

Molecular Structure of the NADH/UDP-glucose Abortive Complex of UDP-galactose 4-Epimerase from *Escherichia coli*: Implications for the Catalytic Mechanism^{†,‡}

James B. Thoden, Perry A. Frey, and Hazel M. Holden*

Institute for Enzyme Research, Graduate School, and Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53705

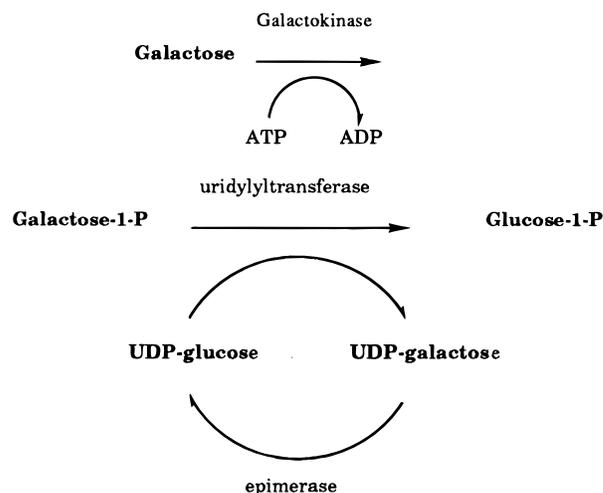
Received January 17, 1996; Revised Manuscript Received February 28, 1996[⊗]

ABSTRACT: UDP-galactose 4-epimerase is one of three enzymes in the metabolic pathway that converts galactose into glucose 1-phosphate. Specifically this enzyme catalyzes the interconversion of UDP-galactose and UDP-glucose. The molecular structure of the NADH/UDP-glucose abortive complex of the enzyme from *Escherichia coli* has been determined by X-ray diffraction analysis to a nominal resolution of 1.8 Å and refined to an *R*-factor of 18.2% for all measured X-ray data. The nicotinamide ring of the dinucleotide adopts the *syn* conformation in relationship to the ribose. Both the NADH and the UDP-glucose are in the proper orientation for a B-side specific hydride transfer from C4 of the sugar to C4 of the dinucleotide. Those residues implicated in glucose binding include Ser 124, Tyr 149, Asn 179, Asn 199, Arg 231, and Tyr 299. An amino acid sequence alignment of various prokaryotic and eukaryotic epimerases reveals a high degree of conservation with respect to those residues involved in both NADH and substrate binding. The nonstereospecificity displayed by epimerase was originally thought to occur through a simple rotation about the bond between the glycosyl C1 oxygen of the 4-ketose intermediate and the β-phosphorus of the UDP moiety, thereby allowing the opposite side of the sugar to face the NADH. The present structure reveals that additional rotations about the phosphate backbone of UDP are necessary. Furthermore, the abortive complex model described here suggests that Ser 124 and Tyr 149 are likely to play important roles in the catalytic mechanism of the enzyme.

In most organisms, the conversion of galactose to glucose 1-phosphate is accomplished by the actions of three enzymes as shown in Scheme 1. Through this pathway galactose is first phosphorylated to galactose 1-phosphate by galactokinase. The second enzyme in the process, galactose-1-phosphate uridylyltransferase, catalyzes the transfer of a uridylyl group from UDP-glucose to galactose 1-phosphate, thereby yielding UDP-galactose and the more metabolically useful glucose 1-phosphate. The third enzyme in the pathway, UDP-galactose 4-epimerase, carries out an epimerization reaction through the transient reduction of NAD⁺. In *Escherichia coli*, all the above-mentioned enzymes are encoded by the *gal* operon. Deficiencies in any one of these enzymes in humans result in the diseased state referred to as galactosemia (Segal, 1995). While the most common form of galactosemia arises from a deficiency in the galactose-1-phosphate uridylyltransferase, epimerase deficiencies have also been described (Segal, 1995).

The epimerase, as isolated from *E. coli*, has been the focus of numerous mechanistic and structural investigations within recent years. The enzyme is a homodimer with each subunit containing 338 amino acid residues and one tightly bound nicotinamide adenine dinucleotide (Wilson & Hogness, 1969;

Scheme 1



Lemaire & Müller-Hill, 1986; Bauer *et al.*, 1992). According to all presently available biochemical data, the reaction mechanism of the enzyme is thought to occur through the abstraction of the C4 hydroxyl hydrogen of the sugar by an enzymatic base and the transfer of a hydride from C4 to NAD⁺ resulting in a 4-ketose intermediate and a reduced cofactor (Frey, 1987). Rotation of the 4-ketose intermediate about the phosphorus–oxygen bond connecting the UDP and the sugar is thought to occur, thereby allowing return of the hydride from NADH to the opposite face of the sugar.

[†] This research was supported in part by grants from the NIH (DK47814 to H.M.H., GM15950 to J.B.T., and GM30480 to P.A.F.).

[‡] Coordinates have been deposited in the Brookhaven Protein Data Bank (filename 1XEL).

* To whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts*, April 1, 1996.

Table 1: Intensity Statistics

	resolution range (Å)							
	overall	30.00–4.00	3.11	2.63	2.32	2.10	1.93	1.80
no. of measurements	69535	10297	8187	10569	10643	10174	9854	9629
no. of independent reflections	39398	3854	4230	5202	5961	6544	6873	6734
completeness of data (%)	95	97	98	99	99	98	93	87
average intensity	1244	1600	1307	495	247	178	117	88
average σ	39.6	45.6	36.3	16.4	13.4	13.8	13.5	13.0
R-factor ^a (%)	2.4	1.5	1.8	2.4	4.1	6.2	10.4	13.7

$$^a R\text{-factor} = (\sum |I - \bar{I}| / \sum I) \times 100.$$

The three-dimensional structures of both the oxidized and reduced forms of epimerase crystallized with UDP bound at the active site are now known to high resolution (Thoden *et al.*, 1996). As typically observed in the NAD⁺-dependent dehydrogenases, the epimerase folds into two distinct domains, with the active site positioned in the cleft formed between these motifs. The N-terminal domain is responsible for NAD⁺ binding while the C-terminal domain provides those amino acid side chains necessary for substrate positioning. The most striking structural difference between the oxidized and the reduced forms of epimerase is the conformation of the nicotinamide ring of the NAD⁺/NADH. In the reduced form, the nicotinamide ring adopts the *anti* conformation while in the oxidized enzyme, the *syn* conformation is observed. The published structures do not implicate a general base within the active site region. The only potential candidates within approximately 5 Å of both the NAD⁺ and the UDP are Asp 31, Asp 58, and Asp 295. These amino acid residues, however, are intimately involved in nucleotide binding and most likely do not play a role in the actual catalytic mechanism of the enzyme. In order to more fully characterize the active site of epimerase and to identify the sugar binding site, we have crystallized and solved the structure of the NADH/UDP-glucose abortive complex of the enzyme. As described here, these studies have allowed for a more detailed description of the active site geometry with regard to sugar binding and have suggested that Ser 124 and Tyr 149 may play critical roles in catalysis.

MATERIALS AND METHODS

Purification and Crystallization Procedures. Recombinant UDP-galactose 4-epimerase was expressed in the *E. coli* strain BL21(DE3), pLysS, carrying the expression plasmid pT7E2 (Swanson & Frey, 1993). The enzyme was purified according to previously described methods (Bauer *et al.*, 1991; Thoden *et al.*, 1996). As isolated, the enzyme contains tightly bound and mostly oxidized pyridine nucleotide. After concentration to 30 mg/mL, the enzyme was incubated with 75 mM dimethylamine/borane complex for 2 h at room temperature. Solid UDP-glucose was then added until a concentration of 15 mM was obtained. This mixture was incubated for 4 h at room temperature after which time additional dimethylamine/borane complex was added to yield a final concentration of 150 mM. Following incubation at 4 °C for approximately 12 h, the reaction mixture was dialyzed against 10 mM potassium phosphate, (pH 8.0) for 24 h at 4 °C. The protein was subsequently concentrated to 50 mg/mL, and additional UDP-glucose was added to a final concentration of 10 mM.

Table 2: Least-Squares Refinement Statistics

resolution limits (Å)	30.0–1.8
R-factor (%) ^a	18.2
no. of reflections used	39401
no. of protein atoms	2705
no. of solvent atoms	479
weighted rms deviations from ideality	
bond length (Å)	0.014
bond angle (deg)	2.1
planarity (trigonal) (Å)	0.007
planarity (other planes) (Å)	0.013
torsional angle (deg) ^b	14.7

^a R-factor = $\sum |F_o - F_c| / \sum |F_o|$, where F_o is the observed structure factor amplitude and F_c is the calculated structure factor amplitude.

^b The torsional angles were not restrained during the refinement.

Large single crystals were grown by the hanging drop method of vapor diffusion against 20% poly(ethylene glycol) 8000, 600 mM NaCl, and 50 mM CHES (pH 9.0) at 4 °C. Crystal growth was generally complete within 5 days with some crystals achieving maximum dimensions of 0.5 mm × 0.5 mm × 0.5 mm. The crystals belonged to the space group P3₂21 with unit cell dimensions of $a = b = 83.8$ Å and $c = 108.4$ Å and one subunit per asymmetric unit. These crystals were isomorphous to those previously described for the NADH/UDP/epimerase complex (Thoden *et al.*, 1996).

X-ray Data Collection and Processing. Prior to X-ray data collection, the crystals were transferred to a solution containing 25% poly(ethylene glycol) 8000, 750 mM NaCl, 20% ethylene glycol, and 50 mM CHES (pH 9.0). Each crystal was subsequently suspended in a thin film of the cryoprotectant mixture using a loop composed of fine surgical thread and flash-cooled to –150 °C in a nitrogen stream.

X-ray data were collected at –150 °C with a Siemens HI-STAR dual area detector system equipped with double-focusing mirrors. The X-ray source was Cu K α radiation from a Rigaku RU200 rotating anode generator operated at 50 kV and 90 mA and equipped with a 300 μ m focal cup. Only one crystal was required for data collection.

The X-ray data were processed according to the procedure of Kabsch (1988a,b) and internally scaled according to the algorithm of Fox and Holmes (1966) as implemented by Dr. Phil Evans. The X-ray data set was 95% complete to 1.8 Å resolution. Relevant X-ray data collection statistics may be found in Table 1.

Computational Methods. The previously determined structure of the epimerase/UDP/NADH complex served as the starting model for the least-squares refinement of the abortive complex with the software package TNT (Tronrud *et al.*, 1987). Ideal stereochemistries for the NADH and UDP moieties were based on the small molecule structural determinations of Reddy *et al.* (1981), Viswamitra *et al.* (1979), and Glasfeld *et al.* (1988). Alternate cycles

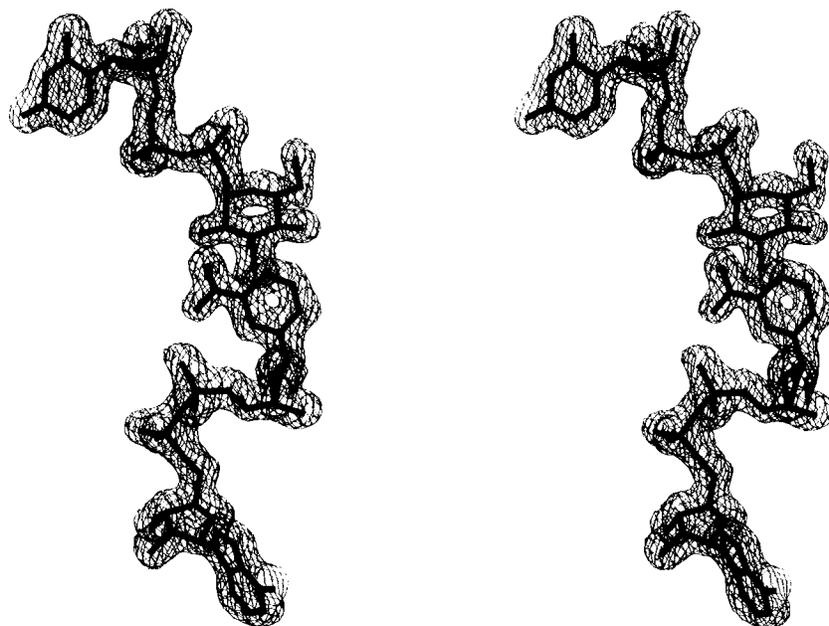


FIGURE 1: Electron density corresponding to NADH and UDP-glucose. The electron density shown was calculated to 1.8 Å resolution with coefficients of the form $(2F_o - F_c)$ where F_o and F_c were the native and calculated structure factor amplitudes, respectively. The map was contoured at 1σ .

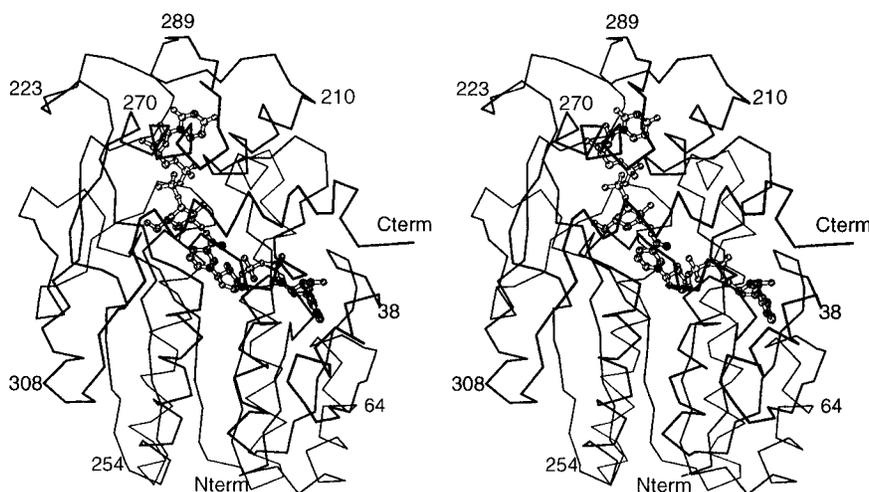


FIGURE 2: α -Carbon trace of one subunit of epimerase with bound substrate and NADH. This figure was prepared with the software package MOLSCRIPT (Kraulis, 1991). Both the NADH and UDP-glucose are displayed in ball-and-stick representation.

of manual model building with the program FRODO (Jones, 1985) and least-squares refinement reduced the R -factor to 18.2% for all measured X-ray data from 30.0 to 1.8 Å. Relevant refinement statistics may be found in Table 2.

Quality of the X-ray Model. Electron density corresponding to the reduced dinucleotide cofactor and the UDP-glucose is shown in Figure 1. As can be seen, the electron density is very well ordered and is representative of the entire map. The only two significantly disordered side chains were Lys 282 and Glu 293. The average B -value for the polypeptide chain backbone atoms was 19.7 Å². A Ramachandran plot of all non-glycinyl main chain dihedral angles revealed only one significant outlier, namely, Phe 178 with $\phi = -93.0^\circ$ and $\psi = -109.5^\circ$. Small peaks of electron density located within approximately 3.2 Å of potential hydrogen-bonding groups were modeled as water molecules. Of the 479 solvent molecules that were located in the electron density map, 148 had temperature factors below 30 Å². The average B -value for the solvent was 39.9 Å². In addition to

these water molecules, two sodium ions, an ethylene glycol moiety, and one di(ethylene glycol) molecule were also included in the model. Both sodium ions were coordinated by oxygen-containing ligands in an octahedral arrangement. These cations were originally modeled as waters but their respective temperature factors refined to anomalously low values.

RESULTS AND DISCUSSION

Abortive complexes of epimerase were originally defined in mechanistic, spectrophotometric, and kinetic studies as complexes of epimerase/NADH with UDP-sugars. These complexes were shown to be formed *in vitro* by the replacement of enzyme-bound UDP-4-ketoglucose in the catalytic intermediate with UDP-sugars (Wee & Frey, 1973; Wong & Frey, 1977). Additionally, abortive complexes have been found as inactive components of the enzyme as purified from *E. coli* (Vanhook & Frey, 1994). The structure of such an abortive complex, namely, the complex of UDP-

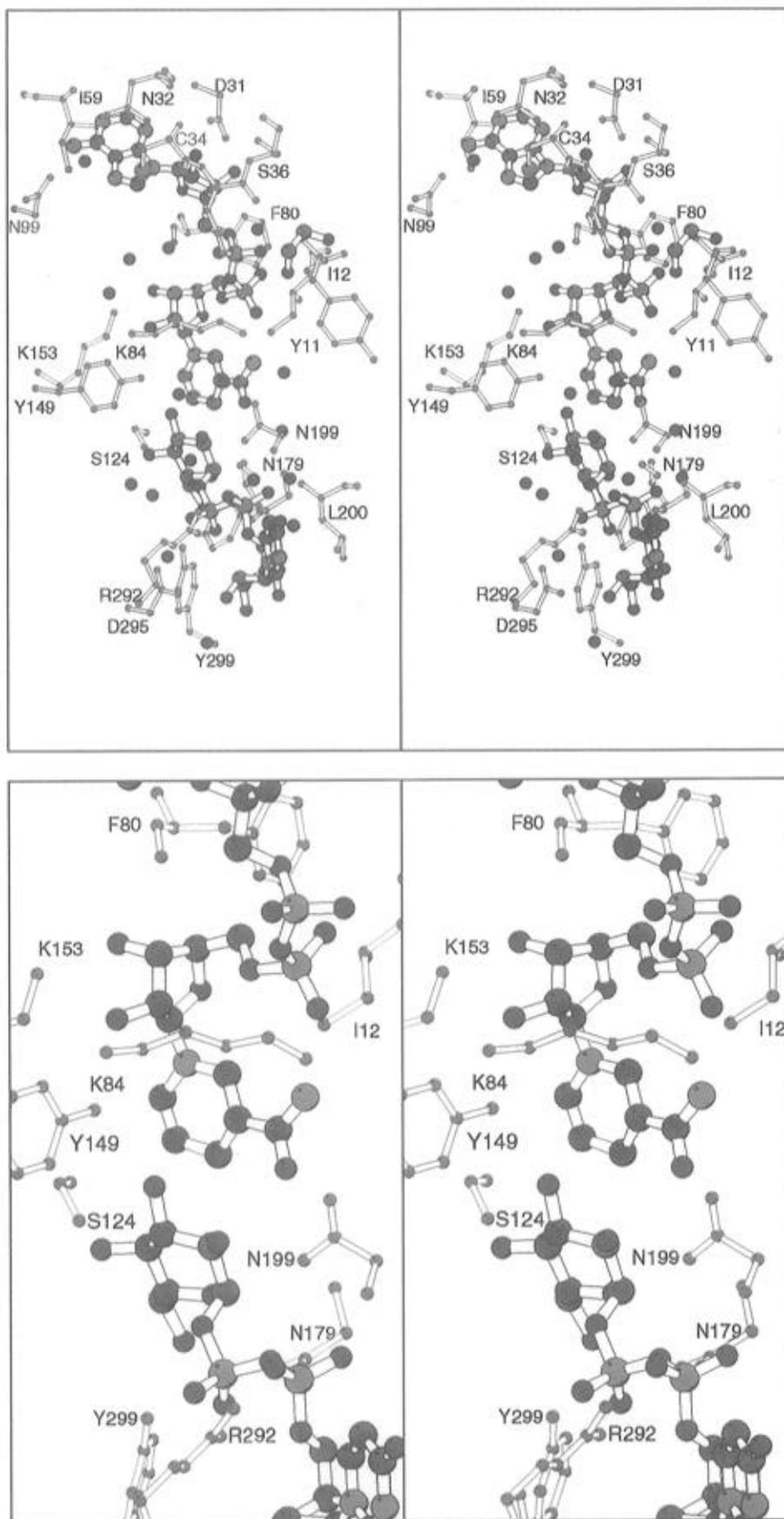


FIGURE 3: Close-up view of the active site for the abortive complex. (a, top) Those amino acid residues that are located within 3.2 Å of the NADH and the UDP-glucose are shown. Ordered water molecules are represented as red spheres. For the sake of clarity, Ala 216, Phe 218, and Arg 231 were omitted from the figure. Their role in the binding of UDP-glucose is indicated in Figure 4a. (b, bottom) Enlarged view of the region near the nicotinamide ring of NADH. The solvent molecules and Phe 178 have been omitted for clearer viewing.

glucose bound to epimerase/NADH, has now been determined to 1.8 Å resolution. From previous studies with UDP

bound at the active site, the position of the nicotinamide ring of the dinucleotide was found to be dependent upon the

oxidation state of the cofactor, with the reduced and oxidized forms adopting the *anti* and *syn* conformations, respectively (Thoden *et al.*, 1996). Since the structure described here is of a reduced form, it was anticipated that the nicotinamide ring would adopt the *anti* conformation. An α -carbon trace for one subunit of the NADH/UDP-glucose complex of epimerase is displayed in Figure 2. As can be seen, however, the *syn* conformation is actually observed. Consequently, the orientation of the nicotinamide ring, in relation to the ribose, is dependent not only upon the oxidation state of the cofactor but also upon the presence or absence of a sugar moiety attached to UDP. Other than the position of the nicotinamide ring, however, the backbone and side chain atoms for the NADH/UDP/epimerase and the NADH/UDP-glucose/epimerase models are superimposed with a root-mean-square (rms) deviation of 0.37 Å. The α -carbon atoms for these models correspond with a root-mean-square deviation of 0.12 Å.

UDP-galactose 4-epimerase is known to transfer a hydride to the B-side of the nicotinamide ring of NAD⁺. Close-up views of the active site and the region immediately surrounding the nicotinamide ring of NADH are displayed in Figure 3. As can be seen, the glucose moiety adopts the chair conformation. Both the UDP-glucose and the NADH are in the proper orientation for a B-side specific hydride transfer from C4 of the sugar to C4 of the nicotinamide. These two atoms are separated by 3.8 Å. Schematics of the potential hydrogen bonds between the nucleotides and the protein are given in Figure 4a,b. As shown in Figure 4a, the hydrogen-bonding pattern observed between the protein and the UDP moiety of the substrate is similar to that observed in the NADH/UDP/epimerase complex (Thoden *et al.*, 1996). Likewise, the interactions between the protein and the NADH, as indicated in Figure 4b, are similar within experimental error to those described for the NAD⁺/UDP/epimerase structure (Thoden *et al.*, 1996). The only major differences occur in the region of the nicotinamide ring. As can be seen in Figure 4c, in the NAD⁺/UDP/epimerase structure, O⁷ of Tyr 177 interacts with both the nicotinamide ring and the ribose whereas in the structure reported here, the side chain hydroxyl group of Tyr 177 lies within hydrogen-bonding distance of Ser 122 and His 243 and does not directly participate in dinucleotide binding. In addition, the carboxamide group of the dinucleotide interacts with Lys 84 in the NAD⁺/UDP/epimerase structure but only with water molecules in the NADH/UDP-glucose/epimerase model. Those amino acid residues responsible for anchoring the glucose to the protein are Ser 124, Tyr 149, Asn 179, Asn 199, and Tyr 299. All of the interactions shown in Figure 4a for these residues are below 3.1 Å.

An amino acid sequence alignment of various prokaryotic and eukaryotic epimerases is given in Figure 5. Those amino acid residues containing side chains that directly interact with the NADH in the *E. coli* epimerase are Asp 31, Asn 35, Ser 36, Lys 84, Asn 99, Tyr 149, and Lys 153 and are highlighted in yellow. Asp 31, which serves as an anchor for the adenine ribose, is highly conserved among all the examples with the exception of the enzyme from *Streptomyces lividans*. Asn 99, which hydrogen bonds to the amino group attached to C6 of the adenine ring, is also conserved except in the enzymes from human and *Rattus norvegicus*. Both Tyr 149 and Lys 153 are absolutely conserved among all the examples

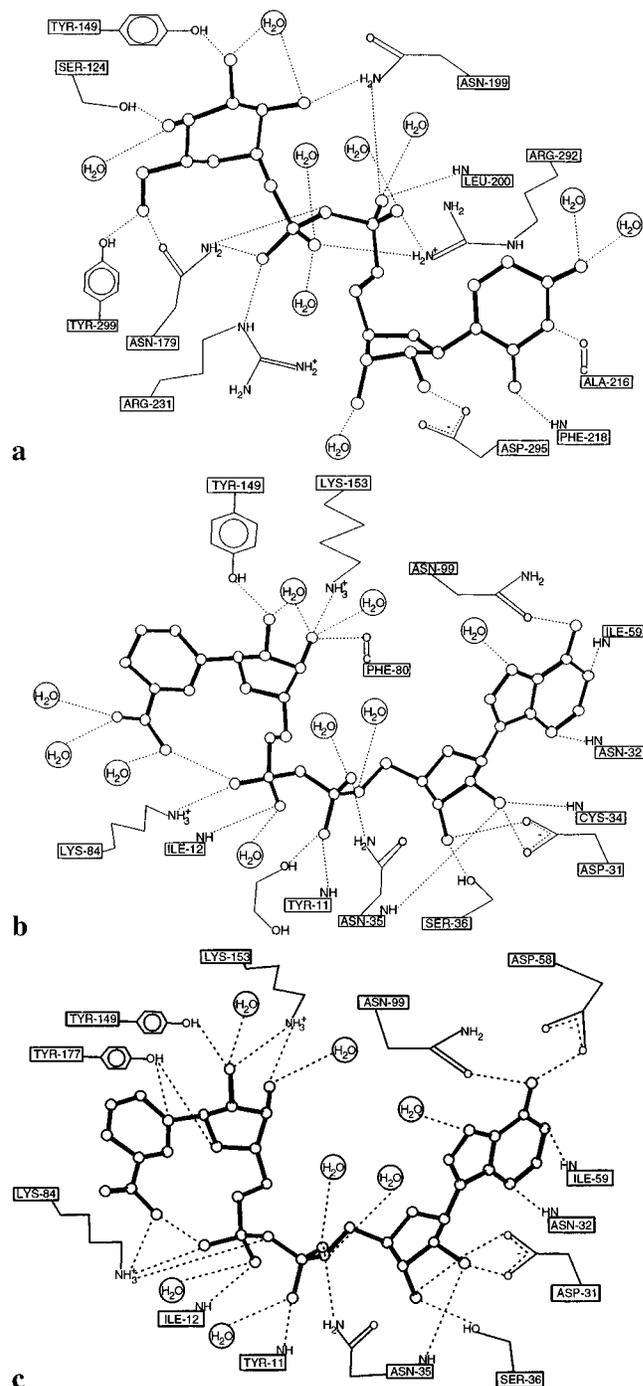


FIGURE 4: Schematic representation of hydrogen bonds between the protein and the nucleotides. Possible interactions within 3.2 Å between epimerase and the UDP-glucose and NADH moieties are shown in (a) and (b), respectively. For comparison purposes, the hydrogen-bonding pattern around the dinucleotide in the NAD⁺/UDP/epimerase structure is shown in (c).

given in Figure 5 except for the protein isolated from *Saccharomyces cerevisiae*. If, however, the amino acid sequence of the *S. cerevisiae* protein is shifted by one residue in this region, then both the tyrosine and lysine residues would align with the other proteins. The amino acid side chains important for binding the UDP-glucose substrate in the *E. coli* protein include Ser 124, Tyr 149, Asn 179, Asn 199, Arg 231, Arg 292, Asp 295, and Tyr 299. Ser 124 is of special interest due to its close contact of 2.6 Å with the 4'-hydroxyl group of glucose. In all the examples listed in Figure 5, there is either a serine or threonine at this position, thereby suggesting the importance of a hydroxyl group in

	31	35	36	
1	MRVLVTGGSGYIGSHTCVQLLQNGHD.VVILDNLCNSKRS.....VLPVIERLGGKHPTFVEGDIRNEALMTEILHDH	72		
3	EKVLVTGGAGYIGSHTVLELLEAGYL.PVVIDNFHNAFRGGGSLPESLRRVQELTGRSVEFEEMDILDQALQRLFKKY	80		
2	KKILVTGGGFIGSHTVVSLLKSGHQ.VVILDNLCNSSIN.....ILPRLKTIQGEIPFYQGDIRDREILRRIFAEN	73		
1	MTVLITGGTGFISHTAVSLVQSGYD.AVILDNLCNSSAA.....VLPRLRQITGRNIPFYQGDIRDCCILRQIFSEH	72		
3	EKVLVTGGAGYIGSHTVLELLEAGYS.PVVIDNFHNSIRGEDSMHPELRRVQELTGRSVEFEEMDILDQAALQHLFKKH	80		
1	MAILVLTGGAGYIGSHVMDRLVEKQGEKVVVDSLVTGHR.....AVH.....PDAIFYQGLSDQDFMRKVFKEN	66		
3	KYLVLTGGAGYIGSHVVAQHLVEAGNE.VVVLHNLSTGPRE.....VCRVPRSSRRHPGRRQVRGR.....I	63		
1	MSILVTGGAGYIGSHTVLSLLQRGDD.VVILDNLSNASRE.....SINRVEKLTGKTATFFEGDILLDRSCLRSVFSAH	72		
1	MAILVTGGAGYIGSHTVVELLNVGKE.VVVDNLCNSSPK.....SLERVKQITGKEAKFYEGDILDRALLQKIFAEN	72		
4	NNILVVTGGAGYIGSHTCLQLAAKGYQ.PVVYDNLNSNGHEE.....FV.....KWGVLEKGDIRDQRLEDEVLARH	67		
10	PRVLVTGGAGYIGSHVHALTDAGIP.AVTIDDLSSAGRE.....AIPAA.....VPLVEGDIGSAELLDLRVMDRH	74		
2	SYLVLTGGAGYIGSHTVLLVNNNGYN.VVVVDNLVNSYD.....VIVRIVLTKRQIPFFKIDLNDHDALDQVFKLY	73		
12	KIVLVTGGAGYIGSHTVVELIENGYD.CVVADNLSNSTYD.....SVARLEVLTKHHIPFYEVLDLDRKLEKVFKEY	83		
5	KYCLVTGGAGYIGSHTVVELCEAGYK.CIVVDNLSNSSYE.....SVARMELTQGEIKFAKIDLCELEPLNKLFFDY	76		
	84	99	124	
73	A.IDTVIHFAGLKAVGESVQKPLEYYDNNVNGTLLRLISAMRAANVKNFIFSS	ATVYGDQPK...IPYVESFPTGTP	145	
81	S.FNAVIHFAGLKAVGESVQKPLDYRVRNLTGTIQLLELMKAHGKVNLFVSS	ATVYGNPQY...LPDEAHTGGC	153	
74	R.IDSVIHFAGLKAVGESVAEPMKYDNNVSGSLVLAEMARAGVFSIVFSS	ATVYDGPQK...VPYTEDMPPGDT	146	
73	E.IESVIHFAGLKAVGESVAEPTKYGGNIVYGLVLAEMARAGVLFKIVFSS	ATVYDGAEK...VPYTEDMRPGDT	145	
81	N.FKAVIHFAGLKAVGESVQKPLDYRVRNLTGTIQLLEIMRAMGVKSLVFSS	ATVYG.KPV...PASGRGPPHRGC	152	
67	PDVDAVIHFAAVSLVGESMEKPLKYFDNNTAGHVKLLVEMNECGVKYIVFSS	AATYGIPEE...IPLETT.P.QNP	139	
64	S.PDGVLFHFAAFSQVGESVVKPEKYWDNNVGGTMMALAMRAGVRRVLFSS	AATYGEPEQ...VPVIVESAPT.RP	135	
73	R.ISAVIHFAGLKAVGASTRQKPLEYYQNNVTGTLVLEEMRSAGVNFIFSS	ATVYGDAP...VPYVETTPIGGT	145	
73	E.INSVIHFAGLKAVGESVQKPTPEYMMNVAGTLVLIQEMKAGVWNVFVSS	ATVYGDPKI...IPITEDCEVGGT	145	
68	K.PRALLHFAAMTEVGESVVKDPAAFYDNNVIGTLLSALAAGIDAFVSS	CATYGLPDS...VPMDES.HKQAP	139	
75	R.VDAMVHFAGSIVVPEVVKPLDYRNRNTANSLTLLGACLRAGIDKVVVSS	AAVYGAPE...VPIREDAPT.VP	146	
74	P.IQAVLHFAALKAVGESTKFPFLNYSNNVGGASLLKVMENNKNIVFSS	ATVYGDATRFENMIPPEHCPTG.P	149	
84	K.IDSVIHFAGLKAVGESTQIPLRYHNNILGTVVLLLELMQYVNVSKFVSS	ATVYGDATRFENMIPPERCPLGPT	160	
77	K.IDSVLHFAGLKAVGESTQIPLTYYPNNVIGTINLLECMKSHDVKLVFSS	ATVYGDATRFENMIPPETCPTG.P	152	
	149	153	179	199
146	QSPYGKSKLHVEQILTDLQKAQPDWSIAL..LRYNPVGAHPSGDMGEDPQGI	PNNLMPYIAQVAVGRRRDSLAIFG	219	
154	TNPYGKSKFFIEEMIRDLCQADKTNVVL..LRYNPVGAHPSGDMGEDPQGI	PNNLMPYVSQVAIGRREALNVFG	227	
147	TSPYGASKSHVERILTDIQKADPRWSMIL..LRYNPVGAHPSGDMGEDPQGI	PNNLMPYICQVAGKLPQLAVFG	220	
146	ANPYGASKAMVERMLTDIQKADPRWSVIL..LRYNPVGAHPSGDMGEDPQGI	PNNLMPYICQVAGSRLPQLSVFG	219	
153	TRPYGKSKFFIEEMIQLDRADTAWNVL..LRYNPVGAHPSGDMGEDPQGI	PNNLMPYVSQVAIGRREALNVFG	226	
140	INPYGESKLMETI...MKWSDQAYGIKYPVPLRYNPVAGANLMLVRLVTRSE..	TLLPILQVAQGVREKIMIFG	210	
136	TNPYGASKLAVDHMITG.EAAAHLGAVS..VPYNPVAGAYGEYGERHDPES...	LLIPLVLQVAQGRREAISVYG	205	
146	TSPYGTSLHVEQILTDIYAKANPEKFTIA..LRYNPVGAHPSGDMGEDPNGI	PNNLMPYIAQVAIGRLEKLGIFG	219	
146	TNPYGTSKYHVEQILRDTAKAEPKFSMTI..LRYNPVGAHPSGDMGEDPNGI	PNNLMPYISQVAIGKLAQLSVFG	219	
140	INPYGRTKWICEQALKDYGLYKGLRSV.I..LRYNPVAGADFEGRIGEWHEPETH.I.	PLAIDAALGRREGFKVFG	211	
147	INPYGASKLHTEQMLRDAGAAHGLRSV.I..LRYNPVAGADPAGRTGQAT.PVAT	ELIKVACQALLGRREPLAIFG	218	
150	TNPYGETKITIENIRDVYANDKSKWKCAI..LRYNPVGAHPSGDMGEDPLGI	PNNLMPYLAQVAIGRREKLSVFG	223	
161	NPYGTAKYALENILNDSKDKSKWKFAL..LRYNPVGAHPSGDMGEDPLGI	PNNLMPYMAQVAVGRREKLYIFG	234	
153	TNPYGKTKLTIEDMMRDLHFSKSFSAI..LRYNPVGAHPSGDMGEDPLGI	PNNLMPYMAQVAVGRREKLYIFG	225	
	231			
220	NDYPTEDGTGVVDYIHVMDLADGHVVAEMEKLANKPG...VHIYNL	GAGVGNVLDVNVAFSKACGKPVNYHFAPR	291	
228	NDYPTEDGTGVVDYIHVVDLAKGHIAALRKLKEQCG...CRIYNL	GTGTGYSVLQMVQAMEKASGKKIPYKVVAR	299	
221	DDYPTPDGTGMVDYIHVMDLAEGHVAAMQAKSNVAG...THLNL	SGRASSVLEIIRAFEAASGLTIPYEVKPR	292	
220	GDYPTPDGTGMVDYIHVMDLAEGHIAAMKAKGGVAG...VHFLN	LGSGRAYSVLEIIRAFEAASGLHIPYRTQPR	291	
227	DDYATEDGTGVVDYIHVVDLAKGHIAALRKLKEQCG...CRIYNL	GTGTGYSVLQMVQAMEKASGKKIPYKVVAR	298	
211	DDYNTPDGTGMVDYVHVPFLDAHLLAVEYLKRGNE...STAFN	LGSGTGFNSLQILEAARKVTGKEIPAEKADR	282	
206	DDYPTPDPRVCAITTSSTPWP...RPTCWCAAAAPGE...HLIC	NLGNMGVSVREVVEVTRRVVTGHPPEIMAPR	275	
220	DDYPTEDGTGVVDYIHVMDLAEGHLKALDHLISAIRG...YKAYN	LGAGKYSVLEMVKAPEKASGCTVAYQISPR	291	
220	SDYDTHDGTGVVDYIHVVDLAVGHKALQRHENDAG...LHIYNL	GTGTGYSVLDMVKAPEKANNITIPYKLEVER	291	
212	TDYDTRDGTGVVDYIHVLDLADAHVRAVDYLLEGGE...SVALN	LGTGTGTTVKELLDATKAVKRPFNIGYAER	283	
219	TDYDTPDGTGCIIDYIHVSDLADAHVLLALHLRRGGG...SLLM	NCGYGRGASREVVRVTLFVSGEQVPATFADR	290	
224	SDYNSKDGTPIDYIHVIDLAKGHIAALNLYLNFHNDK.GLCRE	WNLTGNGSTVFEVFNAFCEAVGKLPFEVVGPR	298	
235	DDYDSRDGTPIIDYIHVVDLAKGHIAALQYLEAYNENEGLCRE	WNLTGSGKSTVFEVYHAFCKASGIDLPYKVTGR	310	
226	DDYDSVDGTPIIDYIHVVDLAKGHIAALKYLEKYAG...TCRE	WNLTGHTTVLQMYRAFCDAIGFNFEVVTAR	298	
	292	295	299	
292	REGVLPAYWADASKADRELNWRVT.RT.LDEMAQCDTWHWQSRHPQGYPD	338	<i>Escherichia coli</i> (Bauer et al., 1991)	
300	REGVVAACYANPSLAQEELGWATA.LG.LDMRCEDLWRWQKQNP	SGFGT	346 <i>Homo sapiens</i> (Daude et al., 1995)	
293	RAGVLACYADPSYTKAQIGWQTO.RD.LTQMMEDESWRWVSN	PNQYDD	339 <i>Neisseria meningitidis</i> (Jennings et al., 1993)	
292	RAGVLACYADPSHTKQQTGWETK.RG.LQMMEDSWRWVSRNP	GRYGD	338 <i>Neisseria gonorrhoeae</i> (Robertson et al., 1993)	
299	REGVVAACYANPSLAHEELGWATA.LG.LDMRCEDLWRWQKQNP	SGLGA	345 <i>Rattus norvegicus</i> (Zeschmick et al., 1990)	
283	PGFDLILASSEKATVVLGKWPQFDN.IEKIIASAWANHSSH	PKGYDD	330 <i>Streptococcus thermophilus</i> (Poolman et al., 1990)	
276	GRDFAVIVASAGTAREKLGWNP.SA.DLAIVSDAWEPLQRRAGQ	319	<i>Streptomyces lividans</i> (Adams et al., 1988)	
292	DGDLAAWADATLADKELNWRVS.RG.IDEMMRDTWNNWQSNP	QGDGDS	337 <i>Ervania amylovora</i> (Metzger et al., 1994)	
292	SGDIATYSDPSLAKELGWVAE.RG.LEKMMQDTWNNWQKNSK	GYRD	338 <i>Haemophilus influenzae</i> (Maskell et al., 1991)	
284	REGSTTVANNDKARQVLGWEPQ.YD.LAAITESAWNHSRRNQ	327	<i>Rhizobium meliloti</i> (Buendia et al., 1991)	
291	PGDPPQVAGADRIREQLGWVPK.HDRLDGIVRSALSWERSLE	QSVGQ	337 <i>Azospirillum brasilense</i> (DeTroch et al., 1994)	
299	DGVLNITANPKRANTEKWKQA.LS.INDACKDLWNWTTKNP	FGFI	345 <i>Pachysolen tannophilus*</i> (Skrzypek & Maleszka, 1994)	
311	RAGVLNITAKPDRAKRELKQWTR.LQ.VEDSKDLWKWTTEN	PFYQL	357 <i>Saccharomyces cerevisiae*</i> (Smits et al., 1994)	
299	DGVLNITAKCDRATNELEWKT.LD.VNKACVDLWKWTTQDN	PFYQI	345 <i>Kluyveromyces lactis*</i> (Webster & Dickson, 1988)	

FIGURE 5: Amino acid sequence alignment of various prokaryotic and eukaryotic epimerases. The alignment was made according to the algorithm of Smith and Waterman (1981) as implemented in the GCG software package (Devereux *et al.*, 1984). Residues involved in binding the NADH and the UDP-glucose are highlighted in yellow and red, respectively. Those epimerase sequences indicated by the asterisks are approximately 700 amino acid residues in length. The N-terminal halves of these enzymes correspond to the epimerase functionalities and were utilized in the alignment. In these larger enzymes, the carboxyl-terminal halves contain a mutarotase activity.

sugar binding. Except for the enzyme from *S. lividens*, Arg 231, which hydrogen bonds to a phosphoryl oxygen of the UDP, is absolutely conserved. Both Arg 292, which provides another hydrogen bond to a phosphoryl oxygen, and Asp 295, which hydrogen bonds to the 2'-hydroxyl group of the adenine ribose, are strictly conserved among all the listed enzymes.

The nonstereospecificity displayed by epimerase is thought to occur through a rotation of the 4-ketose intermediate about the bond connecting the glycosyl oxygen atom and the β -phosphorus atom in the pyrophosphoryl linkage (Kang *et al.*, 1975). In light of the present structural model, however, the movement of the sugar within the active site is probably more extensive than originally envisioned. A simple rotation about the phosphorus-oxygen bond by 180° results in the displacement of C4 of the sugar by 2.6 Å. If, however, other rotations along the phosphate backbone of the UDP are considered, it is possible to maintain the position of C4 next to Ser 124 and still present the opposite face of the ketose intermediate to the nicotinamide ring of NADH. Such rotations, however, move the positions of the α - and β -phosphorus atoms by 0.9 and 1.4 Å, respectively. Obviously the structural determination of the NADH/UDP-galactose/epimerase complex would shed light on this issue. All attempts to produce such an abortive complex with UDP-galactose have been unsuccessful thus far. These experiments have included reduction of the enzyme with dimethylamine/borane complex in the presence of UDP-galactose, UDP, UMP, or TMP and subsequent exchange of these nucleotides with UDP-galactose. In every case the electron density maps unambiguously indicated the presence of UDP-glucose in the active site. Clearly, UDP-glucose binds more tightly in the abortive complex than UDP-galactose, and even though the epimerase has been reduced with dimethylamine/borane, enough residual activity remains to convert UDP-galactose to UDP-glucose.

A key feature of the proposed catalytic mechanism for epimerase is the presence of an active site base that presumably abstracts a proton from the 4'-hydroxyl group of the sugar. Perhaps the most intriguing aspect of this X-ray investigation is the apparent lack of a "classical" base, such as an aspartate, glutamate, or histidine, within the general vicinity of the sugar. Since the structure described here is that of an abortive complex, caution must be applied in its interpretation. As can be seen in Figure 3a, Tyr 149 and Ser 124 are the closest side chains to the 4'-hydroxyl group of glucose. The distance between O' of Tyr 149 and the C4 hydroxyl is 4.3 Å. In the productive NAD⁺/UDP-glucose/epimerase complex, Tyr 149 might move closer to the sugar residue and function as the catalytic base. Indeed, the active site of epimerase is large enough to accommodate such movement. The close approach of Ser 124 to the 4'-hydroxyl group of the sugar (2.6 Å) also suggests that it may play a critical role in the catalytic mechanism of the enzyme.

In summary, the abortive complex of epimerase described here has now defined the binding site for the glucose moiety. Furthermore, these studies have demonstrated that the position of the nicotinamide ring of NADH, in relationship to the ribose, is dependent not only upon the oxidation state but also upon the presence or absence of a sugar moiety attached to UDP. The abortive complex model has suggested that Ser 124 and Tyr 149 may be important for catalysis. Site-directed mutagenesis experiments, kinetic analyses, and X-ray crystallographic studies are currently underway to more fully define the functional roles of these residues.

ACKNOWLEDGMENT

We gratefully acknowledge the helpful discussions of Dr. W. W. Cleland throughout the course of this investigation.

REFERENCES

- Adams, C. W., Fornwald, J. A., Schmidt, F. J., Rosenberg, M., & Brawner, M. E. (1988) *J. Bacteriol.* 170, 203–212.
- Bauer, A. J., Rayment, I., Frey, P. A., & Holden, H. M. (1991) *Proteins: Struct., Funct., Genet.* 9, 135–142.
- Bauer, A. J., Rayment, I., Frey, P. A., & Holden, H. M. (1992) *Proteins: Struct., Funct., Genet.* 12, 372–381.
- Buendia, A. M., Enenkel B., Koeplin R., Niehaus, K., Arnold, W., & Puehler, A. (1991) *Mol. Microbiol.* 5, 1519–1530.
- Daude, N., Gallaher, T. K., Zeschnigk, M., Starzinski-Powitz, A., Petry, K. G., Haworth, I. S., & Reichardt, J. K. V. (1995) *Biochem. Mol. Med.* 56, 1–7.
- DeTroch, P., Keijers, V., & Vanderleydon, J. (1994) *Gene* 144, 143–144.
- Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res.* 12, 387–395.
- Fox, G. C., & Holmes, K. C. (1966) *Acta Crystallogr.* 20, 886–891.
- Frey, P. A. (1987) in *Pyridine Nucleotide Coenzymes: Chemical, Biochemical, and Medical Aspects* (Dolphin, D., Poulson, R., & Avramovic, O., Eds.) pp 461–511, John Wiley & Sons, Inc., New York.
- Glaskfeld, A., Zbinden, P., Dobler, M., Benner, S. A., & Dunitz, J. D. (1988) *J. Am. Chem. Soc.* 110, 5152–5157.
- Jennings, M., van der Ley, P., Wilks, K., Maskell, D., Poolman, J., & Moxon, E. (1993) *Mol. Microbiol.* 10, 361–369.
- Jones, A. T. (1985) *Methods Enzymol.* 115, 157–171.
- Kabsch, W. (1988a) *J. Appl. Crystallogr.* 21, 67–71.
- Kabsch, W. (1988b) *J. Appl. Crystallogr.* 21, 916–924.
- Kang, U. G., Nolan, L. D., & Frey, P. A. (1975) *J. Biol. Chem.* 250, 7099–7105.
- Kraulis, P. J. (1991) *J. Appl. Crystallogr.* 24, 946–950.
- Lemaire, H.-G., & Müller-Hill, B. (1986) *Nucleic Acids Res.* 14, 7705–7711.
- Maskell, D. J., Szabo, M. J., Butler, P. D., Williams, A. E., & Moxon, E. R. (1991) *Mol. Microbiol.* 5, 1013–1022.
- Metzger, M., Bellemann, P., Buyer, P., & Geider, K. (1994) *J. Bacteriol.* 176, 450–459.
- Poolman, B., Royer, T. J., Mainzer, S. E., & Schmidt, B. F. (1990) *J. Bacteriol.* 172, 4037–4047.

- Reddy, B. S., Saenger, W., Mühlegger, K., & Weimann, G. (1981) *J. Am. Chem. Soc.* *103*, 907–914.
- Robertson, B. D., Frosch, M., & vanPutten, J. P. (1993) *Mol. Microbiol.* *8*, 891–901.
- Segal, S. (1995) In *The Metabolic Basis of Inherited Disease* (Scriber, C. R., Beaudet, A. L., Sly, W. S., & Valle, D., Eds.) pp 967–1000, McGraw-Hill, New York.
- Skrzypek, M., & Maleszka, R. (1994) *Gene* *140*, 127–129.
- Smith & Waterman (1981) *Adv. Appl. Math.* *2*, 482–489.
- Smits, P. H. M., deHaan, M., Maat, G., & Grivell, L. A. (1994) *Yeast* *10*, S75–S80.
- Swanson, B. A., & Frey, P. A. (1993) *Biochemistry* *32*, 13231–13236.
- Thoden, J. B., Frey, P. A., & Holden, H. M. (1996) *Biochemistry* *35*, 2557–2566.
- Tronrud, D. E., Ten Eyck, L. F., & Mathews, B. W. (1987) *Acta Crystallogr.* *A43*, 489–501.
- Vanhooke, J. L., & Frey, P. A. (1994) *J. Biol. Chem.* *269*, 31496–31404.
- Viswamitra, M. A., Post, M. L., & Kennard, O. (1979) *Acta Crystallogr.* *B35*, 1089–1094.
- Webster, T. D., & Dickson, R. C. (1988) *Nucleic Acids Res.* *16*, 8192–8194.
- Wee, T. G., & Frey, P. A. (1973) *J. Biol. Chem.* *248*, 33–40.
- Wilson, D. B., & Hogness, D. S. (1969) *J. Biol. Chem.* *244*, 2132–2136.
- Wong, S. S., & Frey, P. A. (1977) *Biochemistry* *16*, 298–305.
- Zeschnigk, M., vonWilcken-Bergmann, B., & Starzinski-Powitz, A. (1990) *Nucleic Acids Res.* *18*, 5289.

BI9601114