Structural Analysis of the Y299C Mutant of *Escherichia coli* UDP-galactose 4-Epimerase

TEACHING AN OLD DOG NEW TRICKS*

James B. Thoden‡, Jenny M. Henderson§§, Judith L. Fridovich-Keil, and Hazel M. Holden‡‡**

From the ‡Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706-1544 and the §§Graduate Program in Nutrition and Health Sciences and the ‡Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia 30322

UDP-galactose 4-epimerase catalyzes the interconversion of UDP-Gal and UDP-Glc during normal galactose metabolism. The mammalian form of the enzyme, unlike its *Escherichia coli* counterpart, can also interconvert UDP-GalNAc and UDP-GlcNAc. One key feature of the epimerase reaction mechanism is the rotation of a 4-ke- topyranose intermediate in the active site. By comparing the high resolution x-ray structures of both the bacterial and human forms of the enzyme, it was previously postulated that the additional activity in the human epimerase was due to replacement of the structural equivalent of Tyr-299 in the *E. coli* enzyme with a cysteine residue, thereby leading to a larger active site volume. To test this hypothesis, the Y299C mutant form of the *E. coli* enzyme was prepared and its three-dimensional structure solved as described here. Additionally, the Y299C mutant protein was assayed for activity against both UDP-Gal and UDP-GalNAc. These studies have revealed that, indeed, this simple mutation did confer UDP-GalNAc/UDP-GlcNAc converting activity to the bacterial enzyme with minimal changes in its three-dimensional structure. Specifically, although the Y299C mutation in the bacterial enzyme resulted in a loss of epimerase activity with regard to UDP-Gal by almost 5-fold, it resulted in a gain of activity against UDP-GalNAc by more than 230-fold.

Within recent years, the so-called short chain dehydrogenase/reductase (SDR)1 superfamily has emerged as an important category of proteins and enzymes with more than 1600 examples being identified from a host of organisms (1). Typically, members of this superfamily catalyze oxidation-reduction type reactions with some functioning, for example, as dehydrogenases, dehydratases, isomerases, or epimerases (2–4). In most cases, enzymes belonging to this superfamily are typically around 250 amino acid residues in length and contain two characteristic signature sequences. The first of these is a YXK motif in which the conserved tyrosine plays a key role in catalysis. The second of the signature sequences is a GXXXGXXG motif, which is located near the cofactor-binding pocket. Recently, on the basis of amino acid sequence homologies, it has been suggested that other members of the SDR superfamily may not function as enzymes or, at least, not in oxidation-reduction type capacities (5). One of these proteins, human TIP30, is a cofactor that enhances human immunodeficiency virus-1 Tat-activated transcription (6). Interestingly, TIP30 is 98% identical to human CCS, a protein associated with suppression of metastasis in small cell lung carcinomas (7). The recent and elegant structural analyses by Stammers et al. (8) have demonstrated that NmrA from *Aspergillus nidulans*, a negative transcriptional regulator, also belongs to the SDR protein superfamily.

One of the first members of the SDR protein superfamily to be extensively studied by high resolution x-ray crystallographic analyses was UDP-galactose 4-epimerase from *Escherichia coli*, which is somewhat larger with 338 amino acid residues per subunit (Ref. 9 and references therein). This dimeric enzyme functions in normal galactose metabolism by catalyzing the interconversion of UDP-Gal and UDP-Glc as indicated in Scheme 1. Within the last three years, the molecular architecture of the human form of the enzyme complexed with NADH and UDP-Glc has also been determined to high resolution (1.5 Å) and has led to the suggestion that Tyr-157 serves as the active site base required for catalytic activity (10). Both the bacterial and human epimerases are B-side-specific enzymes. As indicated in Fig. 1, the polypeptide chain of each subunit of human UDP-galactose 4-epimerase, hereafter referred to simply as epimerase, adopts a distinctly bilobal structure. The N-terminal domain is characterized by a seven-stranded, parallel β-sheet, whereas the smaller C-terminal domain contains six strands of β-sheet with five α-helices. The active site is wedged between the interface of these two domains.

Unlike the bacterial enzyme, human epimerase can interconvert UDP-GalNAc and UDP-GlcNAc as well UDP-Glc and UDP-Gal (11–15). Strikingly, however, the two enzymes are 55% identical with respect to amino acid sequence. A recent structural analysis of the human epimerase complexed with NADH and UDP-GlcNAc demonstrated that accommodation of the additional N-acetyl group at the C2 position of the sugar is accomplished by movement of the side chain of Asn-207 (16). Additionally, from this investigation it was shown that the active site volume of the human protein is ~15% larger than that observed for the *E. coli* enzyme.

The reaction mechanism of epimerase is thought to occur via the abstraction of the 4'-hydroxyl hydrogen of the sugar by the active site tyrosine, transfer of a hydride from C4 of the sugar to C4 of NAD+ (si-face) to generate a 4'-ketopyranose interme-
NADH, and finally rotation of the resulting 4'-keto- 
pyranose moiety in the active site, whereby presenting the op- 
posite face of the sugar to the reduced dinucleotide for 
subsequent hydride transfer back to the UDP-sugar substrate. 
A superposition of the human epimerase/NADH/UDP-GlcNAc 
structure onto the model of the bacterial enzyme with bound 
NADH and UDP-Glc is displayed in Fig. 2. Note that Cys-307 
in the human enzyme is replaced with a tyrosine residue (Tyr- 
299) in the enzyme from E. coli. This substitution of the more 
bulky tyrosine residue in the active site of the bacterial enzyme 
could partly explain the inability of the E. coli enzyme to 
terconvert UDP-GlcNAc and UDP-GalNAc due to lack of the 
space required for rotation of the 4'-keto-pyranose interme-
diate. To test this hypothesis, Tyr-299 in the E. coli enzyme was 
changed to a cysteine residue via site-directed mutagenesis. 
Here we present both enzymatic assays and high resolution 
x-ray crystallographic data demonstrating that replacement of 
Tyr-299 with a cysteine residue confers activity on the E. coli 

**EXPERIMENTAL PROCEDURES**

**Construction of the Plasmid Containing the Y299C Mutation**—The 
plasmid containing the epimerase Y299C mutation was constructed via 
cassette mutagenesis using the plasmid pTZSynE (17). The Y299C 
oligonucleotides were designed as follows: 1) Y299C-top, 5'-GGCGTG-
AAAGTGACCTTCCGCGTGACCTTCCGGCGTGCTGGGCGGACG-3' 
and 2) Y299C-bottom, 5'-CTAGCGTCCGCCAGCACGCCGGAAGGTCACCTTCAC-GE-
GCACG-3' (Invitrogen). These oligonucleotides were utilized to give 
overhangs compatible with SacI and NheI restriction sites at the 5'- 
and 3'-ends. They were boiled for 5 min and then slowly cooled to 4 °C 
to form the double-stranded cassette. The cassette was subsequently 
phosphorylated at the 5'-ends with polynucleotide kinase. The 
pTZSynE plasmid was digested with SacI and NheI, treated with calf 
intestinal phosphatase, and purified with the GeneClean DNA purifi-
cation kit. The mutagenic cassette was ligated into the cut pTZSynE 
plasmid with T4 DNA ligase.

**Protein Purification and Crystallization**—The wild-type and Y299C 
mutant epimerases were purified as previously described (18). Three 
different protein/NADH/UDP-sugar complexes were crystallized for 
this investigation, namely the complex of the wild-type enzyme with 
NADH and UDP-GlcNAc and the complexes of the Y299C protein with 
NADH and either UDP-GlcNAc or UDP-Glc. Note that attempts to 
solve the structures of the Y299C protein with NADH and either UDP-
GalNAc or UDP-Gal resulted in either UDP-GlcNAc or UDP-Glc, re-
spectively, being observed in the resulting electron density maps. 
Clearly the mutant enzyme is catalytically active in the crystalline 
state.

For the preparation of the above complexes, solid UDP-sugar was 
added to either the wild-type or Y299C protein samples to a final 
concentration of 20 mM. The samples were then allowed to stand at 4 °C

![Fig. 1. Ribbon representation of one subunit of human UDP-galactose 4-epimerase complexed with NADH and UDP-Glc. The N- and C-terminal domains are displayed in green and light green, respectively, and the binding positions for the NADH and UDP-Glc molecules are indicated by the ball-and-stick models. Tyr-157 is the catalytic base that abstracts the 4'-hydroxyl hydrogen during the reaction mechanism. All figures were prepared with the software package, MOLSCRIPT (28).](image1)

![Scheme 1](image2)

![Fig. 2. Superposition of the binding pockets for the UDP-sugars in the E. coli and human epimerases. The human and bacterial enzymes are shown in black and gray, respectively. The structure of the bacterial enzyme was solved in the presence of UDP-Glc, whereas that of the human protein was determined in the presence of UDP-GlcNAc. Note that in the bacterial enzyme, Tyr-299 forms a hydrogen bond with either the 6'-hydroxyl of glucose (shown here) or the 2'-hydroxyl group of galactose (9).](image3)
for 3 days. Following this incubation, they were dialyzed against 10 mM potassium phosphate (pH 8.0) and concentrated to 25 mg/ml.

Single crystals of all three complexes were grown by the hanging drop method of vapor diffusion at 4°C. The precipitant was 15–19% polyethylene glycol 8000, 600 mM NaCl, 20 mM UDP-sugar, 20% ethylene glycol, and 100 mM HEPES (pH 7.5). Crystal growth generally required 4–7 days with crystals attaining maximum dimensions of 0.6 × 0.4 × 0.3 mm. All of the crystals belonged to the space group P2₁2₁2 with typical unit cell dimensions of a = b = 83.2 Å and c = 109.5 Å and one molecule per asymmetric unit.

X-ray Data Collection and Processing—Prior to x-ray data collection, the crystals were transferred to a cryoprotectant solution composed of 24% polyethylene glycol 8000, 600 mM NaCl, 20 mM UDP-sugar, 20% ethylene glycol, and 100 mM HEPES (pH 7.5). The crystals were suspended in 20-μm nylon loops and flash-cooled to −150°C in a nitrogen gas stream. All x-ray data sets were collected with a HiStar (Bruker AXS) area detector system using CuKα radiation generated from a Rigaku RU200 rotating anode generator operated at 50 kV and 90 mA and equipped with Supper long mirrors. The x-ray data were processed with SAINT (Bruker AXS, Inc.) and internally scaled with XSCALI-BRE.2 X-ray data collection statistics are presented in Table I.

Structure Determinations and Refinements—All of the structures were solved by Difference Fourier techniques employing the previously determined structure of the bacterial epimerase/NADH/UDP-Glcaborative complex (19) as the starting model. The coordinates for the solvents, UDP-sugars, and/or point mutation were removed from the starting models and rebuilt with TURBO (20) during the course of the least-squares refinements with the software package, TNT (21). Relevant refinement statistics are presented in Table II.

In Vitro Assays—Aliquots of each purified enzyme were stored in 50% glycerol with 4 mM NADH in liquid nitrogen until needed. Assays to determine the level of epimerase activity with respect to UDP-Gal and UDP-GalNAc were performed essentially as previously described (11). Briefly, the conversion of UDP-Gal to UDP-Glc was measured in a 12.5-μl reaction containing 2.5 μl of premix (0.05 μl of UDP-[14C]Gal (Amersham Biosciences), 2 mM cold UDP-Gal, 0.2 mM glycine buffer, pH 8.7), 2.5 μl of 20 mM NAD+, and 7.5 μl of purified protein diluted in Johnston buffer (20 mM HEPES/KOH, pH 7.5, 1 mM diithiothreitol, and 0.3 μg of bovine serum albumin/ml). The amount of protein used per reaction ranged from 0.0001 to 0.005 μg, in order to stay within the predetermined linear range of the assay. Reactions were incubated at 37°C for 30 min and were stopped by boiling at 100°C for 10 min. Following high speed centrifugation for 15 min in a microcentrifuge, 10 μl of the sample was spotted onto a prewashed PEI-Cellulose TLC plate (Baker). After thorough drying, the plate was run for 16–24 h in a solvent containing 1.5 mM Na2B4O7, 5 mM H3BO3, and 20% ethylene glycol. After running, plates were air-dried before being exposed to

---

2 I. Rayment and G. Wesenberg, unpublished data.
storage phosphor screens (Amersham Biosciences) overnight. Images were visualized with a Typhoon 9200 variable mode imager and quantified using ImageQuant software (both from Amersham Biosciences). Percent conversion was determined by dividing the product signal by the total signal and multiplying by 100. This number was converted to specific activity using the formula: ((% conversion/100) × 5.085 nmol of substrate)/(µg of enzyme × 30 min).

The conversion of UDP-GalNAc to UDP-GlcNAc was determined essentially as described above, with the following assay components per 25 µl of reaction: 8.75 µl of premix (0.04 µg of UDP-[14C]GalNAc (ICN), 1.89 nmol UDP-GalNAc, 28.6 mM pyruvate, 286 mM glycine, pH 8.7, 5 µl of 20 mM NAD) and 11.25 µl of protein diluted in Johnston buffer. The amount of protein used in each assay ranged from 0.0005 to 2.0 µg. Assays were performed as for UDP-Gal, with a TLC run-time of 10 h and quantified as described for UDP-Gal. The percentage conversion was translated to specific activity using the following formula: ((% conversion/100) × 16.32909 nmol of substrate)/(µg of enzyme × 30 min).

RESULTS AND DISCUSSION

To confirm the difference in substrate specificities of the wild-type human and E. coli epimerases, as well as to define the catalytic impact of the Y299C mutation in the bacterial enzyme, in vitro enzymatic activity analyses were conducted with purified proteins and each of the substrates in question, UDP-Gal and UDP-GalNAc. With regard to the wild-type human and E. coli enzymes, the results presented in Table III clearly confirm earlier reports (11–15, 22–24) indicating that although human UDP-galactose 4-epimerase demonstrates strong activity against both substrates, the bacterial enzyme is active only against UDP-Gal. Indeed, the wild-type bacterial enzyme is more than 8,000-fold less active against UDP-GalNAc than it is against UDP-Gal. In contrast, the Y299C bacterial enzyme demonstrated significant activity against both substrates (Table III). Although the Y299C mutation in the bacterial enzyme resulted in a loss of epimerase activity with regard to UDP-Gal by almost 5-fold, it resulted in a gain of activity with regard to UDP-GalNAc by 230-fold.

As a control experiment, the first complex to be solved in this study involved the wild-type enzyme. In light of past x-ray crystallographic studies of the binding of UDP-sugar moieties such as UDP-mannose or UDP-4-deoxy-4-fluoro-α-D-galactose to E. coli epimerase (25), it was anticipated a priori that little conformational change would accompany the binding of UDP-GlcNAc in the active site of the wild-type enzyme. In retrospect that assumption was perhaps naive in that the ligand, UDP-GlcNAc, contains a rather bulky group attached to C2 of the hexose. The electron density corresponding to the sugar moiety is indeed disordered in the map, thus indicating that it adopts multiple conformations when bound in the epimerase active site. The electron density corresponding to the UDP portion of the ligand is unambiguous, however. Additionally, the electron density corresponding to the polypeptide chain between Asn-131 and Ile-134 is weak.

There are several regions in the polypeptide chain that change in the presence of UDP-GlcNAc versus UDP-Glc, as can be seen by the superposition displayed in Fig. 3. Indeed, the α-carbons for the epimerase/NADH/UDP-Glc and epimerase/NADH/UDP-GlcNAc models superimpose with a root-mean-square deviation of 0.52 Å. Most of the changes are confined primarily to the regions delineated by Phe-121 to Asn-131, Tyr-177 to Pro-180, Gln-194 to Leu-200, Pro-218 to Arg-231, and Arg-292 to Asp-302. All of these regions contain either side chains or backbone atoms that participate in hydrogen-bonding interactions with the UDP-sugars and, thus, are sensitive to the chemical identity of the ligand positioned in the active site. In the previously determined wild-type epimerase/NADH/UDP-Glc model, for example, O′ of Ser-124 lies within 2.6 Å of the 4′-hydroxyl group of the glucose moiety, while the side chains for both Asn-179 and Tyr-299 form hydrogen bonds with the 6′-hydroxyl group of glucose (9).

The movement of the side chain carboxamide group of Asn-199, as shown in Fig. 3, is especially striking. When UDP-Glc is bound in the active site, both the 2′-hydroxyl group of the sugar and an α-phosphoryl oxygen are hydrogen-bonded to N65 of UDP-GlcNAc.
Asn-199. However, with UDP-GlcNAc bound in the active site, this interaction is missing, and the side chain of Asn-199 rotates toward the nicotinamide ring of NADH. There is a general shift of 0.4 Å of the nucleoside portions of UDP-GlcNAc versus UDP-Glc in the active site. The lack of activity of the E. coli epimerase against UDP-GlcNAc or UDP-GalNAc most likely arises from the following two key factors: (a) the inability of the larger sugar moiety to correctly position into the active site and (b) the inability of the 4′-ketopyranose intermediate to rotate in the active site.

The second structure that was solved in this investigation was that of the Y299C mutant protein complexed with NADH and UDP-GlcNAc. The electron density corresponding to the region surrounding the mutation is well ordered, as displayed in Fig. 4a. In fact, the molecular changes that occur upon substituting a cysteine residue for Tyr-299 are localized in the regions defined by Arg-176 to Asn-179, Ala-216 to Arg-231, and His-287 to Cys-299. The net effect of these changes, as indicated in Fig. 5a, is an opening of the substrate binding cleft, which results in an overall 20% increase in the volume of the active site for the Y299C mutant protein as calculated with the program VOIDOO (26, 27). The α-carbons for the wild-type enzyme with bound UDP-Glc and the Y299C mutant protein with bound UDP-GlcNAc correspond with a root-mean-square deviation of 0.96 Å.

A portion of the electron density map corresponding to the bound UDP-GlcNAc in the Y299C mutant protein is depicted in Fig. 4b. In the case of the wild-type bacterial epimerase, the side chain functional groups of Asn-179, Asn-199, and Tyr-299 form hydrogen-bonding interactions with the hexose sugar moieties of the UDP-Glc or UDP-Gal substrates. Additionally, Oδ of Ser-124 lies within 2.6 Å of the 4′-hydroxyl group of the glucose moiety in UDP-Glc. As can be seen in Fig. 5b, the UDP-GlcNAc ligand is situated in the active site such that these residues are no longer within hydrogen-bonding distance to the hydroxyl groups of the hexose. Indeed, the sugar hydroxyl groups lie only within hydrogen-bonding distance to...
solvent molecules. In the wild-type protein/NADH/UDP-Glc model, the nicotinamide ring of the cofactor is in the syn-conformation as would be expected for a B-side-specific enzyme. In the Y299C mutant protein/NADH/UDP-GlcNAc model, the nicotinamide ring is in the anti-conformation such that the carbonyl oxygen of its carboxamide group lies within 2.5 Å from O\(^\prime\) of Ser-124. The distance between the C4 carbon of the N-acetylglucosamine moiety and the C4 carbon of the nicotinamide ring is 9.4 Å. In the wild-type enzyme/NADH/UDP-Glc complex, this distance is 3.8 Å. The fact that the mutant protein but not the wild-type bacterial enzyme is active against UDP-GalNAc as a substrate argues for the fact that in the mutant protein the UDP-sugar must, at some point, rotate into the proper position for catalysis.

The final structure to be solved in this investigation was that of the Y299C mutant protein complexed with NADH and UDP-Glc. The electron density corresponding to the hexose portion of the UDP-Glc was weak. This is in sharp contrast to that observed for the wild-type enzyme/NADH/UDP-Glc crystalline complex (19). The α-carbons for the wild-type and Y299C pro-
proteins complexed with NADH and UDP-Glc correspond with a root-mean-square deviation of 0.56 Å. For comparison, the α-carbons for the mutant proteins that were complexed with either UDP-Glc or UDP-GlcNAc superimpose with a root-mean-square deviation of 0.22 Å. A close-up view of the Y299C mutant protein active site with NADH and UDP-Glc is shown in Fig. 6. In this complex, the nicotinamide ring of the NADH ligand is in the syn-conformation. The position of the glucose moiety is approximate because of the weakness of the electron density. It is clear, however, that in the Y299C mutant protein the loss of the hydrogen bond between Tyr-299 and the 6′-hydroxyl group of glucose normally observed in the wild-type enzyme results in a movement of the sugar away from the nicotinamide ring. The 5-fold reduction in enzymatic activity of this Y299C mutant protein arises in part because of the loss of this electrostatic interaction, thereby leading to unproductive binding modes for the hexose portion of the UDP-sugar substrate.

In conclusion, by comparing the three-dimensional structures of both the human and bacterial UDP-galactose 4-epimerases, it was possible to predict the necessary single site-specific change that would be required to confer UDP-GlcNAc/UDP-GalNAc interconverting activity on the E. coli protein. This combined structural/functional analysis emphasizes the power in comparing in detail various superfamily members for the rational design of new enzymatic activities on particular protein scaffolds.

Acknowledgments—We thank Sandra Harding for determining the initial crystallization conditions of the Y299C protein/NADH/UDP-GlcNAc complex. We also thank Drs. Dale Edmondson and Paige Newton-Vinson for generously allowing and helping us to use their fermenter.

REFERENCES