The nucleotide-binding site of *Aquifex aeolicus* LpxC

The structure of recombinant *Aquifex aeolicus* UDP-3-O-acyl-N-acetylgulcosamine deacetylase (LpxC) in complex with UDP has been determined to a resolution of 2.2 Å. Previous studies have characterized the binding sites of the fatty-acid and sugar moieties of the substrate, UDP-(3-O-hydroxymyristoyl)-N-acetylgulcosamine, but not that of the nucleotide. The uracil-binding site is constructed from amino acids that are highly conserved across species.

Hydrophobic associations with the Phe155 and Arg250 side chains in combination with hydrogen-bonding interactions with the main chain of Glu154 and the side chains of Tyr151 and Lys227 position the base. The phosphate and ribose groups are directed away from the active site and interact with Arg137, Lys156, Glu186 and Arg250. The orientation of the phosphate-ribose tail is not conducive to catalysis, perhaps owing to the position of an inhibitory Zn$^{2+}$. However, based on the position of uracil revealed in this study and on the previously reported complex of LpxC with an inhibitor, a model is proposed for substrate binding.

1. Introduction

A unique biochemical pathway in Gram-negative bacteria leads to the production of lipopolysaccharide (LPS), a major component of the outer membrane of these organisms. LPS contributes to bacterial resistance to antibiotics (Nikaido, 2003) and hyperstimulation of the host immune system (Cohen, 2002; Lopez-Bojorquez et al., 2004). It has three structural components: core polysaccharide, O-antigen and lipid A, a fatty-acid (FA) substituted phosphorylated glucosamine disaccharide. Lipid A is required for both bacterial viability (Raetz & Whitfield, 2002) and modulation of the host immune system (Galanos et al., 1985).

Several of the enzymes involved in the synthesis of lipid A are essential for the growth and viability of Gram-negative bacteria (Raetz & Whitfield, 2002) and represent targets for chemotherapy. One such enzyme, UDP-3-O-acyl-N-acetylgulcosamine deacetylase (LpxC), catalyzes the first committed step in the synthesis of lipid A (Raetz & Whitfield, 2002). Inhibition of LpxC *in vivo* has led to successful treatment of infection in mice (Kline et al., 2002; Onishi et al., 1996). In both *Escherichia coli* and *Aquifex aeolicus*, LpxC deacetylates UDP-(3-O-hydroxymyristoyl)-N-acetylgulcosamine (Jackman et al., 2001; Young et al., 1995). LpxC is a metalloenzyme that uses Zn$^{2+}$ to activate a nucleophilic attack by water on the amide of the N-acetylgulcosamine (GlcNAc) moiety of the substrate (Fig. 1; Jackman et al., 1999); like other bacterial metalloamidases, a second inhibitory zinc-binding site is present (Jackman et al., 1999; Whittington et al., 2003). LpxC presents an αβ fold with no known mammalian homologues (Whittington et al., 2003). This uniqueness, its essentiality and the previous success in enzyme inhibition make LpxC an attractive target for structure-based drug design.

The elements required by LpxC to bind and catalyze substrate include a zinc-binding motif, UDP-binding and GlcNAc-binding sites and a hydrophobic cavity to bind FA. Structural and biochemical studies have identified HKX(L,F)D as the zinc-binding motif in LpxC (Jackman et al., 2001; Whittington et al., 2003) and several inhibitors containing hydroxamate groups exploit the presence of the divalent cation. Crystallographic studies of LpxC have elucidated a hydro-
phobic FA-binding cavity that stabilizes the structure upon binding FA (Hernick et al., 2005; Whittington et al., 2003). The presence of a longer FA moiety on the GlcNAc results in a higher binding affinity to LpxC and is required for efficient catalysis (Jackman et al., 1999, 2000; Whittington et al., 2003). The inhibitor 1,5-anhydro-2-C-(carboxymethyl-N-hydroxymide)-2-deoxy-3-O-myristoyl-α-glucitol (TU-514; Fig. 1) is a substrate-analogue inhibitor that has been used to characterize the GlcNAc-binding site of LpxC (Coggins et al., 2003, 2005; Gennadios et al., 2006). TU-514 exploits the zinc-coordination and FA-binding and GlcNAc-binding sites of LpxC and the result is an inhibitor with nanomolar affinity for the enzyme (Jackman et al., 2000). Inhibitor design would benefit from including UDP-recognition elements, but presently this binding site remains uncharacterized. Here, we present a crystallographic analysis of a complex of A. aeolicus LpxC with UDP that allows us to describe the molecular recognition of UDP in the nucleotide-binding site of the enzyme.

2. Materials and methods

2.1. Sample preparation

A construct identical to that previously described by Whittington et al. (2003) was used. This produces a truncated version of the enzyme which also contains a cysteine-to-alanine mutation that has been found to reduce aggregation. Briefly, lpxC was amplified from genomic DNA of A. aeolicus (strain VF5) using the primers CATATGGGATTAGAAAGACGGTTAAAGAG and GGATCTTATCAACGGTAGTTTGTTCCTTG, which contained an N-terminal NdeI and a C-terminal BamHI restriction site, respectively, as indicated in bold. The PCR product, which encodes residues 1–271 of the protein, was ligated into pCR-Blunt II-TOPO using the Zero Blunt TOPO Cloning Kit (Invitrogen). The construct, pET21a-lpxC 1–271 of the protein, was verified by DNA sequencing. The experimentally determined weight of the protein was 30 811.6 Da, which agrees with the predicted weight of the protein in the absence of the initial methionine (30 810.3 Da).

The plasmid containing modified lpxC was transformed into Rosetta(DE3) pLysS (Novagen) for expression. A single colony was used to inoculate 10 ml of enriched media (Nair et al., 1991) containing carbenicillin (50 mg l⁻¹) and chloramphenicol (12 mg l⁻¹) and was grown overnight at 310 K. The overnight culture was diluted (1:100) into 1 l fresh medium and grown at 310 K until the A₆₀₀ reached 0.6. Expression was induced at 298 K with 1 mM isopropyl β-D-thiogalactopyranoside and the medium was supplemented with 0.1 mM ZnSO₄ at this point. The cells were harvested approximately 16 h later by centrifugation (Beckman JS-4.2, 3480g at 277 K for 60 min) and stored at 253 K.

The cell pellet was resuspended in 20 ml 20 mM NaCl, 20 mM Tris pH 8 and 50 mM EDTA and lysed using a constant cell disrupter (One Shot model, Constant Systems Ltd). The soluble fraction was separated by centrifugation (Beckman JA-25.50, 48 400g at 277 K for 30 min), passed through a 0.2 µm filter and loaded onto a 5 ml HisTrap Q FF column (GE Healthcare). A linear gradient from 20 to 500 mM NaCl was used to elute the protein. The fractions containing LpxC were pooled, treated with 20 mM EDTA pH 8 and dialyzed into 50 mM NaCl, 20 mM Tris pH 7.7 overnight. The sample was applied onto a 5 ml HisTrap HP column (GE Healthcare) charged with ZnSO₄ and the protein was eluted using a combination of step and linear gradients from 15 to 500 mM imidazole. Fractions containing LpxC were pooled and further purified by gel-filtration chromatography (Superdex 75, GE Healthcare) in 150 mM NaCl, 25 mM HEPES pH 7.5 and 0.5 mM ZnSO₄. 10 mM magnesium acetate was added to the sample, which was subsequently concentrated to 184 µM as determined using a theoretical molar extinction coefficient of 22 900 M⁻¹ cm⁻¹. Purity was ascertained by SDS–PAGE and matrix-assisted laser desorption ionization–time-of-flight mass spectrometry. The experimentally determined weight of the protein was 30 811.6 Da, which agrees with the predicted weight of the protein.

2.2. Crystallization and data collection

Crystals of LpxC were grown by hanging-drop vapour diffusion at 293 K using 1 µl protein solution (184 µM) and 1 µl of a reservoir

![Figure 1](image_url)

(a) The reaction catalyzed by LpxC. (b) The inhibitor TU-514.
solution containing 8% PEG 6000, 100 mM Tris pH 8.5 and 300 mM NaCl. Crystals were harvested and subsequently soaked for 24 h in a pre-equilibrated vapour-diffusion drop initially composed of 0.5 µl 500 mM UDP-N-acetylglucosamine (UDP-GlcNAc) in gel-filtration buffer and 0.5 µl reservoir solution containing 17% PEG 6000, 100 mM Tris pH 8.5 and 150 mM NaCl. The crystals were flash-cooled to ~100 K (Oxford Cryosystems) in a solution consisting of 17% PEG 6000, 100 mM Tris pH 8.5, 150 mM NaCl and 20% glycerol. Data were collected at the European Synchrotron Radiation Facility on beamline ID23-2 using a MAR CCD 225 detector. A total of 80° of data were collected in 1° oscillations at a wavelength of 0.873 Å using an exposure time of 2 s. The data were integrated using MOSFLM (Leslie, 1992) as implemented in CCP4 (Collaborative Computational Project, Number 4, 1994) and then scaled and merged using SCALA (Evans, 1997). Significant radiation damage was evident on this undulator beamline and only the first 60 images were used. Statistics are given in Table 1.

Statistics for the refined structure are presented in Table 1. An overlay of Cα positions for residues 2–266 of our model with the starting model gave an r.m.s.d. of 0.19 Å for chain A and 0.24 Å for chain B.

### Table 1

| Crystallographic statistics. |
|-----------------------------|
| Values in parentheses represent the highest resolution bin of approximate width 0.12 Å. |

| Space group | $P6_1$ |
|-------------|--------|
| Unit-cell parameters (Å) | $a = 101.76$, $c = 124.24$ |
| Resolution range (Å) | 50.9–2.2 |
| No. of observed/unique reflections | $142119/36577$ |
| Wilson B ($Å^2$) | 23.7 |
| Mosaic spread (%) | 0.27 |
| Completeness (%) | 99.0 (98.5) |
| Multiplicity | 3.9 (3.9) |
| $R_{	ext{free}}$ (%) | 10.8 (47.5) |
| $(I/σ(I))$ | 11.3 (3.6) |
| $R$ factor (%) | 16.8 |
| $R_{	ext{free}}$ (%) | 22.5 |
| R.m.s.d. from ideal values, bond lengths (Å) | 0.01 |
| R.m.s.d. from ideal values, bond angles (°) | 1.205 |
| Average $B$ factors ($Å^2$) | 19.9 |
| Main chain | 21.9 |
| Side chain | 34.9 |
| Waters | 69.4 |
| UDP | Other |
| | 28.3 |
| Ramachandran plot analysis (%) | 88.5 |
| Residues in most favoured regions | 11.5 |
| Residues in allowed regions | 0 |
| Residues in disallowed regions | 0.18 |
| Cruickshank DPI† (Å) based on $R_{	ext{free}}$ | 2.65 |
| PDB code | |

† Diffraction-component precision index (Cruickshank, 1999).

2.4. Modelling the substrate

The crystal structure of LpxC with bound inhibitor TU-514 (PDB code 2go4) was overlaid on the structure of LpxC with bound UDP. The substrate was placed at the active site and manipulated using Coot and O (Jones et al., 1991) such that the uracil and pyranose moieties of the substrate overlaid with those of UDP and TU-514, respectively. The geometry of the model was idealized in REFMACS using molecular restraints obtained from PRODRG (Schütte-Kopf & van Aalten, 2004).

3. Results and discussion

3.1. Overall structure

A detailed description of the LpxC structure has been published (Whittington et al., 2003). In brief, LpxC displays an α+β fold containing two domains, each consisting of a five-stranded β-sheets and two α-helices. These domains are sandwiched together with both sets of helices between the sheets (Fig. 2). The active site is located between the domains at the β-sheet interface. The enzyme has been crystallized in the zinc-inhibited form, with two Zn2+ ions present in the active site. A third Zn2+ is located between α-helix C, which flanks the active site, and the loop joining the small βββ domain to the larger β-sheet and is coordinated by residues His58 and His188, a chloride and a water molecule (not shown). A fourth Zn2+ is coordinated by Gly2 and Glu120 of chain A and by Glu90 and His29 of a symmetry-related chain B (not shown). This is a crystallization artefact rather than an indication of oligomerization, as gel-filtration experiments clearly show the protein to be a monomer (data not shown). Electron density was observed in the FA-binding cavity of both chains and was modelled as myristic acid. In chain A the FA directly binds the inhibitory Zn2+, while in chain B the carboxylic acid head group extends away from the active site and is disordered over two

![Figure 2](image-url)

Figure 2

Ribbon diagram of the LpxC–UDP complex. Domains I and II are shown in green and wheat, respectively. Secondary structure is labelled according to the previous assignments (Whittington et al., 2003). UDP is depicted as spheres and myristic acid as sticks. Activity-related Zn2+ ions are depicted as grey spheres, with the inhibitory ion labelled I. Ligand atoms are coloured as follows: C, black; O, red; N, blue; P, yellow. Figs. 2, 3, and 4 were prepared with PyMOL (DeLano, 2002).
positions (not shown). The FA was not added to the crystallization solution and is likely to have been acquired from the expression system. Similar FA binding has been observed in other LpxC structures (Gennadios et al., 2006; Whittington et al., 2003).

3.2. Nucleotide binding

Our current study, the crystal structure of the zinc-inhibited form of LpxC in complex with UDP, provides structural details concerning nucleotide binding for the first time. The nucleotide-binding site is located between strand β3 and helix A' of domain II (Fig. 2). The uracil moiety is well defined in the electron density, with the ribose and phosphates less so (Fig. 3). This clearly correlates with the average B factor observed for each moiety: 58.9, 65.6 and 76.0 Å² for uracil, ribose and diphosphate, respectively. We were unable to model the GlcNAc portion of the UDP-GlcNAc. This segment of the added ligand may be disordered or missing owing to hydrolysis. The presence of the inhibitory Zn²⁺ at the entrance to the active-site cavity (Fig. 2) may preclude GlcNAc binding, leading to nonspecific association and disorder. Interestingly, the use of identical soaking conditions with UDP rather than UDP-GlcNAc resulted in crystal dissolution.

The uracil-binding pocket is mainly formed by residues contributed from β4, the β8-α-β turn and one face of α-A’ (Figs. 2, 3 and 4). The uracil is 10.5 Å from the catalytic Zn²⁺ (measured from O2) and is placed in a hydrophobic patch where it forms a C-H bond with Phe155. This is a highly conserved amino acid; in 241 of 248 assigned LpxC sequences found in a range of databases this position is an aromatic residue (Phe/Tyr/Trp). In the remaining seven sequences, Phe155 is part of a hydrophobic patch that, along with Phe180 and Phe155. This is a highly conserved amino acid; in 241 of 248 assigned LpxC sequences found in a range of databases this position is an aromatic residue (Phe/Tyr/Trp). In the remaining seven sequences, derived from Burkholderia species, a leucine occupies this position. Phe155 is part of a hydrophobic patch that, along with Phe180 and Phe182, is recognized as an affinity determinant for binding to LpxC (Coggins et al., 2005). Both uracil N3 and O2 are within hydrogen-bonding distance of the side-chain hydroxyl of Tyr151, which forms part of the α/β interface of the protein. In alignments of LpxC across 248 species, approximately 50% of all LpxC sequences have phenylalanine, 25% have isoleucine, leucine or valine and 15% have tyrosine at this position. This observation emphasizes the importance of a hydrophobic side chain in defining this part of the uracil-binding pocket. It is possible that in the absence of a tyrosine hydroxyl group a solvent could occupy this site and mediate interactions between the base and the enzyme.

Additional hydrogen-bonding interactions hold the nucleotide in the binding site. O4 accepts a hydrogen bond from the main-chain amide of Glu154; this side chain, which is exposed on the surface of the protein, does not appear to be important in organizing the binding site despite a high level of conservation as either aspartate or glutamate. The uracil O2 accepts a hydrogen bond donated from Lys227 NZ. This residue is conserved in LpxC as either lysine or arginine, where a similar hydrogen-bonding interaction is plausible. Mutagenesis studies (McClerren et al., 2005) show Lys227 to be critical for enzyme activity since a K227A mutant retains only 0.05% activity of the wild-type LpxC. In the structure with the inhibitor TU-514, a water-mediated hydrogen bond links the C-6 hydroxyl of the inhibitor to the side chain of this lysine (Gennadios et al., 2006).

The ribose and phosphates are directed away from the active site in a position that is inconsistent with the substrate orientation required for catalysis. LpxC from A. aeolicus does not deacetylate UDP-GlcNAc (Whittington et al., 2003) and poor penetration of this molecule into the active site of the zinc-inhibited form is likely to be expected. The ribose interacts with the side chains of Arg137 and Glu185. This is a well conserved residue pairing in that a basic lysine is often observed at the corresponding arginine position and the glutamate is retained or replaced by aspartate or glutamine. The α-phosphate interacts with the basic side chains of Arg250 and Lys156 (Fig. 3). The arginine is highly conserved, but not the lysine. The current positions of the ribose and phosphates and interactions with the enzyme may represent a configuration adopted by substrate as it binds in advance of catalysis or alternatively by product as it vacates the catalytic centre.

**Figure 3**

UDP-binding site and the omit difference electron-density map. The map has coefficients [Fo − Fs], αc and is contoured at 3.5σ. Fo represents the observed structure factors, Fs the calculated structure factors and αc the calculated phases from which ligand contributions were omitted. Protein residues within 4 Å of UDP are labelled and putative hydrogen bonds (within 3.6 Å) are shown as black dashed lines. Atoms are coloured as described in Fig. 2, with the Cα atoms of the protein coloured salmon.
3.3. Implications for substrate binding

We propose that the nucleotide-binding site observed in this study represents the position that uracil adopts during substrate recognition and subsequently catalysis. Binding of the base involves highly conserved residues, which when altered influence both binding and enzyme activity (Coggins et al., 2005; McClearen et al., 2005). There is a notable absence of conserved hydrogen bonds between the hexose of the substrate analogue TU-514 and LpxC (Gennadios et al., 2006). Isooxazoline-based inhibitors carrying carboxylate or phosphonate of the substrate analogue TU-514 and LpxC (Gennadios et al., 2006). Isooxazoline-based inhibitors carrying carboxylate or phosphonate head groups bind more tightly to E. coli LpxC if they contain an aliphatic substituent (IC_{50} values of 4 μM with and >450 μM without; Pirrung et al., 2002), suggesting that hexose–protein interactions do not contribute significantly to substrate binding. Therefore, interactions between LpxC and the UDP and FA moieties of the substrate are likely to contribute the most to ordering of the substrate during catalysis. Using our experimentally derived uracil-binding site and a GlcNAc-binding site based on the position of the pyranose in the crystallographic model of LpxC with TU-514, a substrate model can be constructed by varying the torsion angles between the uracil and the ribose, in particular the glycosidic link between the base and ribose, and the ribose and the phosphates while maintaining binding of the FA within the hydrophobic cavity (Fig. 4). Note that the inhibitory Zn^{2+} was removed from the structural template prior to modelling because it prevents access to the catalytic Zn^{2+}. The r.m.s.d. of atoms from uracil common to those of the substrate is 0.03 Å and the r.m.s.d. of the inhibitor pyranose atoms to those of the substrate is 0.08 Å. The model places the functional groups in chemically plausible positions for catalysis, supporting the conclusion that the uracil placement we observe is biologically relevant. Our structure now completes the picture of how LpxC recognizes substrate and imparts information to support a structure-based approach to inhibitor development.

We thank R. Huber for a gift of genomic DNA and D. Flot for assistance at the European Synchrotron Radiation Facility. This work was funded by the Wellcome Trust and BBSRC (Structural Proteomics of Rational Targets).

References

Coggins, B. E., Li, X., McClearen, A. L., Hinds Gaul, O., Raetz, C. R. & Zhou, P. (2003). Nature Struct. Biol. 10, 645–651.
Coggins, B. E., McClearen, A. L., Jiang, L., Li, X., Rudolph, J., Hinds Gaul, O., Raetz, C. R. & Zhou, P. (2005). Biochemistry, 44, 1114–1126.