Biochemistry

 Volume 39, Number 49 December 12, 2000

Accelerated Publications

The Structure of UDP-*N*-Acetylglucosamine 2-Epimerase Reveals Homology to Phosphoglycosyl Transferases[†,‡]

Robert E. Campbell,[§] Steven C. Mosimann,[||] Martin E. Tanner,[*,§] and Natalie C. J. Strynadka[*,||]

Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3, and Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1

Received July 12, 2000; Revised Manuscript Received September 28, 2000

ABSTRACT: Bacterial UDP-*N*-acetylglucosamine 2-epimerase catalyzes the reversible epimerization at C-2 of UDP-*N*-acetylglucosamine (UDP-GlcNAc) and thereby provides bacteria with UDP-*N*-acetylmannosamine (UDP-ManNAc), the activated donor of ManNAc residues. ManNAc is critical for several processes in bacteria, including formation of the antiphagocytic capsular polysaccharide of pathogens such as *Streptococcus pneumoniae* types 19F and 19A. We have determined the X-ray structure (2.5 Å) of UDP-GlcNAc 2-epimerase with bound UDP and identified a previously unsuspected structural homology with the enzymes glycogen phosphorylase and T4 phage β-glucosyltransferase. The relationship to these phosphoglycosyl transferases is very intriguing in terms of possible similarities in the catalytic mechanisms. Specifically, this observation is consistent with the proposal that the UDP-GlcNAc 2-epimerase-catalyzed elimination and re-addition of UDP to the glycal intermediate may proceed through a transition state with significant oxocarbenium ion-like character. The homodimeric epimerase is composed of two similar $\alpha/\beta/\alpha$ sandwich domains with the active site located in the deep cleft at the domain interface. Comparison of the multiple copies in the asymmetric unit has revealed that the epimerase can undergo a 10° interdomain rotation that is implicated in the regulatory mechanism. A structure-based sequence alignment has identified several basic residues in the active site that may be involved in the proton transfer at C-2 or stabilization of the proposed oxocarbenium ion-like transition state. This insight into the structure of the bacterial epimerase is applicable to the homologous N-terminal domain of the bifunctional mammalian UDP-GlcNAc "hydrolyzing" 2-epimerase/ManNAc kinase that catalyzes the rate-determining step in the sialic acid biosynthetic pathway.

The bacterial UDP-*N*-acetylglucosamine 2-epimerase catalyzes the reversible interconversion of UDP-*N*-acetylglucosamine (UDP-GlcNAc)[1] and UDP-*N*-acetylmannosamine (UDP-ManNAc) (Scheme 1) (*1*−*4*). This enzyme is interesting from a mechanistic perspective because it is able to catalyze an NAD^+-independent epimerization at a stereocenter that lacks an acidic proton (*5*, *6*). UDP-GlcNAc 2-epimerase, along with UDP-ManNAc dehydrogenase, provides bacteria with activated forms of both ManNAc and *N*-acetylmannosuronic acid (ManNAcUA) residues for use in the biosynthesis of cell surface polysaccharides (*2*). ManNAc residues are found as components of the antiphagocytic capsular polysaccharide in pathogenic strains of bacteria such as *Streptococcus pneumoniae* types 19F and 19A (*7*, *8*). These capsules are responsible for the virulence of these strains since they coat the bacteria and mask them from the immune system of the host. ManNAc residues are

[†] This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) (graduate fellowship to R.E.C., postdoctoral fellowship to S.C.M., and an operating grant to M.E.T.), the Burroughs Wellcome Foundation (N.C.J.S.), and the Medical Research Council of Canada (MRC) (Scholarship and Operating Grant MT14204 to N.C.J.S.).

[‡] X-ray coordinates have been deposited in the Protein Data Bank (entry 1F6D).

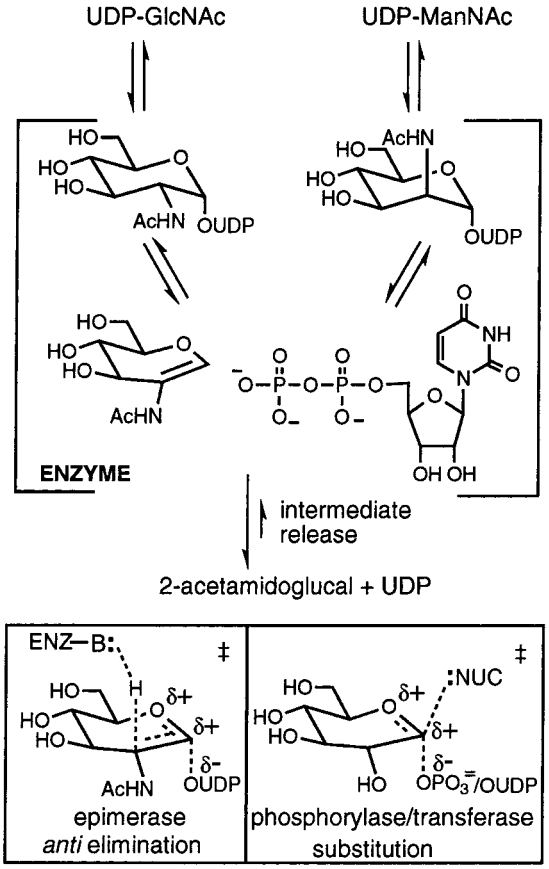
* To whom correspondence should be addressed. N.C.J.S.: phone, (604) 822-0789; fax, (604) 822-5227; e-mail, natalie@byron.biochem.ubc.ca. M.E.T.: phone, (604) 822-9453; fax, (604) 822-2847; e-mail, mtanner@chem.ubc.ca.

[§] Department of Chemistry.

[||] Department of Biochemistry and Molecular Biology.

10.1021/bi001627x CCC: $19.00
Published on Web 11/14/2000

Scheme 1: Mechanism of the Reaction Catalyzed by UDP-GlcNAc 2-Epimerase[a]



[a] The inset at the lower left shows the postulated transition state for the epimerase *anti* elimination. The inset at the lower right shows the postulated transition states for the phosphorylase (first step) and transferase substitutions.

also found in the "linkage unit" that serves to attach teichoic acids to the peptidoglycan in Gram-positive bacteria (*9*, *10*). ManNAcUA residues are found as components of the enterobacterial common antigen (ECA), a surface antigen found in all enteric or "gut bacteria" (*11*, *12*).

The bacterial epimerase is also of interest due to its relationship with the bifunctional mammalian UDP-GlcNAc "hydrolyzing" 2-epimerase/ManNAc kinase (*13*, *14*). This enzyme catalyzes the irreversible formation of free ManNAc and UDP from UDP-GlcNAc via an epimerization at C-2 combined with hydrolysis of the glycosidic bond. It also catalyzes the phosphorylation of ManNAc to generate ManNAc 6-phosphate. ManNAc 6-phosphate is the direct precursor to *N*-acetylneuraminic acid and its derivatives, the sialic acids. The sialic acids are found at the termini of oligosaccharides in a large variety of cell surface glycoconjugates and are key mediators of cell−cell recognition events (*15*, *16*). They play important roles in a variety of processes, including microbial infectivity (*17*), T and B cell activation (*18*), and tumorigenicity (*19*). Recently, it has been found that the UDP-GlcNAc hydrolyzing 2-epimerase catalyzes the rate-determining step in the sialic acid biosynthetic pathway and thus serves as a regulator of cell surface sialylation (*20*). The first 380 residues of the 722-residue mammalian enzyme and the sequence of the full-length *Escherichia coli* epimerase are 22% identical (*21*), indicating that there is an evolutionary link between the two enzymes and a structural analysis of one is pertinent to an understanding of the other.

The mechanism of the bacterial epimerase is unique among sugar nucleotide epimerases in that it involves the elimination and re-addition of UDP (*3*, *4*). An initial *anti* elimination of UDP from UDP-GlcNAc generates the intermediate 2-acetamidoglucal (Scheme 1). A subsequent *syn* addition of UDP gives the product UDP-ManNAc. These steps presumably occur either in an E1 fashion with discrete oxocarbenium ion intermediates or via "E1-like E2" transition states with considerable oxocarbenium ion-like character (see the lower left inset in Scheme 1) (*4*, *5*). Evidence supporting a mechanism in which the proton is removed from C-2 includes the observation that a solvent-derived deuterium is incorporated at C-2 during catalysis and that the epimerization of UDP-[2-^{2}H]GlcNAc is slowed by a primary kinetic isotope effect (*3*, *4*, *22*). The formation of UDP as an enzyme-bound intermediate is also supported by the results of a positional isotope exchange (PIX) experiment. When UDP-GlcNAc containing an ^{18}O label at the anomeric position was incubated with the epimerase, the label was seen to scramble from the bridging (anomeric) position into the nonbridging phosphate positions in a statistical manner. Finally, the direct detection of the intermediates UDP and 2-acetamidoglucal was possible since the enzyme occasionally releases them into solution. In addition, they are more thermodynamically stable than the sugar nucleotides and can therefore be quantitatively generated during extended incubations. The mechanism of the mammalian UDP-GlcNAc hydrolyzing 2-epimerase is likely analogous to that of the bacterial enzyme and involves an *anti* elimination of UDP followed by hydration of the resulting 2-acetamidoglucal intermediate to give ManNAc. The reaction is known to proceed with the incorporation of a solvent-derived proton into the C-2 position of ManNAc (*23*), and follows an ordered reaction mechanism with UDP released first followed by the irreversible formation of ManNAc (*24*). In addition, it appears the enzyme will accept the glycal intermediate from solution and hydrate it to form free ManNAc.

To gain further insight into the mechanism of UDP-GlcNAc 2-epimerase, we have determined the 2.5 Å resolution structure of the selenomethionine-substituted (SeMet) enzyme from *E. coli* by multiple-wavelength anomalous dispersion (MAD) phasing. The most intriguing insight provided by this structure is recognition of the unanticipated structural homology between UDP-GlcNAc 2-epimerase and two enzymes that catalyze phosphoglycosyl transfer, glycogen phosphorylase (GP) and T4 phage β-glucosyltransferase (BGT). This observation has important repercussions for the mechanism of UDP-GlcNAc 2-epimerase and may provide valuable insight into the mechanism of GP for which mechanistic ambiguities remain (*25*).

[1] Abbreviations: UDP, uridine-5′-diphosphoglucose; GlcNAc, *N*-acetylglucosamine; ManNAc, *N*-acetylmannosamine; NAD^+, oxidized nicotinamide adenine dinucleotide; ManNAcUA, *N*-acetylmannosuronic acid; Neu5Ac, *N*-acetylneuraminic acid; CMP, cytosine 5′-monophosphate; GP, glycogen phosphorylase; BGT, T4 phage β-glucosyltransferase; SeMet, selenomethionine; MAD, multiwavelength anomalous dispersion.

MATERIALS AND METHODS

Purification and Crystallization. SeMet UDP-GlcNAc 2-epimerase was expressed according to the procedure of Ramakrishnan (*26*) [V. Ramakrishnan and V. Graziano (http://snowbird.med.utah.edu/~ramak/madms.html)]. The SeMet enzyme was purified as previously reported (*4*) with an additional step of chromatography. Partially purified UDP-GlcNAc 2-epimerase was applied to a column of ceramic hydroxyapatite (Bio-Rad) in 1 mM magnesium chloride containing 10% glycerol and 2 mM dithiothreitol. The column was eluted with a linear gradient of 0 to 200 mM sodium phosphate (pH 6.8) in the loading buffer. To crystallize SeMet UDP-GlcNAc 2-epimerase, hanging drops were constructed by mixing the protein solution (4 mg/mL protein and 2 mM dithiothreitiol) with an equal volume of mother liquor [13% polyethylene glycol 8000, 0.3 M sodium chloride, and 0.1 M sodium citrate (pH 5.25)] and 10 vol % of 0.1 M UDP-GlcNAc. Drops were equilibrated for 24 h at room temperature before microseeding with crushed crystals grown from a higher precipitant concentration. Crystals were visible within 2−7 days and continued to grow for up to 4−6 weeks. All crystals belonged to space group $P2_1$ with the following unit cell dimensions: $a = 89.5$ Å, $b = 94.0$ Å, $c = 100.8$ Å, and $\beta = 109.6°$. A Matthews coefficient of 2.4 Å^3/Da is consistent with the two dimers of UDP-GlcNAc 2-epimerase observed in the asymmetric unit. Crystals could be grown to dimensions in the range of 0.3 mm, though smaller crystals ($\leq$0.1 mm) tended to have lower mosaicity. The wild-type UDP-GlcNAc 2-epimerase suffered from prohibitively poor mosaicity, and satisfactory crystals could not be obtained.

Data Collection and Processing. Prior to data collection, crystals were transferred to a cryoprotectant solution of mother liquor containing 25% glycerol before being flash-cooled in a stream of N_2 (100 K). Diffraction data to 2.5 Å for the MAD experiment with the SeMet UDP-GlcNAc 2-epimerase crystals were collected using a Brandeis Q4 CCD detector mounted at beamline X12C (Brookhaven National Laboratory, Upton, NY). Three data sets (Table 1) were collected on a single crystal at wavelengths corresponding to the selenium absorption edge ($\lambda = 0.9790$ Å), the peak ($\lambda_2 = 0.9787$ Å), and a wavelength remote from the selenium absorption edge ($\lambda_3 = 0.9400$ Å). All diffraction data were processed with DENZO/SCALEPACK (*27*), and selenium atom positions were determined using data to 2.8 Å in SOLVE (*28*) which was successful in identifying 37 of the 40 selenium atoms per asymmetric unit.

Model Building and Structural Refinement. Solvent flattening using DM (*29*) produced a readily interpretable electron density map with well-defined density for 1298 of the 1504 amino acids expected in the asymmetric unit. Noncrystallographic symmetry averaging (DM) provided only moderate improvements in the electron density. Model building was carried out with XTALVIEW (*30*), and all refinement was performed in CNS-XPLOR (*31*). An initial model was built against the most complete monomer and the N- and C-terminal domains individually used for rigid body minimization against the other three monomers. Structural refinement proceeded with multiple cycles of simulated annealing and restrained *B*-factor refinement against the remote data set in CNS-XPLOR. Strict NCS restraints were used in early rounds of refinement but were gradually relaxed and entirely removed by the final round. The refined structure of SeMet UDP-GlcNAc 2-epimerase contained 1489 amino acids, 765 water molecules, 4 UDP molecules, 4 sodium ions, and 4 chloride ions. The model was evaluated with the program PROCHECK (*32*) and is better than average in all statistical indicators of model quality.

Table 1: Data Collection, Phasing, and Refinement Statistics

Data Collection	λ_1	λ_2	λ_3
wavelength (Å)	0.9790	0.9787	0.9400
resolution range (Å)	30−2.50	30−2.50	30−2.50
no. of observations[a]	172475 (7101)	172602 (7007)	163518 (6804)
no. of unique reflcns	51549 (2404)	51929 (2459)	50792 (2386)
completeness (%)	94.7 (89.9)	95.5 (91.8)	93.1 (89.0)
R_{merge} (%)	4.9 (29.3)	6.8 (31.7)	5.1 (20.9)
I/σ	26.7 (3.6)	18.5 (2.6)	21.6 (3.6)

Phasing Statistics[b]		Refinement Statistics[c]	
resolution range (Å)	30−2.8	resolution range (Å)	30−2.5
no. of reflections	37198	no. of reflections	49090
no. of Se sites	37	R_{factor} (%)	19.8
occupancy range	0.24−1.0	R_{free} (test set, %)[d]	27.1
B-factor range (Å^2)	15−60	rmsd for bonds (Å)	0.009
figure of merit	0.68 (0.43)	rmsd for angles (deg)	1.3
		overall *B*-factor (Å^2)	35.8

[a] Values in parentheses refer to the highest-resolution shell which is 2.54−2.50 Å for data collection and 2.80−2.89 Å for phasing. [b] Results from SOLVE were improved when the highest-resolution data was excluded. [c] The model was refined against the remote (λ_3) SeMet data set. [d] R_{free} was calculated on 10% of the reflections randomly omitted from the refinement.

Analysis of Protein Structure. The primary sequence alignment included 29 sequences of UDP-GlcNAc 2-epimerase from both bacterial and mammalian sources that were retrieved from the SWISS-PROT database. The initial alignment was performed with the program CLUSTALW (*33*) and then manipulated to maximize conservation of secondary structural elements as identified by PROMOTIF (*34*). Structural alignment with BGT and GP (PDB entries 1QKJ and 3GPB) was based on visual inspection of α-carbon superpositions that were performed using the LSQ commands (3.8 Å cutoff) in the program O (*35*). Automated analysis of the dimer interface was performed with the protein−protein interaction server of S. Jones and J. M. Thornton (http://www.biochem.ucl.ac.uk/bsm/PP/server). All depictions of UDP-GlcNAc 2-epimerase were created with MOLSCRIPT (*36*).

RESULTS

Tertiary Structure. Bacterial UDP-GlcNAc 2-epimerase is composed of two α/β/α domains that form a deep cleft at the domain interface. A ribbon representation of the overall fold of the enzyme is shown in Figure 1a with secondary structural elements numbered sequentially. The N-terminal domain (residues 1−170 and 360−371) consists of a seven-stranded parallel β-sheet (β1−β7) that is sandwiched between a total of seven α-helices (α1−α6 and α16) with a topology that is similar to the Rossmann dinucleotide binding fold (*37*). The final strand (β7) of the N-terminal domain leads into a region of two α-helices (residues 171−205, α7 and

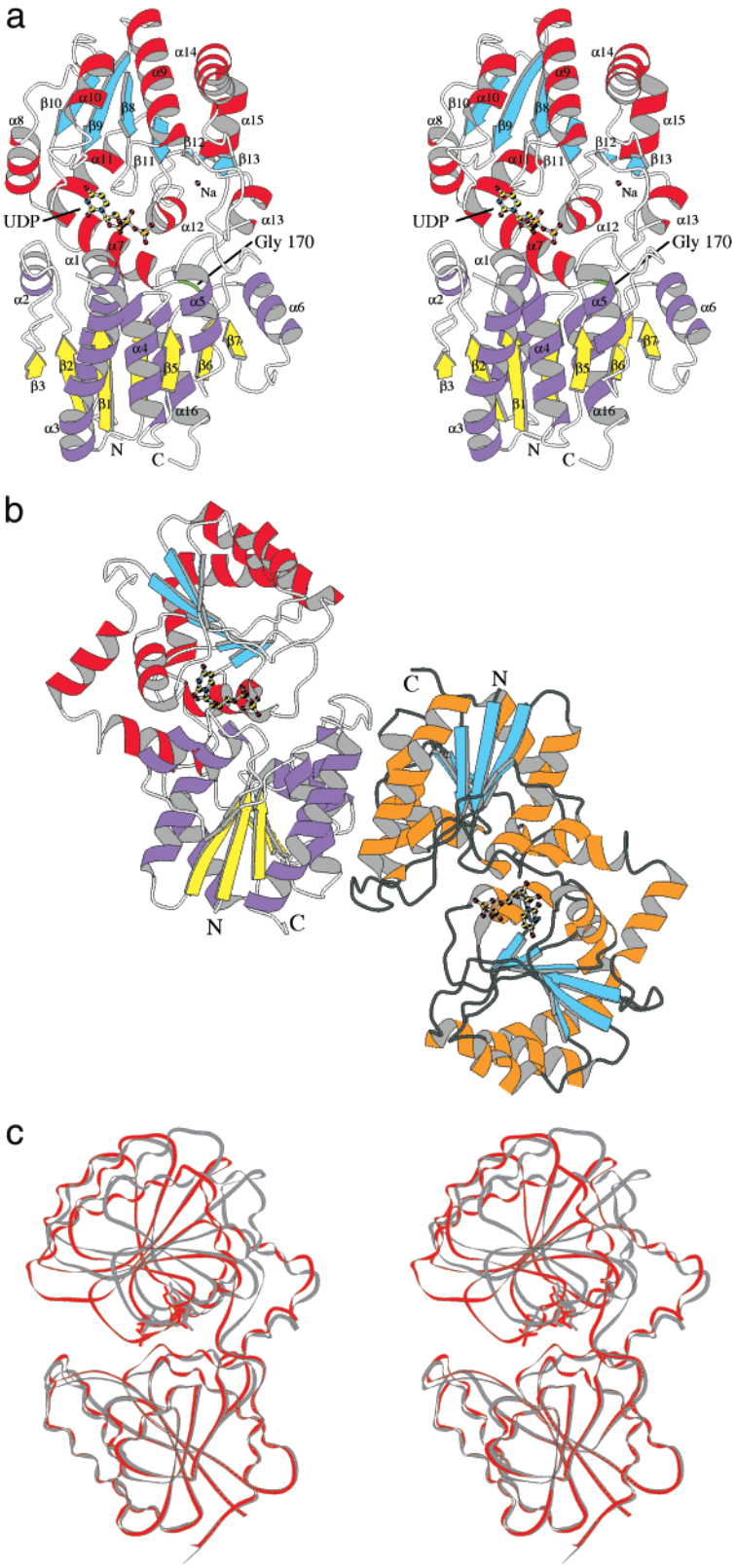


FIGURE 1: Overall structure of UDP-GlcNAc 2-epimerase. (a) Ribbon representation of the closed chain monomer of UDP-GlcNAc 2-epimerase with bound UDP shown as a ball-and-stick representation. The monomer is composed of two similar α/β/α domains that are colored in purple and yellow for the N-terminal domain and red and blue for the C-terminal domain. The strictly conserved Gly170 (green coil) may be a critical component of the interdomain hinge. No density was observed for residues 61−63 of this chain, and thus, there is a break between β3 and α3. (b) The dimer of UDP-GlcNAc 2-epimerase is composed of one open chain and one closed chain that differ by an approximate 10° rigid body rotation between the two domains. The closed chain is colored as described for panel a. (c) Overlap of the C-α atoms of the open (gray) and closed (red) chains of the UDP-GlcNAc 2-epimerase dimer highlighting the interdomain rotation. The N-terminal domain in the open and closed forms has been superimposed in this figure.

α8) that serves to link the two domains. The C-terminal domain (residues 206−359) contains a six-stranded β-sheet (β8−β13) that also has the topology of the Rossmann dinucleotide binding fold and is surrounded by a total of

seven α-helices (α9−α15). The C-terminal end of the C-terminal domain finds the polypeptide chain (residues 356−361) back at the domain interface where it crosses over the interdomain linker at residue 170. Following this crossover is α16 that is packed against the *β*-sheet core of the N-terminal domain and leads to the C-terminus that is approximately 10 Å from the N-terminus. In each of the four monomers in the asymmetric unit, a sodium ion (*B*-factor range of 17−37 Å^2) coordinated by the main chain carbonyls of Pro298, Ser350, and Ala352 was apparent. The sodium ion is distant from the active site, but is located at the C-terminal end of α12, the helix that forms several critical interactions with the UDP moiety of the substrate. A role for sodium in the structural stabilization of UDP-GlcNAc 2-epimerase correlates with a previous report that monovalent metal ions, such as sodium, increase the epimerase activity of the mammalian enzyme (*38*).

A search of the Protein Data Bank with the program DALI (*39*) revealed that the proteins with the greatest degree of structural homology with UDP-GlcNAc 2-epimerase are T4 phage *β*-glucosyltransferase (BGT) (*40*, *41*) and the core domains of glycogen phosphorylase (GP). The conservation of three-dimensional structure between BGT and GP has been previously recognized and discussed in some detail (*42*, *43*). An α-carbon superposition of UDP-GlcNAc 2-epimerase with each of these two enzymes gave a rms deviation of 2.4 Å over 129 and 162 atoms for BGT (Figure 2b) and GP, respectively. When the two domains of UDP-GlcNAc 2-epimerase were treated individually to account for slight changes in domain orientation, the overlap improved in both cases to give a rms deviation of 2.0 Å over 151 and 209 atoms, respectively. With the exceptions of α2 and α8, for every secondary structural element of UDP-GlcNAc 2-epimerase, an analogous element is conserved in BGT, GP, or both. This structural conservation is remarkable when one considers that the degree of sequence identity between UDP-GlcNAc 2-epimerase and each of these two enzymes is less than 10%. Figure 2a shows a structure-based sequence alignment of UDP-GlcNAc 2-epimerase with BGT and GP.

Conserved Residues. A structure-guided sequence alignment of UDP-GlcNAc 2-epimerase (Figure 2a shows 4 of the 29 sequences) revealed 10 strictly conserved residues. Six additional positions exhibit conservation of Glu and/or Asp (Glu12, Glu117, Glu131, Glu132, and Asp175) or Lys and/or Arg (Lys15) in at least 27 of the 29 sequences. Five of these 16 residues can be tentatively assigned a role in stabilization of the tertiary structure (Glu12, Gly94, His115, Gly119, and Glu132). Two residues (Gly170 and Asp175) are found at the interdomain hinge region and are probably involved in facilitating the interdomain rotation. Three residues (Arg10, Ser290, and Glu296) have been confidently assigned a role in binding the UDP portion of the substrate and will be discussed below. The remaining six residues (Lys15, Asp95, Glu117, Glu131, Arg135, and His213) are all found in the vicinity of the active site, close to where the GlcNAc portion of UDP-GlcNAc is expected to bind. While it would be premature to speculate about possible catalytic roles (if any) for the majority of these active site residues, His213 will be discussed below as a possible general acid catalyst. It is also worth noting that a proposed catalytic base of BGT (Glu22) does align with active site Lys15 of UDP-GlcNAc 2-epimerase (*41*). The only published mutagenesis study of UDP-GlcNAc 2-epimerase reportedly identified several histidine residues (including His115) as being important for the catalytic activity and the oligomeric structure of the mammalian enzyme (*21*). Unfortunately, that study neglected to investigate the role of His213, the only conserved histidine residue that is probably involved in the catalytic mechanism.

Quaternary Structure. The asymmetric unit of the crystals of UDP-GlcNAc 2-epimerase contains four copies of the enzyme arranged as two similar copies of the biological homodimer (*2*). As shown in panels b and c of Figure 1, each dimer is composed of one open and one closed monomer that differ by an approximately 10° interdomain rotation. The domains appear to rotate as rigid bodies about the hinge region (residues 169−172 and 356−361) at the domain interface on the backside of the enzyme when oriented as in Figure 1a. The strictly conserved Gly170 (green coil in Figure 1a) appears to be critical for the flexibility of this hinge region. The majority of the 1500 Å^2 dimer interface is composed of the three α-helices (α3−α5). The dimer interface of UDP-GlcNAc 2-epimerase is predominantly hydrophobic and stabilized by nine hydrogen bonds (*44*). On the periphery of the dimer interface are two intersubunit salt bridges (Glu69−Lys79 and the noncrystallographic counterpart) and a potential intersubunit disulfide bridge (Cys72−Cys72) that could contribute to stabilization of the dimer. Although the thiols of Cys72 of adjacent subunits are within 4 Å of each other, $F_o - F_c$ difference density maps provide no evidence of disulfide bridges, possibly due to the inclusion of a reducing agent in the crystallization solution. It is interesting to note that the secondary structural elements that contribute the most to the dimer interface (α3 and α4) are not conserved in the monomeric BGT (see Figure 2a,b). The intersubunit contacts of dimeric GP occur through regions of the protein that are not present in UDP-GlcNAc 2-epimerase (*45*).

Substrate Binding. In the deep interdomain cleft of the closed chains of UDP-GlcNAc 2-epimerase, density corresponding to a well-ordered molecule of UDP (B_{av} = 13 and 21 Å^2 in the two closed chains) is visible, identifying this region as the active site. The active site of the open chain dimer partner also contains density consistent with the uridine moiety; however, there is only very weak difference density in the region where the pyrophosphate moiety of UDP is expected to reside. Efforts to explain this difference density by invoking multiple conformations of the pyrophosphate group were unsatisfactory, and thus in the open chains, UDP has been fixed in a conformation similar to that observed in the closed chains. The observation of UDP, as opposed to UDP-GlcNAc, in the active site is not surprising since prolonged incubation of substrate with UDP-GlcNAc 2-epimerase results in quantitative formation of the thermodynamically favored intermediates, UDP and 2-acetamidoglucal (*3*, *4*). Active site difference density proximal to the *β*-phosphate of UDP in the closed chain was explained with several water molecules that behaved well during the refinement. Attempts to model all the active site density either as UDP-GlcNAc or as both the bound intermediates, UDP and 2-acetamidoglucal, were unsuccessful. The α-carbon superpositions indicate that the position

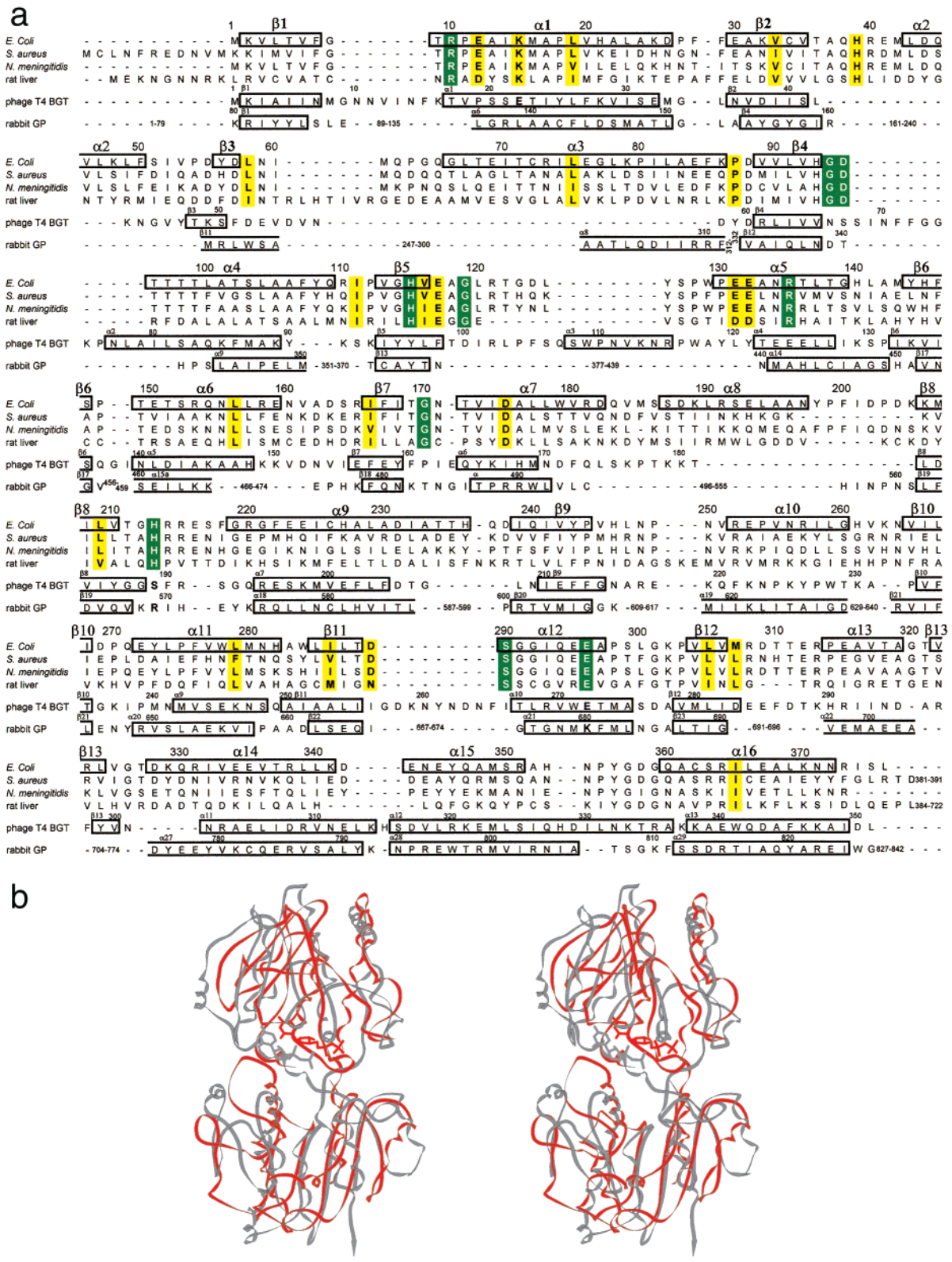


FIGURE 2: (a) Representative sequence alignment of 29 sequences of UDP-GlcNAc 2-epimerase. All α-helix and *β*-strand structures are numbered sequentially as in Figure 1a and indicated by black boxes. Positions in the sequence highlighted with white text on green are strictly conserved, while positions highlighted in black on yellow exhibit strong conservation. The structure-based alignments with GP and BGT are based upon α-carbon superpositions performed in the program O. Residues of either GP or BGT mentioned in the text are in bold. Only those residues of GP identified as being structurally equivalent to residues of UDP-GlcNAc 2-epimerase are shown, and consequently, some secondary structural elements are truncated. Labeling of the secondary structural elements of GP and BGT is consistent with the literature conventions. (b) C-α superposition of BGT with UDP-GlcNAc 2-epimerase. The two domains of each molecule were superimposed individually to account for slight changes in domain rotation between the two structures.

and orientation of the bound UDP are very similar to those observed for UDP in BGT and pyridoxal 5′-phosphate in GP, providing another convincing link between these three enzymes.

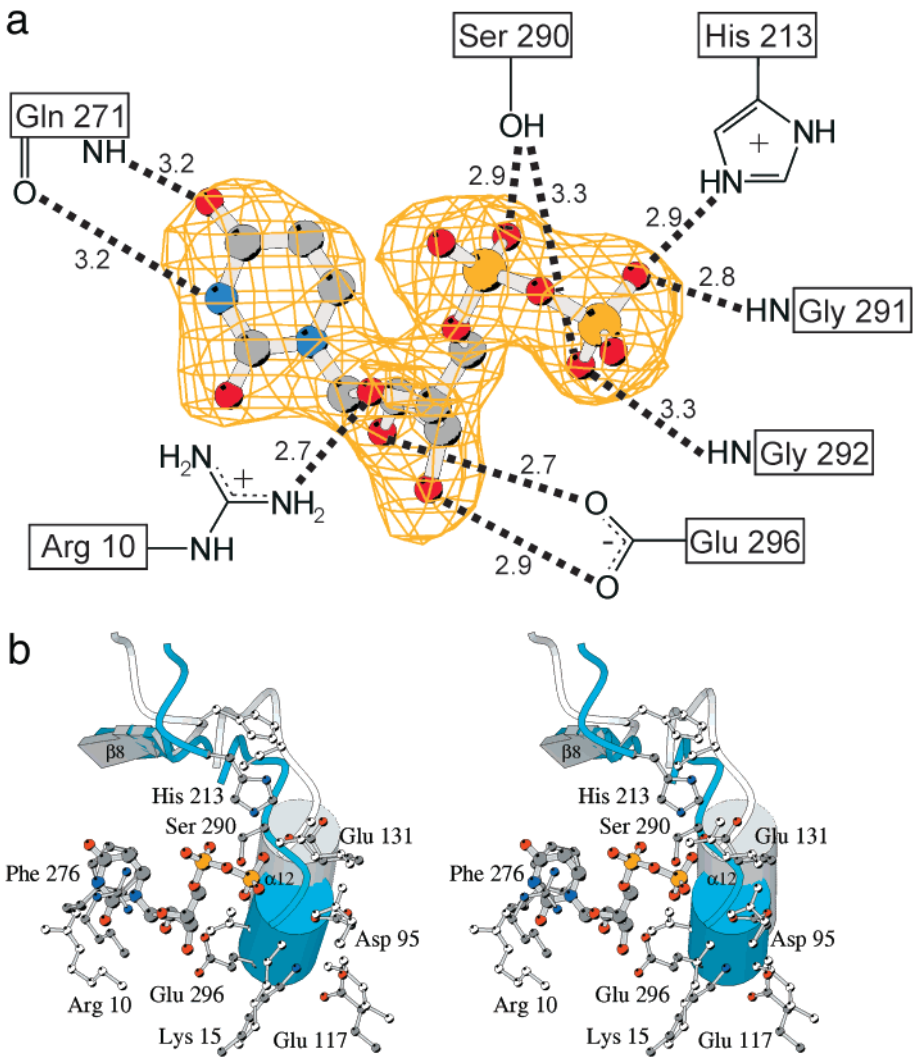


FIGURE 3: (a) Hydrogen bonds (angstroms) involved in binding the UDP portion of the substrate. The density corresponds to an $F_o - F_c$ omit map contoured at 3σ calculated with the final coordinates minus UDP. (b) Superposition of the active sites of both the open and closed chains such that overlap of the uridine groups is maximized. Residues from the closed chain (blue) are represented with colored atoms, while residues from the open chain (gray) are represented with white atoms.

In Figure 3a, the hydrogen bonds involved in binding of the UDP moiety are illustrated. In the closed chain, further stabilization of the bound UDP results from π-stacking of the uracil ring between the phenyl of Phe276 and the guanidinium of Arg10. Arg10 also forms a hydrogen bond with the ribose ring oxygen (2.7 Å) and is within 3.6 Å of the α-phosphate of UDP. In addition to the hydrogen bond with Arg10, the ribose of UDP forms two hydrogen bonds (2.7 and 2.9 Å) between the ribose hydroxyls and the carboxylate of Glu296. The critical interactions for binding of the pyrophosphate group are hydrogen bonds with the hydroxyl of Ser290 (2.9 and 3.2 Å) and the main chain amides of Gly291 (2.8 Å) and Gly292 (3.3 Å) which are found at the N-terminus of α12. His213 also contributes to stabilization of the pyrophosphate group, but as this interaction may be relevant to the catalytic mechanism, it will be further discussed below. The most important differences in the protein−UDP interactions between the open and closed chains are shown in Figure 3b. In the open chain, the 10° interdomain rotation shifts the N-terminus of α12 approximately 3−4 Å away from its position in the closed chain, drawing Glu296 back from the ribose ring and disrupting the majority of the possible hydrogen bonds with UDP. As will be discussed, this conformational change is probably associated with the regulatory mechanism of UDP-GlcNAc 2-epimerase.

DISCUSSION

Despite its widespread distribution in organisms ranging from bacteria to mammals, GP had no known structural analogues prior to the publication of the structure of BGT (*40*). The observed structural homology between BGT and GP prompted some initial speculation about their evolutionary relationship and the suggestion that they are probably the first two members of a glycosyltransferase superfamily that shares a distinctive core structure (*38*, *39*). Interestingly, the work reported here shows that the third member of the GP and BGT superfamily is UDP-GlcNAc 2-epimerase, an enzyme that catalyzes the reversible interconversion of UDP-GlcNAc and UDP-ManNAc. This apparent incongruity is explained when one considers that as with a glycosyltransferase, the epimerization mechanism proceeds via the cleavage of a phosphoglycosyl bond, albeit in an elimination event as opposed to a substitution.

A unique aspect of the bacterial epimerase is the strict regulatory role played by the substrate UDP-GlcNAc which

is required for activity (half-maximal activation occurs at approximately 0.6 mM) (*1*, *2*, *4*). This is most readily observed in the sigmoidal dependence of the reaction rate on the UDP-GlcNAc concentration (Hill coefficient of 2.0) and on the inability of the enzyme to epimerize pure UDP-ManNAc in the absence of UDP-GlcNAc. Our structure shows that each subunit of the homodimeric bacterial UDP-GlcNAc 2-epimerase has a single UDP-GlcNAc binding site, and consequently, allosteric activation through binding to a second regulatory site on the enzyme can be ruled out. The best explanation of the regulatory mechanism is that binding of UDP-GlcNAc to one subunit induces a conformational change across the dimer interface that converts the dimer partner to the catalytically active form. The two forms of UDP-GlcNAc 2-epimerase observed in the X-ray structure differ by a 10° interdomain rotation and several local conformational changes. The interdomain rotation is reminiscent of the well-studied transition between the inactive T state and the active R state of GP (*46*) and the conformational change that occurs upon substrate binding in BGT (*40*). Among the most intriguing differences between the open and closed chains is the movement of the flexible loop region between β8 and α9 (residues 211−218) that contains the strictly conserved His213 (Figure 3b). In the closed chain, the imidazole ring of His213 forms a hydrogen bond (2.9 Å) with the β-phosphate of UDP. In the open chain, the Thr211−Phe218 loop has swung out of the active site, placing the imidazole ring of His213 more than 7 Å away from the β-phosphate. His213 of the closed chain is the only residue that forms a hydrogen bond with the β-phosphate of UDP and could act as a general acid catalyst to activate the UDP of the substrate for elimination. On the basis of this proposal, it is probable that the closed chain of UDP-GlcNAc 2-epimerase is the active subunit and the open chain is the regulatory subunit of the homodimer.

Although UDP-GlcNAc 2-epimerase shares a common overall structure with BGT and the core domains of GP, these three enzymes do not share any conserved catalytic residues. The only residue that is strictly conserved through all sequences of UDP-GlcNAc 2-epimerase and is found in either BGT or GP is Glu296 that aligns with Glu272 of BGT. As in UDP-GlcNAc 2-epimerase, Glu272 of BGT forms hydrogen bonds with the ribose hydroxyls of its substrate, UDP-glucose (*40*). As previously noted, this position also aligns with Lys680 of GP that forms a Schiff base linkage with the pyridoxal 5′-phosphate coenzyme (*42*, *43*). The structure-based alignment with BGT and GP also revealed that His213 aligns with Ser189 of BGT and Arg569 of GP, two residues that have been implicated in binding the phosphate groups of their respective substrates (*40*, *46*). In the transition of GP from the inactive T state to the active R state, Arg569 undergoes a conformational change that results in formation of the phosphate recognition site for glucose 1-phosphate (*46*). As discussed above, His213 undergoes an analogous conformational change during the 10° interdomain rotation and adopts a position in which it can act as a general acid catalyst only in the closed chain.

The detailed mechanism by which UDP-GlcNAc 2-epimerase catalyzes both the *syn* and *anti* elimination of UDP is unknown. It has been suggested that the elimination occurs by either an E1 or an "E1-like E2" mechanism with significant oxocarbenium ion-like character in the transition state (see the lower left inset in Scheme 1) (*4*, *5*). On the basis of this structure, active site residues that likely play key roles in promoting the elimination steps include Asp95, Glu117, and Glu131. They may serve as the acid or base residues involved in proton transfer at C-2 or serve to stabilize an oxocarbenium ion-like transition state via electrostatic catalysis. In addition, His213 is a good candidate for acting as a general acid that protonates the departing UDP. The observation that UDP-GlcNAc 2-epimerase shares a common overall fold with GP and BGT suggests a distant evolutionary relationship exists between these two enzymes. A possible mechanistic link is further supported by experimental evidence for an oxocarbenium ion-like transition state in GP (see the lower right inset in Scheme 1) (*25*, *47*). During the final revision of the manuscript, a paper appeared describing the structure of a GlcNAc transferase, MurG, involved in peptidoglcan biosynthesis (*48*). Superposition of epimerase with MurG shows a marked similarity, suggesting it also evolved from a common ancestor (rms deviation of 2.3 Å for 157 common C-α atoms or, if the two domains in each molecule are superposed individually, an rms deviation of 2.0 Å on 225 common C-α atoms). This notion is particularly appealing since both MurG and the epimerase utilize UDP-GlcNAc as a substrate. It is tempting to speculate that although the primary sequence and catalytic specificity of UDP-GlcNAc 2-epimerase has drastically diverged from that of GP, BGT, and MurG, the ability to stabilize a transition state with oxocarbenium ion character has persisted.

ACKNOWLEDGMENT

We are grateful to the Department of Energy and Robert Sweet at the NSLS at the Brookhaven National Laboratory for access to beamline X12C. We also thank Mark Paetzel for assistance in data collection.

REFERENCES

1. Kawamura, T., Kimura, M., Yamamori, S., and Ito, E. (1978) *J. Biol. Chem. 253*, 3595−3601.
2. Kawamura, T., Ishimoto, N., and Ito, E. (1979) *J. Biol. Chem. 254*, 8457−8465.
3. Sala, R. F., Morgan, P. M., and Tanner, M. E. (1996) *J. Am. Chem. Soc. 118*, 3033−3034.
4. Morgan, P. M., Sala, R. F., and Tanner, M. E. (1997) *J. Am. Chem. Soc. 119*, 10269−10277.
5. Tanner, M. E., and Kenyon, G. L. (1998) in *Comprehensive Biological Catalysis* (Sinnott, M., Ed.) Vol. II, pp 7−41, Academic Press, San Diego.
6. Tanner, M. E. (1998) in *Comprehensive Biological Catalysis* (Sinnott, M., Ed.) Vol. III, pp 76−82, Academic Press, San Diego.
7. Lee, C.-J., Banks, S. D., and Li, J. P. (1991) *Crit. Rev. Microbiol. 18*, 89−114.
8. Morona, J. K., Morona, R., and Paton, J. C. (1997) *Mol. Microbiol. 23*, 751−763.
9. Yokoyama, K., Mizuguchi, H., Araki, Y., Kaya, S., and Ito, E. (1989) *J. Bacteriol. 171*, 940−946.
10. Harrington, C. R., and Baddiley, J. (1985) *Eur. J. Biochem. 153*, 639−645.
11. Meier-Dieter, U., Barr, K., Starman, R., Hatch, L., and Rick, P. D. (1992) *J. Biol. Chem. 267*, 746−753.
12. Kuhn, H.-M., Meier-Dieter, U., and Mayer, H. (1988) *FEMS Microbiol. Rev. 54*, 195−222.
13. Hinderlich, S., Stäsche, R., Zeitler, R., and Reutter, W. (1997) *J. Biol. Chem. 272*, 24313−24318.

14. Stäsche, R., Hinderlich, S., Weise, C., Effertz, K., Lucka, L., Moorman, P., and Reutter, W. (1997) *J. Biol. Chem. 272*, 24319−24324.
15. Traving, C., and Schauer, R. (1998) *Cell. Mol. Life Sci. 54*, 1330−1349.
16. Varki, A. (1997) *FASEB J. 11*, 248−255.
17. Karlsson, K. A. (1995) *Curr. Opin. Struct. Biol. 5*, 622−635.
18. Hennet, T., Chui, D., Paulson, J. C., and Marth, J. D. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 4504−4509.
19. Yogeeswaran, G., and Salk, P. L. (1981) *Science 212*, 1514−1516.
20. Keppler, O. T., Hinderlich, S., Langner, J., Schwartz-Albiez, R., Reutter, W., and Pawlita, M. (1999) *Science 284*, 1372−1376.
21. Effertz, K., Hinderlich, S., and Reutter, W. (1999) *J. Biol. Chem. 274*, 28771−28778.
22. Salo, W. (1976) *Biochim. Biophys. Acta 452*, 625−628.
23. Glaser, L. (1960) *Biochim. Biophys. Acta 41*, 534−536.
24. Sommar, K. M., and Ellis, D. B. (1972) *Biochim. Biophys. Acta 268*, 590−595.
25. Mitchell, E. P., Wither, S. G., Ermert, P., Vasella, A. T., Garman, E. P., Oikonomakos, N. G., and Johnson, L. N. (1996) *Biochemistry 35*, 7341−7355.
26. Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L., and Sweet, R. M. (1993) *Nature 362*, 219−223.
27. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−325.
28. Terwilliger, T. C., and Berendzen, J. (1997) *Acta Crystallogr. D53*, 571−579.
29. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
30. McRee, D. E. (1992) *J. Mol. Graphics 10*, 44−46.
31. Brünger, A. T., Adams, P. D., Clore, G. M., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nigles, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
32. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993*) J. Appl. Crystallogr. 26*, 283−291.
33. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
34. Hutchinson, E. G., and Thornton, J. M. (1996) *Protein Sci. 5*, 212−220.
35. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
36. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
37. Rossmann, M. G., Liljas, A., Branden, C.-I., and Banaszak, L. J. (1975) in *The Enzymes* (Boyer, P. D., Ed.) pp 61−102, Academic Press, New York.
38. Hinderlich, S., Sonnenschein, A., and Reutter, W. (1998) *Biometals 11*, 253−258.
39. Holm, L., and Sander, C. (1993) *J. Mol. Biol. 233*, 123−138.
40. Vrielink, A., Ruger, W., Driessen, H. P. C., and Freemont, P. S. (1994) *EMBO J. 13*, 3413−3422.
41. Morera, S., Imberty, A., Aschke-Sonnenborn, U., Ruger, W., and Freemont, P. S. (1999) *J. Mol. Biol. 292*, 717−730.
42. Artymiuk, P. J., Rice, D. W., Poirrette, A. R., and Willett, P. (1995) *Struct. Biol. 2*, 117−120.
43. Holm, L., and Sander, C. (1995) *EMBO J. 14*, 1287−1293.
44. Jones, S., and Thorton, J. M. (1995) *Prog. Biophys. Mol. Biol. 63*, 31−65.
45. Newgard, C. B., Hwang, P. K., and Fletterick, R. J. (1989) *CRC Crit. Rev. Biochem. Mol. Biol. 24*, 69−99.
46. Barford, D., and Johnson, L. N. (1989) *Nature 340*, 609−616.
47. Street, I. P., Rupitz, K., and Withers, S. G. (1989) *Biochemistry 28*, 1581−1587.
48. Ha, S., Walker, D., Shi, Y., and Walker, S. (2000) *Protein Sci. 9*, 1045−1052.

BI001627X