Human UDP-galactose 4-Epimerase

ACCOMMODATION OF UDP-*N*-ACETYLGLUCOSAMINE WITHIN THE ACTIVE SITE*

James B. Thoden‡, Travis M. Wohlers§¶, Judith L. Fridovich-Keil‖, and Hazel M. Holden‡**

From the ‡Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53705 and the §Graduate Program in Genetics and Molecular Biology and the ‖Department of Genetics, Emory University School of Medicine, Atlanta, Georgia 30322

UDP-galactose 4-epimerase catalyzes the interconversion of UDP-galactose and UDP-glucose during normal galactose metabolism. One of the key structural features in the proposed reaction mechanism for the enzyme is the rotation of a 4′-ketopyranose intermediate within the active site pocket. Recently, the three-dimensional structure of the human enzyme with bound NADH and UDP-glucose was determined. Unlike that observed for the protein isolated from *Escherichia coli*, the human enzyme can also turn over UDP-GlcNAc to UDP-GalNAc and *vice versa*. Here we describe the three-dimensional structure of human epimerase complexed with NADH and UDP-GlcNAc. To accommodate the additional *N*-acetyl group at the C2 position of the sugar, the side chain of Asn-207 rotates toward the interior of the protein and interacts with Glu-199. Strikingly, in the human enzyme, the structural equivalent of Tyr-299 in the *E. coli* protein is replaced with a cysteine residue (Cys-307) and the active site volume for the human protein is calculated to be ∼15% larger than that observed for the bacterial epimerase. This combination of a larger active site cavity and amino acid residue replacement most likely accounts for the inability of the *E. coli* enzyme to interconvert UDP-GlcNAc and UDP-GalNAc.

Enzymes belonging to the short chain dehydrogenase/reductase superfamily are widespread in nature and have been isolated from various sources including mammals, insects, and bacteria. Members of this superfamily catalyze a wide range of biochemical reactions with some displaying dehydrogenase activities, and others acting as dehydratases, isomerases, or epimerases, for example (1–3). These enzymes are typically around 250 amino acid residues in length and contain two characteristic signature sequences. The first of these is a Y*XXX*K motif in which the conserved tyrosine plays a key role in catalysis. The second of the signature sequences is a G*XXX*G*X*G motif, which is located near the cofactor-binding pocket.

UDP-galactose 4-epimerase, the focus of this investigation

and hereafter referred to as epimerase, is a member of the short chain dehydrogenase/reductase superfamily. As outlined in Scheme 1, the enzyme plays a pivotal role in the conversion of galactose to glucose 1-phosphate via the Leloir pathway. The specific function of epimerase in this metabolic pathway is to convert UDP-galactose back to UDP-glucose (step 3). Key features thought to be involved in the reaction mechanism of epimerase are indicated in Scheme 2 and include: 1) abstraction of the 4′-hydroxyl hydrogen of the sugar, most likely by the conserved tyrosine residue contained in the Y*XXX*K motif (4), 2) transfer of a hydride from C4 of the sugar to C4 of NAD^+ leading to a 4′-ketopyranose intermediate and NADH, and finally 3) rotation of the resulting 4′-ketopyranose moiety in the active site, thereby allowing return of the hydride from NADH to the opposite face of the sugar. Extensive x-ray crystallographic analyses with the enzyme obtained from *Escherichia coli* have supported this catalytic mechanism (5–9).

Recently, the structure of human epimerase was determined and refined to 1.5-Å resolution by x-ray diffraction (4). The human protein is a homodimer with each subunit containing 348 amino acid residues. One subunit of the "abortive complex" of the enzyme containing bound NADH and UDP-glucose is shown in Fig. 1. As can be seen, the polypeptide chain folds into two distinct regions: the N-terminal domain (Met-1 to Thr-189) responsible primarily for NAD^+/NADH positioning and the C-terminal motif (Gly-190 to Ala-348) involved in UDP-sugar binding. As expected for an enzyme that requires NAD^+, the N-terminal domain is characterized by a modified Rossmann fold of seven strands of parallel β-sheet flanked on either side by α-helices. Additionally, the characteristic Y*XXX*K motif of the short chain dehydrogenase/reductase superfamily is located in this domain (Tyr-157-Gly-Lys-Ser-Lys-161). The C-terminal portion of the epimerase subunit is composed of six β-strands and five α-helices.

Impairment of human epimerase results in epimerase deficiency galactosemia, a variant form of galactosemia with clinical severity that ranges from apparently benign to potentially lethal (10–13). Epimerase deficiency galactosemia can affect as many as 1/6700 individuals, at least in some ethnic groups (10, 14–15). In addition to catalyzing the interconversion of UDP-galactose and UDP-glucose, the human enzyme is also capable of interconverting UDP-GalNAc and UDP-GlcNAc (16–20). Strikingly, this activity has not been reported for the *E. coli* enzyme. To address the manner in which the human epimerase is able to bind and interconvert UDP-GalNAc and UDP-GlcNAc, the structure of the abortive complex of the protein with bound NADH and UDP-GlcNAc has been solved to 1.5-Å resolution by x-ray crystallographic methods. Here we describe the structure of this ternary complex and compare it to that of the bacterial enzyme.

Scheme 1

(1) galactokinase
(2) uridylyltransferase
(3) epimerase
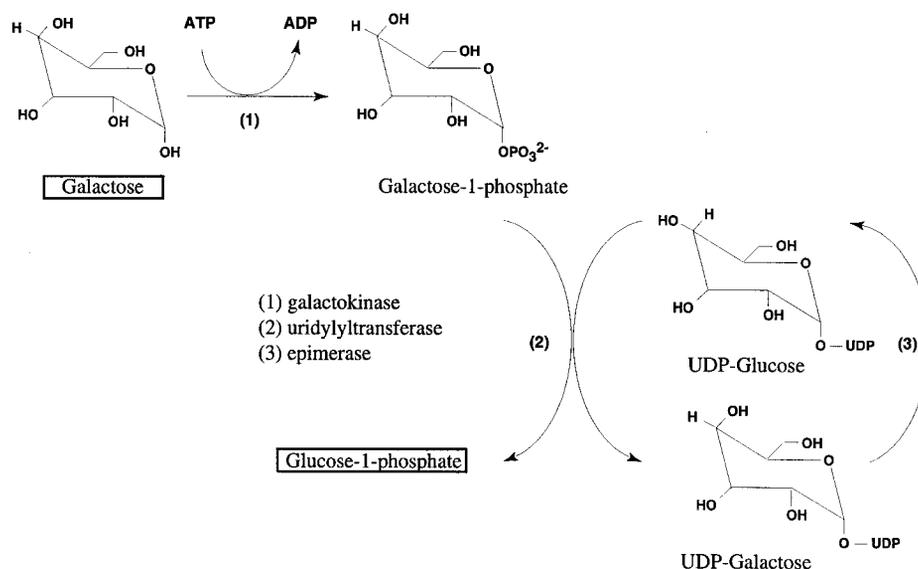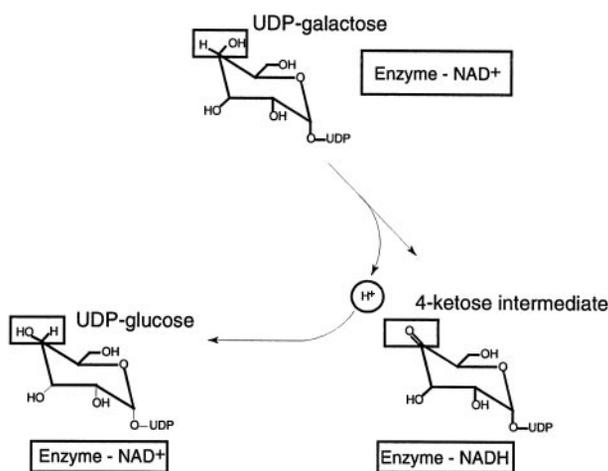
Scheme 2

EXPERIMENTAL PROCEDURES

Crystallization of the Epimerase/UDP-GlcNAc/NADH Abortive Complex—Protein employed for this investigation was purified as described previously (4). The ternary complex was prepared by treating the epimerase (15 mg/ml in the final dialysis buffer) with 5 mM NADH and 20 mM UDP-GlcNAc and allowing the solution to stand for 24 h at 4 °C. Large crystals were grown from 100 mM MES[1] (pH 6.0), 8–9% (w/v) poly(ethylene glycol) 3400, and 75 mM MgCl$_2$ by macroseeding into batch experiments at 4 °C. Typically, the crystals achieved maximum dimensions of 0.1 mm × 0.1 mm × 1.0 mm in ~2–3 weeks.

X-ray Structural Analysis of the Epimerase/UDP-GlcNAc/NADH Abortive Complex—For x-ray data collection, the crystals were first transferred to a synthetic mother liquor containing 15% poly(ethylene glycol) 3400, 75 mM MgCl$_2$, 250 mM NaCl, 5 mM NADH, and 20 mM UDP-GlcNAc buffered at pH 6.0 with 100 mM MES. The crystals were then transferred to a cryoprotectant solution in two steps: first to an intermediate solution containing 20% poly(ethylene glycol) 3400, 500 mM NaCl, and 20% (v/v) methanol and then to a similar solution that had been augmented with 4% (v/v) ethylene glycol. These crystals were suspended in a loop of 20-μm surgical thread and immediately flash-frozen in a stream of nitrogen gas.

The crystals belonged to the space group P2$_1$2$_1$2$_1$ with unit cell dimensions of a = 63.4 Å, b = 88.9 Å, and c = 118.7 Å. The asymmetric unit contained one dimer. A native x-ray data set to 1.5-Å resolution was collected at the Advanced Photon Source (Structural Biology Cen-

---

[1] The abbreviation used is: MES, 2-(*N*-morpholino)ethanesulfonic acid.

ter beamline 19-BM). These data were processed with HKL2000 and scaled with SCALEPACK (21). Relevant x-ray data collection statistics are presented in Table I. The structure was solved via AMORE (22) using the previously determined epimerase/UDP-glucose/NADH structure as the starting model (4). Initial refinement with the package TNT (23) reduced the *R*-factor to ~29.5% at 1.5-Å resolution. Manual adjustment of the model with the program Turbo (24) and subsequent refinement reduced the *R*-factor to 18.5% for all measured x-ray data from 30 to 1.5 Å. Relevant refinement statistics can be found in Table II. Electron density corresponding to the UDP-GlcNAc moiety in subunit II of the asymmetric unit is shown in Fig. 2.

RESULTS AND DISCUSSION

Molecular Structure of the Human Epimerase/UDP-Glc-NAc/NADH Complex—A close-up view of the human epimerase active site with bound UDP-GlcNAc (subunit II) is shown in Fig. 3*a*. Nineteen water molecules are located within 4.0 Å of the NADH and UDP-GlcNAc moieties. Human epimerase is known to be a *B*-side-specific enzyme. As expected, the nicotinamide ring of the NADH is in the *syn*-conformation with its *si*-face oriented toward the sugar ligand. Specifically, the distance between C4 of the UDP-GlcNAc ligand and C4 of the nicotinamide ring of the NADH is 3.0 Å. This distance in the previously determined human epimerase/UDP-glucose/NADH complex is somewhat longer at 3.5 Å (4). As indicated by the *dashed lines*, the 4′-hydroxyl group of the sugar is located at 2.8 Å from O$^\gamma$ of Ser-132 and 3.0 Å from O$^\eta$ of Tyr-157. In the epimerase/UDP-glucose/NADH model, the distance between the 4′-hydroxyl group of the sugar and O$^\gamma$ of Ser-132 is somewhat shorter at ~2.4 Å while the distance between the sugar hydroxyl group and O$^\eta$ of Tyr-157 is comparable at 3.1 Å (4). This close distance of O$^\eta$ of Tyr-157 to the 4′-hydroxyl group of the sugar is suggestive for the role of Tyr-157 as the catalytic base that abstracts the hydrogen during normal catalysis.

A schematic of the hydrogen bonding pattern between the protein and the sugar substrate is displayed in Fig. 3*b*. Eight water molecules lie within hydrogen bonding distance to the ligand. Key amino acid side chains responsible for positioning the sugar within the active site include Ser-132, Tyr-157, Asn-187, Arg-239, Arg-300, and Asp-303. Additionally, the backbone carbonyl or peptidic NH groups of Lys-92, Leu-208, Asn-224, and Phe-226 form hydrogen bonds with the UDP-GlcNAc moiety.

The epimerase/UDP-GlcNAc/NADH abortive complex crystallized with two molecules in the asymmetric unit. Interestingly, the α-carbons for the two subunits superimpose with

FIG. 1. **Ribbon representation of one subunit of human UDP-galactose 4-epimerase.** This figure and Figs. 3–6 were prepared with the program MOLSCRIPT (27). X-ray coordinates utilized for this figure were obtained from the Protein Data Bank (1EK6). Bound UDP-glucose and NADH are displayed in *ball-and-stick* representations. The N-terminal domain, delineated by Met-1 to Thr-189, is shown in *blue*, and the C-terminal motif, formed by Gly-190 to Ala-348, is displayed in *green*. Note that the seventh β-strand in the Rossmann fold is contributed by the C-terminal domain.
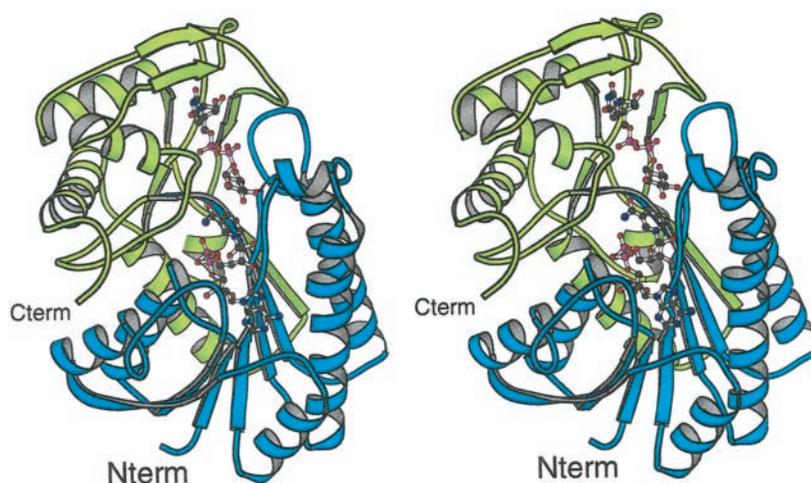
TABLE I

X-ray data collection statistics

Resolution	Independent, reflections	Completeness	Redundancy	Average *I*/average $\sigma(I)$	R_{sym}[a]
Å		%			%
30.0–1.50	104,972	97.0	5.4	26.8	4.7
1.55–1.50	8999	84.6	3.6	2.9	28.6

[a] $R_{\text{sym}} = (\Sigma|I - \bar{I}|/\Sigma I) \times 100$.

[b] Statistics for the highest resolution bin.

TABLE II

Relevant least-squares refinement statistics

Resolution limits (Å)	30.0–1.50
R-factor[a] (overall) %/# rflns	18.5/104,792
R-factor (working) %/# rflns	18.4/96,932
R-factor (free) %/# rflns	22.7/7860
No. of protein atoms	4979[b]
No. of hetero-atoms	1105[c]
Weighted root mean square deviations from ideality	
Bond lengths (Å)	0.011
Bond angles (degrees)	2.10
Trigonal planes (Å)	0.006
General planes (Å)	0.011
Torsional angles (degrees)	15.6

[a] *R*-factor = $(\Sigma|F_o - F_c|/\Sigma|F_o|) \times 100$, where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude.

[b] These include multiple conformations for Glu-63, Gln-75, Thr-134, Lys-291, and Ser-312 in subunit I and Thr-110, Cys-196, Asp-229, Gln-261, and Arg-328 in subunit II.

[c] These include 931 water molecules, 2 NADH molecules, 2 UDP-GlcNAc molecules, 3 chloride ions, and 1 magnesium ion.

quite a large root mean square deviation of 1.1 Å. The major differences between the two polypeptide chains, however, are confined to the N and C termini and to the surface loop delineated by Gly-41 to Gly-45. Shown in Fig. 4 is a close-up view of the superposition of the two active sites contained within the asymmetric unit. Quite unexpectedly, the sugar moieties of the two ligands adopt substantially different orientations. In subunit I, displayed in *red*, C4 of the UDP-GlcNAc group is positioned at 4.6 Å from C4 of the nicotinamide ring of the dinucleotide. Additionally, its 4′-hydroxyl group lies at 3.3 Å from O^γ of Ser 132 and 4.6 Å from O^η of Tyr-157. Due to differences in the torsional angles about the phosphate backbones of the UDP-sugars, the GlcNAc group in subunit I swings out more toward the solvent and the *N*-acetyl group attached to C2 of the sugar adopts two different conformations as depicted in Fig. 4. Clearly these observed differences in sugar binding are in agreement with a catalytic mechanism that requires free rotation of a 4′-ketopyranose intermediate within the active site of the enzyme.

Comparison of the Human Epimerase/UDP-Glucose/NADH and Epimerase/UDP-GlcNAc/NADH Complexes—One of the questions to be addressed by this study is the manner in which the human epimerase can accommodate both UDP-glucose and UDP-GlcNAc. A superposition of the two structures, namely the epimerase/UDP-glucose/NADH (subunit II) and the epimerase/UDP-GlcNAc/NADH (subunit II) abortive complexes, is displayed in Fig. 5. The α-carbons for these two models superimpose with a root mean square deviation of 0.50 Å. Clearly the more bulky *N*-acetyl group attached to C2 of the sugar is accommodated within the epimerase active site by a simple rotation of the carboxamide side chain group of Asn 207.

Comparison of the Human Epimerase/UDP-GlcNAc/NADH, the E. coli Epimerase/UDP-Glucose/NADH, and the E. coli Epimerase/UDP-Galactose/NADH Complexes—By far the most extensive structural studies on UDP-galactose 4-epimerase to date have been conducted on the enzyme from *E. coli*. In one study, the structure of the bacterial enzyme/NADH/UDP-glucose abortive complex was solved in order to address the manner in which a sugar ligand binds within the active site (7). Interestingly, all attempts to produce a similar abortive complex of the *E. coli* enzyme in the presence of UDP-galactose failed. Although crystals could be grown of the putative enzyme/UDP-galactose/NADH species, after x-ray data collection and processing, UDP-glucose rather than UDP-galactose was always observed binding in the active site. It was eventually possible to study the manner in which UDP-galactose binds to the bacterial enzyme, however, by preparing a "double" site-directed mutant protein in which Ser-124 and Tyr-149 were changed to alanine and phenylalanine residues, respectively. The x-ray crystallographic analysis of this "double" site-directed mutant protein demonstrated the manner in which UDP-galactose binds in the active site cleft (9). From these analyses, it was shown that in the *E. coli* epimerase, Asn-179 (which is the structural equivalent to Asn-187 in the human enzyme) serves to hydrogen bond to either the 6′- or the 2′-hydroxyl groups when UDP-glucose or UDP-galactose, respectively, are bound in the active site. Additionally, in the *E. coli* enzyme, Tyr-299 functions in sugar positioning where it also

FIG. 2. **Difference electron density corresponding to the UDP-GlcNAc moiety in subunit II of the asymmetric unit.** The map was contoured at 3σ and calculated with coefficients of the form $(F_o - F_c)$, where F_o was the native structure factor amplitude and F_c was the calculated structure factor amplitude. The UDP-GlcNAc ligand was omitted from the x-ray coordinate file for the electron density map calculation. This figure was produced by the software package BobScript (28).
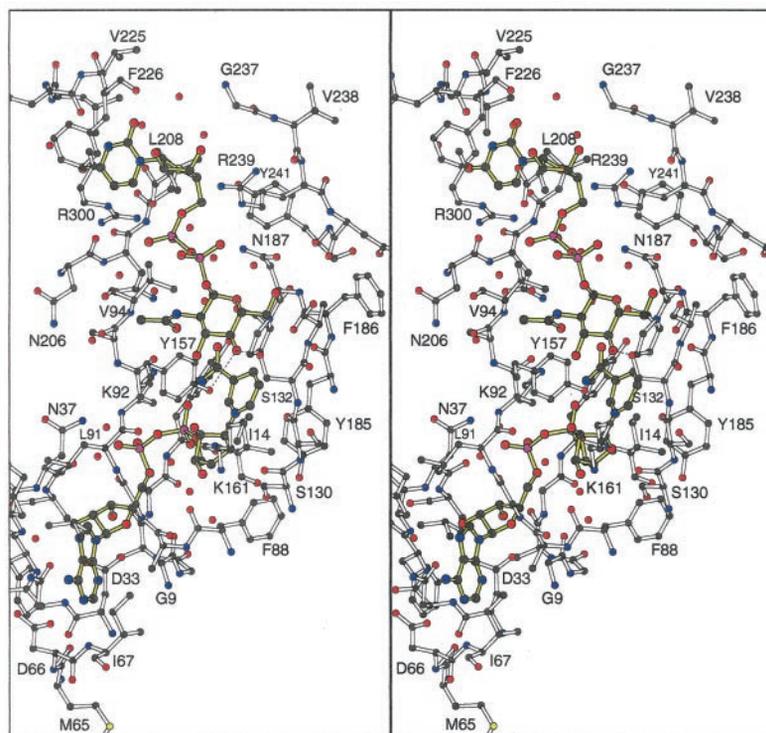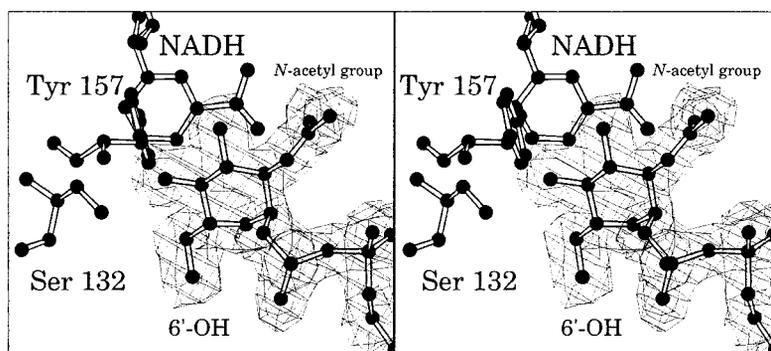
(a)

FIG. 3. **The active site for human epimerase with bound NADH and UDP-GlcNAc.** Those amino acid residues and solvents located within ~4 Å of the ligands are shown in *a*. Both NADH and UDP-GlcNAc are highlighted in *yellow* bonds. Hydrogen bonds between the 4′-hydroxyl group of the sugar and the side chains of Ser-132 and Tyr-157 are indicated by the *dashed lines*. A schematic of the hydrogen bonding pattern around the sugar substrate is given in *b*. The *dashed lines* indicate distances equal to or less than 3.2 Å. The side chain oxygens of Ser-132 and Tyr-157 are separated by 4.0 Å, as indicated by the *dashed line*.
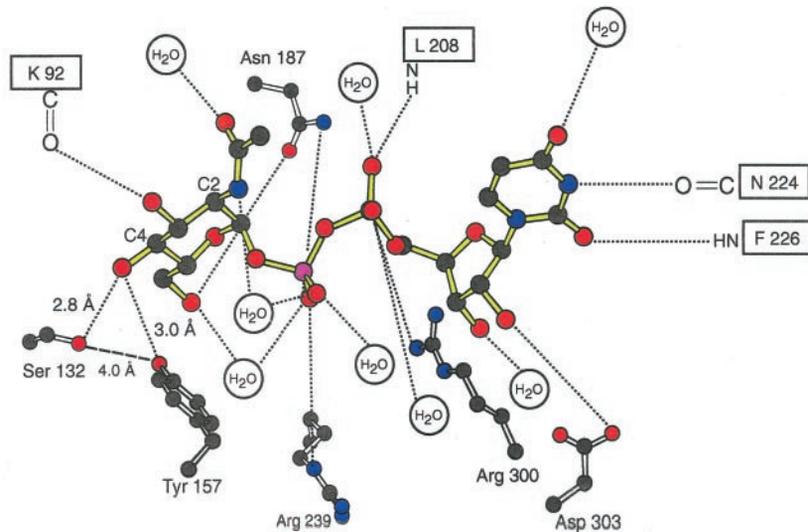
(b)

FIG. 4. **Superposition of the two active sites for the human epimerase dimer contained within the asymmetric unit.** Subunit I in the x-ray coordinate file is depicted in *red*, and subunit II is displayed in *black*. The *asterisk* marks the position of C4 in the sugar moiety bound to subunit I.
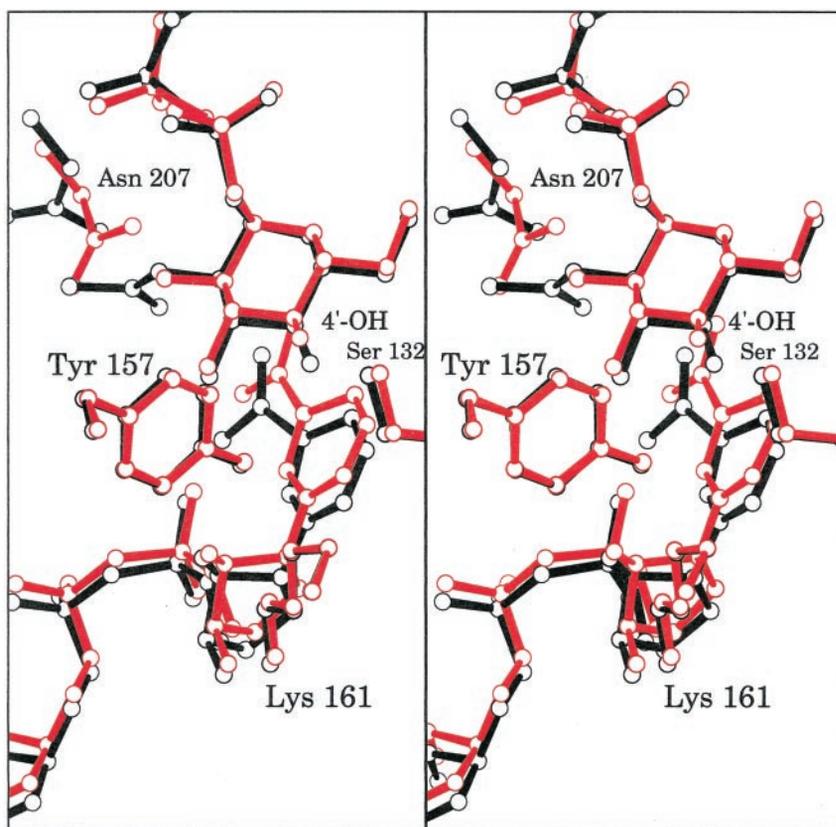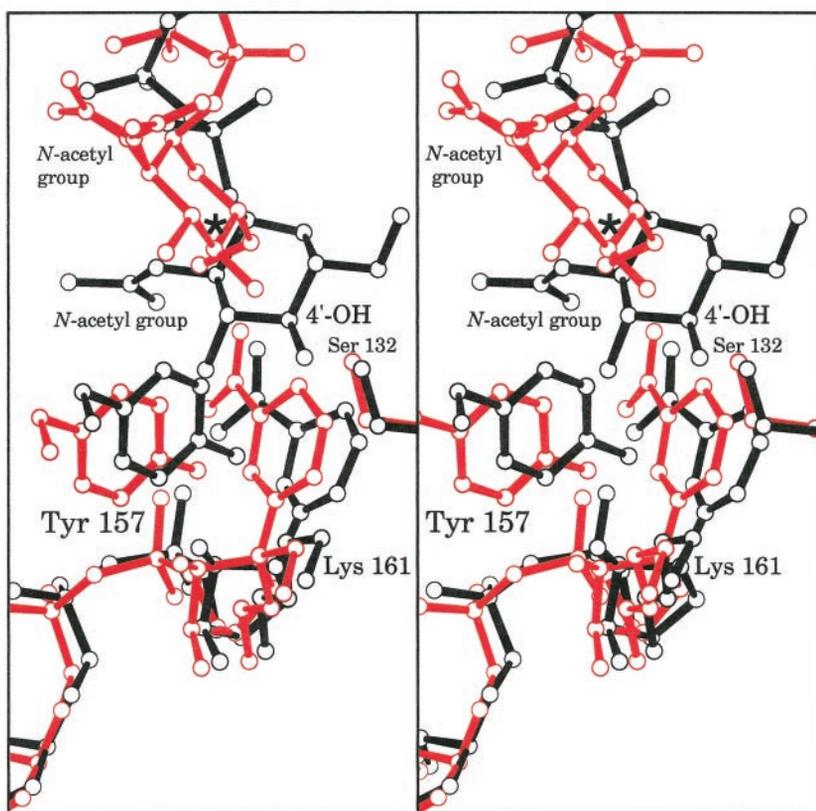

FIG. 5. **Superposition of the active sites for the abortive complexes of human epimerase with either UDP-glucose or UDP-GlcNAc bound.** The abortive complexes with bound UDP-glucose or UDP-GlcNAc are color-coded in *red* and *black*, respectively. Note the movement of the side chain carboxamide group of Asn-207.

hydrogen bonds to either the 6′- or 2′-hydroxyl groups of UDP-glucose or UDP-galactose, respectively. It is particularly noteworthy that Tyr-299 is not conserved in the human enzyme but rather is replaced with a cysteine residue (Cys-307).

A superposition of the *E. coli* epimerase with bound UDP-glucose onto the human epimerase with bound UDP-GlcNAc is shown in Fig. 6. The α-carbons for these two complexes correspond with a root mean square deviation of 1.2 Å. Note that the sugar moieties of the UDP-glucose bound to the *E. coli* enzyme (highlighted in *blue*) and the UDP-GlcNAc bound to the human

FIG. 6. **Superposition of human epimerase/UDP-GlcNAc/NADH structure onto the abortive complex models of the *E. coli* epimerase with bound UDP-glucose or UDP-galactose.** The human protein is shown in *black*, and the *E. coli* protein with bound UDP-glucose is depicted in *blue* and with bound UDP-galactose highlighted in *red*. The *numbering* corresponds to the human epimerase. Note especially the replacement of Cys-307 in the human enzyme by a tyrosine residue (Tyr-299) in the bacterial protein.
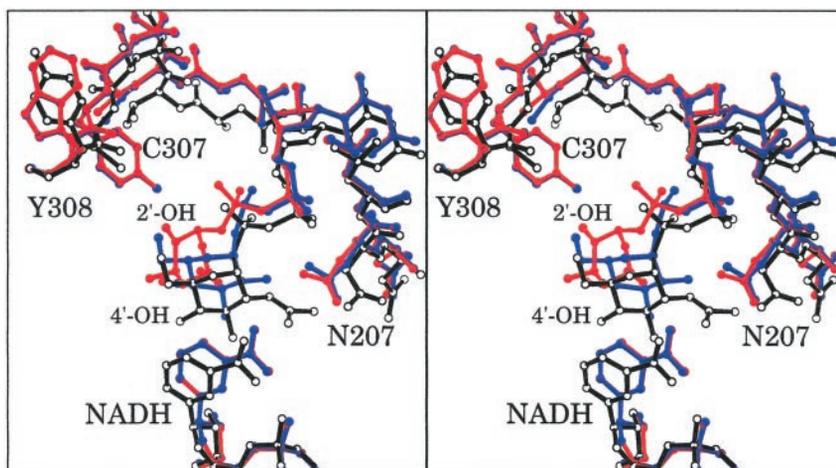
epimerase (outlined in *black*) adopt similar orientations. Also shown in Fig. 6 is the conformation of UDP-galactose when it is positioned in the active site of the *E. coli* enzyme (displayed in *red*). From Fig. 6 it is clear that if UDP-GalNAc binds to the human enzyme in a similar manner to that observed for UDP-galactose in the bacterial protein, the *N*-acetyl group of the sugar ligand will lie quite close to the sulfhydryl group of Cys-307. Attempts to grow crystals of an abortive complex of the human enzyme with NADH and UDP-GalNAc have thus far been unsuccessful. Preparation of the "double" site-directed mutant protein of the human enzyme is presently in progress whereby Ser-132 and Tyr-157 are being changed, respectively, to alanine and phenylalanine residues. An x-ray analysis of this double site-directed mutant protein should clarify the manner in which UDP-GalNAc binds within the active site of the human enzyme.

What is absolutely clear from this investigation is why the human enzyme can interconvert UDP-GlcNAc and UDP-GalNAc while the bacterial enzyme has not been reported to display such activity. Calculations with the program VOIDOO (25, 26) indicate that the active site for the human enzyme, as compared with that of the bacterial protein, is ~15% larger in the region responsible for sugar binding. Furthermore, the replacement of Cys-307 in the human enzyme with the more bulky Tyr-299 in the bacterial protein most likely precludes UDP-GalNAc from binding in the *E. coli* active site in a productive mode. By the judicious use of various site-directed mutant proteins, it should be possible to test this hypothesis. Indeed, experiments are presently under way to construct a form of the *E. coli* epimerase that is catalytically active toward UDP-GlcNAc and UDP-GalNAc.

REFERENCES

1. Jörnvall, H., Persson, B., Krook, M., Atrian, S., González-Duarte, R., Jeffery, J., and Ghosh, D. (1995) *Biochemistry* **34,** 6003–6013
2. Duax, W. L., Griffin, J. F., and Ghosh, D. (1996) *Curr. Opin. Struct. Biol.* **6,** 813–823
3. Duax, W. L., Ghosh, D., and Pletnev, V. (2000) *Vitam. Horm.* **58,** 121–148
4. Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L., and Holden, H. M. (2000) *Biochemistry* **39,** 5691–5701
5. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) *Protein Sci.* **5,** 2149–2161
6. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) *Biochemistry* **35,** 2557–2566
7. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) *Biochemistry* **35,** 5137–5144
8. Thoden, J. B., Gulick, A. M., and Holden, H. M. (1997) *Biochemistry* **36,** 10685–10695
9. Thoden, J. B., and Holden, H. M. (1998) *Biochemistry* **37,** 11469–11477
10. Holton, J. B., Walter, J. H., and Tyfield, L. A. (2000) in *Metabolic and Molecular Basis of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, S. W., Valle, D., Childs, B., Kinzler, K. W., and Vogelstein, B., eds) pp. 1553–1587, McGraw Hill, New York, in press
11. Walter, J. H., Roberts, R. E. P., Besley, G. T. N., Wraith, J. E., Cleary, M. A., Holton, J. B., and MacFaul, R. (1999) *Arch. Dis. Child.* **80,** 374–376
12. Quimby, B. B., Alano, A., Almashanu, S., DeSandro, A. M., Cowan, T. M., and Fridovich-Keil, J. L. (1997) *Am. J. Hum. Genet.* **61,** 590–598
13. Shin, Y. S., Korenke, G. C., Huppke, P., Knerr, I., and Podskarbi, T. (2000) *J. Inherit. Metab. Dis.* **23,** 383–386
14. Alano, A., Almashanu, S., Maceratesi, P., Reichardt, J., Panny, S., and Cowan, T. M. (1997) *J. Invest. Med.* **45,** 191A
15. Gitzelmann, R., Steinmann, B., Mitchell, B., and Haigis, E., (1977) *Helv. Paediatr. Acta* **31,** 441–452
16. Wohlers, T. M., Christacos, N. C., Harreman, M. T., and Fridovich-Keil, J. L. (1999) *Am. J. Hum. Genet.* **64,** 462–470
17. Maley, F., and Maley, G. F. (1959) *Biochim. Biophys. Acta* **31,** 577–578
18. Piller, F., Hanlon, M. H., and Hill, R. L. (1983) *J. Biol. Chem.* **258,** 10774–10778
19. Kingsley, D. M., Kozarsky, K. F., Hobbie, L., and Krieger, M. (1986) *Cell* **44,** 749–759
20. Wohlers, T., and Fridovich-Keil, J. L. (2000) *J. Inherit. Metab. Dis.* **23,** 713–729
21. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
22. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50,** 157–163
23. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. Sect. A.* **43,** 489–501
24. Roussel, A., and Cambillau, C. (1991) in *Silicon Graphics Geometry Partners Directory*, Silicon Graphics, Mountain View, CA
25. Kleywegt, G. J., and Jones, T. A. (1993) *Newsl. Protein Crystallogr,* **29,** 26–28
26. Kleywegt, G. J., and Jones, T. A. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 178–185
27. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
28. Esnouf, R. (1996) *Acta Crystallogr A* **D55,** 938–940