Crystal Structure of WbpP, a Genuine UDP-N-acetylglucosamine 4-Epimerase from Pseudomonas aeruginosa

SUBSTRATE SPECIFICITY IN UDP-HEXOSE 4-EPIMERASES*

Noboru Ishiyama§§, Carole Creuzenet¶¶, Joseph S. Lam**‡‡, and Albert M. Berghuis‡§§‡‡

From the Departments of §§Biochemistry and §§Microbiology & Immunology, McGill University, Montreal, Quebec H3A IA4, the ¶§Department of Microbiology & Immunology, University of Western Ontario, London, Ontario N6A 5C1, and the ¶¶Department of Microbiology, the University of Guelph, Guelph, Ontario N1G 2W1, Canada

Received for publication, February 13, 2004, and in revised form, March 8, 2004
Published, JBC Papers in Press, March 10, 2004, DOI 10.1074/jbc.M401642200

The O antigen of lipopolysaccharide in Gram-negative bacteria plays a critical role in bacterium-host interactions, and for pathogenic bacteria it is a major virulence factor. In Pseudomonas aeruginosa serotype O6 one of the initial steps in O-antigen biosynthesis is catalyzed by a saccharide epimerase, WbpP. WbpP is a member of the UDP-hexose 4-epimerase family of enzymes and exists as a homo-dimer. This enzyme preferentially catalyzes the conversion between UDP-GlcNAc and UDP-GalNAc above UDP-Glc and UDP-Gal, using NAD+ as a cofactor. The crystal structures of WbpP in complex with cofactor and either UDP-Glc or UDP-GalNAc were determined at 2.5 and 2.1 Å, respectively, which represents the first structural studies of a genuine UDP-GlcNAc 4-epimerase. These structures in combination with complementary mutagenesis studies suggest that the basis for the differential substrate specificity of WbpP is a consequence of the presence of a pliable solvent network in the active site. This information allows for a comprehensive analysis of the relationship between sequence and substrate specificity for UDP-hexose 4-epimerases and enables the formulation of consensus sequences that predict substrate specificity of UDP-hexose 4-epimerases yet to be biochemically characterized. Furthermore, the examination indicates that as little as one residue can dictate substrate specificity. Nonetheless, phylogenetic analysis suggests that this substrate specificity is an evolutionary and highly conserved property within UDP-hexose 4-epimerases.

* This work was supported by a Collaborative Health Research Projects Grant 251007-02 from the Natural Sciences and Engineering Research Council of Canada (NSERC) (jointly awarded to C. C., J. S. L., and A. M. B.). Beamline X8C at the NSLS-BNL, Upton, NY was in part supported by a Multi-User Maintenance grant from the Canadian Institutes of Health Research and the NSERC. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 1SB9 and 1SB8) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

¶¶ Recipient of a studentship from the Canadian Cystic Fibrosis Foundation.

†† Holds the Canada Research Chair in Cystic Fibrosis & Microbial Glycobiology.

‡‡ Holds the Canada Research Chair in Structural Biology. To whom correspondence should be addressed: Dept. of Biochemistry and Microbiology & Immunology, 740 Dr. Penfield Ave., Rm. 5202, McGill University, Montreal, Quebec H3A IA4, Canada. Tel.: 514-398-8795, Fax: 514-398-2036; E-mail: albert.berghuis@mcgill.ca.

The lipopolysaccharide (LPS)* component of the outer membrane of Gram-negative bacteria is critical in bacterium-host interactions. For example, LPS is responsible for triggering an innate immune response and plays an important role in septic shock (1, 2). Additionally, in pathogenic bacteria, LPS has been identified as a major virulence factor (3). The structure of LPS is comprised of three domains: Lipid A, core region, and O antigen. Although the first two domains remain buried or less exposed, O antigen is exposed to the surrounding environment and is often used as a target for serum recognition. Variation in the structural composition of O antigen has been linked to various aspects of bacterium-host interactions, including virulence potential (4). In the clinically relevant human pathogen Pseudomonas aeruginosa serotype O6, the O antigen is composed of a tetrasaccharide repeating unit of \(\alpha-L-3-O-3\)-O-acetyl-6-amino-GalNAc-(1\(\rightarrow\)4)-α-L-3-O-3\)-O-acetyl-6-deoxy-2-formamido-D-galacturonic acid-(1\(\rightarrow\)3)-ο-L-3-O-3-amino-2-deoxy-2-formamido-D-glucose-(1\(\rightarrow\)2)-α-L-Rha-(1\(\rightarrow\)) (5). Biosynthesis of this repeating unit is catalyzed by enzymes encoded in the wbp gene cluster (6). For example, two of the saccharide moieties that participate in the tetrasaccharide repeating unit are derived from UDP-GalNAc, which is produced by C4-epimerization of the sugar moiety of UDP-GlcNAc, catalyzed by WbpP (7).

WbpP is a member of the short-chain dehydrogenase/reductase superfamily that contains a diverse collection of oxidoreductases sharing relatively high conservation of structural homology despite low sequence identity (15–30%) (8–10). It is a homologue of UDP-4-epimerase (GalE; \(E. coli\) and \(T. brucei\) enzymes (eGalE, hGalE, and tGalE, respectively) (8, 11, 12), and the reaction mechanism involves a conserved tyrosine that functions as the active site base and a saccharide ring flip when in the oxidized state (Scheme 1). Based on overall sequence similarities with several active site residues shared between GalE and WbpP, it is likely that WbpP will employ a similar reaction mechanism. However, there are significant differences between WbpP and GalE with respect to substrate specificity. Specifically, unlike GalE, WbpP preferentially interconverts N-acetylated UDP-linked galactose and glucose, and hence it has been characterized as a genuine UDP-GlcNAc 4-epimerase (7).
Here we describe three-dimensional crystal structures of WbpP in complex with NAD<sup>+</sup> and either UDP-GalNAc or UDP-Glc, and data from complementary site-directed mutagenesis experiments. These studies are aimed at elucidating the structural basis for the preference of N-acetylated saccharide moieties, observed for WbpP. These results represent the first structure elucidation of a genuine UDP-GlcNAc 4-epimerase, and provide further insights into factors governing substrate specificity for GalE-like UDP-hexose 4-epimerases.

**EXPERIMENTAL PROCEDURES**

**Crystallization and Data Collection—**WbpP, containing an N-terminus His tag, was purified as previously described (7). Crystals of two different ternary complexes of WbpP were grown by hanging drop vapor diffusion. For successful crystallization, the reservoir solution consisted of 0.6 ml of 40–50% (v/v) polyethylene glycol 600 and 0.1 M phosphate-citrate buffer, pH 4.2–5.0, whereas the drop consisted of 2 µl of a solution containing 5 mg/ml WbpP, 3 mM NAD<sup>+</sup>, and either 3 mM UDP-Glc or 3 mM UDP-GlcNAc in 25 mM imidazole buffer, pH 7.4. Note that, although NAD<sup>+</sup> binds tightly to WbpP, and thus likely remains associated with the enzyme during purification, cofactor was provided in the crystallization solution to ensure complete saturation. Crystals with approximate dimensions of 0.3 × 0.2 × 0.1 mm were obtained after 2–4 weeks incubation at 4° or 22°C.

Prior to data collection, crystals were briefly washed in reservoir solution and flash-frozen in liquid nitrogen. Diffraction data were collected at 95 K at the X8C beam line of the National Synchrotron Light Source, Brookhaven National Laboratories, Upton, NY, which was equipped with an ADSC Quantum 4 charge-coupled device detector. For each crystal form (grown in the presence of either UDP-Glc or UDP-GlcNAc) a data set was obtained by exposing one crystal to synchrotron radiation. Data were processed using the HKL suite of programs (17) and showed that the crystals belonged to space group C222<sub>1</sub>, with typical unit cell dimensions of a = 61 Å, b = 96 Å, and c = 140 Å, thus containing one protein molecule per asymmetric unit. Statistics pertaining to the diffraction data are presented in Table I.

**Structure Determination and Refinement—**The crystal structure of WbpP in the presence of NAD<sup>+</sup> and UDP-Glc was solved by molecular replacement using the Crystallography and NMR System software suite (18). As a search model UDP-Gal 4-epimerase from *E. coli* was used (PDB code: 1XEL) (13). Solutions for the rotation and translation functions were readily obtained for the single WbpP molecule in the asymmetric unit. Refinement of the model consisted of successive rounds of reciprocal space refinement and model building using the program O (19). Adjustments and additions to the model were performed based on Fo − Fc-weighted difference electron density maps (2Fo − Fc and Fc − Fo). Refinement was continued until no significant improvement could be obtained as judged by a decrease in the free-R value. During refinement, unambiguous density for the UDP moiety of UDP-Glc could be observed. However, density for the saccharide moiety was weak and discontinuous, likely due to a high mobility for this group.

Therefore the occupancy for the Glc part of UDP-Glc was set to zero (Fig. 1a).

The crystal structure for WbpP grown in the presence of NAD<sup>+</sup> and UDP-GlcNAc was determined by difference Fourier methods, using a partially refined model of WbpP-NAD<sup>+</sup>-UDP-Glc, in which the substrates were removed. Refinement for this crystal structure was performed using analogous procedures as described above. During the course of refinement, density for the entire UDP-linked saccharide substrate could be clearly seen in difference electron density maps, unlike in the WbpP-NAD<sup>+</sup>-UDP-Glc structure. However, the exact nature of the saccharide moiety required assessment, because the observed density was not consistent with the initially added substrate UDP-GlcNAc (Fig. 1b). Based on χ<sub>2</sub>-weighted simulated annealing omit maps and trial occupancy refinements with both UDP-GlcNAc and UDP-GalNAc in either substrate (i.e. catalytically productive conformation in which C4-OH is oriented away from the NAD cofactor) or product (i.e. non-productive conformation in which C4-OH is oriented toward the NAD cofactor) orientations, we concluded that the observed density is consistent with a model in which ~70% of the saccharide moiety is UDP-GalNAc positioned in the catalytically productive substrate conformation. The remaining 30% is likely a mixture of the other three alternative species. However, due to the resolution of the diffraction data, we have included only the UDP-GalNAc productive conformation in the final model, and we will henceforth refer to this model as the WbpP-NAD<sup>+</sup>-UDP-GalNAc structure. Refinement statistics pertaining to the two models are presented in Table I.

**Site-directed Mutagenesis and Substrate Specificity Analysis—**Site-directed mutagenesis was performed using the QuikChange™ procedure (Stratagene, La Jolla, CA). The template for mutagenesis was the His-WbpP-pET vector (7). Two mutants were constructed: A209H and S306Y. After mutagenesis, DNA sequencing was performed on each mutant construct to confirm the presence of the desired mutations and eliminate constructs with unwanted secondary mutations. Protein expression and purification of WbpP variants were performed using analogous procedures as that for wild-type protein (7).

Analysis of kinetic parameters and substrate specificity for the two WbpP variants followed procedures previously used (7). In brief, after purification, enzymes were dialyzed overnight at 4°C in Tris 100 mM, pH 8. Reactions for activity assays were set up with 21 µl of enzyme adjusted to 0.57 or 0.28 mg/ml in Tris 100 mM, pH 8, 3.5 µl of 1 mM Tris buffer, pH 8, and 3.5 µl of 5 mM UDP-linked saccharide substrate for a total reaction volume of 35 µl. The reactions were incubated at 37°C and terminated at the appropriate time by boiling for 6 min. Analysis of substrate conversion was performed by capillary electrophoresis (Beckman model P/ACE MDQ).

**RESULTS AND DISCUSSION**

**Overall Structure of WbpP—**The three-dimensional structure of WbpP is composed of two domains (Fig. 2a). The N-terminal domain (residues 1–192 and 238–264) is the site where the NAD cofactor binds. It possesses a modified Rossmann fold composed of a seven-strand parallel β-sheet flanked by nine α-helices. The C-terminal domain (residues 193–237 and 265–343) holds the UDP-linked hexose substrate and has an αβ motif consisting of four α-helices and four β-strands. Consistent with sequence comparison results, WbpP is structurally similar to UDP-Gal 4-epimerases, such as eGalE, and possesses all the features of a member of the short-chain dehydrogenase/reductase family. The most notable difference between the WbpP and eGalE structures is that residues 3–16 in WbpP form an additional α-helix in the N-terminal domain. Additional differences between WbpP and eGalE are exclusively limited to loops located between secondary structure elements, where there are deletions and insertions observed. Overall, the root mean square difference for Ca atoms of structurally identical residues between WbpP and eGalE is 1.45 Å (291 residues); separating for the N- and C-terminal domains these differences are 1.3 and 1.2 Å, respectively, indicating that hinge motion between the two domains is minimal.

Although the asymmetric unit only contains one WbpP molecule, analytical gel-filtration studies (7) and dynamic light scattering studies (data not shown) indicate that the enzyme exists as a dimer in solution. Analysis of crystal packing con-

![Scheme 1. Epimerization reaction catalyzed by WbpP.](image-url)
tacts suggests that the physiological dimer species is created through the formation of a four-helix bundle with each protomer contributing two α-helices (residues 112–133 and 164–182; Fig. 2b). The interface between the two symmetry-related molecules participating in the four-helix bundle is 2530 Å², indicative of an oligomeric interface (20). Other UDP-hexose 4-epimerases also exist as similar dimers in solution (8).

Table I

| Data set                  | WbpP-NAD⁺-UDP-Glc | WbpP-NAD⁺-UDP-GalNAc |
|---------------------------|-------------------|----------------------|
| **Data collection**       |                   |                      |
| Cell dimensions (Å)       | a = 61.0, b = 95.7, c = 141.2 | a = 60.4, b = 95.9, c = 139.6 |
| Wavelength (Å)            | 1.000             | 0.980                |
| Resolution (Å)            | 2.5               | 2.1                  |
| Completeness (%)          | 97.2 (91.7)       | 96.4 (89.7)          |
| Redundancy*               | 5.5 (5.6)         | 6.4 (6.6)            |
| R cryst(I) (%)            | 9.6 (38.0)        | 8.3 (32.8)           |
| Data with I > 2σ(I) (%)   | 86.1 (61.5)       | 84.1 (63.6)          |
| **Refinement**            |                   |                      |
| R cryst (%)               | 19.5              | 19.0                 |
| R free (%)                | 25.5              | 22.8                 |
| No. protein atoms         | 2654              | 2662                 |
| No. hetero atoms          | 80                | 83                   |
| No. solvent atoms         | 142               | 182                  |
| Deviations bond lengths (Å)| 0.006             | 0.006                |
| Deviations bond angles (°)| 1.3               | 1.4                  |
| Mean B factors protein (Å²)| 33.0              | 22.0                 |
| Mean B factors hetero (Å²)| 32.8              | 21.7                 |
| Mean B factors solvent (Å²)| 35.7              | 27.7                 |

* The values in parentheses are for highest resolution shells.

** Architecture of the Active Site of WbpP—** Examination of the NAD cofactor and UDP-hexose binding sites in WbpP reveal that they are homologous to those found in the human, E. coli and T. brucei GalE crystal structures (Fig. 3). Specifically, all of the residues, which have previously been implicated in catalysis for GalE enzymes, i.e. the SYK triad, are also conserved in WbpP (Ser-142, Tyr-166, and Lys-170). The exact role of these
residues has been thoroughly examined by Holden and co-workers, and our structural data is in complete agreement with their analyses (11, 13–15, 21–23). Here we focus on the notable distinction between WbpP and UDP-Gal 4-epimerases. Specifically, the binding site for the saccharide moiety in WbpP has revealed important modifications compared with hGalE, tGalE, and eGalE. Residue 102 is a glycine in WbpP, whereas in GalE enzymes the corresponding residue is either a lysine or a leucine. The equivalent residue for Ala-209 in WbpP is either asparagine (eGalE and hGalE) or histidine (tGalE). Finally, Ser-306 in WbpP is substituted for a leucine in tGalE, a tyrosine in eGalE, and a cysteine in hGalE. Based on these alterations, WbpP possesses a larger space for the saccharide moiety to reside in when compared with GalE enzymes.

**UDP-Glc Versus UDP-GalNAc Binding in WbpP**—Crystal structures of WbpP were determined in complex with UDP-Glc and UDP-GalNAc. Comparison of the overall fold between the two complex structures reveals that there is a minor rotation between the N- and C-terminal domains of \(-1.4^\circ\), with the UDP-Glc complex being in a slightly more open conformation. Given the accuracy of the crystal structures, the nearly identical crystallization conditions and that the crystals were isomorphous, the observed difference in domain orientation is most likely caused by the difference in substrate.
Concomitant with the difference in domain orientation, noticeable dissimilarities in the active site are also detected between the UDP-Glc- and UDP-GalNAc-bound WbpP complex structures, specifically at the saccharide-binding pocket. The most striking difference is that there is clear and well defined density for the Glc moiety in the WbpP-Glc crystal structure, whereas comparable density for the GalNAc moiety is missing in the WbpP-GalNAc ternary complex (Fig. 1). The GalNAc moiety is in the catalytically competent substrate orientation and is appropriately positioned with respect to the catalytic base (Tyr-166 OH-GalNAc-C4-OH distance is 3.0 Å) and cofactor (NAD+ -C4-GalNAc-C4 distance is 3.2 Å). The absence of well defined density for the Glc moiety in the WbpP active site than is GalNAc. The presumably more disordered nature of the Glc moiety is also mirrored in its immediate environment; specifically, differences in solvent structure within the saccharide-binding pocket are observed between the two WbpP ternary complexes (Fig. 4). Compared with WbpP-GalNAc, the WbpP-Glc structure lacks several ordered solvent molecules in the active site, and other water molecules are displaced. Combined, these differences result in a slightly more closed conformation of the UDP-GalNAc-bound ternary complex, which is reflected in the differences in domain orientation.

Structural Basis for Substrate Specificity in WbpP—WbpP is ~1000-fold more efficient in catalyzing the epimerization between UDP-GlcNAc and UDP-GalNAc than when the non-acetylated substrates UDP-Glc and UDP-Gal are used (7). In contrast eGalE and tGalE are unable to epimerize acetylated UDP-linked saccharides, whereas hGalE is equally capable of catalyzing the epimerization of UDP-GlcNAc/UDP-GalNAc and UDP-Glc/UDP-Gal. Providing a rational explanation for the substrate specificity of eGalE, tGalE, and hGalE is relatively straightforward in that for eGalE and tGalE the saccharide-binding pocket is not large enough to perform the epimerization reaction for acetylated saccharide moieties, and for hGalE the binding pocket is sufficiently large to accommodate both classes of substrates (21). However, analogous reasoning cannot be used to explain the substrate specificity of WbpP.

The differences observed between WbpP and GalE enzymes and between the WbpP-GlcNAc/UDP-GalNAc and WbpP-Glc/UDP-Glc crystal structures provide insight into the structural basis of the preference of WbpP for UDP-GlcNAc/UDP-GalNAc. Compared with tGalE, eGalE, and hGalE, the saccharide-binding pocket is spacious due to the small size of residues 102, 209, and 306 (Gly, Ala, and Ser, respectively; Fig. 3). The additional space in the active site allows for water molecules to participate in the binding of saccharide moieties. Our analysis suggests that the presence of an N-acetyl group in the saccharide moiety enables an ordering of solvent molecules in the active site of WbpP (Fig. 4). This in turn stabilizes the position of the saccharide moiety and places it in the correct location for catalysis. In the absence of an N-acetyl group, the saccharide moiety is less restrained in the WbpP active site, and thus it is, on average, less well positioned for catalysis, resulting in a significant reduction of efficiency for UDP-Glc/UDP-Gal substrates. Therefore, the preference of WbpP for UDP-GlcNAc/UDP-GalNAc over UDP-Glc/UDP-Gal can be explained by orbital steering (24).

To test the validity of our analysis, we constructed two single-site variants of WbpP that were designed to have reduced preference for UDP-linked N-acetylated hexoses by shrinking the size of the saccharide-binding pocket: A209H and S306Y. The results for these mutagenesis studies are shown in Table II. In agreement with our analysis and prediction, the A209H mutant is much less efficient in catalyzing the epimerization of N-acetylated hexose moieties and is in fact more efficient in catalyzing the reaction for non-acetylated hexoses compared with wild-type. Unfortunately, the S306Y mutant proved to be completely inactive irrespective of the substrate tested, thus providing no further insights into the factors governing substrate specificity in WbpP.

Substrate Spectrum in UDP-hexose 4-Epimerases—With the completion of the structure determination of WbpP, crystal structures are now available for three different groups of UDP-hexose 4-epimerases, namely: epimerases that preferentially catalyze the conversion between UDP-Glc and UDP-Gal (eGalE and tGalE, group 1); epimerases that do not show a preference for either UDP-Glc/UDP-Gal or UDP-GlcNAc/UDP-GalNAc (hGalE, group 2); and epimerases that preferentially convert between UDP-GlcNAc and UDP-GalNAc (WbpP, group 3). This has allowed for a comprehensive analysis of the relationship between sequence, three-dimensional structure, and substrate preference. Based on these analyses we have constructed an abstract model to facilitate a better understanding of the factors governing substrate spectrum in UDP-hexose 4-epimerases (Fig. 5). The generic saccharide-binding pocket in groups 1, 2, and 3 UDP-hexose 4-epimerases can be conceptualized as a hexagonal shaped box in which the bottom is formed by the nicotinamide ring of the cofactor and six walls are formed by different regions of the enzymes. The top of the box is open to accommodate the flip of the saccharide ring that is essential to the reaction mechanism. Three of the six walls are conserved and contain the Ser and Tyr residues of the SYK triad and a highly conserved Asn residue (Asn-195 in WbpP); these walls are necessary for chemistry. In our conceptual model the remaining three walls vary between groups 1, 2, and 3 UDP-hexose 4-epimerases, and their exact composition/placement dictates substrate preference. In group 1 enzymes, the walls are constructed such that only non-acetylated UDP-hexoses are able to bind in a manner that allows for catalysis. Analysis of tGalE and eGalE shows that there are at least two ways in
All the equilibrium experiments were performed in 100 mM Tris, pH 8.0, with 0.5 mM of substrates in total volume of 35 ml. 6 μg of enzyme was used for the conversion of acetylated substrates with the total reaction time of 1.5 h. 12 μg of enzyme was used for the conversion of non-acetylated substrates with the total reaction time of 6 h.

| Mutant | Substrate and value of percent conversion |
|--------|-----------------------------------------|
|        | UDP-GlcNac | UDP-GalNac | UDP-Glc | UDP-Gal |
| WT     | 28.4       | 73.3       | 4.0     | 10.0    |
| A209H  | 4.5        | 6.9        | 5.3     | 15.5    |
| S306Y  | 0          | 0          | 0       | 0       |

Table II

Equilibrium data for the substrate conversions by WbpP

Fig. 5. Conceptual model for saccharide-binding pockets in UDP-hexose 4-epimerases. For tGalE, eGalE, hGalE, and WbpP, abstractions of the saccharide-binding pockets are presented, which rationalize the observed substrate spectrum for these UDP-hexose 4-epimerases. For each of the enzymes, the saccharide-binding pocket can be conceived as having up to six side walls composed of protein residues and a bottom wall formed by the nicotinamide ring of the cofactor. The saccharide moieties of UDP-Glc (for tGalE and eGalE) and UDP-GlcNac (for hGalE and WbpP) bound in the catalytically productive conformation are represented as black solid lines. Similarly, the sugar moieties of UDP-Gal and UDP-GalNac bound in the catalytically productive conformations are represented as dashed lines. For the WbpP model, the role of water molecules is signified by spheres. Below each model, pertinent information is provided, i.e. composition of the side walls, substrates preference, and substrate spectrum designation. Note that small letters for key residues indicate residues that effectively remove a wall from the saccharide-binding pocket model.

which the walls can be constructed for this. In tGalE, the saccharide-binding pocket is too small to allow binding of any acetylated UDP-linked hexoses (group 1a). In eGalE, due to the substitution of a histidine for an asparagine, UDP-GlcNac could bind and the 4-keto intermediate could be formed; however, rotation of the sugar moiety can not occur thus preventing formation of UDP-GalNac (group 1b). In group 2 and 3 enzymes, two of the walls are (partly) removed to allow binding of a larger N-acetylated substrate. Furthermore, in group 3 enzymes, ordered solvent molecules become part of the box so as to result in preferential catalysis of UDP-GlcNac and UDP-GalNac substrates.

According to this model, one should be able to alter the substrate spectrum of UDP-hexose 4-epimerases by altering the composition of the walls. The mutagenesis study reported above (Table II), accomplished exactly this. We restored one of the walls absent in WbpP in the A209H variant, partially mimicking tGalE, thereby altering the substrate spectrum from a group 3 profile to a group 2 profile. Thoden et al. (25) have reported an analogous mutagenesis study where they essentially removed one of the walls in eGalE by mutating Tyr-299 to a cysteine thus changing the substrate spectrum from a group 1 profile to a group 2 profile.

The model shown in Fig. 5 should also allow for the prediction of substrate spectrum solely based on sequence. Recently, functional studies of several UDP-hexose 4-epimerases have been performed. For example, WbgU from Pleismononas shigelloides and UAE from Giardia intestinalis have been identified as genuine UDP-GlcNac/UDP-GalNac epimerases (26, 27), and Gne from Yersinia enterocolitica and Bacillus subtilis has been shown to epimerize both UDP-Glc/UDP-Gal and UDP-GlcNac/UDP-GalNac substrates (28, 29). Predicting the
substrate spectrum profile for WbgU, UAE and Gne based on their sequence classifies these enzymes as group 3, group 3, and group 2 UDP-hexose 4-epimerases, respectively, in complete agreement with experimental data.

Evolution of Substrate Spectrum in UDP-hexose 4-Epimerases—Our mutagenesis study, as well as that of Thoden et al. (25), show that substrate specificity in UDP-hexose 4-epimerases is dictated by a very limited number of residues, e.g. a single mutation can change the profile (Table II Ref. 25). This implies that, given appropriate selective pressures, it must be easy to alter the substrate spectrum of UDP-hexose 4-epimerases during evolution. Furthermore, it opens up the possibility that a particular substrate spectrum was re-discovered multiple times during evolution. To examine this possibility, a phylogenetic analysis was performed in conjunction with substrate spectrum prediction based on sequence (Fig. 6). The phylogenetic tree suggests that a major divergence into three clusters occurred very early in evolution. To examine this possibility, a phylogenetic analysis was performed in conjunction with substrate spectrum prediction based on sequence (Fig. 6). The phylogenetic tree suggests that a major divergence into three clusters occurred very early in evolution. To examine this possibility, a phylogenetic analysis was performed in conjunction with substrate spectrum prediction based on sequence (Fig. 6). The phylogenetic tree suggests that a major divergence into three clusters occurred very early in evolution. To examine this possibility, a phylogenetic analysis was performed in conjunction with substrate spectrum prediction based on sequence (Fig. 6). The phylogenetic tree suggests that a major divergence into three clusters occurred very early in evolution. To examine this possibility, a phylogenetic analysis was performed in conjunction with substrate spectrum prediction based on sequence (Fig. 6). The phylogenetic tree suggests that a major divergence into three clusters occurred very early in evolution.

In conclusion, the crystal structure of WbpP represents the first structure of a UDP-hexose 4-epimerase that preferentially catalyzes the conversion between UDP-GlcNAc and UDP-GalNAc. The structural basis for substrate specificity is the result of WbpP better positioning N-acetylated hexose moieties for catalysis than non-acetylated UDP-linked hexoses, as effected by a pliable solvent network in the active site. When the knowledge on the structural basis for substrate specificity is applied to other UDP-hexose 4-epimerases, it is revealed that substrate preference developed very early in the evolution of these enzymes, though re-invention of substrate spectrum has likely occurred multiple times between group 1b and 2 enzymes.

Acknowledgments—We thank Robert Urbanic and Milinda Demendi for technical assistance with protein purification and site-directed mutagenesis.

REFERENCES

1. Downey, J. S., and Han, J. (1998) Front. Biosci. 3, d468–d476
2. Du, X., Paltarakis, A., Silva, M., and Beutler, B. (1999) Blood Cells Mol. Dis. 25, 328–338
3. Heumann, D., and Roger, T. (2002) Clin. Chim. Acta 323, 59–72
4. Lerouge, I., and Vanderleyden, J. (2002) FEMS Microbiol. Rev. 26, 17–47
5. Knirel, Y. A. (1990) Crit. Rev. Microbiol. 17, 273–304
6. Belanger, M., Burrows, L. L., and Lam, J. S. (1999) Microbiology 145, 3505–3521
7. Creuzenet, C., Belanger, M., Wakarchuk, W. W., and Lam, J. S. (2000) J. Biol. Chem. 275, 19060–19067
8. Bauer, A. J., Rayment, I., Frey, P. A., and Holden, M. H. (1992) Proteins 12, 372–381
9. Samuel, J., and Tanner, M. E. (2002) Nat. Prod. Rep. 19, 261–277
10. Allard, S. T., Giraud, M. F., and Naismith, J. H. (2000) Cell. Mol. Life Sci. 58, 1650–1665
11. Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L., and Holden, H. M. (2000) Biochemistry 39, 5689–5701
12. Shaw, M. P., Bond, C. S., Roper, J. R., Gourley, D. G., Ferguson, M. A., and Hunter, W. N. (2003) *Mol. Biochem. Parasitol.* 126, 173–180
13. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) *Biochemistry* 35, 5137–5144
14. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) *Biochemistry* 35, 2557–2566
15. Thoden, J. B., Gulick, A. M., and Holden, H. M. (1997) *Biochemistry* 36, 10685–10695
16. Thoden, J. B., Hegeman, A. D., Wesenberg, G., Chapeau, M. C., Frey, P. A., and Holden, H. M. (1997) *Biochemistry* 36, 2557–2566
17. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* 276, 307–326
18. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grasse-Kanstieve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D* *Biol. Crystallogr.* 54, 905–921
19. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, (1991) *Acta Crystallogr. Sect. A* 47, 110–119
20. Janin, J. (1995) *Biochimie (Paris)* 77, 497–505
21. Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L., and Holden, H. M. (2001) *J. Biol. Chem.* 276, 15131–15136
22. Thoden, J. B., and Holden, H. M. (1998) *Biochemistry* 37, 11469–11477
23. Liu, Y., Thoden, J. B., Kim, J., Berger, E., Gulick, A. M., Rusicka, F. J., Holden, H. M., and Frey, P. A. (1997) *Biochemistry* 36, 10675–10684
24. Mesecar, A. D., Stoddard, B. L., and Koshland, D. E., Jr. (1997) *Science* 277, 202–206
25. Thoden, J. B., Henderson, J. M., Fridovich-Keil, J. L., and Holden, H. M. (2002) *J. Biol. Chem.* 277, 27528–27534
26. Kowal, P., and Wang, P. G. (2002) *Biochemistry* 41, 15410–15414
27. Lopez, A. R., Sener, K., Jarroll, E. L., and van Keulen, H. (2003) *Mol. Biochem. Parasitol.* 128, 51–57
28. Bengoechea, J. A., Pinta, E., Salminen, T., Oertelt, C., Holst, O., Radziejewska-Lebrecht, J., Piotrowska-Sieget, Z., Venho, E., and Skurnik, M. (2002) *J. Bacteriol.* 184, 4277–4287
29. Soldo, B., Scotti, C., Karamata, D., and Lazarevic, V. (2003) *Gene (Amst.)* 319, 65–69
30. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* 24, 946–950
31. Fenn, T. D., Ringe, D., and Petsko, G. A. (2003) *J. Appl. Crystallogr.* 36, 944–947
32. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* 277, 505–524
33. Thompson, J. D., Gibson, T. J., Plewinski, F., Jeannougnin, F., and Higgins, D. G. (1997) *Nucleic Acids Res.* 25, 4876–4882
34. Page, R. D. (1996) *Comput. Appl. Biosci.* 12, 357–358

*Structure of UDP-GlcNAc 4-Epimerase WbpP from P. aeruginosa*