The crystal structure of ADP-L-glycero-D-mannoheptose 6-epimerase: catalysis with a twist

AM Deacon[1], YS Ni[2], WG Coleman Jr[2]* and SE Ealick[1]*

Background: ADP-L-glycero-D-mannoheptose 6-epimerase (AGME) is required for lipopolysaccharide (LPS) biosynthesis in most genera of pathogenic and non-pathogenic Gram-negative bacteria. It catalyzes the interconversion of ADP-D-glycero-D-mannoheptose and ADP-L-glycero-D-mannoheptose, a precursor of the seven-carbon sugar L-glycero-D-mannoheptose (heptose). Heptose is an obligatory component of the LPS core domain; its absence results in a truncated LPS structure resulting in susceptibility to hydrophobic antibiotics. Heptose is not found in mammalian cells, thus its biosynthetic pathway in bacteria presents a unique target for the design of novel antimicrobial agents.

Results: The structure of AGME, in complex with NADP and the catalytic inhibitor ADP-glucose, has been determined at 2.0 Å resolution by multiwavelength anomalous diffraction (MAD) phasing methods. AGME is a homopentameric enzyme, which crystallizes with two pentamers in the asymmetric unit. The location of 70 crystallographically independent selenium sites was a key step in the structure determination process. Each monomer comprises two domains: a large N-terminal domain, consisting of a modified seven-stranded Rossmann fold that is associated with NADP binding; and a smaller α/β C-terminal domain involved in substrate binding.

Conclusions: The first structure of an LPS core biosynthetic enzyme leads to an understanding of the mechanism of the conversion between ADP-D-glycero-D-mannoheptose and ADP-L-glycero-D-mannoheptose. On the basis of its high structural similarity to UDP-galactose epimerase and the three-dimensional positions of the conserved residues Ser116, Tyr140 and Lys144, AGME was classified as a member of the short-chain dehydrogenase/reductase (SDR) superfamily. This study should prove useful in the design of mechanistic and structure-based inhibitors of the AGME catalyzed reaction.

Addresses: [1]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853, USA and [2]Laboratory of Biochemistry and Genetics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA.

*Corresponding authors.
E-mail: see3@cornell.edu
wc3z@nih.gov

Key words: ADP-L-glycero-D-mannoheptose, epimerase, lipopolysaccharide core biosynthesis, multiwavelength anomalous diffraction

Received: 6 January 2000
Revisions requested: 2 February 2000
Revisions received: 17 February 2000
Accepted: 22 February 2000

Published: 12 April 2000

Structure 2000, 8:453–462

0969-2126/00/$ – see front matter

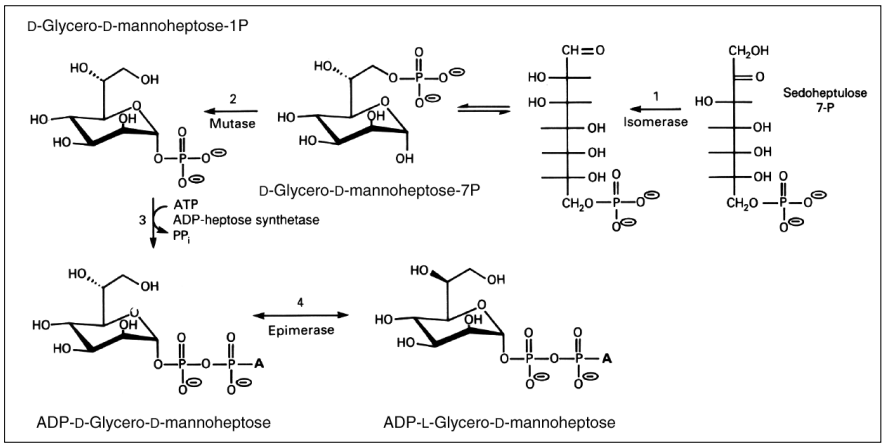
Introduction

Gram-negative bacteria possess two membranes that separate the cytoplasmic compartment from the extracellular environment: the cytoplasmic membrane and the outer membrane. The properties of the cytoplasmic membrane are consistent with those of a typical phospholipid bilayer. The outer membrane is also a bilayered structure: the inner most leaflet is composed primarily of phospholipids, whereas its outer one contains a unique lipid known as lipopolysaccharide (LPS).

The LPS of most Gram-negative bacteria consists of lipid A, core and O-antigen domains. The core oligosaccharide has an outer portion composed of hexoses and *N*-acetylglucosamine and an inner portion composed of the seven-carbon sugar L-glycero-D-mannoheptose (or heptose) and 2-keto-3-deoxyoctulosonic acid. Heptose is a highly conserved component of the LPS core among several genera of enteric and nonenteric bacteria [1].

LPS has an important role in maintaining the structural integrity of the outer membrane by interacting with other components and providing a physical barrier against the entry of deleterious compounds and some bacteriophages [2]. *Escherichia coli* K-12 LPS heptoseless mutants (resulting from a single-site mutation, designated *rfaD70*) display a dramatic reduction in porin proteins and are unable to grow in media containing detergents, bile salts or hydrophobic antibiotics [2,3]. The increased permeability of hydrophobic agents across the outer membrane of heptoseless or other *rfaD* mutants is directly attributed to the loss of the barrier function. Furthermore, the ability of heptoseless or core-defective LPS mutants to survive in body sites of infected hosts is compromised [4]. The LPS biosynthetic pathway is an attractive therapeutic target, as mutants that produce core-defective LPS are less pathogenic and more susceptible to known antibiotics [2–4].

Figure 1



The biosynthetic pathway for ADP-L-glycero-D-mannoheptose. This pathway was first proposed by Eidel and Osborn [5] and has been substantiated and further elaborated by results from several laboratories [2,5,7,8]. Its presentation here using chemical formulae was adopted from a review article by CRH Raetz [6]. (Reprinted with permission from the Annual Review of Biochemistry, Volume 12 © 1990 by Annual Reviews www.AnnualReviews.org.)

ADP-L-glycero-D-mannoheptose 6-epimerase (AGME) catalyzes the interconversion between ADP-D-glycero-D-mannoheptose and ADP-L-glycero-D-mannoheptose, the last step (Figure 1) in the biosynthesis of the precursor of L-glycero-D-mannoheptose [5–8]. In 1983, AGME activity was first demonstrated in extracts of *E. coli* K-12 [7]. In 1990, the gene encoding AGME (*rfaD*) was cloned; the gene product was partially purified and the N terminus of the primary structure of the protein was shown to contain the ADP-binding βαβ fold fingerprint sequence, Gly-Gly-X-Gly-X-X-Gly [8,9]. AGME was purified to homogeneity in 1994 and several physicochemical properties of the enzyme were investigated [10]. The enzyme was shown to be an NADP-dependent epimerase, but also has diminished activity with NAD. The cofactor is tightly bound to AGME and its removal resulted in a loss of enzyme activity and secondary structure. Preliminary crystallization and X-ray studies of AGME were recently reported [11].

The occurrence of the epimerase in several genera of Gram-negative bacteria and its absence in mammalian cells makes this protein a suitable target for new antimicrobial strategies. The X-ray structure determination of AGME was undertaken to better understand the mechanism of catalysis, inhibition and inactivation, and to permit the design of potent and selective mechanism-based inhibitors. The enzyme is a homopentamer. Each monomer is comprised of two domains: a large N-terminal domain consisting of a modified seven-stranded Rossmann fold that is associated with NADP binding; and a smaller α/β C-terminal domain involved in substrate binding. The C-terminal domain appears to position the substrate with respect to the nicotinamide. Subtle variations in the conformation of this domain allow the substrate sugar to twist, thereby allowing the catalytic epimerization reaction to proceed. On the basis of its high structural similarity with UDP-galactose epimerase and the three-dimensional positions of the conserved residues Ser116, Tyr140 and Lys144, AGME is classified as a member of the short-chain dehydrogenase/reductase (SDR) superfamily.

Results and discussion

Structure determination by MAD phasing

AGME exhibits low sequence identity with other enzymes and early attempts to determine the structure by molecular replacement were unsuccessful. The presence of seven methionine residues in each 310 amino acid monomer, however, offered us the possibility of using selenomethionine-based multiwavelength anomalous diffraction (MAD) for phase determination. Preliminary X-ray characterization of AGME [11] showed that it crystallized in space group $P2_1$. The cell dimensions (a = 99.5 Å, b = 109.8 Å, c = 181.5 Å, β = 91.0°) implied that several monomers were present in the asymmetric unit. Early biochemical analysis (i.e., sodium dodecyl sulfate polyacrylamide gel electrophoresis, gel filtration and sedimentation equilibrium studies) of AGME suggested a homohexameric structure [10] and we initially assumed either one or two hexamers were located in the asymmetric unit [11]. Self-rotation searches performed against the native X-ray diffraction data using the program GLRF [12], however, revealed fivefold noncrystallographic symmetry (NCS), indicating that the enzyme was in fact pentameric. In addition, a reasonable Matthews coefficient [13] of 2.7 $Å^3$/Da and a solvent content of 53%, both corresponded to there being two such pentamers in the asymmetric unit. Therefore, 70 crystallographically independent selenium sites had to be located as part of the MAD phasing process.

The selenium substructure was determined with the SnBv2.0 package [14]. The peak wavelength anomalous differences were processed with the DREAR suite [15] to generate difference normalized structure factors (diffE values) [16]. The largest 1400 diffE values were used in SnBv2.0 to generate 14,000 triplet invariants. All trials were conducted on a 500 MHz Digital/Compaq Alpha workstation, taking

20 min per random trial. Each trial was processed for 140 cycles of dual-space refinement. After 236 trials had been processed, one trial was identified that had a significantly lower minimal function value [17] (R_{min} = 0.43) compared to the other 235 trials (R_{min} = 0.55–0.60). This indicated that the 70 selenium atom substructure had been unambiguously determined. All of these initial trials were carried out using an electron-density grid size of 0.9 Å. Later it was found that similar success rates could be achieved using a substantially coarser grid of 2.0 Å. The use of a coarser grid realized an almost eightfold increase in the speed performance of SnB.

Once the selenium atoms had been located their positions were confirmed in several ways. Firstly, in the absence of a second independent SnB solution, five of the top peaks from the initial solution were fed back into the SnB program as a new starting trial. The minimal function quickly converged again to 0.43. The resulting substructure consistently matched the original solution, no matter which five of the top peaks were recycled in this way. Secondly, the fivefold axes for the two pentamers were located from the selenium positions; the axes were parallel to the crystallographic 2_1 axis and matched the previous results from self-rotation searches. Finally, after processing more trials, the peak lists from two completely independent SnB solutions were compared. In total, 65 matching sites were found between these two solutions. These sites were then fed into MLphare [18] for phasing, followed by DM [19] for density modification, according to standard protocols. The resulting electron-density maps were of very high quality even before any NCS averaging had been carried out (Figure 2).

The 70 selenium atom substructure of AGME is the largest substructure that has been used in a *de novo* MAD structure determination [20]. It provided a challenging system to test the robustness of current phasing strategies for selenomethionine-based MAD structure determination and should encourage the use of selenomethionine as the anomalous scatterer to solve even larger structures.

Quality of the model

The final model consisted of ten monomers (two pentamers) of AGME. Four of the monomers were essentially complete, except for three disordered amino acid residues at the C terminus (residues 308–310). Three of the other monomers were also missing residues 265–271, one monomer was missing residues 262–271, one monomer was missing residues 195–207 and 251–271 and the final monomer was missing residues 194–206 and 249–271. The final model also included 996 bound water molecules, ten molecules of NADP and ten molecules of ADP-glucose inhibitor. However, the electron density for the glucose moiety was weak and difficult to interpret in seven of the monomers. In these cases the inhibitor was truncated to an ADP molecule. The refinement converged to a final R factor of 21.2% and a free R factor of 26.2%. The final model exhibited good stereochemical geometry. A Ramachandran plot calculated with PROCHECK [21] showed that 89.3% of the residues were located in the most favored regions and the other 10.7% were in additionally allowed regions.

Overall structure

The AGME monomer measures 35 Å × 35 Å × 65 Å. Architecturally it consists of two domains: a predominantly N-terminal domain including amino acid residues 1–167, 214–236 and 280–292; and a smaller C-terminal domain formed from amino acid residues 168–213, 237–279 and 293–310 (Figure 3a). The N-terminal domain consists of a modified Rossmann fold [22], with a central seven-stranded parallel β sheet, flanked on either side by a total of seven α helices. This domain is associated with the tight binding of the enzymatic cofactor NADP. The C-terminal domain is a globular cluster of two small parallel β sheets and three α helices (Figure 3b). This domain provides the residues that define the specificity for the substrate, which binds in a cleft between the two domains.

The topology of AGME clearly shows that it is a member of the SDR family. The overall structure is very similar to

Figure 2

Electron-density plots of the NADP molecule. (a) Experimental phases without averaging at 3.0 Å resolution. (b) Experimental phases with tenfold averaging at 3.0 Å resolution. (c) Final refined phases at 2.0 Å resolution.

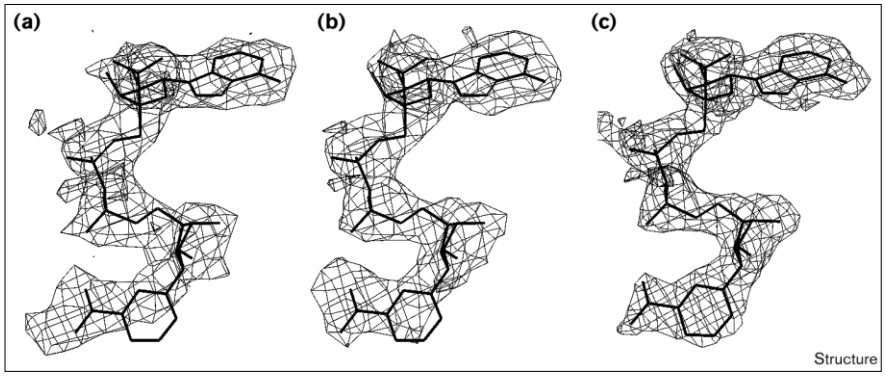
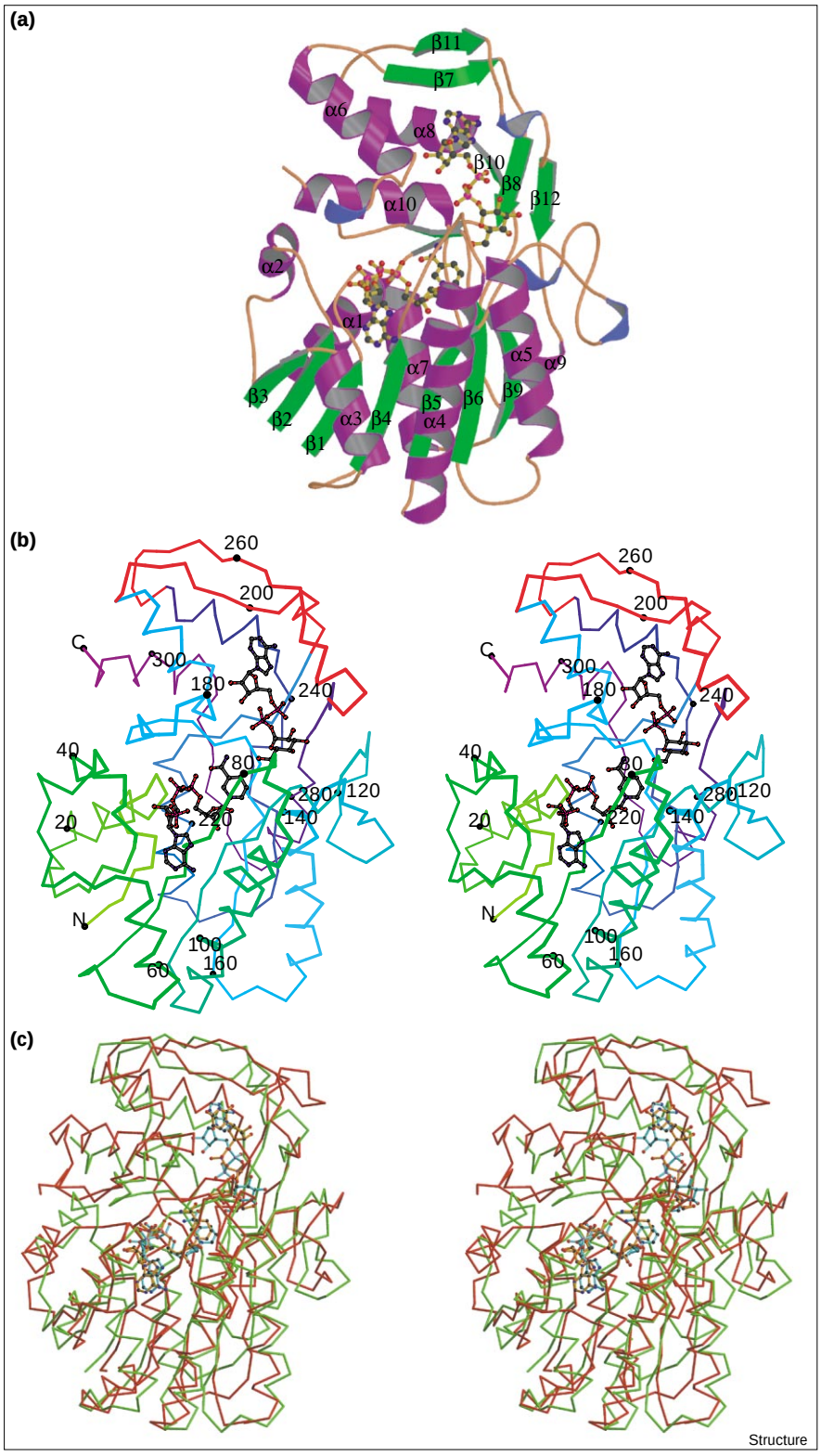


Figure 3



Structure of the AGME monomer. (a) Ribbon diagram of AGME showing its overall fold. The N-terminal domain consist of a seven-stranded Rossmann fold associated with NADP binding and the C-terminal domain consists of a small α/β domain involved in substrate binding. The numbered β strands are represented as green arrows and the α helices as purple spirals; loops are represented as orange pipes. NADP and ADP-glucose are shown in ball-and-stick representation. (b) Stereoview Cα trace of AGME. The N and C termini and every twentieth residue are labeled. The polypeptide chain is color-coded from the N terminus (green) to the C terminus (purple). The red region is disordered in some of the ten monomers in the asymmetric unit. (c) Superimposition of the Cα traces of AGME (green) and UGE (red), showing bound NAD and NADP. NADP and ADP-glucose are shown in cyan; NAD and UDP-phenol are in orange. The illustration was prepared using the program MOLSCRIPT [33].

UDP-galactose 4-epimerase (UGE, Protein Data Bank [PDB] codes 1NAH and 1NAI; Figure 3c) [23], with a root mean square (rms) deviation of 1.88 Å for 238 Cα atoms, and dTDP-D-glucose 4,6 dehydratase (DGD, PDB code

1BXK), with an rms deviation of 1.85 Å for 237 Cα atoms. Similarity is also seen with the structure of GDP-4-keto-6-deoxy-D-mannose epimerase/reductase (GMER, PDB code 1BWS) [24,25], which has a modified Rossmann fold with a six-stranded parallel β sheet (rms deviation of 2.14 Å for 226 Cα atoms). These proteins typically show about 20–30% sequence conservation.

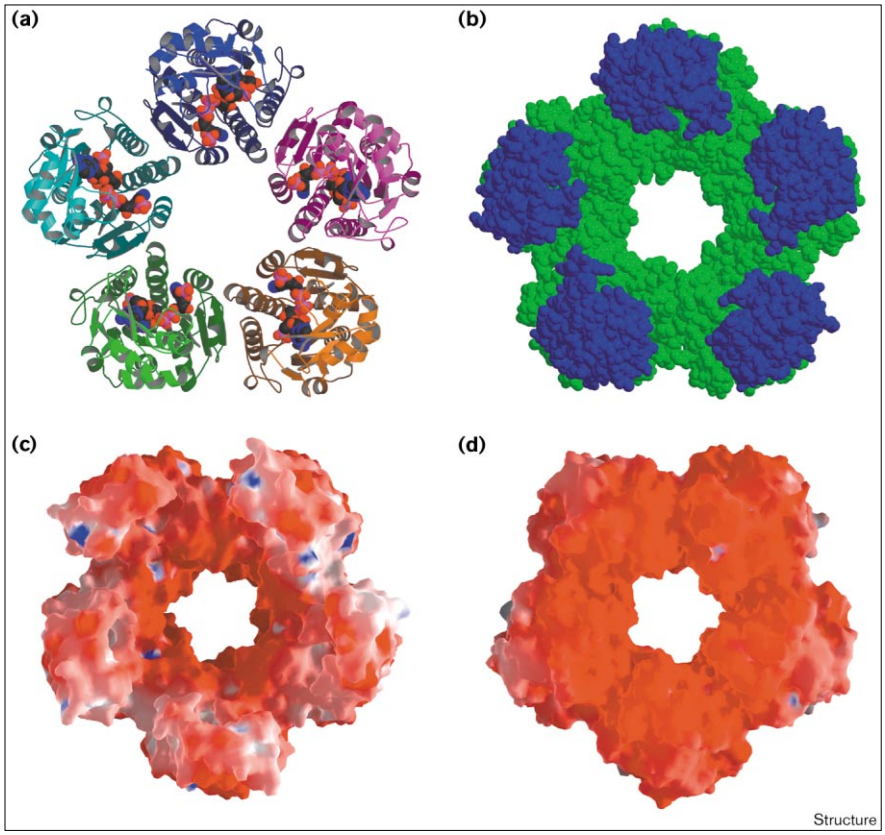
Quaternary structure and intersubunit contacts

AGME is a homopentamer with the individual subunits related by a local fivefold symmetry axis (Figure 4a). NAD(P)-dependent enzymes are often either dimers or tetramers. In particular, many members of the SDR family are dimers [25] and share a common dimer interface motif, in which the long α4 and α5 helices from the N-terminal domain of each subunit pack together to form a four-helix bundle [24]. This interface results in the C-terminal substrate-binding domains being located on opposite sides of the dimer. In the AGME pentamer the subunit interface is also formed solely from the N-terminal domains. However, the contributing residues are from helix α2 and strand β3 of one subunit and helices α4 and α5 of the adjacent subunit. This interface buries 1750 Å^2 surface area per monomer, compared with a total monomer surface area of 11,250 Å^2. The local fivefold symmetry results in the C-terminal domains of the individual subunits being arranged on the same face of the pentamer, with the substrate-binding cleft facing towards the interior of the pentamer (Figure 4b).

Nevertheless, there are no interactions between the C-terminal domains of the individual subunits. The AGME pentamer measures 70 Å on each side, with a channel down the center. The channel measures 50 Å across between residues from the C-terminal domain and narrows to 30 Å across at the base, between residues from the N-terminal domain, where the subunits are tightly packed (Figure 4b). The disordered residues that were omitted from the final model (described above) were all located in the C-terminal domain, mainly in β strands β7 and β11 and their connecting loops. This region is involved in rather weak crystal-packing contacts with the same region from an adjacent pentamer. The precise contacts vary from monomer to monomer. The most well ordered monomers are involved in more intimate crystal contacts with a few direct hydrogen bonds, whereas the more disordered monomers lack any significant crystal contacts in the C-terminal domain. These weak interpentamer crystal contacts contributed to the significant anisotropic diffraction that was observed for crystals of AGME.

An electrostatic surface calculated for AGME is characterized by a large negatively charged area on the base of

Figure 4



Structure of the AGME pentamer. (a) Ribbon representation of the AGME pentamer with space-filling representations of NADP and ADP-glucose. (b) Corey, Pauling and Koltun (CPK) space-filling representation of the AGME pentamer (N-terminal domain in green; C-terminal domain in blue). (c,d) van der Waals surface representation of AGME. The surface is colored according to electrostatic potential: blue for positive and red for negative. The bottom surface (d) is very negative. The top view (c) is less interesting electrostatically, but does show the overall shape with the small substrate-binding domain arranged on top of the larger NADP-binding domain. The figure was prepared using the program GRASP [34].

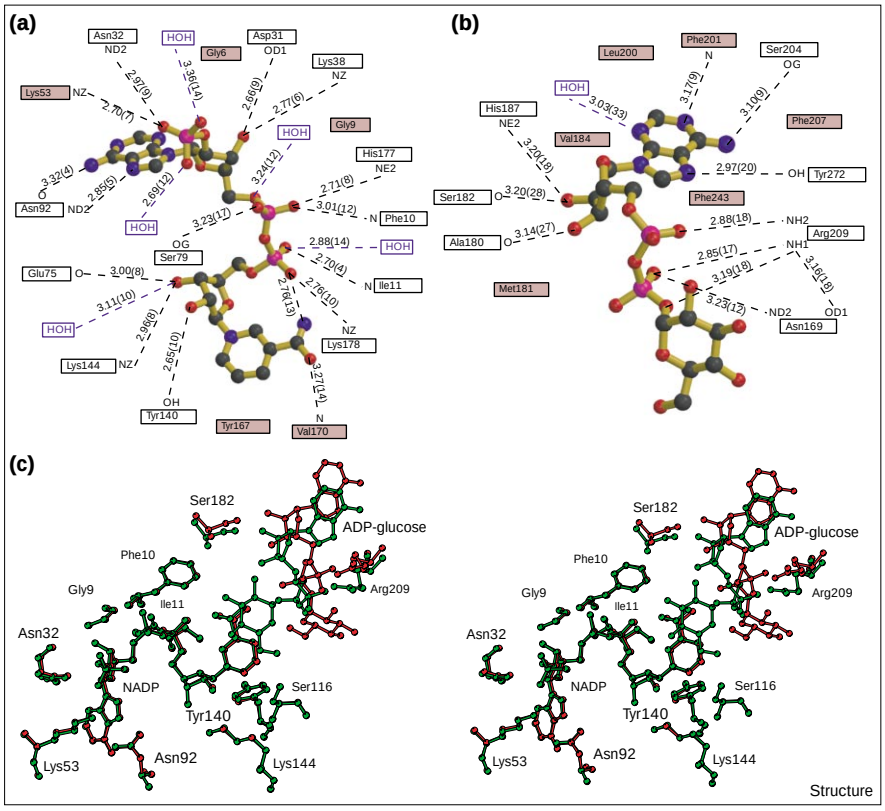
the pentamer (residues Asp26, Glu64, Glu65, Asp68, Glu70, Glu106, Glu108, Glu132, Glu157 and Glu226) and also around the interior surface at the bottom of the pentamer (residues Asp52, Glu54, Asp55, Asp85, Asp91 and Glu99; Figure 4c).

NADP-binding site

The NADP-binding site is located at the topological switch point of the N-terminal domain modified Rossmann fold. The binding site is characterized by the fingerprint sequence Gly-Gly-X-Gly-X-X-Gly, which forms a tight loop between β1 and α1 (residues 6–12). The NADP is tightly bound in an extended conformation with an average distance of 12.90 Å (range 12.71–13.12 Å over the ten monomers) between the adenine C6 and the nicotinamide C2 position (Figure 5a). This is shorter than the distances observed in other similar structures (13.91 Å for UGE, 14.39 Å for DGD and 15.47 Å for GMER). This more closed conformation allows AGME to accommodate NADP rather than NAD, with extra specificity for the phosphate being provided by Lys53 and Asn32 and two tightly bound water molecules. In addition, the loop between Asp31 and Tyr50, which closes over the cofactor like a lid, is in a more open conformation than compared with UGE (which binds NAD). This region provides extra specificity for NADP, while maintaining the same tight binding of the cofactor as UGE. This lid region is missing in GMER. AGME exhibits a loss of secondary structure but the active enzyme can be reconstituted on addition of fresh cofactor. GMER, on the other hand, shows little change in structure between the NADP-bound and free forms [24,25]. The importance of the fingerprint sequence associated with cofactor binding is demonstrated by a single-site mutation in this region [2,3]. The mutation, designated *rfaD70*, results in a Gly6→Ser substitution. The mutation of glycine to serine alters the size and flexibility of the NADP-binding motif, resulting in an inactive enzyme. Consequently, the *rfaD70* mutation results in heptoseless LPS structures.

The adenine base is firmly locked in position through hydrophobic stacking interactions between the entire length of the Lys53 sidechain on one side and the mainchains of Gly76 and Ala77 on the other (Figure 5a). Furthermore, hydrogen bonds coordinating the carbonyl oxygen and the OD1 of Asn92 with the N6 and N7 of the adenine base cause a distortion in the longest α helix (α4), which results in four amino acid residues in a single turn of the helix. The carbonyl oxygen of Asn92 does not participate in the regular α-helical hydrogen-bonding pattern. Consequently, the carbonyl oxygens of Met90 and Asp91 are hydrogen bonded to the mainchain nitrogen atoms of amino acids five residues ahead in the chain. A similarly distorted

Figure 5



The cofactor- and substrate-binding sites of AGME. (a) The NADP-binding site. (b) The ADP-glucose-binding site. Potential hydrogen bonds are indicated by dashed lines with distances given in Å, the numbers in brackets indicate the standard deviation of the equivalent distances for all ten AGME molecules. Shaded boxes indicate hydrophobic contacts. Atoms are shown in standard colors: carbon, black; nitrogen, blue; oxygen, red; phosphorus, magenta. (c) Stereoview of the active site, two AGME molecules are superimposed (red and green). The glucose moiety of the substrate (ADP-glucose) adopts two different conformations. For clarity only selected residues are shown.

feature is also present in the structure of UGE, however, in GMER the helix appears to be normal and the coordination of the adenine at this point is somewhat different.

Finally, an oxidized cysteine residue (Cys78) is involved in coordinating the adenosine ribose 2′-phosphate. The cysteine appears to be a fully occupied sulfenic acid in all ten monomers. In some of the monomers it coordinates directly, whereas in others the coordination is via a bridging water molecule.

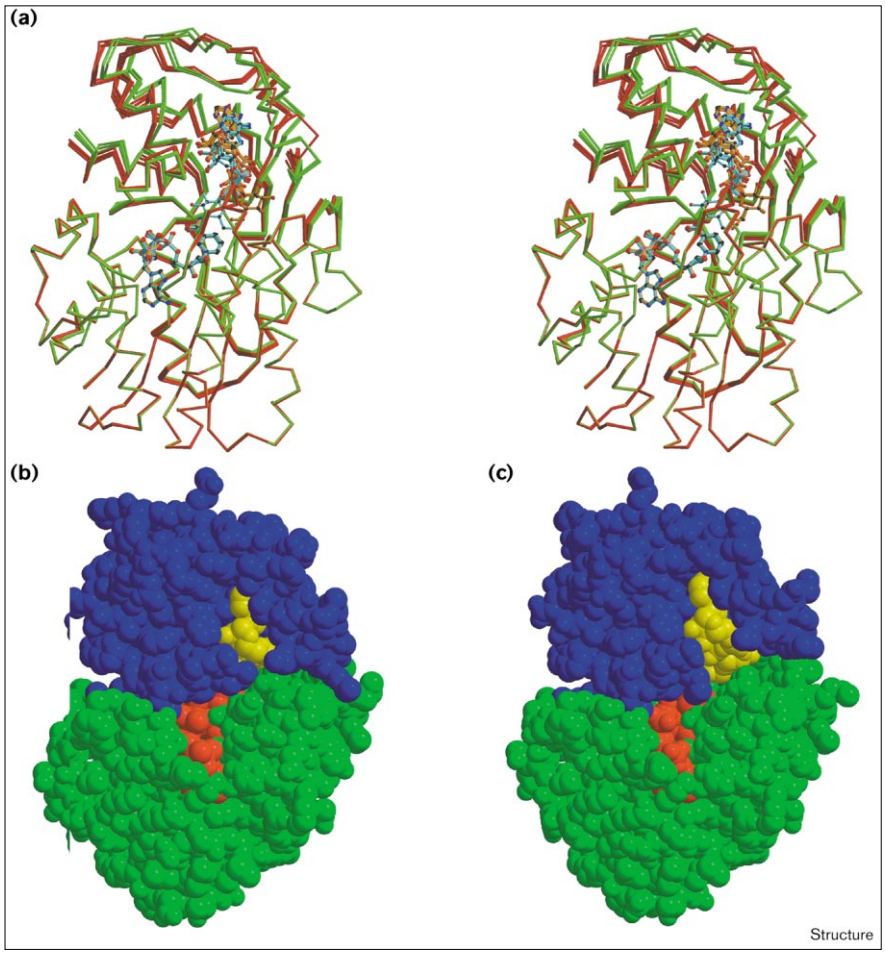
The average atomic temperature factor of the NADP molecules is 28.2 Å^2. The temperature factor at the nicotinamide end, however, is consistently higher (average of 33.2 Å^2) than at the more tightly bound adenine end (average of 24.2 Å^2). There is also more variation in the average temperature factors of the nicotinamide moiety of the ten monomers. This suggests that there is more flexibility at this end of the cofactor, and as a result the electron density is less well defined. The density was still of sufficient quality, however, to allow the unambiguous assignment of the orientation of the nicotinamide. The nicotinamide moiety adopts the *syn* conformation with respect to the ribose, similar to the findings for UGE in complex with UDP-sugars [26]. This is consistent with a nicotinamide B side specific pro-S hydride transfer during catalysis.

Substrate-binding site

The ADP-glucose-binding site is characterized by strong hydrophobic interactions at the adenine base. In particular, stacking interactions sandwich the base between Phe201 and Phe243 (Figure 5b). In addition, the pyrophosphate is pinned in place through hydrogen bonds to Arg273. There appear to be far fewer interactions between the protein and the glucose moiety. As a result, the electron density for the glucose moiety showed more disorder and was difficult to interpret. Although significant electron density was present in all ten monomers, it could only be satisfactorily interpreted to assign the orientation of the glucose in three of these. The average temperature factor of the ADP-glucose (55.4 Å^2) is higher than that of the NADP cofactor (28.2 Å^2) and shows significantly more variation from monomer to monomer. Furthermore, for the three molecules that have the sugar included in the refined model

Figure 6

Comparison of AGME monomers. (a) Stereoview superposition of Cα traces of ten monomers of AGME. The C-terminal domain adopts one conformation in five of the molecules (red) and an alternative conformation in the other five molecules (green). (b,c) CPK space-filling model of two monomers of AGME comparing the different binding of the glucose moiety of ADP-glucose. ADP-glucose (yellow) is more exposed in (c); NADP (red) is uniformly bound in both monomers. The substrate domain (domain B) is in blue and the NADP-binding domain (domain A) is in green.



the ADP-glucose analog has significantly higher temperature factors at the glucose sugar position (71.6 Å^2) than compared with the adenine base (45.2 Å^2) .

The glucose moiety appears in two significantly different orientations in these three molecules, represented primarily by a rotation of approximately 125° in the torsion angle formed from the pyrophosphate linking oxygen, the β-phosphate, the glycosyl oxygen and the glucose C1 position. This twist effectively rotates the β-phosphate O3 onto the O1 position. There is also an additional smaller rotation around the torsion angle formed from the β-phosphate, the glycosyl oxygen, the glucose C1 and the glucose C2 positions. Even though the sugar is not well enough defined in the other seven monomers for it to be included in the refined model, there is nevertheless significant electron density at either the O1 or the O3 position in each molecule, which gives some indication of the more likely glucose orientation. In both orientations the C6 of the glucose is located near to the C4 of the NADP nicotinamide moiety (4.13 Å and 3.96 Å for the molecules in one orientation and 3.27 Å for the molecule in the other orientation).

Domain conformation and catalytic mechanism

The catalytic mechanism of AGME appears to be consistent with the one proposed for UGE [26]. The key active-site residues identified in the structural and biochemical analysis of UGE are generally conserved within members of the SDR family, and are also conserved in AGME (Figure 5c). In particular, Ser116, Tyr140 and Lys144 of AGME are all located in structurally similar positions when compared with the equivalent residues in UGE. All these residues are in close proximity to the ADP-glucose and Ser116 and Tyr140 are 4.5 Å and 4.8 Å from the C4 position of the nicotinamide group, respectively.

In analogy to UGE, where it was found that UDP-galactose and UDP-glucose bound to an inactive mutant of the enzyme in different orientations [26], we observe differences in the orientation of the inhibitor ADP-glucose in the ten monomers of the asymmetric unit. In the case of AGME, however, the active enzyme is used but with a substrate analog that is slightly smaller (i.e., one carbon less). Furthermore, in AGME the orientation of the glucose moiety appears to be correlated with the relative conformation of the C-terminal domain with respect to N-terminal domain. The ten monomers of AGME superimpose with an rms deviation in the range 0.089–0.256 Å if only the first 160 Cα atoms (i.e., all residues from the N-terminal domain) are used in the calculation. However, if all the available Cα atoms are used then the rms deviation increases significantly (0.236–0.965 Å).

There appear to be two distinct conformations for the C-terminal domain (Figure 6a). These conformations seem to be dictated by variations in the crystal-packing freedom of the individual monomers. Five of the monomers are clearly seen to adopt a more closed conformation, whereas the other five monomers adopt a slightly more open conformation. The monomers where the ADP-glucose is in a similar orientation belong to one of these species, whereas the other monomer, with the ADP-glucose in an alternative orientation, belongs to the other species. Even in the cases where the sugar is missing from the final refined model there is a clear correlation between the relative domain conformation and the location of the uninterpreted sugar electron density (i.e., either close to the O1 or the O3 position). The ADP-glucose is clearly more solvent exposed in the second, more open conformation (Figure 6b). A potential cycling between open and closed conformations may also facilitate substrate binding and product release.

Despite the observed differences in conformation and ADP-glucose binding of the individual monomers, there does not appear to be any cooperativity between them with respect to substrate binding and catalysis. As was discussed earlier, there are no interactions between the substrate-binding domains of the individual subunits (Figure 3b), an observation that is consistent with other members of the SDR family where the substrate-binding sites are on opposite sides of the dimer. It appears that the C-terminal domain simply acts to position the substrate with respect to the nicotinamide, and subtle variations in the conformation of this domain allow the substrate sugar to twist thereby allowing the catalytic epimerization reaction to proceed.

Biological implications

Bacterial resistance to antibiotic therapies represents a major challenge to clinicians worldwide. Both intrinsic and acquired factors can cause antimicrobial resistance. The intrinsic resistance of Gram-negative bacteria to oligopeptides and hydrophobic molecules such as antibiotics is attributed to the low permeability of the outer membrane. Lipopolysaccharide (LPS), lipid, and protein are the components of the outer membrane. Characteristics of LPS determine the barrier function of the outer membrane; severe truncation of the LPS structure results in reduced growth rate, increased temperature sensitivity and hypersensitivity to detergents and hydrophobic antibiotics. ADP-L-glycero-D-mannoheptose 6-epimerase (AGME) catalyzes the interconversion of ADP-D-glycero-D-mannoheptose and ADP-L-glycero-D-mannoheptose. This epimerization reaction is the last of four enzymatic steps required in the biosynthesis of the precursor of the seven-carbon sugar L-glycero-D-mannoheptose (heptose). Heptose is an obligatory component of the LPS core domain. AGME mutants are not capable of LPS inner-core biosynthesis resulting in the loss of the intrinsic resistance of Gram-negative bacteria. Thus, AGME represents a suitable target for the design of a novel class of antimicrobial agents.

In this study, we determined the crystal structure of *Escherichia coli* K-12 AGME by multiwavelength anomalous diffraction (MAD) phasing of the selenomethionyl form of the enzyme. This is the first reported three-dimensional structure for any LPS inner-core biosynthetic enzyme. A model comprising ten monomers of the enzyme, each with tightly bound NADP and less tightly bound ADP-glucose, has been refined to 2.0 Å resolution. The model provides the first structural information for elucidating the mechanism of catalysis, inhibition and inactivation. The results of this study should be useful for the design of potent and selective mechanism-based inhibitors of AGME.

Materials and methods

Protein purification

AGME was overproduced in *E. coli* strain BL21 (DE3) pLysS (Novagen) transformed with the plasmid-borne *E. coli* K-12 *rfaD* gene (i.e., pCG6) [5]. Cells were grown in luria broth medium [5] aerobically at 37°C to $OD_{600} \approx 0.8$, at which time isopropyl-β-D-thiogalactoside (IPTG) was added to 0.4 mM. AGME was then purified to apparent homogeneity using purification protocols previously described [10,11].

Incorporation of selenomethionine and crystallization

To prepare selenomethionyl epimerase protein, plasmid pCG6 containing the epimerase gene [8] was transformed into the methionine auxotrophic *E. coli* strain BL41 (DE3) pLysS (kindly supplied by JR Horton). The transformed bacteria were grown in 10 l of LeMaster medium [27] supplemented with 0.05 g of selenomethionine (SeMet). Cells were harvested 3 h after induction with 0.4 mM IPTG and the SeMet protein was purified as described above, except that all buffers contained 3 mM β-mercaptoethanol and were purged continuously with helium. The level of SeMet substitution (~ 100%) in the purified AGME was estimated by amino acid composition analysis. Native AGME crystals were grown from 2.0 M ammonium sulfate and 2% PEG400 at pH 7.5 [11]. They formed in space group $P2_1$, with cell dimensions a = 99.5 Å, b = 109.8 Å, c = 181.5 Å, β = 91.0°. Crystals of the SeMet enzyme were grown under similar conditions to the native protein, although the best crystals were obtained at a slightly lower pH (7.2–7.3). The crystals were smaller, typically $0.15 \times 0.1 \times 0.05$ mm and more sensitive to manipulation and changes in temperature. These crystals belonged to the same space group, with cell dimensions a = 99.2 Å, b = 110.9 Å, c = 180.9 Å, β = 91.0°.

Data collection and reduction

MAD data were collected using an ADSC Quantum 4 CCD detector, at three wavelengths around the Se K-edge on station F2 at the Cornell High Energy Synchrotron Source (CHESS). Data were collected including an inverse beam sweep for the f″ maximum and the f′ minimum wavelengths, to guarantee the collection of all the anomalous pairs. A remote wavelength was collected below the Se K-edge, where the anomalous signal is absent. Cryoprotection of AGME crystals was achieved by gradually equilibrating them against solutions of mother liquor mixed with increasing concentrations of glycerol. At an optimal concentration of 25% glycerol the crystals were frozen at 100K. The crystals diffracted anisotropically. The diffraction extended to approximately 2.25 Å resolution along the b axis, whereas only weak diffraction was observed beyond 3.0 Å resolution along the c and a axes. Native crystals also diffracted anisotropically. Data were collected from a single crystal at 100K to 2.0 Å resolution on station F1 at CHESS. All data were processed with Mosflm [28] and scaled with Scala [29].

Structure determination

The positions of the selenium atoms were determined using the program SnB v2.0 [14]. The anomalous differences at the peak wavelength were converted into normalized diffE magnitudes with the DREAR program suite [15]. In all, 65 of the 70 independent selenium positions were located. The noncrystallographic symmetry fivefold axes (parallel to the crystallographic 2_1 axis) of the AGME pentamers were located from the selenium positions and were confirmed with locked self-rotation searches with the program GLRF [10]. Heavy-atom refinement and phase calculations were carried out at 3.0 Å resolution using MLphare [18], followed by solvent flattening with DM [19]. Clearly interpretable electron-density maps were obtained without using any form of NCS averaging. Subsequently a mask was generated from the experimental bones with MAMA [30] and NCS matrices were calculated from the selenium positions in each monomer. These matrices were then improved with the program IMP [31]. Full NCS averaging was performed with DM. The initial correlation between the monomers was about 0.45 in DM.

Table 1

Summary of X-ray diffraction data.

	Se MAD*		Monochromatic†	
Data type	Inflection‡	Peak‡	Remote‡	Native
Wavelength (Å)	0.9793	0.9791	0.99	0.919
Resolution (Å)	3.0	3.0	3.0	2.0
No. observations	391,088	391,325	196,899	573,756
No. unique reflections	72,470	72,504	71,662	242,238
Completeness (%)				
overall	92.2	92.3	91.4	93.3
last shell§	92.2	93.1	90.9	78.6
R_{sym} (%)#				
overall	4.8	4.8	4.3	6.6
last shell§	8.6	7.2	7.3	21.8
Average I/σ I				
overall	31.1	31.5	25.2	11.7
last shell§	13.4	15.5	12.2	3.0

*Data collected at the Cornell High Energy Synchrotron Source (CHESS) beamline F2 using an ADSC Q4 detector. †Data collected at CHESS beamline F1 using an ADSC Q4 detector. ‡The inflection point and peak wavelengths were collected in inverse beam mode, whereas the remote wavelength was collected at the low-energy side of the Se edge where there is little anomalous signal and as a result no inverse beam data were collected. §Last resolution shell for CHESS F2 MAD data, 3.16–3.00 Å; last resolution shell for CHESS F1 native data, 2.11–2.00 Å. #$R_{sym} = \Sigma\Sigma i|Ii-<I>|/\Sigma<I>$, where <I> is the mean intensity of the N reflections.

Table 2

Phasing statistics, refinement statistics and model quality.

Phasing	
FOM before solvent flattening	0.473
FOM after solvent flattening (no averaging)	0.804
FOM after solvent flattening (with averaging)	0.851
Refinement	
Resolution (Å)	25.0–2.0
Number of atoms (average B value (Å^2))	
protein (mainchain)	11,898 (29.1)
protein (sidechain)	11,716 (30.5)
water	995 (33.7)
ligand	783 (38.8)
R factor (%)	21.2
R_{free} (%)	26.2
Rms deviation from ideal geometry	
bonds (Å)	0.012
angles (°)	1.386

FOM, figure of merit

Model building and refinement

An initial model of a single subunit, amounting to 88% of the amino acid sequence, was built with the program O [31]. The known location of the selenium sites (including two consecutive amino acid residues) allowed the unequivocal assignment of the position and direction of the peptide chain. There were three sections of uninterpretable electron density corresponding to residues 193–207, 250–272 and 306–310. Clear density corresponding to the NADP and to the ADP-glucose was readily identified in the experimental maps. The known NCS operators were used to generate the two complete pentamers. Subsequent refinement was carried out using all data from 25.0–2.0 Å resolution and a bulk-solvent correction was applied with X-PLOR [32]. The application of overall anisotropic scaling parameters resulted in a significant drop in both the conventional R factor and the free R factor (by about 6% each) and proved crucial in improving the quality of the electron-density maps. The relative temperature factors between the b direction and the a and c directions was about 20 $Å^2$. The refinement consisted of several rounds of simulated annealing followed by manual rebuilding in O. Several residues were added to the starting model. Additionally a molecule of NADP was fitted to each monomer.

The quality of the electron density for the catalytic inhibitor ADP-glucose varied significantly between the individual subunits, in particular the glucose tended to be disordered. Strict NCS restraints of 500 kcal/mol were used in the initial stages of the refinement. As the model improved the restraints were relaxed gradually as judged by the free R factor. At no stage in the refinement did the free R factor indicate that NCS restraints could be completely removed. NCS restraints of 100 kcal/mol were still used for the bulk of the final model, although restraints were removed from a few regions of the polypeptide chain. Water molecules were added at positions with $> 3\sigma$ density in the $F_o–F_c$ maps according to reasonable geometry criteria. The final round of positional and B-factor refinement yielded an R factor of 21.2% and a free R factor of 26.2%. The final model contains ten monomers of AGME each with a tightly bound molecule of NADP and a more disordered molecule of ADP-glucose (in some molecules the glucose moiety was too disordered to be modeled into the electron density so only the ADP was included).

Accession numbers

The coordinates for the AGME model have been deposited with the Protein Data Bank with accession code 1EQ2.

Acknowledgements

SEE and AMD would like to acknowledge the generous support of grant GM-46733 from the National Institutes of Health. Research at CHESS is supported by grant DMR-9311772 from the National Science Foundation and the MacCHESS research resource is funded by grant RR-01646 from the National Institutes of Health.

References

1. Adams, G.A., Quadling, C. & Perry, M.B. (1967). D-Glycero-D-mannoheptose as a component of lipopolysaccharides from Gram-negative bacteria. *Can. J. Microbiol.* 13,1605-1613.
2. Coleman, W.G., Jr. & Leive, L. (1979). Two mutations which affect the barrier function of the *Escherichia coli* K-12 outer membrane. *J. Bacteriol.* 139, 899-910.
3. Vaara, M. (1993). Outer membrane permeability barrier to azithromycin, clarithromycin, and roxithromycin in Gram-negative enteric bacteria. *Antimicrob. Agents Chemother.* 37, 354-356.
4. Taylor, P.W. (1983). Bactericidal and bacteriolytic activity of serum against Gram-negative bacteria. *Microbiol. Rev.* 47, 46-83.
5. Eidel, L. & Osborn, M.J. (1974). Phosphoheptose isomerase, first enzyme in biosynthesis of aldoheptose in *Salmonella typhymurium. J. Biol. Chem.* 249, 5642-5648.
6. Raetz, C.R.H. (1990). Biochemistry of endotoxins. *Annu. Rev. Biochem.* 59, 129-170.
7. Coleman, W.G., Jr. (1983). The *rfaD* gene codes for ADP-L-glycero-D-mannoheptose 6-epimerase, an enzyme required for lipopolysaccharide core biosynthesis. *J. Biol. Chem.* 259, 1985-1990.
8. Pegues, J.C., Chen, L., Gordon, A., Ding, L. & Coleman, W.G., Jr. (1990). Cloning, expression, and characterization of the *Escherichia coli* K-12 *rfaD* gene. *J. Bacteriol.* 172, 4652-4660.
9. Wierenga, R.K., Terpstra, P. & Hol, W.G.J. (1986). Prediction of the occurrence of the ADP-binding βαβ-fold in proteins using an amino acid sequence fingerprint. *J.Mol. Biol.* 187, 101-107.
10. Ding, L., Seto, B.L., Ahmed, S.A. & Coleman, W.G., Jr. (1994). Purification and properties of the *Escherichia coli* K-12 NAD-dependent nucleotide diphosphosugar epimerase, ADP-L-glycero-D-mannoheptose 6-epimerase. *J. Biol. Chem.* 269, 24384-24390.
11. Ding, L., *et al.,* & Coleman, W.G., Jr. (1999). Crystallization and preliminary X-ray diffraction studies of the lipopolysaccharide core biosynthetic enzyme ADP-L-glycero-D-mannoheptose 6-epimerase from *E. coli* K-12. *Acta Crystallogr. D* 55, 685-688.
12. Tong, L. & Rossmann, M.G. (1990). The locked rotation function. *Acta Crystallogr. A* 46, 783-792.
13. Matthews, B.W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* 33, 491-497.
14. Weeks, C.M. & Miller, R. (1999). The design and implementation of SnB version 2.0. *J. Appl. Crystallogr.* 32, 120-124.
15. Blessing R.H., Guo D.Y. & Langs D.A. (1996). Statistical expectation value of the Debye–Waller factor and E(hkl) values for macromolecular crystals. *Acta Crystallogr. D* 52, 257-266.
16. Blessing, R.H. & Smith, G.D. (1999). Difference structure factor normalization for heavy atom or anomalous scattering substructure determinations. *J. Appl. Crystallogr.* 32, 664-670.
17. Hauptman, H.A. (1991). A minimal principle in the phase problem. In *Crystallographic Computing 5: From Chemistry to Biology*. (Moras, D., Podjarny, A.D. & Thierry, J.C. eds), pp. 324-332, IUCr and Oxford University Press, Oxford, UK.
18. Otwinowski, Z. (1991). Maximum-likelihood refinement of heavy atom parameters. In *Proceedings of the CCP4 Study Weekend: Isomorphous Replacement and Anomalous Scattering.* (Wolf, W., Evans, P.R. & Leslie, A.G.W., eds), pp. 80-88, SERC Proceedings, Daresbury Laboratories, Warrington, UK.
19. Cowtan, K.D. & Main, P. (1996). Phase combination and cross validation in iterated density-modification calculations. *Acta Crystallogr. D* 52, 43-48.
20. Deacon, A.M. & Ealick, S.E. (1999). Se-based MAD phasing: setting the sites on larger structures. *Structure* 7, R161-R166.
21. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283-291.
22. Rossmann, M.G., Moras, D. & Olsen, K.W. (1974). Chemical and biological evolution of a nucleotide-binding domain. *Nature* 250, 194-199.
23. Thoden, J.B., Frey, P.A. & Holden, H.M. (1996). Crystal structures of the oxidized and reduced forms of UDP-galactose 4-epimerase isolated from *Escherichia coli. Biochemistry* 35, 2557-2566.
24. Rizzi, M., *et al.,* & Bolognesi, M. (1998). GDP-4-keto-6-deoxy-D-mannose epimerase/reductase from *Escherichia coli*, a key enzyme in the biosynthesis of GDP-L-fucose, displays the structural characteristics of the RED protein homology superfamily. *Structure* 6, 1453-1465.
25. Somers, W.S., Stahl, M.L. & Sullivan, F.X. (1998). GDP-fucose synthetase from *Escherichia coli*: structure of a unique member of the short-chain dehydrogenase/reductase family that catalyzes two distinct reactions at the same active site. *Structure* 6, 1601-1612.
26. Thoden, J.B. & Holden, H.M. (1998). Dramatic differences in the binding of UDP-galactose and UDP-glucose to UDP-galactose 4-epimerase from *Escherichia coli*. *Biochemistry* 37, 11469-11477.
27. Hendrickson, W.A., Horton, J.R. & LeMaster, D.M. (1990). Selenomethionyl protein produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* 9, 1665-1672.
28. Leslie, A.G.W., Brick, P. & Wonacott, A.T. (1986). An improved package for the measurement of oscillation photographs. *Daresbury Laboratory Information Quart. Protein Crystallogr.* 18, 33-39.
29. Evans, P.R. (1993). Data reduction. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing.* pp. 114-122.
30. Kleywegt, G.J. & Jones, T.A. (1994). Halloween ... mask and bones. In *From First Map to Final Model*. (Bailey, S., Hubbard, R. & Waller, D., eds), pp. 59-66, SERC Daresbury Laboratory, Warrington, UK.
31. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* 47, 110-119.
32. Brünger, A.T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science* 235, 458-460.
33. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24, 946-950.
34. Nicholls, A., Sharp, K. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11, 281-296.