John Badger,^{a*} Barbara
 Chie-Leon,^a Cheyenne Logan,^a
 Vandana Sridhar,^a Banumathi
 Sankaran,^b Peter H. Zwart^b and
 Vicki Nienaber^a

^aZenobia Therapeutics Inc., 505 Coast
 Boulevard South, Suite 111, La Jolla, CA 92037,
 USA, and ^bBerkeley Center for Structural
 Biology, Physical Biosciences Division,
 Lawrence Berkeley Laboratory, 1 Cyclotron
 Road, Berkeley, CA 94720, USA

Correspondence e-mail:
 john@zenobiatheapeutics.com

Received 26 November 2012
 Accepted 28 November 2012

PDB References: LpxD, 4e75; 4e79

Structure determination of LpxD from the lipopolysaccharide-synthesis pathway of *Acinetobacter baumannii*

Acinetobacter baumannii is a Gram-negative bacterium that is resistant to many currently available antibiotics. The protein LpxD is a component of the biosynthetic pathway for lipopolysaccharides in the outer membrane of this bacterium and is a potential target for new antibacterial agents. This paper describes the structure determination of apo forms of LpxD in space groups $P2_1$ and $P4_322$. These crystals contained six and three copies of the protein molecule in the asymmetric unit and diffracted to 2.8 and 2.7 Å resolution, respectively. A comparison of the multiple protein copies in the asymmetric units of these crystals reveals a common protein conformation and a conformation in which the relative orientation between the two major domains in the protein is altered.

1. Introduction

Acinetobacter baumannii is a Gram-negative pathogenic bacterium that is inherently resistant to most currently available antibiotics and which poses a significant threat to immune-compromised hospital patients (Dent *et al.*, 2010; Bassetti *et al.*, 2008). The enzymes involved in the biosynthetic pathway of the lipopolysaccharides that are included in the bacterial outer membranes are potentially attractive drug targets for the treatment of *A. baumannii* infections. The major components of these lipopolysaccharides are lipid A (a glucosamide-based saccharolipid) and KDO (3-deoxy-D-manno-oct-2-ulosonic acid). The synthesis of KDO₂-lipid A occurs in the cytoplasm, where the first three enzymes in the pathway, LpxA, LpxC and LpxD, are all soluble proteins (Raetz & Whitfield, 2002). Although LpxC is a validated target for antibacterial drug discovery, LpxA and LpxD do not appear to have been subjected to as much attention. The Binding Database (<http://www.bindingdb.org/bind/index.jsp>) contains no entries for compounds in complex with LpxD and the only significant ligands in crystal structures of LpxD in the Protein Data Bank (PDB) are uridine diphosphate *N*-acetylglucosamine, 2-(*N*-morpholino)-ethanesulfonic acid and palmitic acid (Table 1). LpxD inhibitors are potentially attractive as antibacterial agents because there are no closely similar proteins in humans. For example, in a comparison of the *A. baumannii* LpxD sequence with human protein sequences in the UniProtKB database the most similar protein (DCTN5) is less than half the length of LpxD and shows sequence identity at only 29% of amino-acid positions.

Crystal structures of LpxD have been determined for the proteins from *Pseudomonas aeruginosa* (Badger *et al.*, 2011), *Escherichia coli* (Bartling & Raetz, 2009) and *Chlamydia trachomatis* (Buetow *et al.*, 2007). With the exception of the structure from *P. aeruginosa*, which appears to be stabilized by an N-terminal tag that extends from one protein molecule into the enzymatic site of a neighboring protein molecule, these structures were all solved at moderate resolution. The LpxD sequences from this set of bacteria are 34–42% identical to the *A. baumannii* sequence, resulting in a general topological similarity but still allowing a potentially significant differentiation in the ligand-binding properties of different species.

The biological unit of LpxD is a trimer and all of the known crystal structures contain either one copy of the protein in the asymmetric unit, where the trimer is generated by crystallographic symmetry operations, or contain complete trimeric units in the asymmetric unit. The structure of LpxD from *C. trachomatis* containing uridine

Table 1

LpxD crystal structures in the Protein Data Bank.

Structure similarity is based on *BLAST* alignment using the RCSB PDB query interface over the given number of amino acids (Align). Only ligands larger than ten non-H atoms are listed. UD1, uridine diphosphate *N*-acetylglucosamine; MES, 2-(*N*-morpholino)-ethanesulfonic acid; PLM, palmitic acid.

Type	Species	Identity (%)	Similarity (%)	Align	PDB code	Resolution (Å)	Ligands
LpxD	<i>P. aeruginosa</i>	42	61	345	3pm0	1.30	—
LpxD	<i>E. coli</i>	39	58	352	3eh0	2.60	—
LpxD	<i>C. trachomatis</i>	34	53	357	2iua	2.70	MES, PLM
					2iu9	3.10	UD1, PLM
					2iu8	2.20	UD1, PLM

diphosphate *N*-acetylglucosamine (Buetow *et al.*, 2007; PDB entry 2iu9) illustrates the general architecture and the location of the enzymatic site in this protein. LpxD structures consist of a globular N-terminal domain that is loosely connected to a tightly wound β -strand domain. These domains are followed by a C-terminal helical extension that bundles together helices from the three molecules in the trimer. The enzymatic binding sites are sandwiched between the β -sheet domain from one molecule (including an associated loop) and the N-terminal domain from a neighboring molecule.

2. Methods

2.1. Cloning, expression and purification

Wild-type LpxD from *A. baumannii* (UniProt ID B0VMV2) was cloned into a proprietary vector containing an N-terminal 6 $\times$ His tag that is cleavable by TEV protease using the primers 5'-TAT ATA GGA TCC AAA GTG CAA CAA TAT CGT-3' and 5'-TAT ATA CTC GAG TCA TTT ACG CAA ATT AAA AG-3'. The restriction-enzyme sites in this vector were *KpnI* and *XhoI*. The sequence of the product was verified. The product was expressed in *E. coli* BL21 (DE3) cells and these cells were grown at 310 K to an OD₆₀₀ of ~0.6. The cells were induced with 1 mM IPTG at 298 K overnight, harvested and stored at 193 K until use.

The purification of the target protein was performed using a three-column system. The cell biomass was lysed by sonication in 50 mM Tris–sodium chloride pH 7.8, 500 mM sodium chloride, 10% glycerol, 20 mM imidazole, 5 mM β -mercaptoethanol (buffer A) plus one Roche Complete Protease Inhibitor Tablet and 20 000 units of Benzonase. The target protein was extracted by binding to Ni²⁺-charged IMAC resin and was eluted with 250 mM imidazole. The peak fractions were pooled and cleaved with 3 mg TEV overnight in buffer A at 277 K. The cleaved protein was then run over Ni²⁺-charged IMAC resin and the flowthrough was collected. Highly aggregated protein was eliminated by size-exclusion chromatography (S-200) in 10 mM Tris–sodium chloride pH 8, 500 mM sodium chloride, 1 mM DTT. At this size limit, the protein that is assembled into trimers in solution is included in the fraction collected for concentration. The protein in the collected fraction was concentrated to 25 mg ml⁻¹. Compared with the native protein sequence, the sequence of the construct used in crystallization contains three additional amino acids at the N-terminus.

2.2. Crystallization and data collection

One crystal form of the cleaved *A. baumannii* LpxD (LpxD-1) was grown by the hanging-drop vapor-diffusion method by mixing 2 μ l protein at 12 mg ml⁻¹ in 10 mM Tris pH 8, 500 mM sodium chloride, 1 mM DTT buffer with 2 μ l 24% polyethylene glycol (PEG) 3350 and

Table 2Crystal properties and refinement statistics for the crystal structures of LpxD from *A. baumannii*.

Values in parentheses are for the highest resolution shell. The percentage of amino acids in the allowed region of the Ramachandran plot is defined according to Kleywegt & Jones (1996) and was calculated by the *PDB Validation Server* (<http://validate.rcsb.org/>).

Crystal	LpxD-1	LpxD-2
PDB code	4e75	4e79
Space group	<i>P</i> ₂ ₁	<i>P</i> ₄ ₃ ₂ ₂
Unit-cell parameters		
<i>a</i> (Å)	106.7	64.7
<i>b</i> (Å)	209.6	64.7
<i>c</i> (Å)	107.6	463.8
β (°)	119.3	—
Beamline	21ID-D, APS	21ID-D, APS
Resolution (Å)	2.84	2.66
Wavelength (Å)	0.9794	0.9794
<i>R</i> _{merge}	0.113 (0.585)	0.120 (0.572)
Completeness (%)	99.5 (99.5)	99.2 (97.1)
Observed reflections	268582 (25959)	270436 (24582)
Unique reflections	94647 (9423)	29759 (2889)
$\langle I/\sigma(I) \rangle$	8.7 (1.8)	16.9 (4.2)
Multiplicity	2.8 (2.8)	9.1 (9.5)
Resolution range	48.15–2.85	49.60–2.66
Total reflections for refinement (work + free)	89870	28213
No. of reflections used for <i>R</i> _{free}	4744	1508
<i>R</i> _{work}	0.234	0.234
<i>R</i> _{free}	0.280	0.300
Total protein atoms	15088	7709
Total ligand atoms	0	0
Total waters	0	13
Average <i>B</i> factor (Å ²)	52.2	43.0
R.m.s. deviation from ideal		
Bond lengths (Å)	0.012	0.011
Bond angles (°)	1.444	1.352
Ramachandran plot (% of residues in allowed regions)	97.1	97.1

0.2 M ammonium formate at 293 K. These crystals appeared after 2 d. A second crystal form (LpxD-2) with a distinct morphology was obtained by mixing the same protein solution with 19% polyethylene glycol (PEG) 3350 and 0.4 M magnesium acetate. Both types of crystal were harvested and transferred into a cryoprotectant solution made up of 20%(v/v) ethylene glycol in crystallization buffer and were flash-cooled in liquid nitrogen for data collection.

Data were collected from both types of crystal using a Rayonix MX-300 detector on beamline 21ID-D at the Advanced Photon Source (APS). For the LpxD-1 crystal a total of 180 diffraction images were collected using an oscillation range of 0.75° per image. For the LpxD-2 crystal a total of 240 diffraction images were collected with an oscillation range of 0.5° per image.

2.3. Structure solution and refinement

Data sets from both LpxD-1 and LpxD-2 crystals were processed and reduced to merged intensities using the *HKL-2000* software (Otwinowski & Minor, 1997; Table 2). These two data sets were obtained from the same synchrotron shift; therefore, parallel molecular-replacement trials were performed on the two crystal forms with the *Phaser* program (McCoy *et al.*, 2007) using the protein monomer from *P. aeruginosa* (Badger *et al.*, 2011; PDB entry 3pm0) as the search model. Despite the relatively high degree of noncrystallographic symmetry, the calculations using the LpxD-2 data set gave the highest scoring solution, with six molecules in the asymmetric unit assembled as two trimers. The structure corresponding to data set LpxD-1 was subsequently solved using a refined monomer from LpxD-2 as the search model, and the solution contained three molecules in the asymmetric unit.

The refinement of both of these structures was performed using the *REFMAC5* program (Murshudov *et al.*, 2011) via the *MIExpert*

Table 3

Comparison of equivalent C α positions (Å) for the six independent monomers in LpxD-1 (copies *A*, *B*, *C*, *D*, *E* and *F*) after structure superposition.

	LpxD-1 (<i>A</i>)	LpxD-1 (<i>B</i>)	LpxD-1 (<i>C</i>)	LpxD-1 (<i>D</i>)	LpxD-1 (<i>E</i>)	LpxD-1 (<i>F</i>)
LpxD-1 (<i>A</i>)	—	0.40	0.39	0.25	0.40	0.36
LpxD-1 (<i>B</i>)		—	0.43	0.37	0.21	0.44
LpxD-1 (<i>C</i>)			—	0.44	0.37	0.19
LpxD-1 (<i>D</i>)				—	0.40	0.42
LpxD-1 (<i>E</i>)					—	0.39

graphical interface. Interactive model building employed the *MIFit* software (<http://code.google.com/p/mifit/>; Table 2). For both the LpxD-1 and LpxD-2 structures the initial model-rebuilding cycles utilized the expansion of monomer models by noncrystallographic symmetry, but later in the refinement process the protein copies were independently refitted. Additional effort was needed to refit one of the monomers in the LpxD-1 structure owing to a large deviation from the conformation observed in the other two proteins. Non-crystallographic symmetry restraints were not found to be necessary in the refinement of structures from either the LpxD-1 or LpxD-2 crystals. Final rebuilding and refinement cycles were guided by automated tests applied after each set of refinement cycles to locate specific amino acids that misfitted density, contained *cis*-peptide bonds, contained ϕ - ψ angles corresponding to Ramachandran plot outliers or contained abnormal side-chain rotamer angles, strained covalent geometries or van der Waals violations.

3. Results and discussion

3.1. Structure description

In the LpxD-1 crystal, there are two protein trimers in the asymmetric unit. Out of the complete sequence of 357 amino acids, the electron density was traced through positions 5–341 in each of the six protein monomers. Relatively weak regions of electron density include the linker sequence between domains (amino acids 98–103) and the final extension at the C-terminus (amino acids 334–341). The full matrix of structure superpositions for the six individual proteins

Table 4

Comparison of equivalent C α positions (Å) for the three independent monomers in LpxD-2 (copies *A*, *B* and *C*) after structure superposition.

	LpxD-2 (<i>A</i>)	LpxD-2 (<i>B</i>)	LpxD-2 (<i>C</i>)
LpxD-2 (<i>A</i>)	—	1.09	3.10
LpxD-2 (<i>B</i>)		—	2.75

in this crystal shows very little structural variation across the protein monomers, but gives a slight indication that protein copies *A*, *B* and *C* from one trimer match best to protein copies *D*, *E* and *F*, respectively, from the second trimer (Table 3, Fig. 1).

The LpxD-2 crystal form contains one protein trimer in the asymmetric unit. In this case, the model was traced through amino-acid positions 5–350, 6–355 and 5–352 for protein units *A*, *B* and *C*, respectively. As in LpxD-1, the electron density in the linker region is relatively weak, but the additional amino acids that are visible at the C-terminus are more clearly in an α -helical conformation than the more limitedly visible C-terminal extension in the LpxD-1 form.

The protein trimer in the LpxD-2 crystal is much less symmetric than either of the two trimers in the asymmetric unit of the LpxD-1 crystal. In the LpxD-2 crystal the overall conformation of protein copy *C* is very significantly different from copies *A* and *B*, with r.m.s. deviations across all equivalent C α atoms of ~ 3 Å (Table 4). More detailed structure-superposition calculations, in which the aligned regions were restricted to specific domains, show that the structural non-equivalence of protein copy *C* compared with protein copies *A* and *B* is almost entirely owing to a shift in the relative orientations of the two major domains and the C-terminal helix (Fig. 2). Internally, the domain structures remain highly conserved. For example, comparing just the C α atoms in the β -sheet domains (amino acids 104–107) of protein copies *A* and *C* the r.m.s. deviation is only 0.38 Å. Comparing just the C α atoms in the N-terminal domains (amino acids 5–97) of protein copies *A* and *C* the r.m.s. deviation is only 0.42 Å. Similarly, if a representative protein copy from the LpxD-1 crystal form (protein *A*) is superimposed onto copies *A*, *B* and *C* of the LpxD-2 crystal form the r.m.s. deviations over all equivalent C α atoms are 0.91, 0.70 and 2.47 Å, respectively. Overall, it appears that protein

Figure 1

(*a*) Structure superposition of the six protein copies in the crystal asymmetric unit of LpxD-1 (*A*, red; *B*, yellow; *C*, blue; *D*, purple; *E*, pale blue; *F*, brown). (*b*) Structure superposition of the three protein copies in the crystal asymmetric unit of LpxD-2 (*A*, yellow; *B*, green; *C*, blue).

Figure 2

Final likelihood-weighted electron-density map for the LpxD-2 crystal structure showing the electron density around the helical bundle at the C-termini of protein copies *B* and *C*. The helix from protein copy *C* is on the left and the helix from protein copy *B* is on the right of the image. The map is contoured at $1\times$ the r.m.s. density fluctuation of the map.

copies *A* and *B* in the LpxD-2 crystal closely reproduce the conformation found across all six proteins in the LpxD-1 crystal, but that protein copy *C* in the LpxD-2 crystal is significantly displaced from this standard conformation.

3.2. Ligand-binding sites

Structure superpositions of LpxD from *C. trachomatis* containing uridine diphosphate *N*-acetylglucosamine (Buetow *et al.*, 2007; PDB

entry 2iu8) onto the structures obtained from both the LpxD-1 and LpxD-2 crystals were performed in order to assess the accessibility of the enzymatic binding sites to compound soaking, as might be performed in cocrystallography experiments. In both crystal forms these sites appear to be accessible to small molecules that could be competitive with the naturally occurring ligands. Preliminary experiments with the LpxD-2 crystal form have now shown that the binding of compounds with a molecular weight of >200 Da can be determined using these types of experiment (data not described here).

The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health, National Institute of General Medical Sciences and the Howard Hughes Medical Institute. The Advanced Light Source is supported by the Director, Office of Basic Energy Sciences of the US Department of Energy under Contract No. DE-AC02-05CH11231. We thank Rick Walter and Gina Ranieri (Shamrock Structures) for data collection at the Advanced Photon Source.

References

- Badger, J., Chie-Leon, B., Logan, C., Sridhar, V., Sankaran, B., Zwart, P. H. & Nienaber, V. (2011). *Acta Cryst.* **F67**, 749–752.
- Bartling, C. M. & Raetz, C. R. H. (2009). *Biochemistry*, **48**, 8672–8683.
- Bassetti, M., Righi, E., Esposito, S., Petrosillo, N. & Nicolini, L. (2008). *Future Microbiol.* **3**, 649–660.
- Buetow, L., Smith, T. K., Dawson, A., Fyffe, S. & Hunter, W. N. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 4321–4326.
- Dent, L. L., Marshall, D. R., Pratap, S. & Hulet, R. B. (2010). *BMC Infect. Dis.* **10**, 196.
- Kleywegt, G. J. & Jones, T. A. (1996). *Structure*, **4**, 1395–1400.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Raetz, C. R. H. & Whitfield, C. (2002). *Annu. Rev. Biochem.* **71**, 635–700.