous difference Fourier map calculated with all X-ray data between 30.0 and 5.0 Å resolution by use of the solvent-modified MIR phases (Kabsch, 1985). A rotational matrix and translational vector relating the irons were calculated and refined by use of the program MUNCHKINS [developed by I. Rayment and G. Wesenberg as described in Rypniewski et al. (1991)]. Subsequently, an averaged MIR electron density map was calculated with all data to 2.7 Å resolution.

Table 3: Phase Calculation Statistics

	resolution (Å)							
	100−9.19	−5.86	−4.60	−3.91	−3.45	−3.13	−2.88	−2.68
no. of observations	1289	2338	2934	3320	3455	3401	3208	2331
figure of merit	0.71	0.66	0.55	0.50	0.48	0.47	0.41	0.29
phasing power[a]								
uranyl actetate								
centric	1.18	1.33	0.94	1.01	0.96	1.04	1.00	0.97
acentric	1.41	1.69	1.41	1.35	1.29	1.26	1.19	1.15
CMDNP								
centric	0.97	0.70	0.59	0.61	0.56	0.60	0.67	0.68
acentric	1.11	1.02	0.83	0.82	0.82	0.86	0.95	0.95
NH_4 bromo osmate								
centric	1.29	1.14	0.90	0.70	0.75	0.79	0.79	0.0
acentric	1.66	1.56	1.09	0.99	1.01	0.99	1.00	0.0

[a] Phasing power is the ratio of the root-mean-square heavy atom scattering factor amplitude to the root-mean-square lack of closure error.

This map was the basis for defining a molecular envelope for a single subunit. Initial phases were improved by cyclical molecular averaging at 2.7 Å resolution for 15 cycles (Bricogne, 1976).

The initial averaged electron density map contained several identifiable strands of β-pleated sheet whose polarity was affirmed by connectivity to five well-defined α-helices. These regions were fitted with a polyalanine model by use of the graphics program FRODO (Jones, 1978). The initial subunit model consisted of 178 alanines and a single iron atom. This model was returned to the crystallographic setting and refined by least squares with the program TNT (Tronrud et al., 1987). A refined rotational matrix and translational vector were obtained by the algorithm of Kabsch (1977), followed by further molecular averaging. An additional 120 alanines and a zinc atom were included in the model, followed by additional crystallographic least squares refinement. Two ensuing cycles of averaging were based on electron density maps generated by combining the MIR phases and those of the model according to the SIGMA A algorithm (Read, 1986). The phase-combined averaged maps allowed the assignment of most amino acid side chains according to the nucleotide sequences of Lemaire and Mueller-Hill (1986) and Cornwell et al. (1987). After additional crystallographic least squares refinement, a phase-combined map was generated in the crystallographic setting with all X-ray data to 2.6 Å resolution. This map revealed the positions of three regions previously missing in the averaged maps. These regions included residues His^{11} through Asn^{15}, residues Ala^{29} through Ala^{36}, and residues Asp^{85} through Arg^{98}. Due to crystal contacts, residues Ala^{29} through Ala^{36} can be seen only in subunit I. Electron density corresponding to residues Ala^{29} through Gln^{43} is absent in subunit II. Least squares refinement of this model resulted in an R-factor of 20.1% for all data to 2.6 Å resolution. Further least squares refinement utilized all X-ray data to 1.8 Å resolution, which originated from a data set collected at the Stanford Synchrotron Radiation Laboratory. Electron density maps with the coefficients $2F_o - F_c$ and $F_o - F_c$ were used in further manual adjustments of the model and the assignment of solvent and two nucleotides. Presently the model contains 393 solvent molecules (B-factors ≤85 Å^2) and a single uridine nucleotide in the active site of each subunit. Ideal bond lengths and geometries for the nucleotide were obtained from the crystal structure of Viswamitra et al. (1975). In addition, there are four modified cysteines in the structure corresponding to the covalent attachment of β-mercaptoethanol (BME) to Sγ of cysteines 99, 160, and 272 of subunit I and cysteine 160 of subunit II. BME was present during the protein purification (Wedekind et al., 1994).

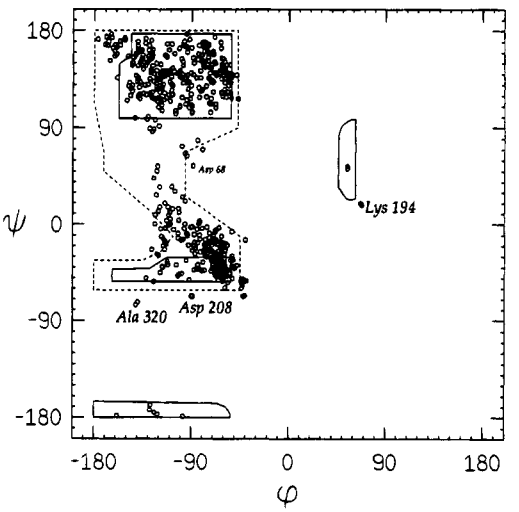


FIGURE 1: Ramachandran plot of the main-chain non-glycinyl dihedral angles for the dimeric enzyme model. Fully allowed φ, ψ values are shown by solid enclosures. Partially allowed regions are enclosed by broken lines. Some outliers can be attributed to hydrogen bonding. Asp 208 forms a turn through hydrogen bonding to the NH of Gly 210. The dihedral angles of Asp 68 are distorted by an interaction of Oδ1 to the NH group of Val 74 and a water-mediated contact to Oγ1 of Thr 72 in a crystallographically related subunit; this interaction occurs in subunit I only. The dihedral angles of Ala 320 are outliers in both subunits due to the formation of a six-membered ring by the carboxylate group of Glu 321 and the guanidinium group of Arg 12. Lys 194 is in a left-handed helical turn. The need for flexibility at this position is indicated by the existence of Gly at this position in nine other uridylyltransferase sequences.

The refined structure has an R-factor of 19.1% for all data to 1.8 Å resolution. The root-mean-square deviation from ideal geometry is 0.018 Å for bond lengths, 2.2° for bond angles, and ≤0.007 Å for coplanar groups. A Ramachandran plot of all main-chain non-glycinyl dihedral angles is given in Figure 1.

RESULTS AND DISCUSSION

Electron Density. A representative portion of electron density calculated with all X-ray data between 50 and 1.8 Å resolution is shown in Figure 2. Beginning at residue 2, the electron density was contiguous throughout each polypeptide chain with the exception of a single region in each subunit.

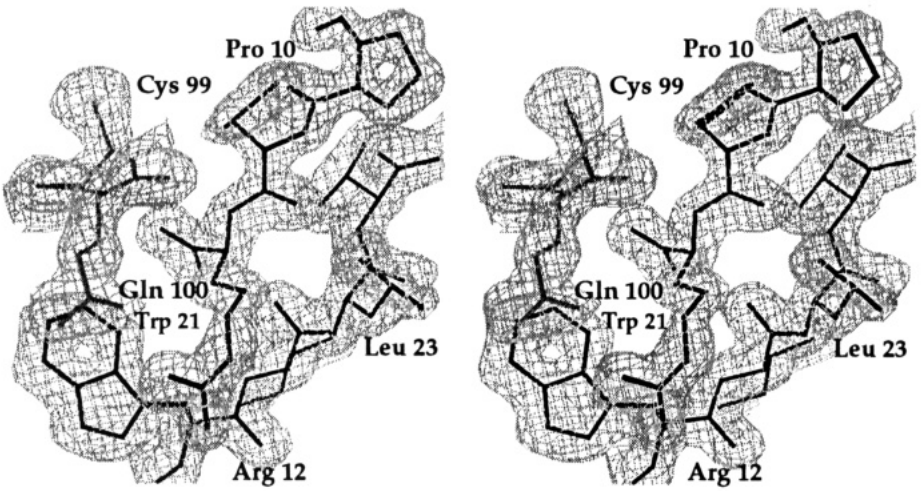


FIGURE 2: Representative portion of a $2F_o - F_c$ electron density map at 1.8 Å resolution calculated with phases from the refined model. The map is contoured at 1.5σ (dark) and 6.5σ (light) for Sγ of Cys 99. Adjacent strands of antiparallel sheet at the subunit interface are shown. Sheets β1, β2, and β14 of subunit I and sheets β4 and β5 of subunit II form the "mini-sheet" of the quaternary fold. Black lines indicate the model for subunit I. Gray lines indicate subunit II. Residues from subunit I include His 11 and Arg 12 of strand β1 and residues Trp 21, Ile 22, Leu 23, and Val 24 of strand β2. Residues from subunit II include Cys 99 and Gln 100 of strand β5. The hydrogen bond between the Arg 12 guanidinium group and the Oδ1 of Gln 100 is typical of the subunit interface. This figure and Figure 6a were generated by FROST (G. Wesenberg, University of Wisconsin).

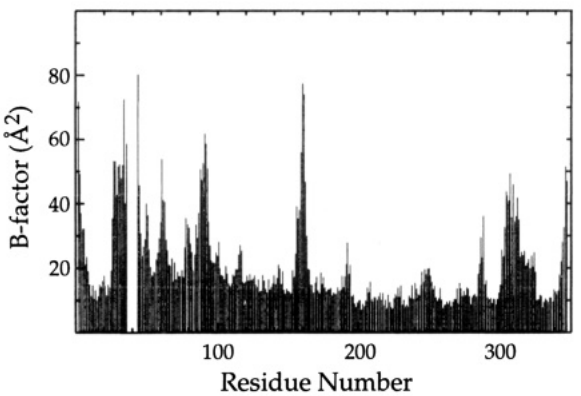


FIGURE 3: Mean temperature factor *vs* amino acid residue for subunit I main-chain atoms. Residues 2 through 36 and residues 44 through 348 are shown. Residue 2 is the first visible amino acid in both subunits.

This break extends from residues Ala[36] through Gln[43] of subunit I and residues Ala[29] through Gln[43] of subunit II. The variable extent of this break can be attributed to the asymmetric crystal packing environment. This effect accounts for the nonsymmetric reactivity of heavy atom compounds at the protein surface (Table 2) and the elimination of conformational flexibility for Gly[347] and Val[348], which are present only at the C-terminus in subunit I. In this manner, crystal packing places the indole ring from Trp[35] of subunit I against a crystallographically related molecule, absent in the dyad-related environment of subunit II. The result is a larger break beginning at Ala[29] of subunit II rather than Gln[37] as seen in subunit I.

A plot of the mean main-chain *B*-factors for subunit I is shown in Figure 3. Atoms that exceed average main-chain *B*-factors are mostly in external loops. The average *B*-factor for all protein atoms of subunits I and II, respectively, is 23.3 Å^2 and 25.6 Å^2. The average solvent *B*-factor is 37.1 Å^2. Waters were assigned such that they were within 3.6 Å of at least one hydrogen-bonding partner. It is notable that the loop extending from Gly[159] through Asn[162] exhibits unusually high temperature factors in both subunits, despite contacts made by several side-chain and backbone atoms to the phosphoryl moieties of the respective nucleotides. Most likely this is due to the modification of Cys[160] by BME. This modification blocks substrate access to the nucleophile His[166], which appears to culminate in excessive flexibility at the phosphoryl binding site. This observation is supported by the partial disorder of an adjacent loop that contains Arg[31]. Collectively these regions may have implications in substrate recognition and binding in the unmodified state.

Subunit Fold. Figure 4a shows an α-carbon trace for subunit I of the uridylyltransferase. The connectivity may be described as $\beta_6\alpha\beta_2\alpha_2\beta_3\alpha\beta_2\alpha\beta\alpha$. Assignments for secondary structural elements are provided in Table 4 and Figure 4b. Helical and sheet elements are mostly in agreement with those assignments made with the algorithm of Kabsch and Sander (1983). Each subunit may be viewed as an independently folding single domain composed of 6 α-helices and 13 strands of antiparallel sheet. Strand β14 is the only strand that pairs in a parallel orientation with an adjacent strand (Figure 4b). The subunit core is composed of nine strands of antiparallel β-sheet lacking an additional orthogonal layer typical of true antiparallel β-barrels. Compensation for the missing second layer results from the insertion of two amphipathic helices, α1 and α4, into the concave face of the twisted sheet (Figure 4b). The intermingling of 11 aromatic amino acids among numerous other hydrophobic residues is comparable to the helix−sheet interfaces of the "open-faced β-sandwich" structures described by Richardson (1981). Additional helix and sheet elements contribute to the interactions that fortify the subunit interface. Among these latter elements is a mini-sheet formed by strands β1, β2, and β14 and strands β4′ and β5′ from the symmetry-related subunit. Approximately one-quarter of the subunit residues are involved in turns, mostly type I and type III (Table 4). Loops arising from the necessary strand crossovers contribute significantly in the formation of the active site (Figures 4 and 5).

It has been proposed that enzymes of the same operon may be related by a common evolutionary ancestor (Horowitz, 1965). However, the three-dimensional structure of the uridylyltransferase bears little resemblance to that of UDP-

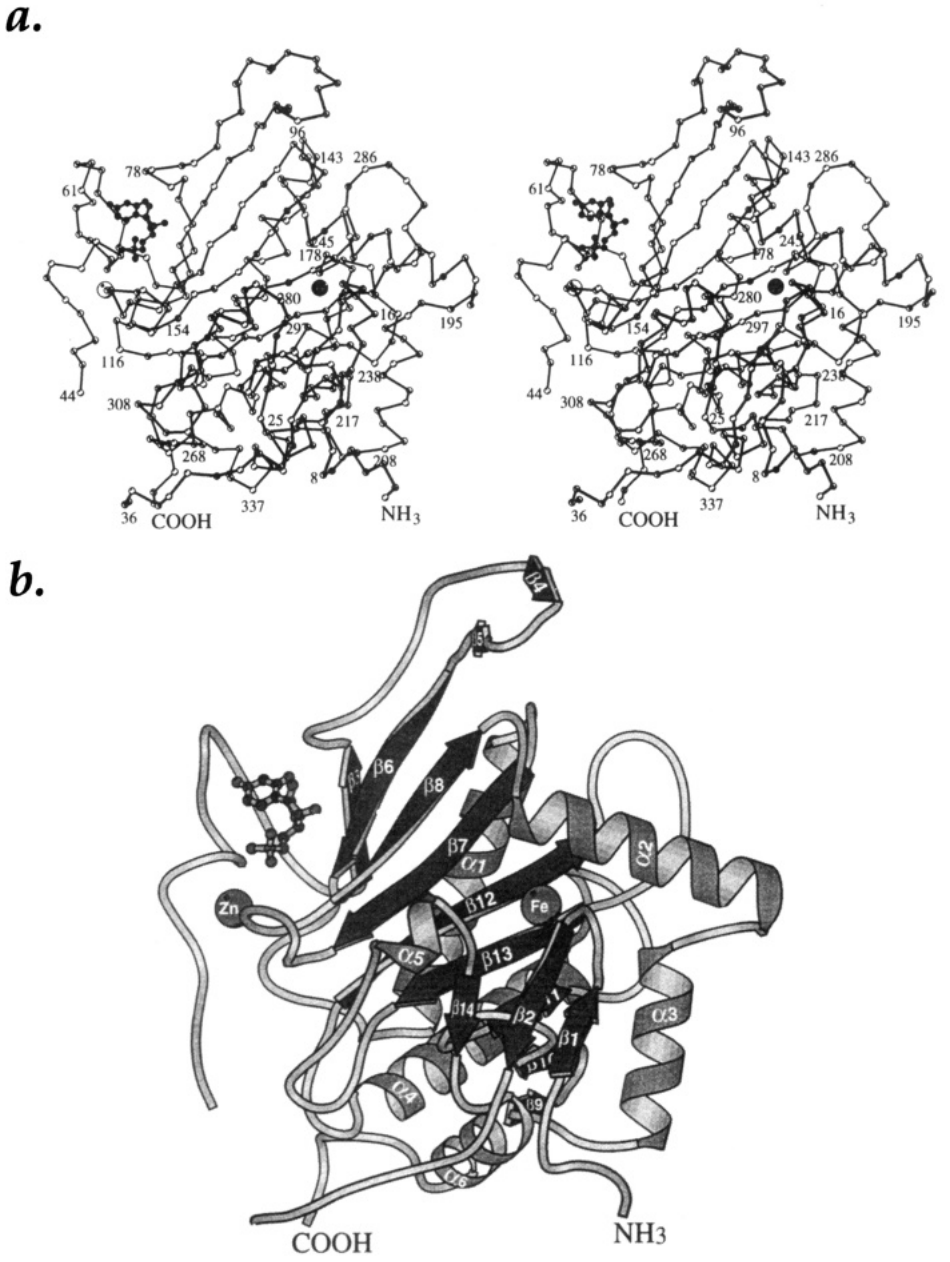


FIGURE 4: Schematic representations of monomeric galactose-1-phosphate uridylyltransferase. Subunits are comprised of 13 strands of antiparallel sheet and 6 α-helices. The view is along the crystallographic *z*-axis. (a) α-Carbon trace of subunit I. The UMP nucleotide (ball and stick) is located among strands β3, β6, and β8, adjacent to a zinc atom (light sphere). The iron atom (dark sphere) is buried in the protein interior. (b) Ribbon diagram delineating the nomenclature for secondary structure elements. This figure and Figures 5, 6, 7b, and 8b were generated by the use of MOLSCRIPT (Kraulis, 1991).

galactose 4-epimerase (Bauer et al., 1992), despite their use of a common substrate. Vorgias et al. (1991) previously discounted such a relationship based upon comparisons of genes from the *gal* operon of *E. coli*. Although the open-faced β-sandwich fold has been described previously (Richardson, 1981), the connectivity of the uridylyltransferase appears unique.

Quaternary Fold. The uridylyltransferase is composed of two chemically identical polypeptide chains. The molecular dimensions of the dimer are approximately 55 × 71 × 60 Å^3. Figure 5 emphasizes the extensive subunit interface in which a surface area of approximately 3500 Å^2 is buried. [A search probe of 1.4 Å was used according to Lee and Richards (1971).] This value is comparable to that reported for other dimers of similar molecular weight (Janin et al., 1988). The central cavity between the subunits is approximately 5 × 8 Å^2 in cross section and 20 Å in depth. This cavity readily accommodates water molecules, which are involved in hydrogen-bonding networks across the cavity interior. The protein interface includes a substantial number of hydrophobic residues as well. The expanse of the buried subunit interface is due in part to three structural elements that protrude from each subunit into the dyad-related molecule (Figure 5). In particular, strands β4 and β5 extend 17 Å outward from the subunit core to engage in the formation of a five-stranded mini-sheet with strands β1, β2, and β14 of the symmetry-related subunit. Figure 2 illustrates the proximity of these intersubunit strands. Helix α5 protrudes into the symmetry-related subunit at the molecular interface, providing a number of hydrophobic interactions (Figure 6b). These structural features are described in greater detail in the next section.

Iron Binding Site. Ruzicka et al. (1995) describe both iron and zinc associated with gal-1-P uridylyltransferase. In order

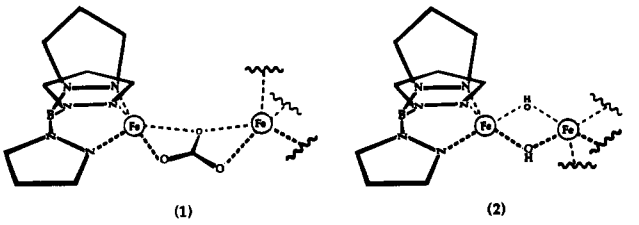
Table 4: Secondary Structural Elements

amino acid	structural element[a]	amino acid	structural element[a]
His 11–Tyr 14	β-strand	Asn 177–Gln 193	α-helix
Asn 15–Thr 18	type I turn	Met 197–Asp 208	α-helix
Gln 20–Val 24	β-strand	Gly 209–Thr 212	type I turn
Ser 25–Arg 28	~type VIII turn	Val 213–Glu 215	β-strand
Asp 49–Cys 52	type I turn	Thr 216–Trp 219	type I turn
Cys 52–Cys 55	~type I turn	Leu 220–Val 223	β-strand
Cys 55–Asn 58	type II turn	Pro 224–Ala 227	type I turn
Arg 60–Gly 63	type I turn	Trp 229–Glu 232	~type I turn
Thr 72–Thr 76	β-strand	Thr 233–Leu 236	β-strand
Asn 77–Ala 80	type VIII turn	Pro 237–His 240	type VIII turn
Met 83–Thr 86	~type I turn	Arg 243–Asp 246	type III turn
Glu 91–His 93	β-strand	Asp 249–Asp 267	α-helix
Asp 94–Met 97	type III turn	Asn 268–Gln 271	~type I turn
Arg 98–Ser 101	β-strand	Pro 275–Ala 283	β-strand
Arg 103–Cys 110	β-strand	Asn 290–Trp 293	type III turn
Asp 114–Lys 117	type I turn	Leu 295–Pro 301	β-strand
Leu 119–Leu 122	type I turn	Ser 306–Val 309	~type I turn
Val 124–Leu 141	α-helix	Gly 315–Leu 319	α-helix
Gly 142–Tyr 145	~type III turn	Thr 322–Arg 324	β-strand
Trp 147–Lys 154	β-strand	Ala 328–Ala 337	α-helix
Ala 156–Gly 159	type I turn	His 342–Glu 345	type III turn
His 166–Asn 172	β-strand		

[a] Turns defined in Creighton (1993).

to discern the identities of the metals in the protein structure, a native anomalous difference Fourier map was calculated. The resulting electron density map revealed a single spherical peak per subunit at contour levels greater than 6.0σ where no other discernible features were present. (Peaks were related by a distance of approximately 30 Å across the molecular dyad.) On the basis of the greater magnitude of f'', the imaginary anomalous dispersion component (measured at a wavelength of 1.54 Å) for iron compared to zinc (MacGillavry & Rieck, 1968), these peaks were attributed to iron. However, the metal stoichiometry at these sites indicates there is likely to be approximately 0.7 mol of iron/mol of "iron" binding site with the remainder occupied by zinc (Ruzicka et al., 1995). Since the anomalous dispersion signal was observed only at the iron site, it is likely that only this site demonstrates a mixed occupancy in accordance with a stoichiometry of ~1.2 mol of zinc/mol of enzyme subunit (1995). However, since the exact metal composition of the crystals was not determined, each site was defined as either iron or zinc with unit occupancy. No attempt was made to refine the contents of the iron site with partial zinc and iron occupancies. Likewise, no restraints were applied to the iron– and zinc–ligand distances.

Figure 6a shows representative electron density calculated with the coefficients $F_o - F_c$. Phases were from a refined model in which the iron ligands, Glu[182], His[281], His[296], and His[298], were removed from the model beginning at the α-carbon prior to refinement and map calculation. The latter ligands of the uridylyltransferase are comparable to those of soybean lipoxygenase-1 (Boyington et al., 1993), which utilizes three histidines and the carboxyl group of its C-terminus to coordinate iron. However, coordination to iron by the uridylyltransferase is bidentate with respect to the carboxylic acid ligand, resulting in pentacoordination. Although protein structures with bidentate carboxylato coordination to iron have been described previously for ribonucleotide reductase (Nordlund et al., 1990) and transferrin (Day et al., 1993), the latter structures exhibit hexacoordinate ligation. Studies of binuclear iron and zinc centers formed by the model compound tris(pyrazolyl) borate reveal the possibilities for pentacoordinate metal geometry with either a bis(hydroxo) or μ-carbonato group as the bridging unit. X-ray crystallographic studies of the μ-carbonato iron tris(pyrazolyl) borate complex demonstrate the existence of a pentacoordinate iron in which six nitrogens and three oxygens act as the ligands in a binuclear center (Kitajima et al., 1993). A portion of this structure is schematized (**1**) to



(1) (2)

indicate its likeness to the uridylyltransferase (Figure 6a). A superposition analysis of the uridylyltransferase iron ligands and those of **1** results in an rms deviation of 0.33 Å; a single iron atom, three nitrogen atoms, and the carboxylato atoms from each respective structure were compared. Superpositions were by the algorithm of Kabsch (1978). The bidentate carboxylato group causes the iron geometries of both the uridylyltransferase and compound **1** to differ substantially from square-pyramidal. Compound **2** represents a more "ideal" square pyramid in which the μ-carbonato unit is replaced by two bridging hydroxides (Kitajima et al., 1993). Independent superpositions of compound **2** iron and its four coplanar metal ligands upon spatially equivalent atoms from compound **1**, and upon the uridylyltransferase ligands result in rms deviations of 0.32 and 0.43 Å, respectively. In contrast, when the equivalent five atoms of the uridylyltransferase are superimposed upon compound **1**, the rms deviation is 0.18 Å; this excludes axial Nδ1 of His[281]. In light of model compound studies, the pentacoordinate iron observed in the uridylyltransferase appears reasonable, although its geometry deviates substantially from an ideal square pyramid.

Gal-1-P uridylyltransferase has been isolated in completely active forms containing either 2 mol of iron or 2 mol of zinc per enzyme subunit (Ruzicka et al., 1995). The plasticity in the metal binding sites is not unexpected given the similar ionic radii for zinc and iron and the identities of the protein ligands (Figures 6 and 7); sulfur and nitrogen are common ligands to both zinc and iron (Coucouvanis et al., 1984; Baltusis et al., 1980; Thomas et al., 1958; Vallee & Auld, 1990; Schwabe & Klug, 1994). Coordination by three histidyl ligands and one bi- or unidentate carboxylato ligand appears to be more common to zinc than iron. The structure fuculose-1-phosphate aldolase (Dreyer & Schulz, 1993) describes symmetric carboxylato coordination to zinc with oxo distances of 2.0 Å. In contrast, the structure of bicarbonate bound to carbonic anhydrase shows a unidentate carboxylato group with oxo–zinc distances of 1.8 and 3.1 Å (Kumar & Kannan, 1994). Studies of a zinc–carboxylato tris(pyrazolyl) borate complex, similar to **1**, indicate both uni- and bidentate coordination is possible in the same binuclear center (Kitajima et al., 1993). Oxo–zinc distances range from 2.0 to 2.6 Å in the unidentate complex and 2.0 to 2.3 Å in the bidentate complex (1993). The latter asymmetric oxo–metal distances observed in the zinc– and iron–carboxylato complexes of **1** resemble the distances

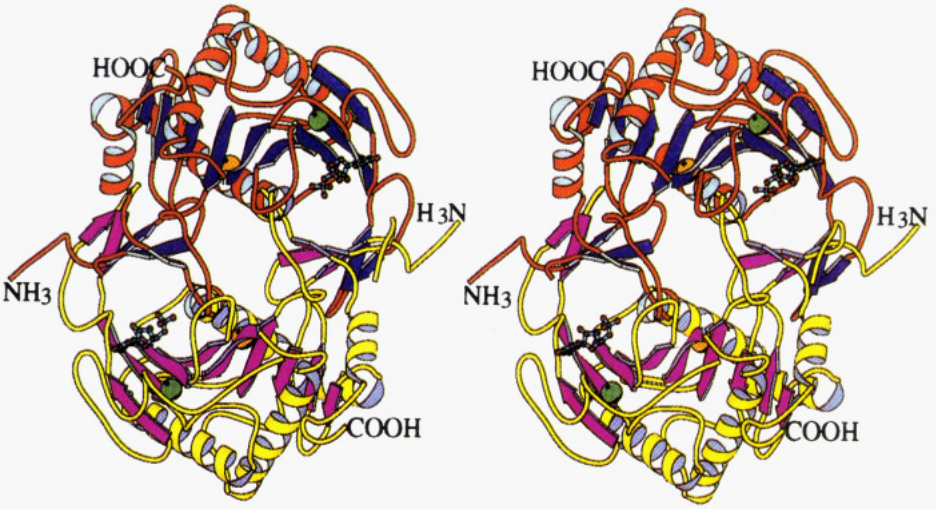


FIGURE 5: Stereo ribbon representation of the galactose-1-phosphate uridylyltransferase dimer. Subunit I is shown as yellow helices and violet sheets; subunit II is shown as red helices and blue sheets. Iron atoms (orange) and zinc atoms (green) are indicated by spheres. The nucleotides are shown as ball-and-stick models in the respective active sites. Strands β1, β2, and β14, helix α5, and strands β4 and β5 (left to right, subunit I) protrude significantly from the subunit core to engage in intermolecular contacts. Breaks in the model extend from residue 37 through residue 43 of subunit I, and from residue 29 through residue 43 of subunit II. (Residue 346 is the last visible amino acid of subunit II.)

observed for iron coordination by the uridylyltransferase (Figure 6a). These coordination schemes explain the observation that the iron site can accommodate either metal and that zinc is compatible with either unidentate or bidentate oxo coordination in the appropriate environment. However, given the probable mixed occupancy at the iron site and the absence of conformational disorder in this region (Figure 6a), an enzyme with two zincs per subunit is likely to maintain bidentate oxo coordination by Glu[182].

The metal ligands at the iron site reside on strands β12 and β13 in a mostly hydrophobic environment at the subunit interface (Figure 6b). The half-barrel architecture of this protein requires that the strands that contribute iron ligands be hydrophobic on both faces. This supports the formation of hydrophobic interactions at the subunit interface and in the half-barrel interior with helices α1 and α4 (Figure 5). Figure 6b shows several hydrophobic side chains contributed by helix α5 of the dyad-related subunit at the molecular interface. The consensus sequence for this helix is Val-Gly-(Phe/Tyr)-Glu-Met-Leu. The conservation[2] of these amino acids suggests that α5 contributes significantly to the fortification of the molecular interface. The carbonyl oxygen of Gly[315] of the *E. coli* enzyme caps the start of α5 by formation of a bifurcated hydrogen bond with the amide hydrogens of residues Met[318] and Leu[319]. Hydrogen bonds between the carbonyl oxygens of Leu[319] and Ala[320] with Nϵ2 of His[281] and the backbone amide of Glu[178] of the barrel exterior provide additional conformational stability to this region.

Helix α2 also participates in the formation of the molecular interface (Figures 5 and 6b). This helix does not contribute a significant number of hydrophobic side chains to the "core" strands surrounding the iron binding site. Instead, the bidentate coordination to iron by Glu[182] appears to be a major stabilizing feature in this region. It is notable that the least conserved amino acid sequence of the uridylyltransferase, that of *Streptomyces lividans*, does not possess an amino acid equivalent to Glu[182]. Residues corresponding to helix α5 and strands β4 and β14 of the subunit interface mini-sheet are absent as well. These observations suggest the uridylyltransferase from *S. lividans* may not form a dimer comparable to that of *E. coli* or that it exists in a monomeric state.

Zinc Binding Site. The metal ligands at the zinc site are shown in Figure 7. The conformation of the peptide backbone required for coordination of the cysteine thiolates to zinc is comparable to the rubredoxin "knuckle" motif common among transcription factors (Schwabe & Klug, 1994). Such Cys-X-X-Cys segments form type I turns in which the *i*th Sγ thiolate (Cys[52]) is hydrogen bonded to the backbone amide hydrogens of residues $i + 2$ and $i + 3$ (residues 54 and 55 of the uridylyltransferase, respectively). Additional hydrogen bonds occur between Cys[52] Sγ and Asn[162] Nδ2 and between Cys[55] Sγ and the backbone NH of Asp 49. This extensive bifurcated hydrogen-bonding network is made possible by the formal negative charge on the sulfur ligands. It is a common feature of cysteine ligands bound to metal ions. Collectively these interactions suitably position residues Phe[53] and Leu[54] of the knuckle at the back side of the uracil and ribose rings as shown in Figure 8b. In higher organisms, including mammals, the residue equivalent to Cys[52] of *E. coli* is absent. In these organisms, the sequence Asn-Pro-Leu-Cys prevails in place of Cys-X-Leu-Cys. It has been noted that asparagine is favored at position *i* of type I turns and that proline favors position $i + 1$ (Richardson & Richardson, 1988). In this way, the mammalian enzymes may allow Asn Oδ (*i*) to hydrogen bond with the backbone NH ($i + 2$) in a turn, thereby replacing the need for a zinc-bound thiolate. The mammalian amino acid sequences also differ from those of bacteria and yeast in the substitution of a serine for the zinc ligand equivalent to His[115] of *E. coli*. Such sequence differences may indicate that the uridylyltransferase from higher organisms does not utilize a metal binding site analogous to that of yeast and bacteria. Presently, there is no structural information to indicate how this might be accomplished.

[2] Conserved residues were determined on the basis of comparisons of amino acid sequences from the following sources: *E. coli*, *F. neoformans*, *H. influenzae*, *S. lividans*, *S. typhimurium*, *K. lactis*, *S. cerevisiae*, *C. elegans*, *M. musculus*, *R. norvegicus*, and *H. sapiens*.

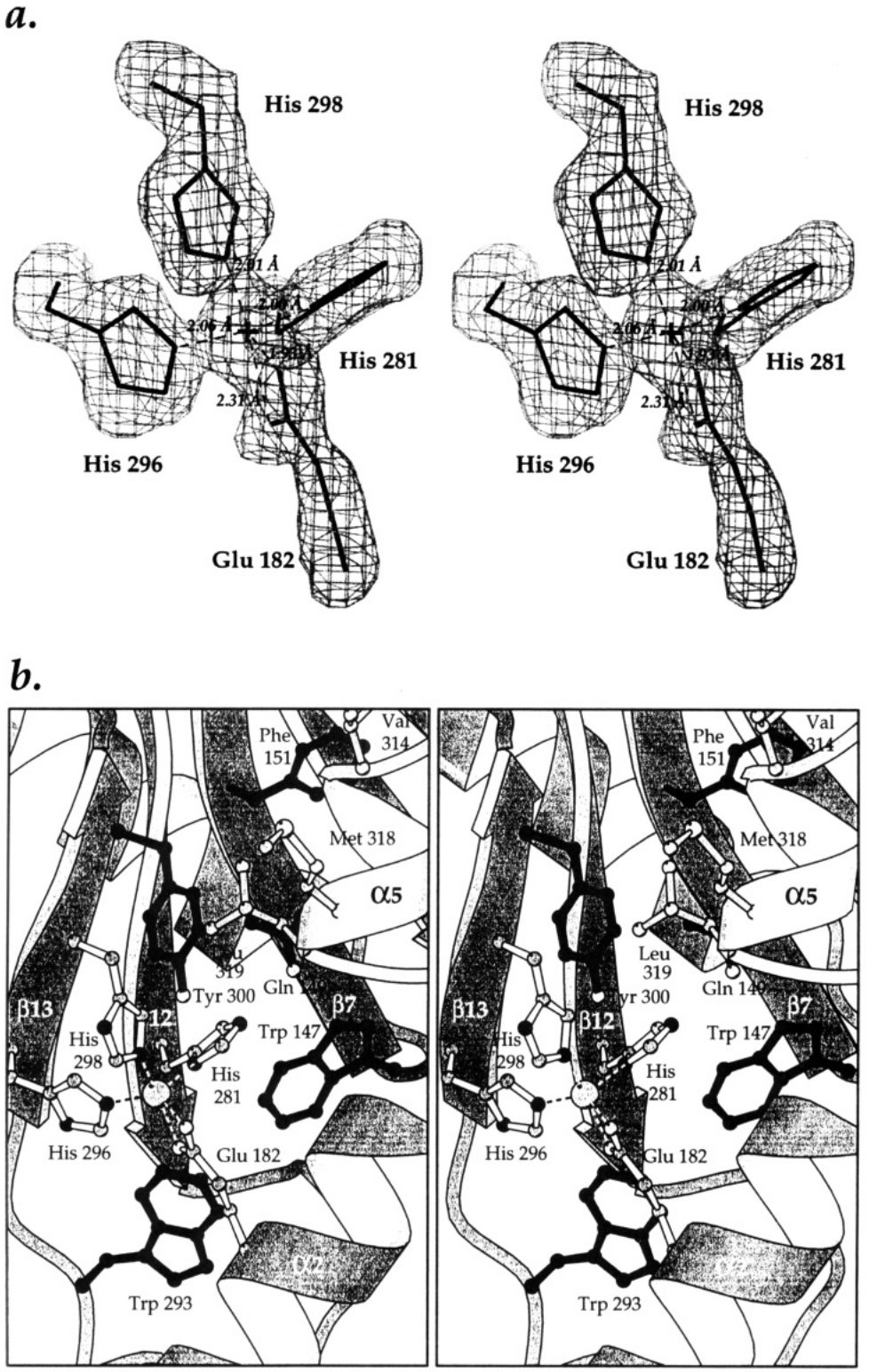


FIGURE 6: Stereoview of the iron binding site. (a) Electron density calculated with the coefficients $F_o - F_c$. The side-chain ligands and iron atom were excluded from the phase calculation. The structure was refined without these atoms to remove model bias in the phases used to calculate this difference density. The electron density map is contoured at 3σ. Iron atoms were identified as single peaks $>6\sigma$ in a native anomalous difference Fourier maps calculated with X-ray data between 5 and 30 Å resolution. The coordination geometry resembles a distorted square pyramid with His 281 as the axial ligand. Ligand distances accompanying broken lines are the average of both subunits. (b) Schematic view of the protein environment surrounding the iron atom. Hydrophobic residues are black (ball and stick). Residues from α5 of the symmetry-related subunit (gray) project toward the iron binding pocket contributing to the hydrophobic subunit interface.

His^{115} and His^{164} complete the tetrahedral zinc coordination sphere. The importance of these residues in *E. coli* has been demonstrated by mutagenesis studies. Field et al. (1989) observed significant losses in enzyme activity for the $His^{115}Asn$ substitution. On the basis of the structure (Figure 7), this substitution is expected to destabilize the zinc binding site, which serves to position the catalytic nucleophile His^{166}. The $His^{164}Asn$ and $His^{164}Gly$ substitutions resulted in complete inactivation of the enzyme (Field et al., 1989; Kim et al., 1990). The severity of the latter His^{164} mutations is most likely attributable to their direct connectivity to His^{166} compared to His^{115}. Since the active site nucleophile resides

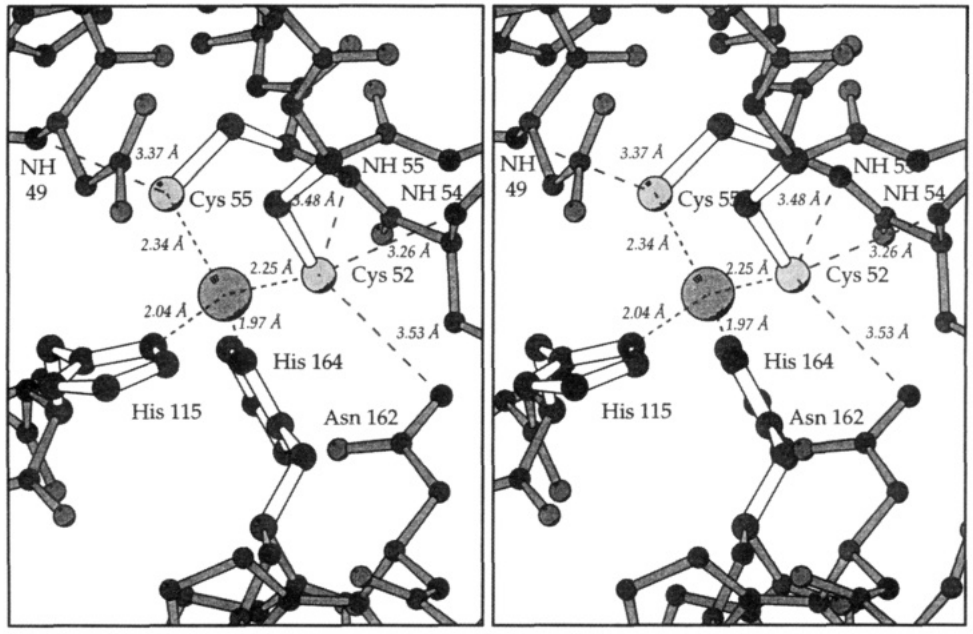


FIGURE 7: Stereoview of the zinc binding site. Ball and sticks with open bonds designate zinc ligands Cys 52, Cys 55, His 115, and His 164. The ligand coordination geometry forms a regular tetrahedron indicated by short broken lines. Coordination distances are the average from both subunits. Possible hydrogen bonds from donor thiolates to acceptor amide hydrogens are indicated by long broken lines. Distances are the average of both subunits.

at the end of a loop (Figure 4a) flanked by Gly[159] and Gly[167], it may require additional constraints to be effective in catalysis. This may be accomplished by the active site sequence His-Pro-His, which is conserved among all known uridylyltransferases. The existence of a metal ligand followed by a proline restricts the backbone dihedral angles in the enzyme from *E. coli* such that the carbonyl oxygen of His[164] is hydrogen bonded to Nδ1 of His[166]. This causes Nϵ2 of the imidazole ring to project toward the electrophilic α-phosphorus of the nucleotidyl donor substrate. His[115] and His[164] appear to play a role in positioning the nucleophile, whereas the "rubredoxin knuckle" organizes peripheral loop elements involved in nucleotide binding and recognition.

The Active Site. An $F_o - F_c$ electron density map for UDP of subunit II is shown in Figure 8a. Although the nucleoside portion of the electron density is well defined, the β-phosphate is disordered partially. Subunit I, in contrast, shows no indications of the β-phosphate. Since the original substrate analog in the crystallization medium was phenyl-UDP (Wedekind et al., 1994), phenol and phenol−phosphate appear to have been lost; no evidence at the level of 0.5σ exists for any of the latter missing atoms. This suggests that the substrate analog was hydrolyzed during the crystallization at pH 5.9. Differences in the electron density at the position of the respective β-phosphate groups are presumably due to the asymmetry of the crystal packing environment (Table 2 and Figure 5). In subunit II, the guanidinium group of Arg[31] provides an intersubunit contact to the β-phosphate (Figure 8b). This interaction is not seen for the β-phosphate of subunit I. Of the two nucleotides, UDP in subunit II has a higher occupancy than UMP in subunit I based upon their respective average *B*-factors of 56.4 and 61.5 Å^2. For this reason, the nucleotide in subunit II has been emphasized in the discussion.

In general, the active sites of the respective subunits are identical within the coordinate error of the model. The subunit α-carbons are superimposed with an rms deviation of 0.26 Å. The aromatic ring of subunit I Phe[53] represents the only significantly disordered residue of the protein model in either active site. This residue is not strictly conserved and has been substituted by proline in several mammalian amino acid sequences. In contrast, neighboring residue Leu[54] is highly conserved, as are the other residues shown in Figure 8b. The number of strong hydrogen bonds to the uracil and ribose portions of the nucleotide accounts for the enzyme specificity at this site. In addition, hydrophobic side chains of residues Leu[54], Phe[74], Phe[79], and Val[108] are within 3.5 Å of the nucleotide. A comparison of the uridylyltransferase to UDP-gal-4-epimerase (Thoden and Holden, in preparation) reveals different strategies for binding the nucleoside moiety of the substrate. The active site of the epimerase contains numerous hydrophobic residues, which form a pocket. Relatively few electrostatic interactions originate from side chains. In contrast, the uridylyltransferase binds the nucleotide in an open cleft spanning sheets β3, β6, and β8 (Figure 8a). This fissure utilizes the side chains of Asn[77] and Asp[78] as well as the amide hydrogens of the polypeptide backbone (Val[61] and Asn[78]) to bind the nucleoside (Figure 8b). It is interesting to compare substrate binding with that of NDP kinase, which forms a phosphohistidyl intermediate in its catalytic mechanism. NDP kinase is devoid of hydrogen bond interactions to its nucleotide base, although a dozen hydrogen bonds are made with the ribose and pyrophosphoryl moieties (Williams et al., 1991). These observed structural differences explain the greater base specificity of the uridylyltransferase compared to NDP kinase.

The present conformation of the nucleotide-bound uridylyltransferase represents an inactive state of the enzyme. Yang (1978) has described previously the presence of a reactive sulfhydryl group whose modification is capable of inactivating the enzyme. The BME-bound Sγ of Cys[160] (Figure 8b) fits this description and explains the need for reducing agents in the enzyme assay procedure reported by Wong and Frey (1974*a*). The modified Cys precludes approach of the electrophilic α-phosphorus to Nϵ2 of the nucleophile. Consequently, the α-phosphorus is restricted to a distance

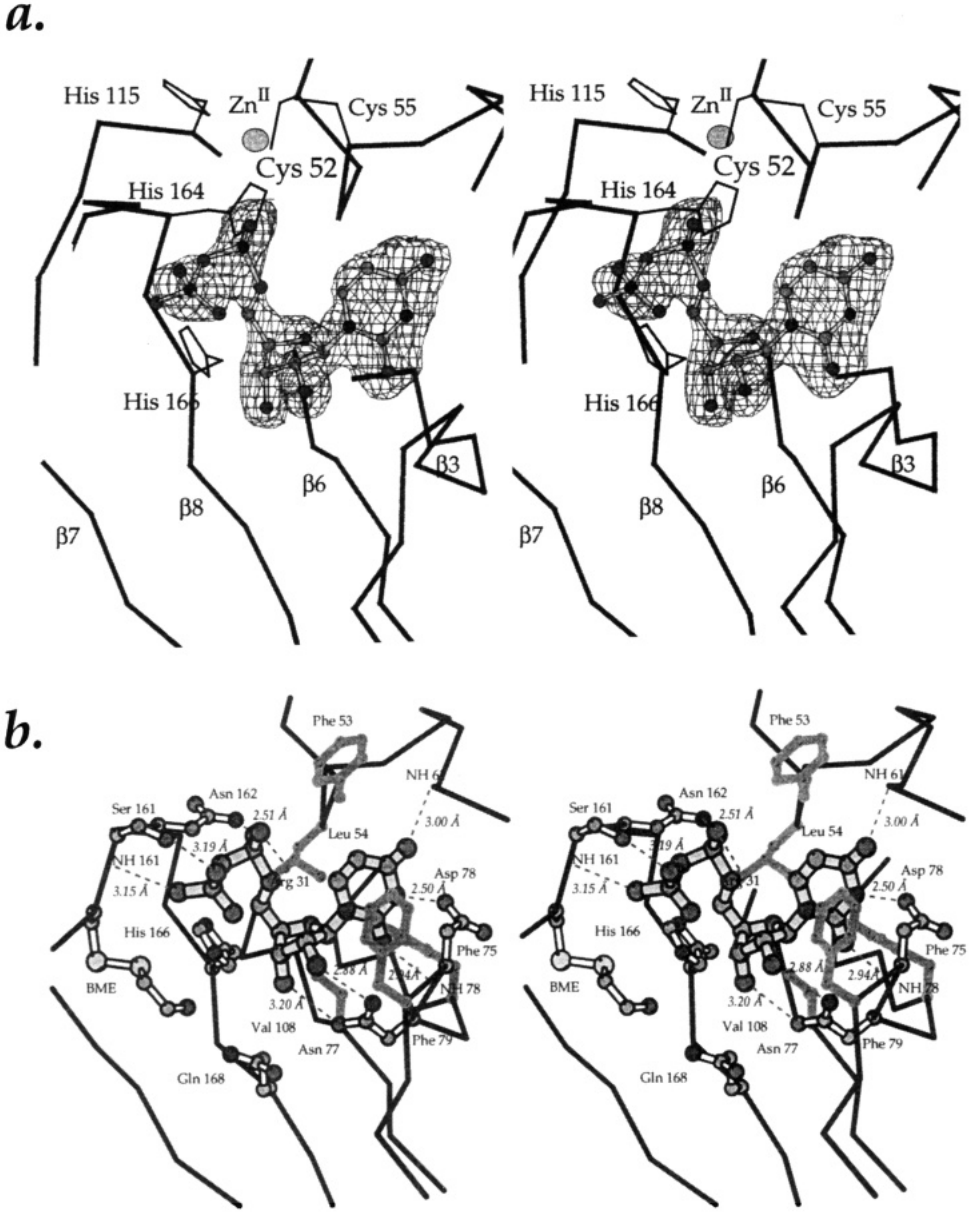


FIGURE 8: Stereo close-up view of the nucleotide binding site. (a) Electron density calculated with the coefficients $F_o - F_c$. Phases were from a refined model in which UDP was removed prior to refinement and map calculation. The electron density is contoured at 3σ. His 166 is 5.5 Å from the α-phosphorus of the nucleotide. (b) Schematic view of the nucleotide interactions in the active site of subunit II. Conserved residues are shown as ball-and-stick models. Residues with white bonds indicate polar residues; residues in gray are hydrophobic. Potential hydrogen bonds are depicted by broken lines. Distances shown are the average from both subunits with the exception of residues 31, 161, and 162. Cys 160 is modified by BME. Panel a was generated by use of the program MOLVIEW (Smith, 1993).

of approximately 6.4 Å from Nε2 of His[166]. In its present conformation, the nucleoside permits a rotation about the C4′−C5′ ribose bond that would position the α-phosphorus to within 2.5 Å of the nucleophile (if Cys[160] was not modified). The resulting α-phosphorus would be positioned adjacent to Gln[168]. Additional "catalytic" conformational changes in the loop comprised of Cys[160], Ser[161], and Asn[162] are anticipated. These residues are conserved and are likely to play a role in pyrophosphoryl recognition (Figure 8b). During catalysis at least one of these hydrogen bond "donor" groups may serve to stabilize the charge buildup on the β-phosphorus during the transition state.

A sequence comparison among the uridylyltransferases from bacteria, yeast, and higher organisms, in the context of the present structure, suggests that there may be differences in the local tertiary fold of these enzymes. Differences are exemplified by the absence of suitable zinc ligands equivalent to Cys[52] and His[115] in the respective amino acid sequences of mammals. The work of Reichardt et al. (1992) showed that mutation of Ser[135] to a Leu had no adverse effect on enzyme activity even though this is the residue equivalent to His[115] in *E. coli*. In contrast, activity losses were observed for the conserved His[115]Asn mutation in the enzyme from *E. coli* (Field et al., 1989). This implies that subtle yet important differences are to be anticipated between the bacterial and mammalian enzymes. Such differences must be considered when inferring the outcome of a human galactosemic mutant mapped onto the global structure of the bacterial enzyme. Even so, these galactosemic mutations provide important information about the role of individual amino acids in the catalytic function of the enzyme.

Of the known human galactosemic mutations, three correspond to residues that interact directly with the nucleotide as seen in the structure of the enzyme from *E. coli*. These

mutations include Leu^{74}Pro (Leu^{54}), Ala^{81}Thr (Val^{61}), and Gln^{188}Arg (Gln^{168}) Elsas et al., 1994; Reichardt et al., 1991, 1992); corresponding residues from the *E. coli* enzyme are denoted parenthetically. The role of Ala^{81} appears to be the donation of its backbone NH group to O4 of the uracil ring (see Val^{61} of Figure 8b). This role is supported by the observation that Ala^{81} is not conserved. Attenuation of enzyme activity at this position is expected to result from an additional hydrogen bond donated by Oγ1 of the mutant Thr. The significance of Leu^{74} may be explained by a van der Waals contact with the uracil and ribose moieties in the active site of the enzyme from *E. coli* (Figure 8b). This packing role cannot be supported by the corresponding proline mutation. The importance of Asn^{188}Arg (Reichardt et al., 1991; Elsas et al., 1994) appears less obvious from the structure of the enzyme from *E. coli*. Fridovich-Keil and Jinks-Robertson (1993) documented the complete inactivation of the human enzyme containing the Gln^{188}Arg mutation when this altered protein was overexpressed in yeast. The proximity of this residue to the α-phosphorus (Figure 8b) and its strict conservation make this residue a strong candidate for participation in the chemical steps leading to the formation and stabilization of the uridylyl-enzyme. A complete description of the galactosemic mutations mapped onto the enzyme from *E. coli* will be described elsewhere.

The position of the hexose binding site is presently unknown. Wong and Frey (1974a) have proposed that enzymes which bind isosteric and electronically equivalent substrates evolve with single or generously overlapping substrate binding sites. The location of the nucleotide implies that a single hexose binding site exists that can accommodate either glc-1-P or gal-1-P. The possibility of two binding sites is precluded by the protein fold, which restricts access to a single face of the histidine nucleophile. The most reasonable location for the hexose site is the cavity formed at the juncture of the barrel exterior and the edge of the mini-sheet (Figure 5). The sugar phosphate might interact with the aromatic and polar side chains of the strictly conserved sequence Phe^{151}-Glu-Asn-Lys/Arg-Gly^{155} on strand β7. Residues Phe^{151} and Asn^{153} of the strand project into this crevice, which begins at the amide side chain of Gln^{168} (Figure 8b). This pocket proceeds toward Trp^{170} and includes residues contributed by helix α5 of the symmetry-related subunit as well as the side chains of Arg^{28} and Phe^{151}. Presently, there is no direct structural evidence to substantiate the participation of any residues in sugar binding. Studies are in progress to isolate the hexose-bound enzyme and to identify the residues required in the formation and stabilization of the uridylyl-enzyme intermediate.

ACKNOWLEDGMENT

We thank Dr. Frank J. Ruzicka and Dr. James B. Thoden for helpful discussions during the preparation of the manuscript. In addition, we express thanks to Professor Ted Baker for his helpful comments regarding the zinc binding site.

REFERENCES

Baltusis, L. M., Karlin, K. D., Rabinowitz, H. N., Dewan, J. C., & Lippard, S. J. (1980) *Inorg. Chem. 19*, 2627−2632.

Bauer, A. J., Rayment, I., Frey, P. A., & Holden, H. M. (1992) *Proteins: Struct., Funct., Genet. 12*, 372−382.

Boyington, J. C., Gaffney, B. J., & Amzel, L. M. (1993) *Science 260*, 1482−1486.

Bricogne, G. (1976) *Acta Crystallogr. A32*, 832−847.

Cornwell, T. L., Adhya, S. L., Reznikoff, W. S., & Frey, P. A. (1987) *Nucleic Acids Res. 15*, 8116.

Coucouvanis, D., Salifoglou, A., Kanatzidis, M. G., Simopoulos, A., & Papaefthymiou, V. (1984) *J. Am. Chem. Soc. 106*, 6081−6082.

Creighton, T. E. (1993) in *Proteins: Structures and Molecular Properties, 2nd ed.*, pp 225−227, Freeman, New York.

Dale, G. L., & Popjak, G. (1976) *J. Biol. Chem. 251*, 1057−1063.

Day, C. L., Anderson, B. F., Tweedie, J. W., & Baker, E. N. (1993) *J. Mol. Biol. 232*, 1084−1100.

Delmer, D. P. (1983) *Adv. Carbohydr. Chem. Biochem. 41*, 105−153.

Delmer, D. P. (1987) *Annu. Rev. Plant Physiol. 38*, 259−290.

Dreyer, M., & Schulz, G. E. (1993) *J. Mol. Biol. 231*, 549−553.

Elsas, L. J., Dembure, P. P., Langley, S., Paulk, E. M., Hjelm, L. N., & Fridovich-Keil, J. (1994) *Am. J. Hum. Genet. 54*, 1030−1036.

Field, T. L., Reznikoff, W. S., & Frey, P. A. (1989) *Biochemistry 28*, 2094−2099.

Fridovich-Keil, J. L., & Jinks-Robertson, S. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 398−402.

Haller, K. J., Johnson, P. L., Feltham, R. D., Enemark, J. H., Ferraro, J. R., & Basile, L. J. (1979) *Inorg. Chim. Acta 33*, 119−130.

Hester, L. S., & Raushel, F. M. (1987) *J. Biol. Chem. 262*, 12092−12095.

Horowitz, N. H. (1965) in *Evolving Genes and Proteins*, pp 15−23, Academic Press, New York.

Janin, J., Miller, S., & Chothia, C. (1988) *J. Mol. Biol. 204*, 155−164.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268−272.

Kabsch, W. (1978) *Acta Crystallogr. A32*, 922−923.

Kabsch, W. (1988a) *J. Appl. Crystallogr. 21*, 67−71.

Kabsch, W. (1988b) *J. Appl. Crystallogr. 21*, 916−924.

Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577−2637.

Kalckar, H. M. (1960) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 19*, 984−990.

Kim, J., Ruzicka, F., & Frey, P. A. (1990) *Biochemistry 29*, 10590−10593.

Kitajima, N., Hikichi, S., Tanaka, M., & Moro-oka, Y. (1993) *J. Am. Chem. Soc. 115*, 5496−5508.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

Kumar, V., & Kannan, K. K. (1994) *J. Mol. Biol. 241*, 226−232.

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

Leloir, L. (1971) *Science 172*, 1299−1303.

Lemaire, H. G., & Mueller-Hill, B. (1986) *Nucleic Acids Res. 14*, 7705−7711.

MacGillavry, C. H., & Rieck, G. D., Eds. (1968) in *International Tables for X-ray Crystallography*, Vol. 3, pp 213−216, The Kynoch Press, Birmingham.

Markus, H. B., Wu, J. W., Boches, F. S., Tedesco, T. A., Mellman, W. J., & Kallen, R.G. (1977) *J. Biol. Chem. 252*, 5363−5369.

Nordlund, P., Sjöberg, B.-M., & Eklund, H. (1990) *Nature 345*, 593−598.

Rayment, I., Johnson, J. E., & Suck, D. (1977) *J. Appl. Crystallogr. 10*, 365.

Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.

Reichardt, J. K. V., Packman, S., & Woo, S. L. C. (1991) *Am. J. Hum. Genet. 49*, 860−867.

Reichardt, J. K. V., Levy, H. L., & Woo, S. L. C. (1992) *Biochemistry 31*, 5430−5433.

Richardson, J. (1981) *Adv. Protein Chem. 34*, 167−329.

Richardson, J. S., & Richardson, D. C. (1988) *Science 240*, 1648−1652.

Rossmann, M. G. (1960) *Acta Crystallogr. 13*, 221−226.

Ruzicka, F. J., Wedekind, J. E., Kim, J., Rayment, I., & Frey, P. A. (1995) *Biochemistry 34*, 5610−5617.

Rypniewski, W. R., Breiter, D. R., Benning, M. M., Wesenberg, G., Oh, B.-H., Markley, J. L., Rayment, I., & Holden, H. M. (1991) *Biochemistry 30*, 4126−4131.

Saito, S., Ozutsumi, M., & Kurahashi, K. (1967) *J. Biol. Chem. 242*, 2362−2368.

Schwabe, J. W. R., & Klug, A. (1994) *Nature Stuct. Biol. 1*, 345−349.

Segawa, T., & Fukasawa, T. (1979) *J. Biol. Chem. 254*, 10707−10709.
Smith, T. J. (1993) *J. Appl. Crystallogr. 26*, 496−498.
Terwilleger, T. C., & Eisenberg, D. (1983) *Acta Crystallogr. A39*, 813−817.
Thomas, J. T., Robertson, J. H., & Cox, E. G. (1958) *Acta Crystallogr. 11*, 599−604.
Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489−501.
Viswamitra, M. A., Seshadri, T. P., Hosur, M. V., Post, M. L., & Kennard, O. (1975) *Acta Crystallogr. A31*, S45.
Vorgias, C. E., Lemaire, H.-G., Wilson, K. S. (1991) *Protein Expression Purif. 2*, 330−338.
Wang, B. C. (1985) *Methods Enzymol. 115*, 90.
Wedekind, J. E., Frey, P. A., & Rayment, I. (1994) *Acta Crystallogr. D50*, 329−331.
Wong, L.-J., & Frey, P. A. (1974a) *Biochemistry 13*, 3889−3894.
Wong, L.-J., & Frey, P. A. (1974b) *J. Biol. Chem. 249*, 2322−2324.
Wu, J. W., Tedesco, T. A., Kallen, R. G., & Mellman, W. J. (1974) *J. Biol. Chem. 249*, 7038−7040.
Yang, S.-L. L. (1978) Ph.D. Dissertation, The Ohio State University, Columbus, OH.
Yang, S.-L. L., & Frey, P. A. (1979) *Biochemistry 18*, 2980−2984.

BI951107F