BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

Structures of *Bacteroides fragilis* uridine 5′-diphosphate-*N*-acetylglucosamine (UDP-GlcNAc) acyltransferase (BfLpxA)

Alice Ngo,[a] Kai T. Fong,[a] Daniel L. Cox,[b] Xi Chen[a]* and Andrew J. Fisher[a,c]*

[a]Department of Chemistry, University of California, One Shields Avenue, Davis, CA 95616, USA, [b]Department of Physics, University of California, One Shields Avenue, Davis, CA 95616, USA, and [c]Department of Molecular and Cellular Biology, University of California, One Shields Avenue, Davis, CA 95616, USA. *Correspondence e-mail: xiichen@ucdavis.edu, ajfisher@ucdavis.edu

Received 10 January 2015
Accepted 16 February 2015

Edited by R. McKenna, University of Florida, USA

Keywords: left-handed β-helix; lipid A; LpxA; UDP-*N*-acetylglucosamine acyltransferase.

PDB references: BfLpxA, ligand-free, 4r36; with UDP-GlcNAc bound, 4r37

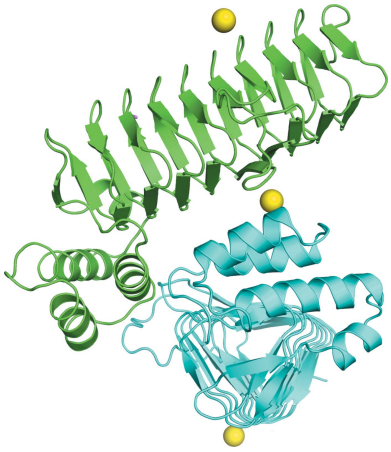
Supporting information: this article has supporting information at journals.iucr.org/d

Uridine 5′-diphosphate-*N*-acetylglucosamine (UDP-GlcNAc) acyltransferase (LpxA) catalyzes a reversible reaction for adding an *O*-acyl group to the GlcNAc in UDP-GlcNAc in the first step of lipid A biosynthesis. Lipid A constitutes a major component of lipopolysaccharides, also referred to as endotoxins, which form the outer monolayer of the outer membrane of Gram-negative bacteria. Ligand-free and UDP-GlcNAc-bound crystal structures of LpxA from *Bacteroides fragilis* NCTC 9343, the most common pathogenic bacteria found in abdominal abscesses, have been determined and are presented here. The enzyme crystallizes in a cubic space group, with the crystallographic threefold axis generating the biological functional homotrimer and with each monomer forming a nine-rung left-handed β-helical (LβH) fold in the N-terminus followed by an α-helical motif in the C-terminus. The structure is highly similar to LpxA from other organisms. Yet, despite sharing a similar LβH structure with LpxAs from *Escherichia coli* and others, previously unseen calcium ions are observed on the threefold axis in *B. fragilis* LpxA to help stabilize the trimeric assembly.

1. Introduction

Uridine 5′-diphosphate-*N*-acetylglucosamine (UDP-GlcNAc) acyltransferase (LpxA) catalyzes a reversible reaction for adding an *O*-acyl group to the GlcNAc in UDP-GlcNAc in the first step of lipid A biosynthesis (Fig. 1). Although LpxA is present in all Gram-negative bacteria, its substrate preference towards specific acyl chains varies for different species. LpxA from *Bacteroides fragilis* NCTC 9343 is most likely to prefer the 15-carbon saturated fatty acid-acyl carrier protein since the structure of its lipid A was determined to have a C_{15} fatty-acyl group at the 3-OH position of the GlcNAc in UDP-GlcNAc (Berezow *et al.*, 2009).

Lipid A is a highly acylated glycophospholipid that is a major component and serves as the hydrophobic anchor of lipopolysaccharides (LPSs) for most Gram-negative bacteria (Raetz *et al.*, 2007). LPSs are often referred to as endotoxins because of their immune-activating effects on mammalian host systems (Masoudi *et al.*, 2014). The development of detoxified forms of lipid A could provide vaccine adjuvants and immunostimulants. For example, Toll-like receptor (TLR)-activating lipid IVa and synthetic mimetics have been developed to decrease the risk of inflammatory side effects (Saitoh *et al.*, 2004; Hashimoto *et al.*, 2002). Gram-negative bacteria also have the ability to modify the structures of lipid A, which may be related to their evolutionary survival or manipulation of



the host innate immune response, but the mechanism is not well understood. One example has been shown for *Salmonella typhimurium*, which can modify its lipid A by adding an amino-arabinose to the 4′-phosphate of lipid A, which resulted in conferring resistance to the antimicrobial peptide polymyxin (Gunn *et al.*, 1998).

Bacteria strongly rely on the integrity of the amphipathic outer membrane for survival. The opportunistic *B. fragilis* is an anaerobic microbe that constitutes only 1–2% of the human gastrointestinal tract flora; however, it is the predominant isolate in abdominal abscesses and soft-tissue infections and has the highest polysaccharide capsule diversity among Gram-negative bacteria (Pumbwe *et al.*, 2006; Patrick *et al.*, 2010). In addition, cases of lethal multidrug-resistant strains of *B. fragilis* have been reported (Katsandri *et al.*, 2006; Wareham *et al.*, 2005) and efforts to identify the antimicrobial resistant genes have been published (Sydenham *et al.*, 2015). Our historic reliance on antibiotics for combating bacterial infections has led to the evolution of antibiotic-resistant pathogenic bacteria. Therefore, there is an urgent need to develop new drugs for the treatment of bacterial infectious diseases. One approach is to design inhibitors specific for the enzymes involved in the biosynthetic pathway of LPS synthesis belonging to Gram-negative bacteria. LpxA, which catalyzes the first step in lipid A biosynthesis (Coleman & Raetz, 1988), is an excellent target for the development of antimicrobial drugs.

LpxA also has unique structural features. It belongs to the left-handed parallel β-helix (LβH) protein superfamily which has an LβH structural domain containing multiple imperfect tandem-repeated six-residue motifs (Raetz & Roderick, 1995). In addition to LpxA, the LβH domain has also been identified in another acyltransferase, LpxD, involved in lipid A biosynthesis (Buetow *et al.*, 2007; Bartling & Raetz, 2009; Badger *et al.*, 2011), γ-class carbonic anhydrases (Kisker *et al.*, 1996; Jeyakanthan *et al.*, 2008) and other acyltransferases

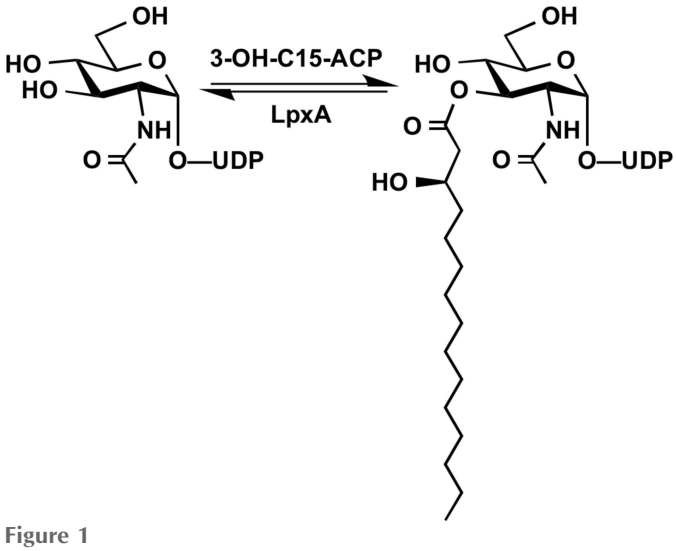


Figure 1
The reversible UDP-GlcNAc acylation reaction catalyzed by *B. fragilis* LpxA. Abbreviations: 3-OH-C15-ACP, 3-hydroxypentadecanoic fatty acid-acyl carrier protein; LpxA, UDP-GlcNAc acyltransferase.

including tetrahydrodipicolinate (THDP) *N*-succinyltransferases (Beaman *et al.*, 1997), xenobiotic acetyltransferases (Beaman *et al.*, 1998), bifunctional glucosamine-1-phosphate (GlcN-1-P) acetyltransferase and *N*-acetylglucosamine-1-phosphate (GlcNAc-1-P) uridyltransferase (GlmU; Brown *et al.*, 1999), another UDP-sugar *N*-acetyltransferase from *Campylobacter jejuni* (PglD; Rangarajan *et al.*, 2008), antibiotic acetyltransferases (Sugantino & Roderick, 2002; Kehoe *et al.*, 2003), serine acetyltransferases (Olsen *et al.*, 2004), the bacterial sialic acid-containing polysaccharide *O*-acetyltransferases OatWY (Lee *et al.*, 2009) and NeuO (Schulz *et al.*, 2011), galactoside acetyltransferases (Wang *et al.*, 2002) and maltose acetyltransferases (Lo Leggio *et al.*, 2003). Recently, the first triangular-fold LβH domain has been structurally confirmed in a human protein, a p27 subunit (PDB entry 3tv0; Z. S. Derewenda, U. Derewenda, A. Kowalska & M. Zheng, unpublished work) that acts as part of the dynactin protein complex involved in the *in vivo* function of a microtubule (MT)-based motor for mitosis (Yeh *et al.*, 2013). Other human proteins that have been predicted, but not confirmed, to contain the LβH fold are dynactin p25 subunit, eukaryotic translation initiation factor 2B ε- and γ-subunits and GDP-mannose pyrophosphorylases A and B (Choi *et al.*, 2008). A significant structural difference between *Escherichia coli* LpxA (EcLpxA) and the p27 subunit is the lack of an α-helical cap in the C-terminus of the p27 domain. In addition, p27 does not form a homotrimer as is found in LpxA, but instead exists as a heterodimer with a different protein, a p25 subunit. Since this newly determined crystal structure of the human LβH protein possesses structural similarity to LpxA, a potential antimicrobial drug target, it is essential to identify structural differences between the human and bacterial proteins and to test drugs that are selective against bacterial LpxAs (Yeh *et al.*, 2013).

The crystal structures of LpxAs from the pathogenic bacteria *E. coli* (Raetz & Roderick, 1995; Ulaganathan *et al.*, 2007), *Helicobacter pylori* (Lee & Suh, 2003), *Acinetobacter baumannii* (Badger *et al.*, 2012), *Leptospira interrogans* (Robins *et al.*, 2009), *Burkholderia thailandensis* (Baugh *et al.*, 2013) and *C. jejuni* (PDB entry 3r0s; Center for Structural Genomics of Infectious Diseases, unpublished work), as well as the plant *Arabidopsis thaliana* (Joo *et al.*, 2012), have been reported. Although these LpxAs share a high degree of sequence homology and display similar three-dimensional structures, the LpxAs from different species vary in their selectivity towards acyl chains of various lengths and structures. For example, *E. coli* LpxA (EcLpxA) prefers the 14-carbon fatty acid (*R*)-3-hydroxymyristoyl chain, *Chlamydia trachomatis* LpxA prefers the 14-carbon myristoyl chain and *H. pylori* LpxA (HpLpxA) prefers the 16-carbon (*R*)-3-hydroxypalmitoyl chain (Sweet *et al.*, 2001; Lee & Suh, 2003). The difference in the acyl-chain length and structure in the lipid A structures of *B. fragilis* and *B. thetaiotaomicron* may contribute to the varied degrees of toxicity of these two closely related species (Berezow *et al.*, 2009). Here, we present the crystal structures of ligand-free and UDP-GlcNAc-bound antibiotic-resistant *B. fragilis* NCTC 9343 LpxA (BfLpxA). Its

Table 1
BfLpxA production information (the N-terminal His_6 tag in the fusion protein comes from the pET-15b vector).

Source organism	*B. fragilis* NCTC 9343
DNA source	Genomic DNA of *B. fragilis* NCTC 9343 (from ATCC)
Forward primer†	5′-GATCCATATGATAAGTCCCTTAGCGTATATT-CCCTTAGCGTATATTCATCCCGAA-3′
Reverse primer‡	5′-CGCGGATCCTTACCTGATAATTCCCCGTTCC-ATAATTCCCCGTTCCGAGTTGCGAAT-3′
Cloning vector	pET-15b
Expression vector	pET-15b
Expression host	*E. coli* BL21(DE3)
Complete amino-acid sequence of the construct produced§	MGSSHHHHHHSSGLVPRGSHMISPLAYIHPEAKI-GENVEIAPFVYIDRNVVIGDNNKIMANANILY-GSRIGNGNTIFPGAVIGAIPQDLKFKGEESTA-EIGDNNLIRENVTINRGTAAKGRTIVGNNNXL-MEGVHVAHDALIGNGCIVGNSTKMAGEIIIDD-NAIISANVLMHQFCRVGGYVMIQGGCRFSKDI-PPYIIAGREPIAYSGINIIGLRRRGFSNEIIE-NIHNAYRIIYQSGLNTSDALTKVEAEVPASPE-IEYIVDFIRNSERGIIR

† The NdeI restriction site is underlined. ‡ The BamHI restriction site is underlined. § The first 20 amino acids in the N-terminus introduced by the pET-15b vector are underlined.

Table 2
Crystallization.

Method	Vapor diffusion
Plate type	Cryschem plate
Temperature (K)	294
Protein concentration (mg ml^{-1})	7.5
Buffer composition of protein solution	20 m*M* Tris–HCl pH 7.5
Composition of reservoir solution	28% PEG 600, 200 m*M* calcium acetate, 100 m*M* sodium cacodylate pH 6.5
Volume and ratio of drop	2 µl:2 µl (protein:reservoir)
Volume of reservoir (µl)	500

structural similarities to and differences from EcLpxA are examined.

2. Materials and methods

2.1. Cloning, overexpression and purification of BfLpxA

Genomic DNA of *B. fragilis* NCTC 9343 [American Type Culture Collection (ATCC), Manassas, Virginia, USA] was used as the template for polymerase chain reaction (PCR) to amplify the target gene *lpxA* (NCBI accession No. CAH06570). The primers used are shown in Table 1. PCR was performed in a reaction mixture (50 µl) containing genomic DNA (100 ng), forward and reverse primers (0.2 µ*M* each), 1× Herculase II buffer, dNTP mixture (1 m*M*) and 5 U (1 µl) Herculase II DNA polymerase (Agilent Technologies, Santa Clara, California, USA). The reaction was subjected to 30 cycles of amplification at an annealing temperature of 50°C. The resulting PCR product was digested with NdeI and BamHI restriction enzymes and inserted into pET-15b expression vector pre-digested with the same restriction enzymes to construct a plasmid to express N-terminally His_6-tagged recombinant BfLpxA. BfLpxA was overexpressed in *E. coli* BL21 (DE3) cells. The cells were grown at 37°C in Luria–Bertani medium containing ampicillin (100 mg l^{-1}) until the $OD_{600\,nm}$ of the culture reached 0.6. Expression of the protein was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 m*M* followed by incubation of the culture at 37°C for 3 h. To harvest the cells, the culture was centrifuged at 4000 rev min^{-1} for 2 h and the cell pellet was resuspended in lysis buffer (100 m*M* Tris–HCl pH 8.0, 0.1% Triton X-100; 20 ml per litre of cell culture) followed by freezing at −80°C overnight. Cell lysis was performed by thawing the frozen cells at 37°C and adding lysozyme (0.3 mg ml^{-1}) and DNaseI (18 µg ml^{-1}), followed by incubation at 37°C for 1 h with shaking at 190 rev min^{-1}. The crude lysate was centrifuged at 12 000 rev min^{-1} for 30 min and the supernatant was collected and applied onto nickel–NTA resin that had been pre-equilibrated with ten column volumes of binding buffer

Table 3
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Protein sample	Ligand-free	UDP-GlcNAc-bound
Data-collection statistics		
X-ray source	Beamline 7-1, SSRL	Beamline 7-1, SSRL
Wavelength (Å)	1.12709	1.12709
Space group	$P4_132$	$P4_132$
Unit-cell parameters (Å, °)	$a = b = c = 149.05$, $\alpha = \beta = \gamma = 90$	$a = b = c = 149.40$, $\alpha = \beta = \gamma = 90$
Monomers in asymmetric unit	2	2
V_M (Å^3 Da^{-1})	2.49	2.51
Solvent content (%)	50.7	51.1
Resolution (Å)	1.90 (1.94–1.90)	1.90 (1.95–1.90)
No. of reflections	477344 (29337)	321337 (21128)
No. of unique reflections	45045 (2836)	45186 (2959)
Completeness (%)	99.9 (99.0)	99.9 (99.6)
R_{merge}† (%)	5.4 (55.4)	5.2 (35.7)
$\langle I/\sigma(I)\rangle$	35.3 (4.9)	30.0 (5.6)
Refinement statistics		
Resolution (Å)	1.90	1.90
No. of reflections ($F > 0$) used in refinement	42699	42852
R factor‡ (%)	15.5	14.7
R_{free}‡ (%)	19.2	18.5
R.m.s.d., bond lengths (Å)	0.018	0.019
R.m.s.d., bond angles (°)	1.228	1.798
No. of atoms		
Protein	3897	3925
Ca^{2+}	3	4
Acetate	8	8
PEG	43	35
UDP-GlcNAc	—	39
Water	453	517
Average B factors (Å^2)		
Protein	25.2	20.0
Ca^{2+}	24.9	26.4
Acetate	26.5	24.1
PEG	42.4	42.2
UDP-GlcNAc	—	32.1
Water	35.5	31.4
Ramachandran plot statistics§		
Most favourable (%)	97	96
Allowed (%)	3	4
Disallowed (%)	0	0
PDB code	4r36	4r37

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean of i observations $I_i(hkl)$ of reflection hkl. ‡ R factor and $R_{free} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}| \times 100$ for 95% of the recorded data (R factor) or for 5% of the data, which were omitted from refinement (R_{free}). § Ramachandran plot statistics are from the PDB submission validation report.

(50 m*M* Tris–HCl, 0.5 *M* NaCl, 5 m*M* imidazole pH 7.5) in a 10 × 1.5 cm column. The bound BfLpxA was washed with eight column volumes of binding buffer and eight column volumes of washing buffer (50 m*M* Tris–HCl, 0.5 *M* NaCl, 35 m*M* imidazole pH 7.5). The protein was eluted with elution buffer (50 m*M* Tris–HCl, 0.5 *M* NaCl, 200 m*M* imidazole pH 7.5). Fractions containing protein were pooled and dialyzed against 20 m*M* Tris–HCl buffer pH 7.5 using 10 kDa dialysis membrane tubing. The concentration of the purified protein was determined with a bicinchoninic acid (BCA) Protein Assay Kit (Pierce Biotechnology, Rockford, Illinois, USA) using bovine serum albumin (BSA) as the protein standard. The absorbance of the samples was measured in a 96-well microplate at 562 nm using a BioTek Synergy HT Multi-Mode Microplate Reader. Generally, each litre of expression medium produced 90 mg purified protein.

2.2. Crystallization and X-ray data collection

Initial screening for crystallization conditions was performed using the Wizard Classic 1, Classic 2, Cryo 1 and Cryo 2 crystallization screens (Rigaku). Optimized

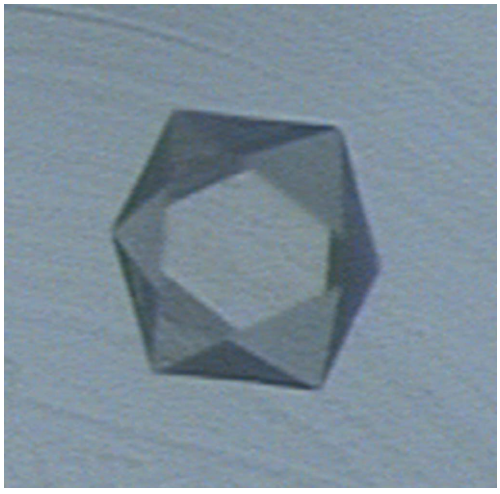


Figure 2
A crystal of ligand-free recombinant BfLpxA. The crystal is approximately 150 µm in diameter.

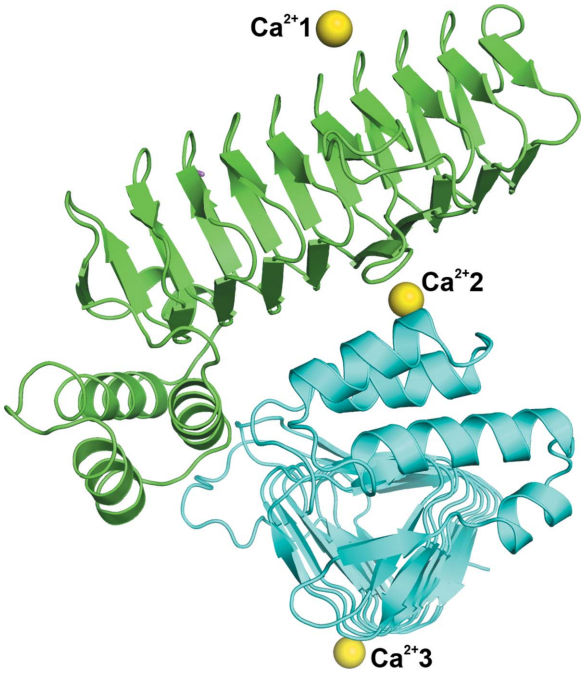
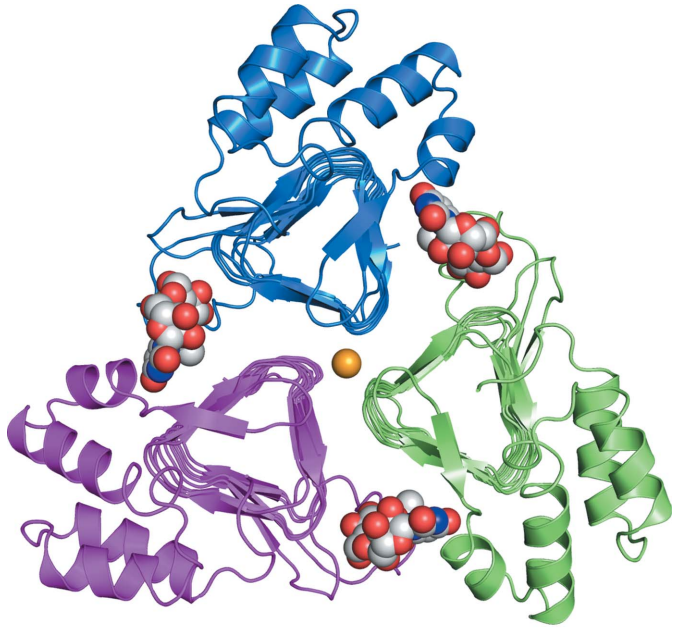
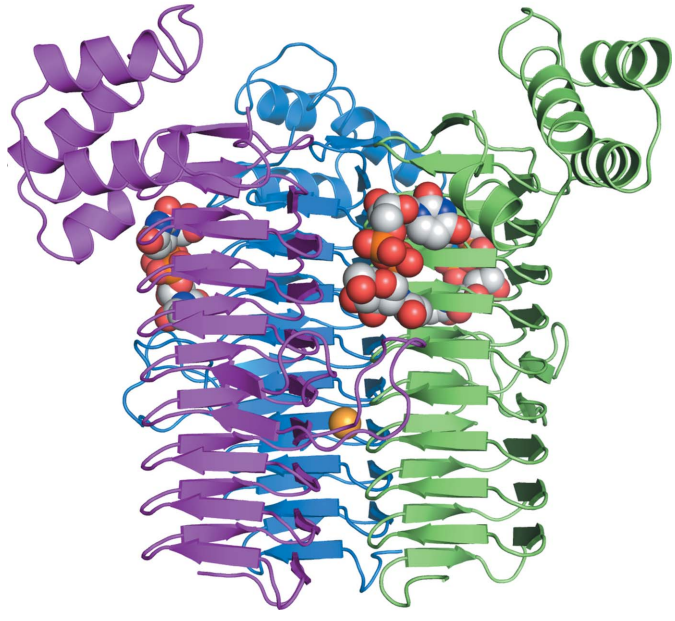


Figure 3
Two subunits of BfLpxA in an asymmetric unit with three calcium cations (gold). Two calcium ions (1 and 3) are situated on the crystallographic threefold axis, which resides in the center of the functional trimer.



(*a*)



(*b*)

Figure 4
BfLpxA functional trimer with bound UDP-GlcNAc (white space-filling atoms) and a Ca^{2+} cation (gold) on the crystallographic threefold axis. The view in (*a*) is down the crystallographic threefold axis of the trimer generated by the *A* subunit; that in (*b*) shows the same trimer rotated 90° so that the crystallographic threefold axis is vertical.

crystallization was performed by the sitting-drop vapor-diffusion method using 24-well Cryschem Plates (Hampton Research) at 294 K (Table 2). The sitting drop was prepared by mixing 2 µl reservoir solution with 2 µl protein solution (7.5 mg ml^{-1}) and was equilibrated against 0.5 ml reservoir solution. The condition that resulted in the highest quality crystals was 28% PEG 600, 200 m*M* calcium acetate, 100 m*M* sodium cacodylate pH 6.5. To generate the binary complex of BfLpxA with UDP-GlcNAc, the crystal was soaked with 15 m*M* UDP-GlcNAc for 30 min prior to crystal cryocooling. Crystals were flash-cooled directly from the mother liquor.

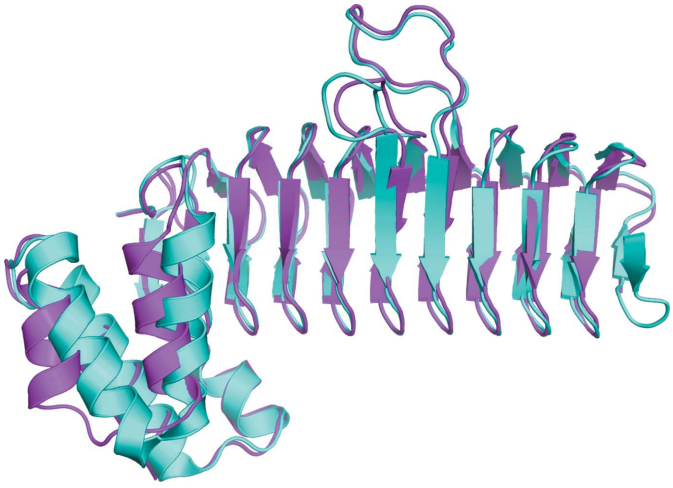


Figure 5
BfLpxA (magenta) overlaid with EcLpxA (cyan). The r.m.s.d. of BfLpxA with EcLpxA is 0.950 Å for 243 C^{α}-atom pairs.

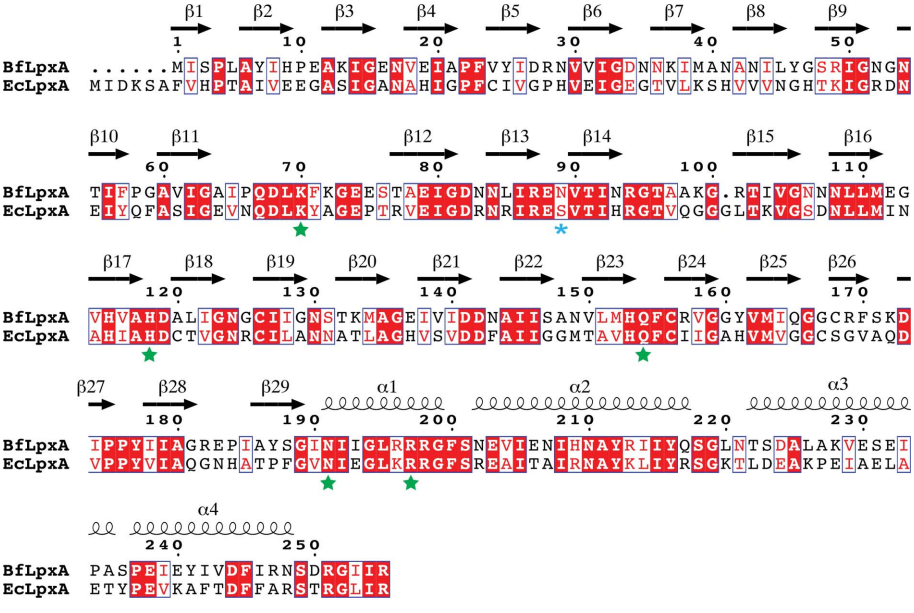


Figure 6
Protein sequence alignment of BfLpxA (GenBank accession No. Q5LH16) and EcLpxA (GenBank accession No. P0A722), which share 42% identity. Secondary-structure elements are drawn above the sequence. Residues that hydrogen-bond to the UDP-GlcNAc substrate are indicated with a green star, while Asn89, which coordinates the calcium ion at the threefold axis, is marked by a blue asterisk (the numbers reflect the BfLpxA sequence). This figure was generated using *ESPript* 3.0 (Robert & Gouet, 2014).

X-ray diffraction data were collected on beamline 7-1 at Stanford Synchrotron Radiation Lightsource (SSRL) using an ADSC Q315r CCD detector. The crystals were exposed for 4 s for each 0.3° crystal rotation. Diffraction intensity data were integrated with *XDS* (Kabsch, 2010) and scaled with *AIMLESS* (Evans & Murshudov, 2013; Winn *et al.*, 2011).

2.3. Structure solution and refinement

The ligand-free structure was solved by molecular replacement with *Phaser* (McCoy, 2007) using the previously determined structure of EcLpxA (PDB entry 1lxa; Raetz & Roderick, 1995). The model was built using the molecular-modeling program *Coot* (Emsley & Cowtan, 2004) and was refined using *PHENIX* (Adams *et al.*, 2010). The UDP-GlcNAc-bound structure was solved by initially refining the ligand-free structure (without any solvent molecules) against diffraction data from UDP-GlcNAc-soaked crystals and using *Coot* to build in the UDP-GlcNAc substrate. Numerous rounds of refinement with subsequent model adjustment and solvent addition were then carried out for the UDP-GlcNAc-bound binary structure. Both the ligand-free and the UDP-GlcNAc-bound binary structures were refined to 1.90 Å resolution. The final data-collection and refinement statistics are listed in Table 3.

3. Results and discussion

3.1. Crystallization of BfLpxA and soaking with UDP-GlcNAc

BfLpxA was crystallized using 28%(*w*/*v*) PEG 600, 200 m*M* calcium acetate, 100 m*M* sodium cacodylate pH 6.5. The crystals (Fig. 2) belonged to the cubic space group $P4_132$, with unit-cell parameter a = 149.05 Å (Table 3). This results in two monomers per asymmetric unit (V_M = 2.49 Å^3 Da^{-1}, 50.7% solvent content; Fig. 3; Matthews, 1968). The ligand-free structure was solved by molecular replacement using the EcLpxA structure (PDB entry 1lxa) as a phasing model. Each monomer of the asymmetric unit is situated at the crystallographic threefold axis, generating the homotrimeric structure (Fig. 4). The homotrimeric structure of BfLpxA is similar to the previously solved crystal structures of other LpxAs. Soaking the crystals in UDP-GlcNAc (15 m*M*) for 30 min prior to flash-cooling and data collection resulted in strong and well defined electron density corresponding to the entire UDP-GlcNAc molecule bound to the surface of the β-helix (Fig. 4).

3.2. BfLpxA apo structure

The BfLpxA apo structure is a homotrimer with a nine-rung left-handed β-helical (LβH) fold in the

N-terminus with two loops extending from rungs three and four followed by an α-helical motif in the C-terminus (Fig. 4). The crystal structure is highly similar to and aligns well with that of EcLpxA (PDB entry 1lxa). The EcLpxA structure consists of 18 amino acids per rung and six amino acids per turn in the LβH region, as shown in the structural alignment of the monomer with BfLpxA (Fig. 5). BfLpxA is six amino acids shorter than EcLpxA and its N-terminus is found in the second short β-strand when the sequence is aligned with that of EcLpxA (Fig. 6). Each complete rung consists of three short β-strands constructed from imperfect hexapeptide repeats. The six-residue sequence motif in BfLpxA is generally (I/V)-*XXX*-(A/V/N/C)-*X*, where the first position I/V is Ile or Val and forms an internal facing hydrophobic ladder and the fifth position is generally occupied by a small residue (Ala, Val, Ser, Thr or Cys) also facing inside the LβH. Sequence-homology alignment using *Clustal Omega* yields 41.96% identity between EcLpxA and BfLpxA (Fig. 6). Superposition of EcLpxA with BfLpxA in *PyMOL* yields a root-mean-square difference (r.m.s.d.) of 0.950 Å for 243 C^{α}-atom pairs.

3.3. BfLpxA complex with UDP-GlcNAc

Previous reports on EcLpxA have shown that its active site is located at the subunit–subunit interface, where UDP-GlcNAc binds (Ulaganathan *et al.*, 2007). This is also the case for BfLpxA (Fig. 7). The 1.90 Å resolution crystal structure of BfLpxA soaked with UDP-GlcNAc revealed the substrate bound to only one monomer of the asymmetric unit because the active site of the other monomer was blocked by crystal contacts. The uracil base of the UDP-GlcNAc is sandwiched between Met163 and Ile193 from the α-helical region of one subunit (subunit *A*1). Additionally, two amino-acid residues, Asn191 and Arg197 of subunit *A*1, interact with the UDP

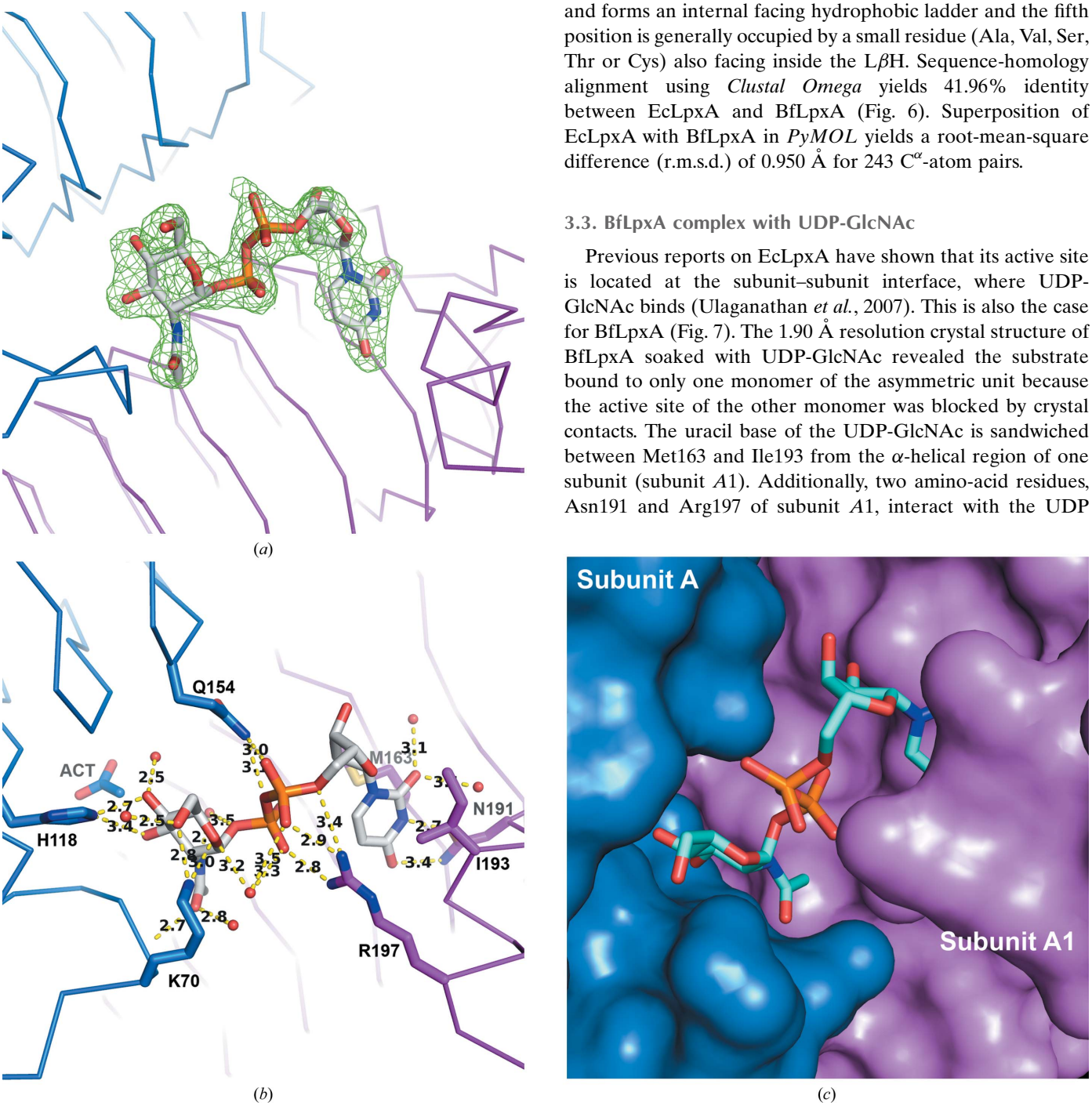


Figure 7
Binding of UDP-GlcNAc at the interface between two monomers within the active trimer. (*a*) Simulated-annealed OMIT map contoured at 2.5σ showing the unbiased positive $F_o - F_c$ electron density around the UDP-GlcNAc substrate. Side chains are omitted for clarity. (*b*) Amino-acid residues Arg197 and Asn191 from subunit *A*1 (magenta) of BfLpxA interact with the uracil and diphosphate of UDP-GlcNAc. Amino-acid residues Lys70, His118, Leu69 and Gln154 from subunit *A* (blue) of BfLpxA interact with the α-phosphate and GlcNAc of UDP-GlcNAc. Also shown are ordered solvent molecules (red spheres) that interact with UDP-GlcNAc, and the acetate ion (ACT). (*c*) Surface rendering at the interface with UDP-GlcNAc positioned between subunit *A* and the threefold-related partner subunit *A*1 of the BfLptA trimer.

(Fig. 7). Asn191 hydrogen-bonds to N3 and O4 of the uracil and Arg197 interacts with three phosphate O atoms: one each from the α- and β-phosphates and the O atom between the α-phosphate and the ribose (Fig. 7). Gln154 extends from coil 8 of the neighboring threefold-related subunit (subunit *A*) and interacts with an O atom of the α-phosphate as well as the bridging O atom between the α- and β-phosphates (Fig. 7). Three amino acids from subunit *A* interact with the GlcNAc moiety as follows. His118 is the general base that generates the nucleophile on the 3-hydroxyl group and also appears to interact with the 4-hydroxyl group of GlcNAc in UDP-GlcNAc (Fig. 7). This is similar to the general base-catalyzed mechanism of His125 of EcLpxA; however, His125 in EcLpxA does not interact with the 4-hydroxyl of GlcNAc in UDP-GlcNAc (Ulaganathan *et al.*, 2007). Leu69 and Lys70 both extend from loop 1 (from subunit *A*), where the main-chain N atom of Leu69 interacts with the *N*-acetyl group of UDP-GlcNAc and N^{ζ} of Lys70 interacts with the ring O atom and 6-hydroxyl group of GlcNAc in UDP-GlcNAc. BfLpxA shares many of the conserved residues, Asn198, Arg204, Gln161, His125, Lys76 and Leu75, found in EcLpxA (Ulaganathan *et al.*, 2007), which are analogous to Asn191, Arg197, His118, Lys70 and Leu69 in BfLpxA. Arg204 of EcLpxA forms hydrogen bonds to O2 and N3 of the uracil (Ulaganathan *et al.*, 2007). However, in BfLpxA the corresponding Arg198 is not within contact range of the uracil. Also, Arg204 of EcLpxA only interacts with the α-phosphate (Ulaganathan *et al.*, 2007), whereas the corresponding Arg197 in BfLpxA interacts with O atoms on both the α- and β-phosphates.

3.4. Ca^{2+}, acetate and PEG in the BfLpxA structures

Extended tubular electron density was found between the α-helix of one subunit and the loops of a symmetry-related subunit, which we modeled as a polyethylene glycol (PEG) molecule. The PEG molecule resides on the outer face of the homotrimer opposite to the active binding site of UDP-GlcNAc (Fig. 8). This electron density is not likely to be an acyl-chain substrate since it is not located near the active cleft and no lipids or detergents were used in crystallization. The observed PEG comes from the crystallization solution, and its location suggests that it plays an important role in the formation and maintenance of the crystal lattice.

Strong electron density in the shape of a flat triangle observed near the UDP-GlcNAc active site at the interface

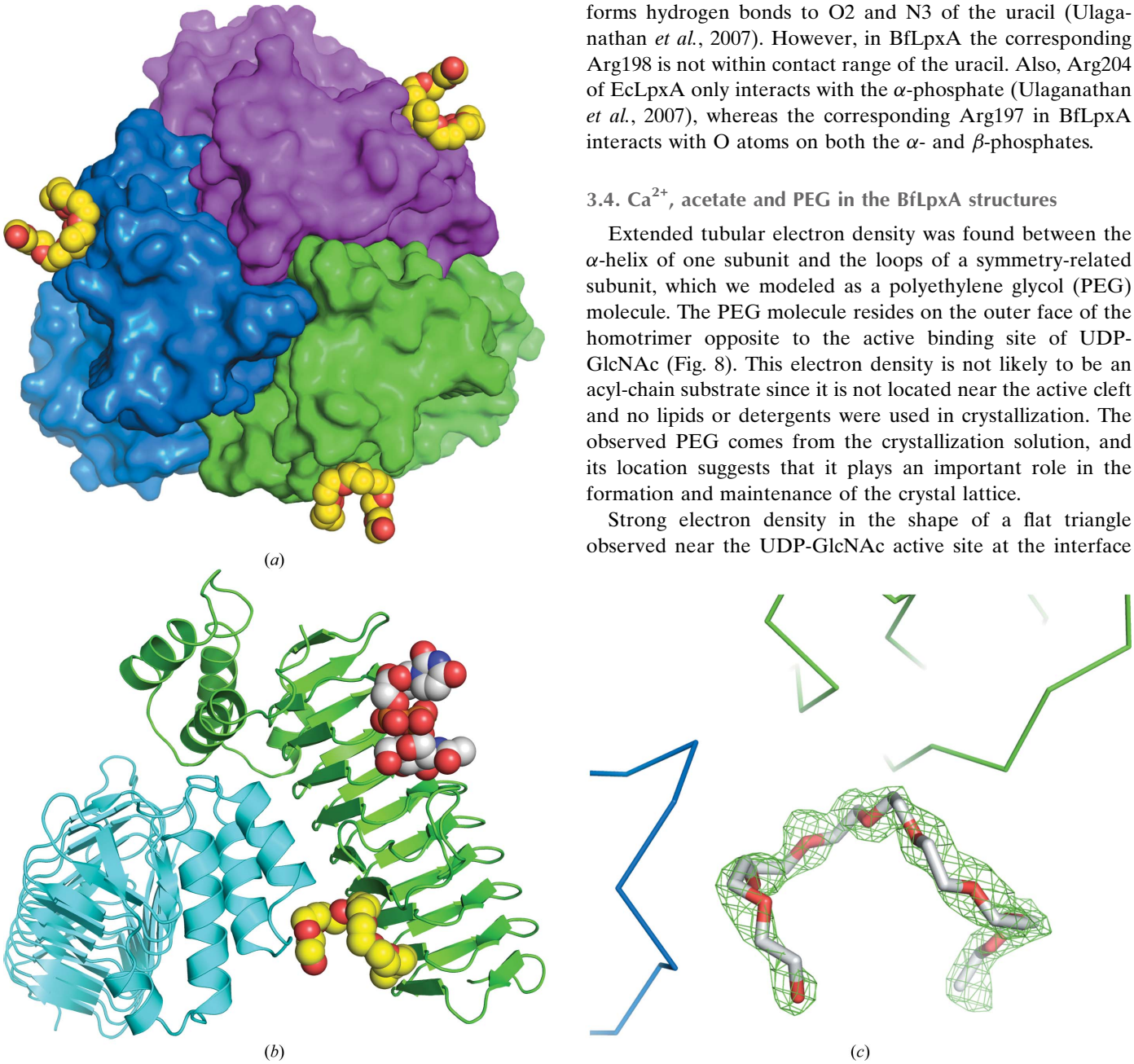


Figure 8
(*a*) BfLpxA trimer surface rendering showing the position of bound PEG (space-filling representation with yellow C atoms). The view is looking down the LβH domain along the crystallographic threefold axis. (*b*) BfLpxA in the asymmetric unit showing bound PEG (space-filling representation with yellow C atoms) and UDP-GlcNAc (space-filling representation with white C atoms). (*c*) Simulated-annealed OMIT map contoured at 2.5σ showing the unbiased positive $F_o - F_c$ electron density around the PEG molecule bound between the two subunits in the asymmetric unit.

between two subunits was modeled as an acetate ion, which was present in the crystallization condition. This acetate ion is seen in both active sites in the asymmetric unit, and is observed even in the absence of UDP-GlcNAc substrate. The acetate ion forms hydrogen bonds to $N^{\varepsilon 2}$ of His115 and $O^{\varepsilon 1}$ of Gln67 of one subunit and $N^{\delta 2}$ of Asn130 from another trimer-related subunit. When superimposed on EcLpxA with the bound product UDP-3-*O*-(*R*-3-hydroxymyristoyl)-GlcNAc, the acetate molecule aligns near C3 of the 3-hydroxymyristoyl-GlcNAc fatty-acyl chain (Fig. 9). The acetate O atom that hydrogen-bonds to Gln67 and His115 in BfLpxA superimposes almost exactly on the O atom of the 3-hydroxyl group, which may reflect the preference of LpxAs towards 3-hydroxy-containing fatty-acyl chains in the substrate.

Strong electron density situated on the crystallographic threefold axis for each trimeric complex was modeled as a calcium cation (with an occupancy of 0.33; Fig. 10; $Ca^{2+}1$ and $Ca^{2+}3$ in Fig. 3). This ion is coordinated by $O^{\delta 1}$ of Asn89 from each monomer in the trimer with a ligand distance of 2.3 Å. Additionally, three water molecules (one from each asymmetric unit) are at 2.4 Å from the calcium cation. Finally, a more distant water ligand is also situated on the crystallographic threefold axis and is 2.8 Å away from the calcium cation to complete the coordination with the geometry of a capped trigonal biprism (Glusker, 1991).

While no previous LpxA structures contained a divalent cation at the threefold axis, the presence of a divalent ion situated on the threefold axis of other left-handed β-helical folds is observed in *N*-acetylglucosamine-1-phosphate uridyltransferase (GlmU) structures (Mochalkin *et al.*, 2007, 2008; Sulzenbacher *et al.*, 2001; Olsen *et al.*, 2007; Doig *et al.*, 2014; Olsen & Roderick, 2001; Jagtap *et al.*, 2013; Kostrewa *et al.*, 2001) and that of γ-class carbonic anhydrase (Jeyakanthan *et al.*, 2008).

The BFLpxA structure also contains another calcium ion ($Ca^{2+}2$ in Fig. 3), which is located between the α-helical region of subunit *B* and rungs 4–5 of the LβH of subunit *A*. This calcium is coordinated by seven water molecules and is involved in mediating crystal packing because residues from subunits *A* and *B* hydrogen-bond to the water ligands of the calcium. The crystallization buffer contained 0.2 *M* calcium acetate, and the exact role of Ca^{2+} in the structure and function of BfLpxA is unknown.

4. Conclusion

In conclusion, BfLpxA shares a similar active site with known LpxAs such as that from *E. coli*. Their crystal structures appear to align well despite their selectivity towards acyl chains of different lengths and structures. The active site is composed of two neighboring subunits in which one subunit interacts with the UDP and the partner subunit interacts with the GlcNAc moiety with some shared interactions on the diphosphate. His118 is the general base that generates the nucleophile on the 3-hydroxyl group and also appears to interact with the 4-hydroxyl group of the GlcNAc in UDP-GlcNAc. This is the first crystal structure of LpxA with calcium cations bound. Asn89 of BfLpxA, which coordinates the

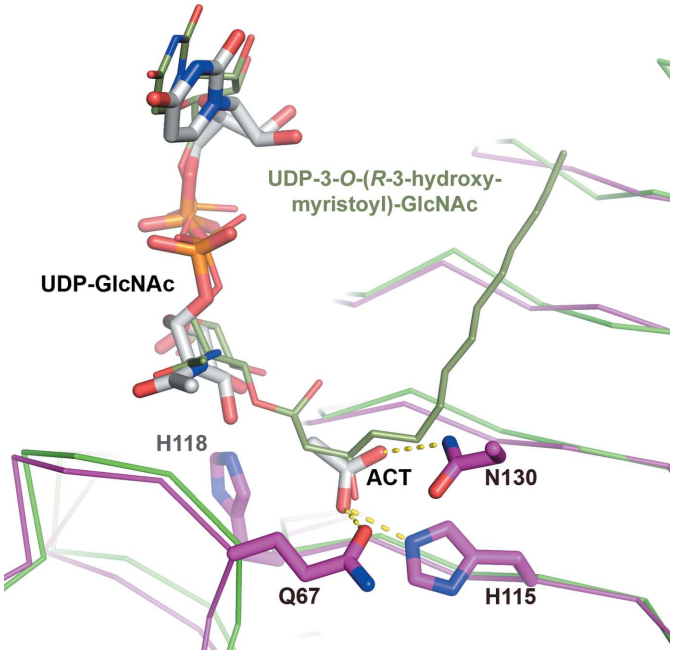


Figure 9
The acetate ion (ACT; stick representation with white C atoms) found in BfLpxA (magenta) aligns near C3 of the hydroxymyristoyl chain of the product (olive-green sticks) bound to EcLpxA (green). Gln67 and His115 interact with the acetate along with Asn130 from a trimer-related subunit. The UDP-GlcNAc bound to BfLpxA is also shown in stick representation with white C atoms.

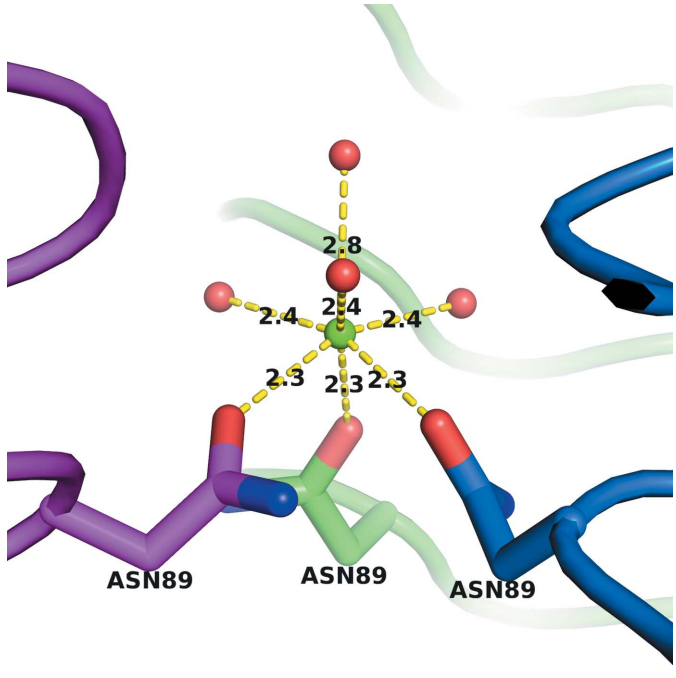


Figure 10
Calcium ligation at the trimer interface of BfLpxA. The calcium ion (small green sphere) is situated on the crystallographic threefold axis (vertical in the plane of the figure) and is ligated by Asn89 from each subunit and three water ligands (small red spheres). A fourth water molecule also lies on the crystal threefold axis above the calcium ion. Numbers represent distances in Å.

calcium ion, is a serine in EcLpxA and may explain the absence of calcium in this enzyme. The calcium cations may come from the calcium acetate buffer used in the crystallization conditions, but their exact role, if any, is unknown.

Acknowledgements

This work was supported by the UC-Davis Campus Research Investments in Science and Engineering (RISE) Program and National Institutes of Health grants R01HD065122 and R01GM094523. Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource, a Directorate of SLAC National Accelerator Laboratory and an Office of Science User Facility operated for the US Department of Energy Office of Science by Stanford University. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393) and the National Center for Research Resources (P41RR001209). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Badger, J., Chie-Leon, B., Logan, C., Sridhar, V., Sankaran, B., Zwart, P. H. & Nienaber, V. (2011). *Acta Cryst.* F**67**, 749–752.
Badger, J., Chie-Leon, B., Logan, C., Sridhar, V., Sankaran, B., Zwart, P. H. & Nienaber, V. (2012). *Acta Cryst.* F**68**, 1477–1481.
Bartling, C. M. & Raetz, C. R. (2009). *Biochemistry*, **48**, 8672–8683.
Baugh, L. *et al.* (2013). *PLoS One*, **8**, e53851.
Beaman, T. W., Binder, D. A., Blanchard, J. S. & Roderick, S. L. (1997). *Biochemistry*, **36**, 489–494.
Beaman, T. W., Sugantino, M. & Roderick, S. L. (1998). *Biochemistry*, **37**, 6689–6696.
Berezow, A. B., Ernst, R. K., Coats, S. R., Braham, P. H., Karimi-Naser, L. M. & Darveau, R. P. (2009). *Microb. Pathog.* **47**, 68–77.
Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreulx, D., Cambillau, C. & Bourne, Y. (1999). *EMBO J.* **18**, 4096–4107.
Buetow, L., Smith, T. K., Dawson, A., Fyffe, S. & Hunter, W. N. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 4321–4326.
Choi, J. H., Govaerts, C., May, B. C. H. & Cohen, F. E. (2008). *Proteins*, **73**, 150–160.
Coleman, J. & Raetz, C. R. (1988). *J. Bacteriol.* **170**, 1268–1274.
Doig, P., Boriack-Sjodin, P. A., Dumas, J., Hu, J., Itoh, K., Johnson, K., Kazmirski, S., Kinoshita, T., Kuroda, S., Sato, T. O., Sugimoto, K., Tohyama, K., Aoi, H., Wakamatsu, K. & Wang, H. (2014). *Bioorg. Med. Chem.* **22**, 6256–6269.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Glusker, J. P. (1991). *Adv. Protein Chem.* **42**, 1–76.
Gunn, J. S., Lim, K. B., Krueger, J., Kim, K., Guo, L., Hackett, M. & Miller, S. I. (1998). *Mol. Microbiol.* **27**, 1171–1182.
Hashimoto, M., Kirikae, F., Dohi, T., Adachi, S., Kusumoto, S., Suda, Y., Fujita, T., Naoki, H. & Kirikae, T. (2002). *Eur. J. Biochem.* **269**, 3715–3721.
Jagtap, P. K., Verma, S. K., Vithani, N., Bais, V. S. & Prakash, B. (2013). *J. Mol. Biol.* **425**, 1745–1759.
Jeyakanthan, J., Rangarajan, S., Mridula, P., Kanaujia, S. P., Shiro, Y., Kuramitsu, S., Yokoyama, S. & Sekar, K. (2008). *Acta Cryst.* D**64**, 1012–1019.
Joo, S. H., Chung, H. S., Raetz, C. R. & Garrett, T. A. (2012). *Biochemistry*, **51**, 4322–4330.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Katsandri, A., Papaparaskevas, J., Pantazatou, A., Petrikkos, G. L., Thomopoulos, G., Houhoula, D. P. & Avlamis, A. (2006). *J. Clin. Microbiol.* **44**, 3465–3467.
Kehoe, L. E., Snidwongse, J., Courvalin, P., Rafferty, J. B. & Murray, I. A. (2003). *J. Biol. Chem.* **278**, 29963–29970.
Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G. & Rees, D. C. (1996). *EMBO J.* **15**, 2323–2330.
Kostrewa, D., D'Arcy, A., Takacs, B. & Kamber, M. (2001). *J. Mol. Biol.* **305**, 279–289.
Lee, B. I. & Suh, S. W. (2003). *Proteins*, **53**, 772–774.
Lee, H. J., Rakić, B., Gilbert, M., Wakarchuk, W. W., Withers, S. G. & Strynadka, N. C. J. (2009). *J. Biol. Chem.* **284**, 24501–24511.
Lo Leggio, L., Dal Degan, F., Poulsen, P., Andersen, S. M. & Larsen, S. (2003). *Biochemistry*, **42**, 5225–5235.
Masoudi, A., Raetz, C. R., Zhou, P. & Pemble, C. W. IV (2014). *Nature (London)*, **505**, 422–426.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J. (2007). *Acta Cryst.* D**63**, 32–41.
Mochalkin, I., Lightle, S., Narasimhan, L., Bornemeier, D., Melnick, M., VanderRoest, S. & McDowell, L. (2008). *Protein Sci.* **17**, 577–582.
Mochalkin, I., Lightle, S., Zhu, Y., Ohren, J. F., Spessard, C., Chirgadze, N. Y., Banotai, C., Melnick, M. & McDowell, L. (2007). *Protein Sci.* **16**, 2657–2666.
Olsen, L. R., Huang, B., Vetting, M. W. & Roderick, S. L. (2004). *Biochemistry*, **43**, 6013–6019.
Olsen, L. R. & Roderick, S. L. (2001). *Biochemistry*, **40**, 1913–1921.
Olsen, L. R., Vetting, M. W. & Roderick, S. L. (2007). *Protein Sci.* **16**, 1230–1235.
Patrick, S., Blakely, G. W., Houston, S., Moore, J., Abratt, V. R., Bertalan, M., Cerdeño-Tárraga, A. M., Quail, M. A., Corton, N., Corton, C., Bignell, A., Barron, A., Clark, L., Bentley, S. D. & Parkhill, J. (2010). *Microbiology*, **156**, 3255–3269.
Pumbwe, L., Skilbeck, C. A. & Wexler, H. M. (2006). *Anaerobe*, **12**, 211–220.
Raetz, C. R., Reynolds, C. M., Trent, M. S. & Bishop, R. E. (2007). *Annu. Rev. Biochem.* **76**, 295–329.
Raetz, C. R. & Roderick, S. L. (1995). *Science*, **270**, 997–1000.
Rangarajan, E. S., Ruane, K. M., Sulea, T., Watson, D. C., Proteau, A., Leclerc, S., Cygler, M., Matte, A. & Young, N. M. (2008). *Biochemistry*, **47**, 1827–1836.
Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
Robins, L. I., Williams, A. H. & Raetz, C. R. (2009). *Biochemistry*, **48**, 6191–6201.
Saitoh, S., Akashi, S., Yamada, T., Tanimura, N., Kobayashi, M., Konno, K., Matsumoto, F., Fukase, K., Kusumoto, S., Nagai, Y., Kusumoto, Y., Kosugi, A. & Miyake, K. (2004). *Int. Immunol.* **16**, 961–969.
Schulz, E. C., Bergfeld, A. K., Ficner, R. & Mühlenhoff, M. (2011). *PLoS One*, **6**, e17403.
Sugantino, M. & Roderick, S. L. (2002). *Biochemistry*, **41**, 2209–2216.
Sulzenbacher, G., Gal, L., Peneff, C., Fassy, F. & Bourne, Y. (2001). *J. Biol. Chem.* **276**, 11844–11851.
Sweet, C. R., Lin, S., Cotter, R. J. & Raetz, C. R. (2001). *J. Biol. Chem.* **276**, 19565–19574.
Sydenham, T. V., Soki, J., Hasman, H., Wang, M. & Justesen, U. S. (2015). *Anaerobe*, **31**, 59–64.
Ulaganathan, V., Buetow, L. & Hunter, W. N. (2007). *J. Mol. Biol.* **369**, 305–312.
Wang, X.-G., Olsen, L. R. & Roderick, S. L. (2002). *Structure*, **10**, 581–588.
Wareham, D. W., Wilks, M., Ahmed, D., Brazier, J. S. & Millar, M. (2005). *Clin. Infect. Dis.* **40**, e67–e68.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yeh, T.-Y., Kowalska, A. K., Scipioni, B. R., Cheong, F. K. Y., Zheng, M., Derewenda, U., Derewenda, Z. S. & Schroer, T. A. (2013). *EMBO J.* **32**, 1023–1035.