PROTEINS: Structure, Function, and Genetics 12:372–381 (1992)

The Molecular Structure of UDP-Galactose 4-Epimerase From *Escherichia coli* Determined at 2.5 Å Resolution

Alan J. Bauer,[1] Ivan Rayment,[1] Perry A. Frey,[1] and Hazel M. Holden[2]

Departments of [1]Biochemistry and [2]Chemistry, the Institute for Enzyme Research, University of Wisconsin, Madison, Wisconsin 53705

ABSTRACT **UDP-galactose 4-epimerase catalyzes the conversion of UDP-galactose to UDP-glucose during normal galactose metabolism. The molecular structure of UDP-galactose 4-epimerase from *Escherichia coli* has now been solved to a nominal resolution of 2.5 Å. As isolated from *E. coli*, the molecule is a dimer of chemically identical subunits with a total molecular weight of 79,000. Crystals of the enzyme used for this investigation were grown as a complex with the substrate analogue, UDP-benzene, and belonged to the space group $P2_12_12_1$ with unit cell dimensions of a = 76.3 Å, b = 83.1 Å, c = 132.1 Å, and one dimer per asymmetric unit. An interpretable electron density map calculated to 2.5 Å resolution was obtained by a combination of multiple isomorphous replacement with six heavy atom derivatives, molecular averaging, and solvent flattening.**

Each subunit of epimerase is divided into two domains. The larger N-terminal domain, composed of amino acid residues 1–180, shows a classic NAD^+ binding motif with seven strands of parallel β-pleated sheet flanked on either side of α-helices. The seventh strand of the β-pleated sheet is contributed by amino acid residues from the smaller domain. In addition, this smaller C-terminal domain, consisting of amino acid residues 181–338, contains three strands of β-pleated sheet, two major α-helices and one helical turn. The substrate analogue, UDP-benzene, binds in the cleft located between the two domains with its phenyl ring in close proximity to the nicotinamide ring of NAD^+. Contrary to the extensive biochemical literature suggesting that epimerase binds only one NAD^+ per functional dimer, the map clearly shows electron density for two nicotinamide cofactors binding in symmetry-related positions in the dimer. Likewise, each subunit in the dimer also binds one substrate analogue.

Key words: protein structure, X-ray crystallography, NAD binding domain, galactose metabolism, nonstereospecific hydride transfer

INTRODUCTION

UDP-galactose 4-epimerase, as shown in Figure 1, is the third enzyme in the Leloir pathway for the conversion of galactose to glucose.[1] The enzyme, as isolated from *Escherichia coli*, is a dimer of identical subunits with a total molecular weight of 79,000.[2] Since the initial report of epimerase activity in 1949,[3] a great deal of biochemical research has been directed toward understanding the reaction catalyzed by the enzyme. In 1964, Wilson and Hogness[4] showed that the *E. coli* protein contains noncovalently bound pyridine nucleotide, nicotinamide adenine dinucleotide (NAD^+). Their investigations, as well as those of subsequent researchers, demonstrated that the tightly bound pyridine nucleotide is involved in a transient oxidation/reduction reaction in which a hydride is removed from the 4 position of the substrate sugar and then returned to the opposite face of the enzyme-bound 4-ketopyranose intermediate as shown schematically in Figure 2. Such a mechanism would involve a nonstereospecific return of the same hydride to an enzyme-bound intermediate to form either product or substrate. Epimerase isolated by the techniques of Wilson and Hogness invariably showed the presence of a single molecule of NAD^+ per dimer, and reductive inactivation of epimerase[5] in the presence of uridine nucleotide and reducing sugar produced only a single equivalent of NADH during extended reaction times.[6]

Research interest in epimerase has focused on several unique features of the enzyme, including the nonstereospecificity of hydride return from the B-side of the nicotinamide ring,[7] the binding of a single molecule of NAD^+ to two identical monomers,[8] and the ability of various reducing sugars to inactivate the enzyme in the presence of uridine nucleotides.[9] We initiated a crystallographic study of epi-

Received June 13, 1991; revision accepted August 1, 1991.

Address reprint requests to Dr. Hazel M. Holden, Department of Chemistry, Institute for Enzyme Research, University of Wisconsin, 1710 University Avenue, Madison, WI 53705.

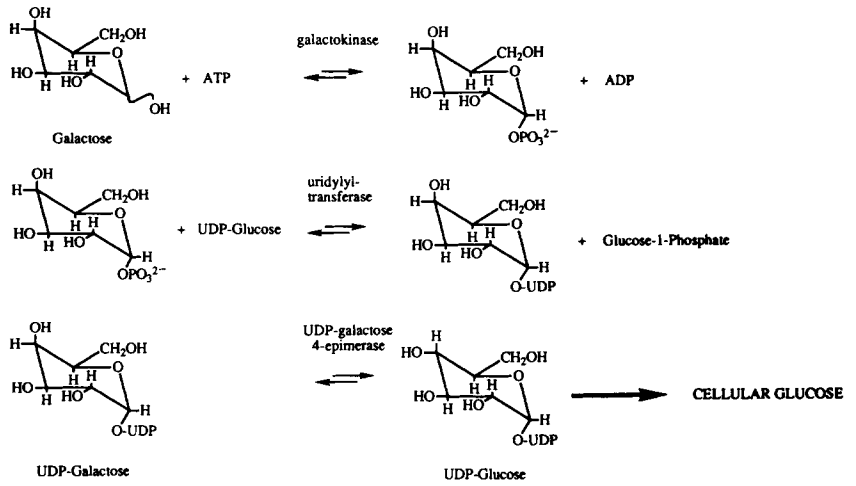

Fig. 1. Leloir pathway for the conversion of galactose to glucose. In this metabolic pathway, galactose is first converted to galactose-1-phosphate by the action of galactokinase and then to UDP-galactose by the enzyme, uridylyltransferase. Subsequently, epimerase, the enzyme described in this report, converts UDP-galactose to UDP-glucose as shown.

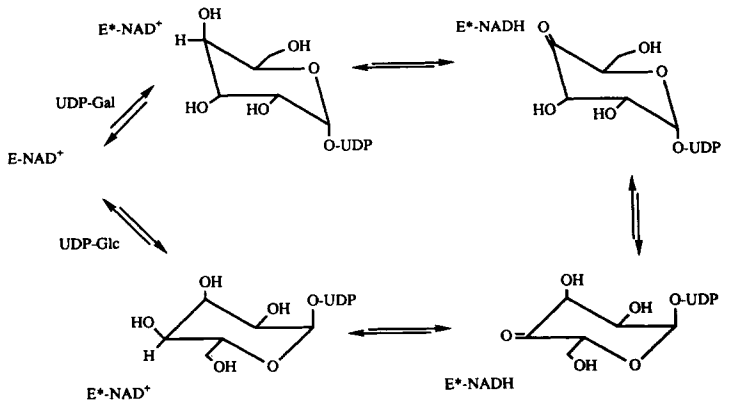

Fig. 2. Possible mechanism for epimerase action. From Frey.[2] In this postulated mechanism, there is a transient oxidation/reduction reaction in which a hydride from the four position of the sugar is transferred stereospecifically to the B-side of NAD^+ resulting in a putative 4-ketopyranose intermediate. Subsequently, there is a rotation of the 4-ketoglucose moiety around the glycosyl oxygen β-phosphorus bond thus altering the face of the intermediate which is available for hydride return.

merase isolated from *E. coli* in order to address some of the structural issues associated with this enzyme.[10] Our goal was to understand how identical subunits could irreversibly bind a single coenzyme molecule, to identify the structural features of the active site that allow for nonspecific reduction of the cofactor and nonstereospecific hydride return during normal catalysis, and to examine the subunit: subunit interactions that prevent dissociation of the holoenzyme. Here we describe the molecular structure of epimerase, complexed with a substrate analogue, UDP-benzene, determined to a nominal resolution of 2.5 Å.

MATERIALS AND METHODS

Biochemical Methods

The purification of epimerase from *E. coli* was routinely carried out by the method of Bauer et al.[10] Crystals of epimerase, complexed with a substrate analogue UDP-benzene, were grown from 6.5% (w/v) polyethylene glycol 8000, 50 mM sodium succinate, 200 mM NaCl, 5 mM NaN_3 solutions, pH 6.0, by microbatch experiments at 4°C as previously described.[10] Crystallization was generally complete within 6 weeks. The crystals belonged to the orthorhombic space group, $P2_12_12_1$, with unit cell dimensions of a = 76.3 Å, b = 83.1 Å, c = 132.1 Å and one dimer per asymmetric unit.[10]

For heavy atom derivative searches, crystals were transferred to a synthetic mother liquor, buffered at pH 7.0 with 50 mM sodium HEPES, and containing 2 mM UDP-benzene, 150 mM NaCl, 13% w/v PEG 8000, and various heavy metals. Six isomorphous heavy atom derivatives were readily prepared using 20 mM $(CH_3)_3PbO_2CCH_3$, 5 mM K_2PtCl_4, 1 mM CH_3HgCl, 2 mM $K_2OsO_4 \cdot 2\ H_2O$/1

mM pyridine, 3 mM ammonium bromosmate, and 3 mM $SmCl_3$.

X-Ray Data Collection and Processing

Three-dimensional X-ray data sets were collected from the native crystals and from the six heavy-atom derivatives with a Siemens-Xentronics 1000 area detector system and processed with the data reduction package XDS.[11] Crystals used for X-ray data collection had typical dimensions of 0.4 mm × 0.5 mm × 1.2 mm. The X-ray source was nickel-filtered copper Kα radiation from a Rigaku RU200 rotating anode operated at 50 kV and 50 mA. X-ray data collection schemes employed overlapping omega scans about the *b*-axis, where the angular width of a given frame was 0.1–0.2°. Only one crystal was required for collection of the unique X-ray data to 2.8 Å resolution. For higher resolution X-ray data, however, several crystals were required, although no crystal remained in the X-ray beam for more than 26 hr.

During the initial stages of data collection, it was found that the crystals had a tendency to slip in the capillary tube. In order to prevent such slippage, the epimerase crystals were first mounted in quartz capillary tubes, the excess mother liquid removed and then a small volume of a 0.2% solution of poly(vinylformal) in 1,2-dichloroethane was drawn over the crystal to form a very thin plastic film. The excess poly(vinylformal) solution was then removed from the capillary, and the film was allowed to dry prior to sealing of the tube. This method has been previously described.[12]

The derivative X-ray data sets were placed on the same scale as the native X-ray data set by a "local" scaling procedure developed in this laboratory by G. Wesenberg, W. Rypniewski, and I. Rayment. With this method the scale for a particular reflection is computed from the neighboring reflections in a volume defined by a sphere or a rectangular prism. The relative contribution of a reflection to a scale factor is weighted according to the distance of the reflection from that which is to be scaled. Relevant data processing and scaling statistics may be found in Tables I and II. The native X-ray data set contains 94% of the total theoretical number of observations to 2.5 Å resolution.

Computational Methods

The positions of the heavy atom binding sites were determined by inspection of difference Patterson maps calculated from 30–5 Å resolution and placed on a common origin by use of appropriate difference Fourier maps. The positions, occupancies, and thermal parameters of the heavy atom sites were refined with the origin-removed Patterson-function correlation method[13,14] and these refined parameters may be found in Table III.

Protein phases based on the six heavy atom derivatives were calculated with the program HEAVY.[14] Relevant phase calculation statistics may be found in Table IV. An initial electron density map, using centroid protein phases based on the six heavy atom derivatives and a structure factor weighting scheme based on the figure of merit, was calculated with X-ray data from 30.0 Å to 5.0 Å. This map, plotted on transparencies and then stacked on thin Plexiglass sheets, clearly showed the positions and molecular envelopes of the two subunits within the asymmetric unit. While it was possible to identify various secondary structural features in the electron density map calculated to a resolution of 2.5 Å, the course of the polypeptide chain could not be unambiguously traced.

Since there were two molecules in the asymmetric unit related by a local twofold rotation axis, it was possible to improve the quality of the protein phases by the techniques of molecular averaging and solvent flattening. The initial rotation matrix relating the two subunits in the asymmetric unit was determined from the positions of the mercury and platinum heavy atom binding sites as listed in Table II. This matrix was subsequently optimized by the program MUNCHKINS (developed in the laboratory by G. Wesenberg and I. Rayment). An averaged electron density map, based on this matrix, was then used to create a molecular envelope for subsequent refinement of the protein phases by iterative molecular averaging and solvent flattening.[15] The initial averaging was performed at 3.0 Å resolution for 15 cycles. After the first cycle, the initial protein phases based on the six isomorphous heavy atom derivatives, were discarded. The structure factor weighting algorithm used in the averaging process was of the form

$$w = e^{-(|Fo|-|Fc|)/|Fo|}$$

where |Fo| was the observed structure factor amplitude and |Fc| was the calculated structure factor amplitude from electron density map inversion.[16] The resolution was extended to 2.5 Å in 0.1-Å increments by including the isomorphous heavy atom derivative phases for each new shell of X-ray data for the first cycle, followed by 10 cycles of refinement at the corresponding resolution. The final *R*-factor between the calculated structure factors from the averaged electron density map and the observed X-ray data was 17.4%. The resulting electron density map confirmed the choice of hand of the heavy atom constellations in that the α-helices were right-handed.

A model of the polypeptide chain was built into the averaged electron density map calculated to 2.5 Å resolution with the molecular modeling program FRODO[17] on an Evans and Sutherland PS390. The amino acid sequence, based on the gene sequence,[18] was used in the fitting process. At the time of the initial fitting process, it was believed, based on ex-

TABLE I. Intensity Statistics for the Native and Derivative Crystals

	Resolution range														
	∞ to 9.64	6.83	5.58	4.84	4.33	3.95	3.66	3.42	3.23	3.06	2.92	2.79	2.69	2.59	2.50
Native															
Average intensity	3,962	2,790	1,663	2,329	2,680	2,110	1,738	1,241	855	605	430	295	220	150	135
Average σ	174	146	114	157	191	161	148	124	101	91	77	67	60	42	33
R factor*	0.027	0.034	0.044	0.045	0.047	0.051	0.057	0.067	0.079	0.100	0.119	0.147	0.176	0.182	0.174
$(CH_3)_3PbOOCCH_3$															
Average intensity	2,089	1,352	846	1,167	1,394	1,108	966	710	510	341	255	281	—	—	—
Average σ	93	62	49	70	84	73	72	65	62	56	57	59	—	—	—
R factor*	0.023	0.028	0.036	0.037	0.037	0.043	0.051	0.064	0.083	0.116	0.154	0.145	—	—	—
K_2PtCl_4															
Average intensity	4,541	3,280	1,936	2,516	2,760	2,073	1,698	1,200	778	585	462	301	286	274	173
Average σ	263	231	184	228	274	250	227	205	173	159	127	60	65	66	59
R factor*	0.035	0.045	0.062	0.061	0.063	0.078	0.087	0.108	0.142	0.172	0.162	0.131	0.148	0.158	0.216
CH_3HgCl															
Average intensity	2,818	1,892	1,226	1,699	1,860	1,408	1,143	796	537	391	281	180	134	89	91
Average σ	127	107	102	134	166	135	128	109	90	80	65	44	42	41	39
R factor*	0.027	0.038	0.053	0.052	0.057	0.065	0.074	0.091	0.114	0.138	0.158	0.182	0.225	0.314	0.297
$K_2OsO_4 \cdot 2H_2O$/pyridine															
Average intensity	2,668	2,957	1,755	1,313	1,656	1,850	2,078	1,733	1,391	1,251	990	697	555	437	405
Average σ	152	124	108	96	109	117	130	121	119	109	100	96	90	88	89
R factor*	0.033	0.028	0.042	0.048	0.046	0.044	0.043	0.051	0.060	0.064	0.075	0.102	0.120	0.146	0.158
Ammonium bromosmate															
Average intensity	4,089	3,589	1,985	1,387	1,699	1,678	1,762	1,393	923	830	587	431	360	263	222
Average σ	258	155	130	124	152	160	175	173	163	163	167	157	161	154	152
R factor*	0.028	0.029	0.045	0.062	0.064	0.070	0.072	0.090	0.131	0.145	0.210	0.261	0.316	0.403	0.449
Samarium chloride															
Average intensity	3,052	2,056	1,047	1,559	1,842	1,503	1,308	970	681	482	345	211	—	—	—
Average σ	101	67	52	72	87	84	72	65	55	49	45	44	—	—	—
R factor*	0.020	0.021	0.033	0.031	0.032	0.037	0.039	0.048	0.058	0.074	0.099	0.148	—	—	—

*R factor = $\Sigma|I-\bar{I}|/\Sigma I$.

TABLE II. *R* Factors Between Native and Heavy Atom Derivative X-Ray Data Sets

Heavy atom derivative	Resolution (Å)	*R* factor* (%)
$(CH_3)_3PbOOCCH_3$	30.0–2.8	19.7
K_2PtCl_4	30.0–2.6	26.3
CH_3HgCl	30.0–2.5	19.3
$K_2OsO_4 \cdot 2H_2O$/1mM pyridine	30.0–3.0	18.0
Ammonium bromosmate	30.0–3.0	29.3
Samarium chloride	30.0–2.9	18.8

*$R = \Sigma||F_N| - |F_H||/\Sigma|F_N|$, where F_N is the native structure factor amplitude, and F_H is the derivative structure factor amplitude.

tensive biochemical evidence, that epimerase binds only one molecule of NAD^+ per dimer. It was therefore quite surprising that the electron densities corresponding to both the NAD^+ cofactor and the substrate analogue were clearly visible in the averaged map at the same contour level as the polypeptide backbone and the side chains. Accordingly, if there were only one cofactor and substrate analogue per dimer, the averaged electron density for these molecules should have been no greater than one-half that observed for the protein portion.

The molecular model of the "averaged" subunit was placed back into the unit cell and subjected to two cycles of alternating least squares refinement with the computer package TNT[19] and manual adjustment of the model. It is important to note that only the polypeptide chain was included in the initial refinement and that the models for the substrate analogue and the nicotinamide cofactor were not included. The initial *R*-factor was 49.5% for all measured X-ray data from 30.0 Å to 2.5 Å, where

TABLE III. Refined Heavy Atom Parameters*

Derivative	Site no.	Relative occupancy	x	y	z	B	Location
$(CH_3)_3PbOOCCH_3$	1	0.42	0.871	0.227	0.247	40.1	Glu 320 Subunit 1
	2	1.01	0.036	0.387	0.256	12.2	Asp 271 Subunit 1
K_2PtCl_4	1	1.19	0.668	0.226	0.080	19.9	His 25, Met 1, Met 247 Subunit 1
	2	0.92	0.123	0.591	0.815	22.4	His 25, Met 1, Met 247 Subunit 2
	3	0.75	0.271	0.033	0.154	18.9	Cys 34 Subunit 2
	4	0.54	0.529	0.293	0.015	29.2	Cys 34 Subunit 1
CH_3HgCl	1	0.60	0.014	0.247	0.000	27.9	Cys 34 Subunit 1
	2	0.74	0.128	0.140	0.213	23.1	Cys 280 Subunit 1
	3	0.63	0.772	0.499	0.853	26.6	Cys 34 Subunit 2
	4	0.65	0.680	0.833	0.802	24.8	Cys 280 Subunit 2
$K_2OsO_4 \cdot 2H_2O$/pyridine	1	0.84	0.522	0.269	0.013	27.0	Near Cys 34 Subunit 1
	2	0.92	0.266	0.024	0.150	19.9	Near Cys 34 Subunit 2
Ammonium bromosmate	1	0.80	0.526	0.254	0.008	22.3	Near Cys 34 Subunit 1
	2	0.91	0.258	0.374	0.030	24.5	His 72 Subunit 1
	3	1.09	0.270	0.032	0.148	15.1	Near Cys 34 Subunit 2
	4	0.66	0.657	0.187	0.206	32.6	Arg 331 Subunit 2
	5	0.65	0.527	0.098	0.235	33.0	His 72 Subunit 2
	6	0.64	1.000	0.120	0.235	34.5	Gln 323 Subunit 1
Samarium chloride	1	0.52	0.204	0.339	0.085	39.2	Near Arg 291, Asp 71, Glu 67 Subunit 1
	2	0.35	0.558	1.000	0.173	40.5	Near Leu 4, Ile 74 and the carbonyl oxygen of Thr 76 Subunit 2
	3	1.16	0.965	0.885	0.244	16.2	Asp 271 Subunit 1

*x, y, and z are the fractional atomic coordinates; B is the thermal factor in A^2.

$$R = \Sigma|F_{obs} - F_{calc}|/\Sigma|F_{obs}|$$

and was reduced to 32.9%. At this stage, it was obvious that each subunit composing the dimer contained both a substrate analogue and a nicotinamide cofactor. Models were thus built into the electron density and 16 more cycles of least squares refinement and model building reduced the *R*-factor to 18.4% with root-mean-square (rms) deviations from "ideal" geometry of 0.016 Å for bond lengths, 3.2° for bond angles, and 0.009 Å for groups of atoms expected to be coplanar. Thus far, 363 solvent molecules have been located in the electron density map. Details of the refinement will be presented elsewhere upon completion at 2.5 Å resolution.

RESULTS AND DISCUSSION

A representative portion of the electron density map, at the present stage of least squares refinement to 2.5 Å resolution, is shown in Figure 3. The initial map, derived from the combination of multiple isomorphous replacement and molecular averaging, was of sufficient quality to permit complete chain tracing, starting from the N-terminal methionine residue and ending with the C-terminal aspartate residue. Both the N-terminal and C-terminal regions were well ordered. There were, however, two areas of weak electron density occurring in the surface loops delineated by amino acid residues 129–143 and 219–229. Both regions begin and end with glycine residues that appear weak in electron density. Other than these two regions, the electron densities corresponding to the amino acid side chains were strong and well ordered. Electron densities for the pyridine nucleotides and substrate analogues in the dimer were also very clear, with the pyrophosphate densities amongst the highest features in the map as shown in Figure 4.

Both NAD^+ and the substrate analogue bind to epimerase in an extended conformation. Recent data from ^{31}P-NMR analysis of NAD^+ bound to UDP-galactose 4-epimerase demonstrated that the binding of UMP to the substrate site perturbs the phosphorus chemical shifts of NAD^+.[20] This indicated that UMP-binding induces an enzyme conformational change that alters the environment around the pyrophosphoryl moiety of NAD^+. However, the possibility that the perturbations could result from a direct interaction between UMP and NAD^+ at the active site could not be ruled out. The three-dimensional structure of epimerase described here clearly shows that NAD^+ and UDP-benzene are bound in adjacent sites with the phenyl ring of UDP-benzene

TABLE IV. Phase Calculation Statistics

	Resolution range							
	∞ −8.86	5.64	4.42	3.76	3.32	3.01	2.77	2.58
No. of reflections	1465	2483	3141	3617	4050	4285	4316	3063
Figure of merit*	0.75	0.67	0.59	0.47	0.40	0.31	0.18	0.12
Phasing power†								
$[(CH_3)_3PbOOCCH_3]$								
Acentric reflections	1.47	1.40	1.02	1.05	1.07	1.22	1.20	—
Centric reflections	0.68	0.75	0.73	0.72	0.79	0.77	0.85	—
(K_2PtCl_4)								
Acentric reflections	1.84	1.68	1.29	1.13	1.27	1.42	1.63	1.58
Centric reflections	1.09	1.12	0.75	0.79	0.84	0.81	0.98	0.83
(CH_3HgCl)								
Acentric reflections	2.01	1.70	1.22	1.13	1.14	1.21	1.42	1.47
Centric reflections	1.25	1.01	0.80	0.77	0.79	0.80	0.87	0.83
$(K_2OsO_4 \cdot 2H_2O$/pyridine)								
Acentric reflections	1.51	1.60	1.27	1.14	1.16	1.08	—	—
Centric reflections	0.95	0.97	0.79	0.80	0.80	0.74	—	—
(Ammonium bromosmate)								
Acentric reflections	1.33	1.41	1.08	1.08	1.04	1.02	—	—
Centric reflections	0.86	0.94	0.66	0.67	0.65	0.69	—	—
(Samarium chloride)								
Acentric reflections	1.72	1.57	1.20	1.13	1.19	1.38	1.55	—
Centric reflections	1.02	0.99	0.75	0.77	0.75	0.84	0.87	—

*Figure of merit is the weighted mean of the cosine of the error in phase angle as described by Dickerson et al. (1961).[25]
†Phasing power is the ratio of the root-mean-square heavy atom scattering factor amplitude to the root-mean-square lack of closure error.

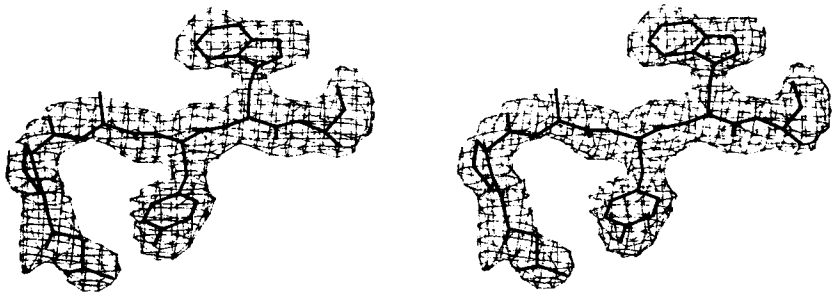

Fig. 3. Representative portion of the electron density map. The electron density shown in stereo here was calculated with X-ray data from 30.0 Å to 2.5 Å and coefficients of the form $(2F_o\text{-}F_c)$, where F_o is the observed structure factor amplitude and F_c is the calculated structure factor amplitude. The electron density map was contoured at 1σ. Those amino acids built into the electron density shown are Leu 296, Pro 297, Ala 298, Tyr 299, Trp 300, and Ala 301.

approaching the nicotinamide ring of NAD^+. The two molecules do not otherwise interact directly. UMP moieties of UMP and UDP-benzene probably bind to the same site, so that UMP would interact even less with NAD^+. It is therefore unlikely that a direct interaction between NAD^+ and UMP led to the changes in phosphorus chemical shifts for NAD^+. This effect was very likely mediated through a protein conformational change. Kinetic evidence for a UMP-induced conformational change was first presented by Kalckar and co-workers, who observed the UMP-dependent reductive inactivation of UDP-galactose 4-epimerase by glucose and other sugars.[9]

The amino acid sequence derived from the structural gene by Lemaire and Muller-Hill[18] was used in the fitting process. For the most part, the published amino acid sequence corresponded well with the electron density map. Difficulty was encountered, however, in fitting the sequence with the observed electron density between amino acid residues 293–311. For example, according to the published amino acid sequence, the electron density shown in Figure 3 should have corresponded to Phe–Arg–Pro–Thr–Gly–Arg. Since it was obvious there was a problem, the structural gene for epimerase in this area was resequenced.[21] The correct amino acid sequence for this segment is Leu–Pro–Ala–Tyr–

Fig. 4. Electron density corresponding to the coenzyme and substrate analogue. The electron density map shown in stereo was calculated and contoured as described in the legend to Figure 3 and corresponds to molecules associated with subunit 1. The electron density for these molecules in subunit 2 is equally good. As can be seen, both the coenzyme and UDP-benzene bind to epimerase in an extended conformation.

Trp–Ala and, as can be seen in Figure 3, corresponds well with the electron density. The entire epimerase gene is being resequenced at this time.

A ribbon drawing of an epimerase subunit is shown in Figure 5 with an α-carbon trace of the molecule depicted in Figure 6. A list of the secondary structural elements is given in Table V. Epimerase is a roughly spherical molecule with an overall diameter of approximately 50 Å. As can be seen from Figures 5 and 6, the molecular architecture of epimerase is divided into two domains. The N-terminal domain, composed of approximately the first 180 amino acid residues, displays a classic NAD^+ binding fold with six strands of parallel β-pleated sheet flanked on either side by α-helices. The seventh strand (amino acid residues 256–261) in the NAD^+ binding fold is contributed by the smaller C-terminal domain. This smaller domain, which is involved in substrate binding, also contains a three-stranded mixed β-pleated sheet and two small helices. As can be seen in Figure 6, the cleft formed at the juncture of the two domains provides the binding sites for the NAD^+ and the substrate analogue.

A closeup view of the active site is shown in Figure 7. Only those amino acid residues within 3.5 Å of atoms of the coenzyme and the substrate analogue are displayed. Since the electron density for NAD^+ is approximately twofold symmetric, the initial model for the cofactor was fitted into the map according to that observed in lactate dehydrogenase.[22] With this conformation, the nicotinamide ring lies in close proximity to the phenyl ring of the substrate anlogue. Specifically, the reactive carbon of the nicotinamide ring is approximately 3 Å from the phenyl ring of the substrate anlogue which would be the sugar portion of the natural substrate. Since epimerase transfers a hydride to the B-side of the nicotinamide ring, it was anticipated that the nicotinamide ring would be in the syn conformation relative to the ribose, as observed in B-side specific dehydrogenases.[23] At this stage of the structural analysis, however, the NAD^+ model clearly fits to the electron density in the anti- rather than syn-conformation.

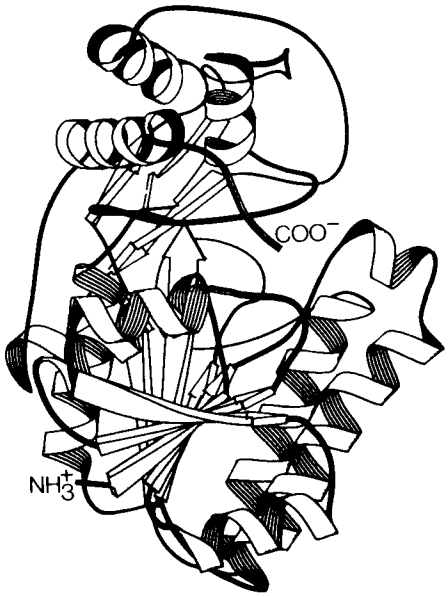

Fig. 5. Ribbon drawing of one subunit of epimerase, generated using software kindly provided by Dr. J.P. Priestle. Strands of β-pleated sheet are shown as arrows, while α-helices are shown as coils. Each epimerase subunit contains 10 strands of β-pleated sheet, and 11 α-helices.

The binding pocket contains seven aromatic residues (Tyr 11, Phe 80, Tyr 149, Tyr 177, Phe 178, Phe 218, and Tyr 299) and nine aspartate or asparagine residues (Asp 31, Asn 32, Asn 35, Asp 58, Asn 99, Asn 179, Asn 198, Asn 199, and Asp 295). There are two lysine residues in the active site, both of which interact with the coenzyme. The side chain of Lys 84 forms a salt bridge with one of the phosphate oxygens, while the side chain nitrogen of Lys 153 is within hydrogen-bonding distance of one of the hydroxyl groups of the nictotinamide mononucleotide ribose. There are, also, two serine residues, Ser 122 and Ser 124, that are involved in the binding of the NAD^+. The one arginine residue in the pocket, Arg 292, interacts with one of the phosphate oxygens of the substrate analogue. While a detailed analysis of the hydrogen bonding pattern in the active site must await completion of the least-squares refine-

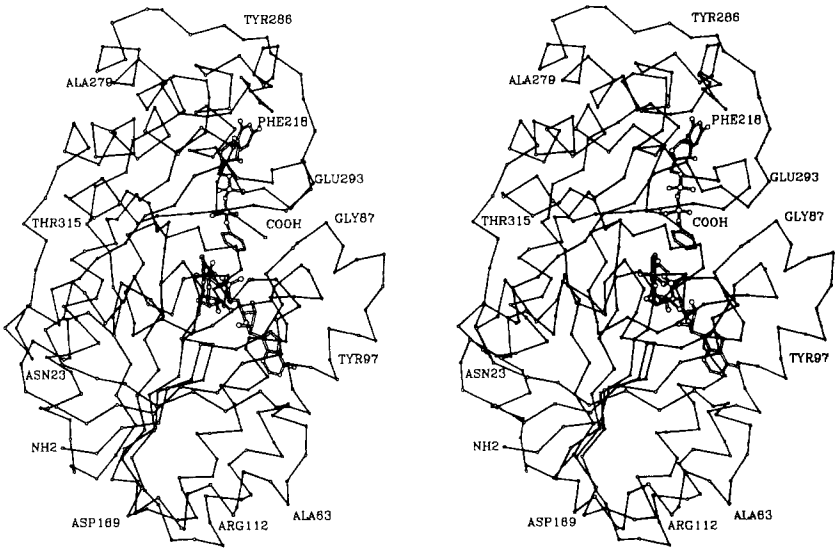

Fig. 6. Stereo drawing of one subunit of epimerase. For the sake of simplicity, only the positions of the backbone α-carbons are represented. The NAD^+ cofactor and the UDP-benzene are shown as atomic models. Various amino acid residues are labeled to aid the reader in following the course of the polypeptide chain. The plot was generated with the software package PLUTO, originally written by Dr. Sam Motherwell and modified for proteins by Dr. Eleanor Dodson and Dr. Phil Evans.

TABLE V. Secondary Structure Pattern of Epimerase

Residue no.	Type of secondary structure
2–6	β-sheet
11–22	α-helix
26–31	β-sheet
39–48	α-helix
50–56	β-sheet
62–72	α-helix
76–80	β-sheet
86–91	α-helix
93–113	α-helix
118–122	β-sheet
148–166	α-helix
171–179	β-sheet
200–208	α-helix
229–233	β-sheet
237–251	α-helix
256–261	β-sheet
265–267	β-sheet
269–280	α-helix
298–302	β-sheet
306–309	α-helix
318–331	α-helix

ment of the model, it should be noted that all the oxygens attached to the phosphorus atoms in both the coenzyme and the substrate analogue are within approximately 3 Å to backbone amide nitrogens or nitrogens in asparagine, lysine, and arginine side chains.

The two subunits composing the functional dimer of UDP-galactose 4-epimerase are related by a two-fold rotational axis as shown in Figure 8. The symmetry-related NAD^+ cofactors are at, a minimum, 22 Å apart. In each subunit, there are two helices delineated by amino acid residues 93–113 and 148–166 that run parallel to one another and through symmetrical contacts form the dimer interface. There are several hydrophobic amino acid residues lining the dimer interface, including Leu 94, Tyr 97, Val 101, Leu 105, Ile 159, and Leu 163. Likewise, numerous hydrogen-bonding interactions occur along the subunit–subunit interface and include amino acid residues Tyr 97, Asp 98, Asn 102, Arg 106, and Asp 162.

At the time this X-ray crystallographic analysis was initiated, amino acid sequence comparisons failed to predict any three-dimensional structural similarities between epimerase and other NAD^+ binding proteins, such as the dehydrogenases. Also, unlike the dehydrogenases, epimerase from *E. coli* does not bind NAD^+ reversibly. When the coenzyme is removed from epimerase *in vitro*, the enzyme becomes denatured (P.A. Frey, unpublished data). As the structural model of epimerase was being built, however, it became apparent that the molecular architecture of this enzyme is very similar to that observed with the dehydrogenases. All these proteins have large and small domains involved in coenzyme and substrate binding, respectively. The structural manifestations as to why the *E. coli* epimerase does not bind NAD^+ reversibly but rather precipitates when the coenzyme is removed will be addressed once the model has been completely refined.

Perhaps the most unanticipated result from this X-ray crystallographic investigation was the ap-

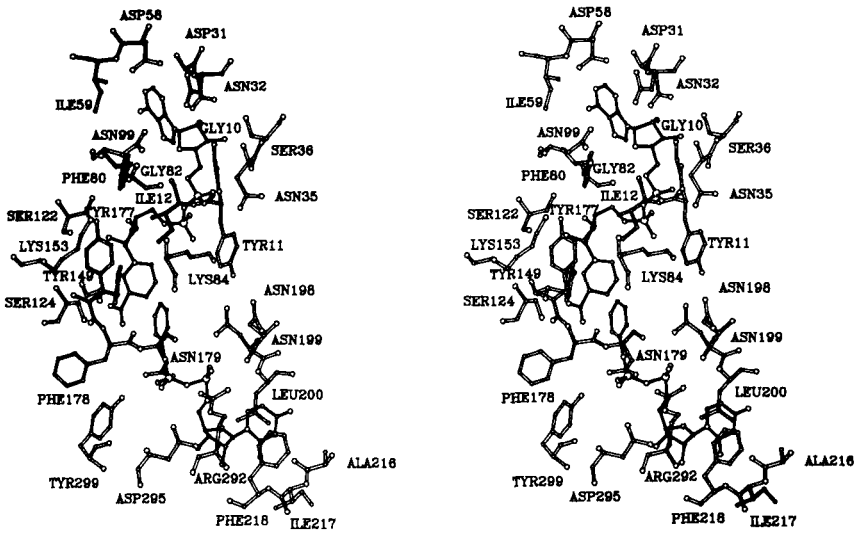

Fig. 7. Closeup view of the active site. Those amino acid residues within approximately 3.5 Å of the NAD^+ cofactor and the UDP-benzene substrate analogue are shown in stereo.

Fig. 8. α-carbon trace of the epimerase dimer. Shown in stereo is the functional dimer of epimerase as seen in the crystalline lattice. The dyad relating one subunit to another lies in the plane of the paper along the horizontal direction and running between the interface of the two molecules.

pearance of electron density for two NAD^+ molecules appearing in symmetry-related positions in the dimer. All previously published reports on epimerase isolated from *E. coli* stated that the protein binds only one cofactor per dimer. In fact, one of the most intriguing questions was how could an enzyme of identical subunits noncovalently bind a single molecule of NAD^+? Recent phosphorus analysis of the enzyme, however, has now suggested that, as observed in the crystalline lattice, there are two molecules of NAD^+ per holoenzyme in solution (A.J. Bauer and P.A. Frey, unpublished data). The question as to why all past biochemical evidence suggested only one NAD^+ bound per dimer is currently being examined.

In conclusion, the molecular conformation of

UDP-galactose 4-epimerase complexed with UDP-benzene has now been solved to a resolution of 2.5 Å. Like a typical dehydrogenase, the molecule folds into a two-domain structure with one domain responsible for the binding of the NAD^+ cofactor and the other for the binding of the substrate analogue. Each subunit in this dimeric enzyme binds one NAD^+ and one substrate analogue in a symmetry-related manner. In light of the preliminary model presented here, there are many interesting biochemical questions that can now be addressed concerning the mode of epimerase action. For example, how does the enzyme facilitate the nonstereospecific hydride transfer between β-NADH on the enzyme and the 4-keto hexopyranose intermediate as indicated in Figure 2? Is the intermediate free to rotate within the active site, again as indicated in Figure 2? Which base removes the C-4 hydroxyl proton to initiate catalysis, and how does this base act on the hydroxyl protons of both glucose and galactose? The future goal of this research is to address these various structural questions by a combination of X-ray crystallographic studies with various forms of the enzyme, including the free enzyme and the enzyme complexed with different substrate analogues, and site-directed mutagenesis experiments. X-ray coordinates for the positions of the α-carbons have been deposited in the Brookhaven Protein Data Bank.[24]

ACKNOWLEDGMENTS

This research was supported in part by grants from the NIH (GM30480 to P.A.F., GM351865 to I.R., and GM30982 to H.M.H.). A.J.B. was supported by an NIH predoctoral training grant (GM07215). H.M.H. is an Established Investigator of the American Heart Association. We would like to thank Stewart Loh for help with the gene sequencing.

REFERENCES

1. Caputto, R., Leloir, L.F., Trucco, R.E., Cardini, C.E., Paladini, A.C. Enzymatic transformations of galactose into glucose derivatives. J. Biol. Chem. 179:497–498, 1949.
2. Frey, P.A. Complex pyridine nucleotide-dependent transformations. In: "Pyridine Nucleotide Coenzymes: Chemical, Biochemical, and Medical Aspects." Vol. 2B. Dolphin, D., Poulson, R., Avramovic, O. eds. New York: Wiley, 1987:461–511.
3. Caputto, R., Leloir, L.F., Cardini, C.E., Paladini, A.C. Isolation of the coenzyme of the galactose-phosphate-glucose phosphate transfer. J. Biol. Chem. 184:333–350, 1950.
4. Wilson, D.B., Hogness, D.S. The enzymes of the galactose operon in *Escherichia coli:* Purification and characterization of uridine diphosphogalactose 4-epimerase. J. Biol. Chem. 239:2469–2481, 1964.
5. Bhaduri, A., Christensen, A., Kalckar, H.M. Enhanced fluorescence of 4-epimerase elicited by 5′ uridine nucleotides. Biochem. Biophys. Res. Commun. 21:631–637, 1965.
6. Bertland, A.U., Seyama, Y., Kalckar, H.M. Concerted reduction of yeast uridine diphosphate galactose 4-epimerase. Biochemistry 10:1545–1551, 1971.
7. Nelsestuen, B.L., Kirkwood, S. The mechanism of action of the enzyme uridine diphosphoglucose 4-epimerase. J. Biol. Chem. 246:7533–7543, 1971.
8. Wilson, D.B., Hogness, D.S. The enzymes of the galactose operon in *Escherichia coli.* J. Biol. Chem. 244:2132–2136, 1969.
9. Bertland, A., Bugge, B., Kalckar, H.M. Fluorescence enhancement of uridine diphosphogalactose 4-epimerase induced by specific sugars. Arch. Bioch. Biophys. 116:280–283, 1966.
10. Bauer, A.J., Rayment, I., Frey, P.A., Holden, H.M., The isolation, purification, and preliminary crystallographic characterization of UDP-galactose 4-epimerase from *Escherichia coli.* Proteins 9:135–142, 1991.
11. Kabsch, W. Automatic indexing of rotation diffraction patterns. J. Appl. Crystallogr. 21:67–71, 1988.
12. Rayment, I., Johnson, J.E., Suck, D.J. A method for preventing crystal slippage in macromolecular crystallography. Appl. Crystallogr. 10:365, 1977.
13. Rossmann, M.G. The accurate determination of the position and shape of heavy-atom replacement groups in proteins. Acta Crystallogr. 13:221–226, 1960.
14. Terwilliger, T.C., Eisenberg, D. Unbiased three-dimensional refinement of heavy-atom parameters by correlation of origin-removed Patterson functions. Acta Crystallogr. A39:813–817, 1983.
15. Bricogne, G. Methods and programs for direct space exploitation of geometric redundancies. Acta Crystallogr. A32:832–847, 1976.
16. Rayment, I. Molecular replacement method at low resolution: optimum strategy and intrinsic limitations as determined by calculations in icosohedral virus models. Acta Crystallogr. A39:102–116, 1983.
17. Jones, T.A. Interactive Computer Graphics: FRODO. Methods Enzymol. 115:157–171, 1985.
18. Lemaire, H.G., Muller-Hill, B. Nucleotide sequence of the *gal* E gene and the *gal* T gene of *E. coli.* Nuc. Acids. Res. 14:7705–7711, 1986.
19. Tronrud, D.E., Ten Eyck, L.F., Matthews, B.W. An efficient general-purpose least-squares refinement program for macromolecular structures. Acta Crystallogr. A43: 489–501, 1987.
20. Konopka, J.M., Halkides, C.J., Vanhooke, J.L., Gorenstein, D.G., Frey, P.A. UDP-galactose 4-epimerase. Phosphorus-31 nuclear magnetic resonance analysis of NAD^+ and NADH bound at the active site. Biochemistry 28: 2645–2654, 1989.
21. Bauer, A.J. The structure of UDP-galactose 4-epimerase from *Escherichia coli* at 2.5 Å resolution. Ph.D. thesis, University of Wisconsin, Madison, 1991.
22. Grau, U.M., Trommer, W.E., Rossman, M.G. Structure of the active ternary complex of pig heart lactate dehydrogenase with *S*-lac-NAD at 2.7 Å resolution. J. Mol. Biol. 151:289–307, 1981.
23. Wu, Y.D., Houk, K.N. Theoretical evaluation of conformational preferences of NAD^+ and NADH: An approach to understanding the stereospecificity of NAD^+/NADH-dependent dehydrogenases. J. Am. Chem. Soc. 113:2353–2358, 1991.
24. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr., Brice, M.D., Rogers, J.R., Kennard, O., Shimanouchi, T., Tasumi, M. The protein data bank: A computer based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.
25. Dickerson, R.E., Kendrew, J.C., Strandberg, B.E. The crystal structure of myoglobin: Phase determination to a resolution of 2 Å by the method of isomorphous replacement. Acta. Crystallogr. 14:1188–1195, 1961.