

Structure of the essential *Haemophilus influenzae* UDP-diacetylglucosamine pyrophosphohydrolase LpxH in lipid A biosynthesis

Jae Cho, Chul-Jin Lee[†], Jinshi Zhao, Hayley E. Young and Pei Zhou*

In most Gram-negative pathogens, the hydrolysis of UDP-2,3-diacetylglucosamine to generate lipid X in lipid A biosynthesis is catalysed by the membrane-associated enzyme LpxH. We report the crystal structure of LpxH in complex with its product, lipid X, unveiling a unique insertion lid above the conserved architecture of calcineurin-like phosphoesterases. This structure reveals elaborate interactions surrounding lipid X and provides molecular insights into the substrate selectivity, catalysis and inhibition of LpxH.

The outer leaflet of the outer membrane of Gram-negative bacteria is enriched with lipopolysaccharide (LPS), which is anchored into the membrane through a hexa-acylated disaccharide lipid called 'lipid A'^{1,2}. Lipid A serves as a protective barrier to shield bacterial cells from the damage of antibiotics and detergents³ and is the active component of endotoxin that elicits potent host responses to bacterial infection.

The biosynthesis of lipid A in *Escherichia coli* involves nine constitutive enzymes of the Raetz pathway (Supplementary Fig. 1)¹. The first membrane-associated reaction of the pathway, the cleavage of the pyrophosphate group of UDP-2,3-diacetylglucosamine (UDP-DAGn) to form lipid X, is carried out by LpxH in β - and γ -proteobacteria⁴, by LpxI in α -proteobacteria⁵ and by LpxG in Chlamydiae (Fig. 1a)⁶. LpxH and LpxG are unique members of the metal-dependent calcineurin-like phosphoesterase (CLP) family, although they share limited sequence similarity⁶. LpxI, on the other hand, is structurally and mechanistically unrelated to LpxH and LpxG (refs 7, 8). LpxI uses water to attack the β -phosphate^{5,7}, whereas LpxH and LpxG use water to attack the α -phosphate on the common UDP-DAGn substrate^{4,6}. Among these three enzymes, LpxH is most widespread, functioning in ~70% of Gram-negative bacteria and the vast majority of Gram-negative pathogens.

As constitutive lipid A biosynthesis is required for bacterial viability and pathogenesis, essential lipid A enzymes, including LpxH (ref. 9), are attractive targets for the development of novel antibiotics against Gram-negative pathogens. Recently, an LpxH-targeting antibiotic has been reported¹⁰, opening the arena for targeted development of LpxH inhibitors. Despite the molecular cloning and biochemical characterization of LpxH over a decade ago^{4,9}, its structure has not yet been elucidated, leaving fundamental questions about this membrane-associated, essential lipid pyrophosphatase unanswered.

To provide molecular insights into the catalysis and inhibition of LpxH, we crystallized *Haemophilus influenzae* LpxH (HiLpxH) in complex with its product, lipid X (Supplementary Table 1). The structure of LpxH (Fig. 1b,c) reveals a conserved core architecture shared with the CLP enzymes, such as PP-1 (ref. 11). However, there are major distinctions. The core domain of LpxH consists of two tightly packed central β -sheets, instead of three β -sheets in

PP-1. These two prominent β -sheets consist of a total of eleven β -strands arranged in a mixed parallel and anti-parallel fashion. Starting with the N-terminus in the middle of β -sheet 1, the backbone forms a β 1- α 1- β 2- α 2- β 3- α 3- β 4 scaffold, which then crosses over to β -sheet 2 to form two antiparallel β -strands (β 5 and β 6). Following this, there is an elongated, triangle-shaped insertion domain formed by two long helices (α 1' and α 3') connected by a short helix (α 2'). This insertion (dubbed the 'lid' domain in Fig. 1b,c) functions similarly to the lid of the pitfall trap of a pitcher plant to engulf lipid X. Following the insertion domain, there is another helix (α 4) that packs against β -sheet 2. The α 4 helix is followed by three mixed orientation β -strands (β 7, β 8 and β 9) before the backbone crosses back to β -sheet 1 to form the last two antiparallel β -strands (β 10 and β 11). The active site of LpxH is located between the insertion lid and the core domain (Fig. 1b,c).

Within the catalytic site of LpxH, electron densities of two metal ions can be identified. As LpxH is most active in the presence of Mn^{2+} , and because our previous electron paramagnetic resonance (EPR) study has revealed the presence of a di-manganese cluster in LpxH (ref. 12), these two densities were interpreted as manganese ions. Of the two manganese ions, the first ion (Mn1) is coordinated by side chains of D42, N80, H115 and H196, while the second ion (Mn2) is coordinated by side chains of D9, H11 and D42, with D42 bridging the di-manganese cluster (Fig. 1d). The short distances between these residues and the di-manganese cluster are consistent with their suggested role within the signature metal chelating motifs (D9 and H11 of the DXH motif, D42 of the GDXXD motif, N80 of the GNRD motif, H115 of the HXD motif, and H196 of the GHXH motif) of the CLP enzymes (Supplementary Fig. 2). Substitutions of these residues by alanine (D9A, H11A, D42A, H115A and H196A) dramatically reduced the specific activity of LpxH by 5,000- to 200,000-fold¹².

As LpxH was crystallized in the absence of any added lipid A intermediates, we were surprised to observe lipid X, the product of LpxH, sandwiched between the core domain and the insertion lid of LpxH, with its glucosamine-1-phosphate headgroup sitting immediately above the di-manganese cluster within the active site (Fig. 1e). To verify that the lipid molecule co-crystallized with LpxH was indeed lipid X, we subjected the crystallization sample to single-phase Bligh–Dyer extraction¹³ followed by reversed-phase liquid chromatography mass spectrometry (RPLC-MS) analysis. Consistent with the predicted mass for *E. coli* lipid X ($[M-H]^-$ m/z of 710.42), a singly charged species with $[M-H]^-$ at m/z 710.44 was observed as a discrete peak eluting from the column (Supplementary Fig. 3a,b). Tandem mass spectrometry (MS/MS) analysis also yielded components consistent with

Department of Biochemistry, Duke University Medical Center, Research Drive, Durham, North Carolina 27710, USA. [†]Present address: Cell Biology and Metabolism Program, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA. *e-mail: peizhou@biochem.duke.edu

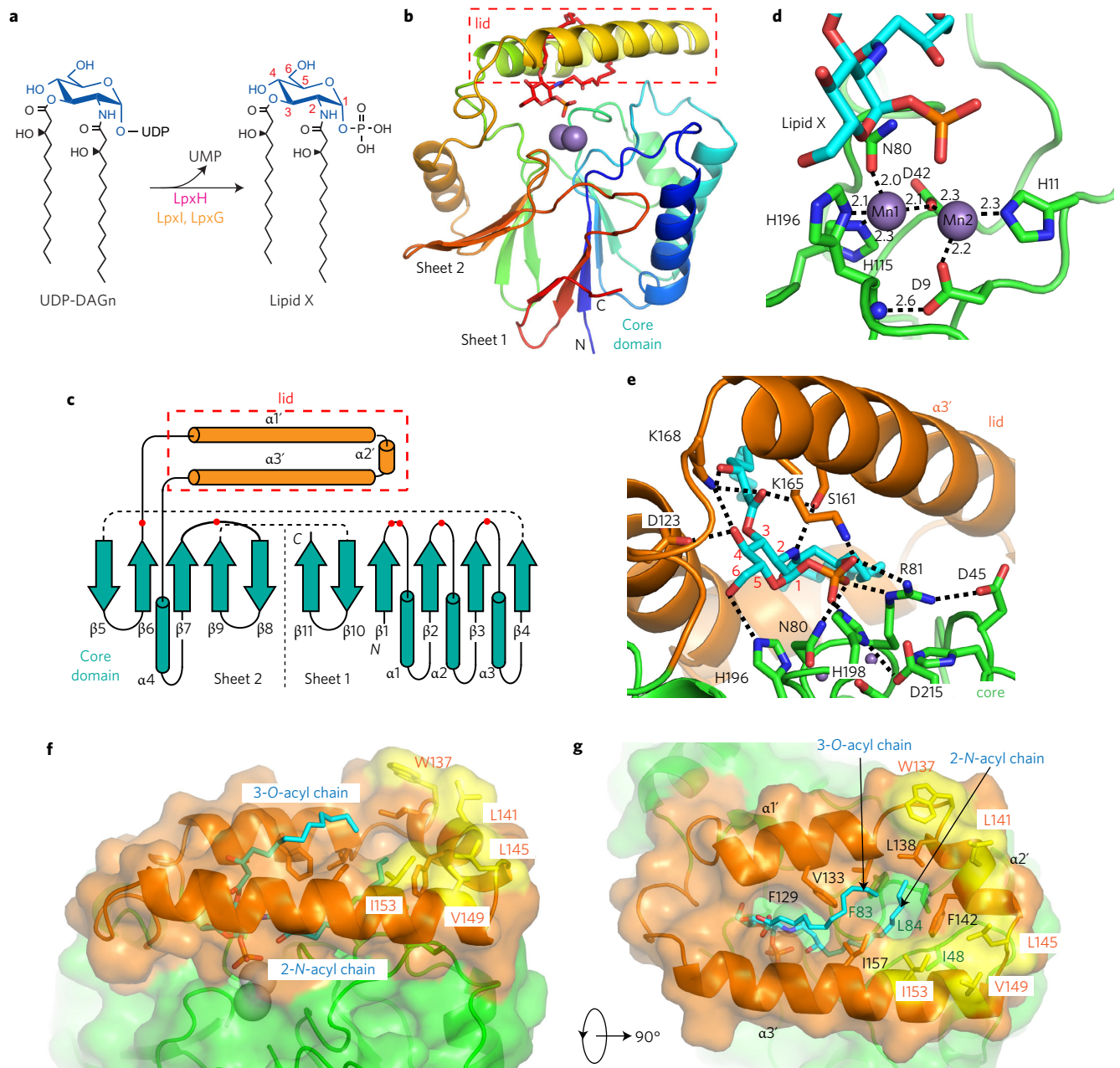


Figure 1 | Structure of the LpxH-lipid X complex. **a**, Cleavage of the pyrophosphate group of UDP-2,3-diacylglucosamine (UDP-DAGn) to form lipid X and UMP by LpxH (magenta) and its functional orthologues LpxI and LpxG (orange). **b**, Ribbon representation of LpxH, with blue to red colours corresponding to the N- to C-terminus. Lipid X is shown as a stick model, and the di-manganese cluster is shown as spheres. The insertion 'lid' is indicated by the dashed rectangle. **c**, Topology diagram of LpxH. The core domain is denoted in green, and the insertion lid is in orange. Locations of metal-binding residues are denoted as red dots. **d**, Coordination of the di-manganese cluster in LpxH. Manganese ions are shown as spheres. Side chains of manganese chelating residues are shown as stick models, and their distances to the manganese ions are labelled in Å. Lipid X is shown as a stick model to illustrate its location in the active site. **e**, Recognition of the glucosamine-1-phosphate headgroup of lipid X by LpxH. LpxH residues from the core domain are in green, and those from the insertion lid are in orange. Side chains of residues that interact with lipid X through polar interactions are shown as stick models. Lipid X, in cyan, is also shown as a stick model. Positions of the glucosamine ring are numbered in red. **f, g**, Side (**f**) and top (**g**) views of the LpxH-lipid X complex in the surface representation, illustrating the engulfment of lipid X by LpxH within the hydrophobic chamber. The backbone of LpxH is depicted as the ribbon diagram. Lipid X and its acyl-chain-interacting residues are denoted as stick models. Lipid X is in cyan, and the core domain and insertion lid of LpxH are in green and orange, respectively. Clustered hydrophobic surface residues in the insertion lid not involved in lipid X interaction are shown as stick models and are in yellow, revealing a potential surface area for membrane association. The structure is a representative model derived from the best-quality crystal.

secondary fragmentation of lipid X (Supplementary Fig. 3c), confirming co-purification of lipid X with the overexpressed LpxH and suggesting a high flux of lipid X and UDP-DAGn, despite their low steady-state levels.

The electron density of lipid X is well defined for the glucosamine-1-phosphate headgroup, the entire 2-*N*-linked β -hydroxymyristoyl chain, and the first half of the 3-*O*-linked β -hydroxymyristoyl chain (Supplementary Fig. 4). The electron density for the remaining part

of the 3-*O*-linked acyl chain, albeit weak, is visible at lower contour levels. The glucosamine-1-phosphate adopts a chair conformation, with the 1-phosphate group in the axial configuration and the remaining substituents of the hexose ring in the equatorial configuration. An extensive array of LpxH residues can be identified that interact with every polar group of the glucosamine-1-phosphate headgroup. The 1-phosphate group in particular is very well recognized; it forms two salt bridges with R81 of the GNRD motif, which itself is anchored by D45 from the GDXXD motif of a neighbouring loop. Mutation of R81A decreased the specific activity of LpxH by 7,000-fold compared to the wild-type enzyme¹², consistent with its critical contribution to the recognition of the substrate. The 1-phosphate group is additionally recognized through hydrogen bonds with the manganese-chelating residue N80 and H198 of the GHXH motif, the latter of which is anchored by a hydrogen bond with D215. In addition to interacting with residues of the core domain, the 1-phosphate group of lipid X also forms a salt bridge with K165 from the α 3' helix of the insertion lid and is engaged in a hydrogen bond with the β -hydroxyl group of its 2-*N*-linked β -hydroxymyristoyl chain. The nitrogen atom of the 2-*N*-linked acyl chain is fixed in space by a hydrogen bond with S161 of the α 3' helix from the insertion lid. Recognition of polar groups at the 3- and 4-positions is achieved entirely within the insertion lid. K168 at the end of the α 3' helix forms three hydrogen bonds with the β -hydroxyl group and the carbonyl oxygen of the 3-*O*-linked β -hydroxymyristoyl chain as well as the 4-hydroxyl group of the glucosamine ring. The carbonyl group of the 3-*O*-linked acyl chain additionally forms a hydrogen bond with S161, whereas the 4-hydroxyl group is further recognized by D123 of the insertion lid, which also serves as the cap of the insertion helix α 1'. The last polar group of the glucosamine ring, the 6-hydroxyl group, forms a hydrogen bond with the manganese-chelating residue H196. With the exception of K165, all of the residues directly interacting with the 1-phosphate group are strictly conserved (Supplementary Fig. 2). As the level of UDP-DAGn only represents 0.005% of the total bacterial phospholipids¹⁴, such an elaborate array of specific interactions surrounding the glucosamine-1-phosphate headgroup, a molecular entity shared by lipid X and UDP-DAGn, bestows LpxH with the extraordinary ability to scavenge its substrate UDP-DAGn from an overwhelming amount of competing phospholipids in the bacterial inner membrane.

The two acyl chains of lipid X are recognized differently. The 2-*N*-linked β -hydroxymyristoyl chain is nearly completely buried within a hydrophobic chamber at the interspace between the core domain and the insertion lid (Fig. 1f). Lining the hydrophobic chamber are conserved residues from the core domain (I48, F83 and L84) and from the insertion lid (F129 and V133 from the α 1' helix, L138 and F142 from the α 2' helix, and I157 from the α 3' helix) that form extensive interactions with the 2-*N*-linked β -hydroxymyristoyl chain (Fig. 1g). The 3-*O*-linked β -hydroxymyristoyl chain, however, does not interact with the core domain. It rises rapidly through an open area above the active site, with its remaining acyl chain gliding over the surface of the insertion lid and forming hydrophobic interactions with F129, V133 of the α 1' helix, and I157 of the α 3' helix. A significant portion of the 3-*O*-linked acyl chain is surface exposed and is probably surrounded by detergent molecules. Similarly, the insertion lid is decorated with a large number of surface-exposed hydrophobic residues (W137, L141, L145, V149 and I153) that do not interact with lipid X. These residues probably facilitate the association of the insertion lid of LpxH with detergent micelles *in vitro* and with the bacterial inner membrane *in vivo*.

Recognition of lipid X by LpxH is distinct from its functional orthologue, LpxI, which is structurally unrelated to LpxH (ref. 7). LpxI consists of an N-terminal lipid-binding domain with a central β -sheet decorated with helices on both sides and a C-terminal catalytic domain⁷. Lipid X recognition by LpxI is achieved by

encapsulating both acyl chains of lipid X in the lipid-binding domain, but with limited interactions of the glucosamine-1-phosphate group. Binding of UDP-DAGn by LpxI induces a conformational switch that brings the catalytic domain close to the lipid-binding domain for specific recognition of the substrate and catalysis. Whether binding of UDP-DAGn by LpxH similarly induces a conformational change remains to be established.

Incubation of UDP-DAGn with LpxH in H₂¹⁸O results in cleavage of the pyrophosphate linkage and incorporation of a solvent molecule into UMP (Supplementary Fig. 5)⁴. HiLpxH is most active at slightly alkaline pH values and exhibits a sharp decrease in its catalytic activity at acidic pH (ref. 12). Fitting the data to a single-lobed pH curve reveals a general base with a pK_a of ~6.6, consistent with the presence of a catalytic histidine residue in the active site¹². After excluding active-site histidine residues involved in manganese coordination, we propose that H198, which is located at the far side of Mn2, serves as the catalytic general base to mediate a metal ion-activated water molecule for in-line attack of the α -phosphate group of UDP-DAGn to incorporate ¹⁸O into the UMP moiety (Supplementary Fig. 5).

Lipid A biosynthesis is required for the viability of virtually all Gram-negative bacteria. Nearly all of the inhibitor development efforts thus far have focused on cytosolic lipid A enzymes, such as LpxC (refs 15–17). However, targeting late-stage lipid A enzymes such as LpxH may be uniquely advantageous, as inhibition of LpxH results in accumulation of toxic lipid intermediates in the bacterial inner membrane^{9,10}, providing an independent mechanism of bacterial killing.

Recently, a small-molecule inhibitor of *E. coli* LpxH has been reported, and four resistance mutations to the inhibitor (G48D, L84R, F141L and R149H) in *E. coli* LpxH have been isolated¹⁰. By mapping these residues onto the structure of HiLpxH, we show that all of the corresponding residues in this orthologue, G49, I85, F142 and R150 (Supplementary Fig. 2), are clustered together around the terminus of the 2-*N*-linked β -hydroxymyristoyl chain of lipid X (Supplementary Fig. 6). Such an observation suggests that the reported LpxH inhibitor probably targets the hydrophobic chamber between the insertion lid and the core domain and competes with the 2-*N*-linked acyl chain of the substrate UDP-DAGn for LpxH inhibition.

The prevalence of antibiotic resistance in clinically important Gram-negative pathogens demands novel therapeutic options that target previously unexploited enzymes such as LpxH in the lipid A biosynthetic pathway. The structural elucidation of LpxH in complex with its product lipid X provides a first glimpse of the mechanism and inhibition of LpxH catalysis and sets the stage for future lead optimization campaigns based on structural insights.

Methods

Sample preparation and protein crystallization. The expression construct for HiLpxH has been described previously¹², and contains the full-length enzyme followed by a tobacco etch virus (TEV) site and a C-terminal His₁₀-tag. BL21(DE3) cells transformed with the HiLpxH expression vector were grown in the LB medium at 37 °C until the optical density at 600 nm reached 0.5. The cells were induced with 1 mM isopropyl β -D-thiogalactopyranoside (IPTG) at 30 °C for 5 h, collected by centrifugation, and then kept frozen at –80 °C until purification. The cell pellet from each litre of culture was resuspended in 30 ml lysis buffer containing 200 mM NaCl and 20 mM HEPES (pH 8.0). Resuspended cells were lysed with a French Press. Cell debris was removed by low-speed centrifugation at 10,000g for 40 min. HiLpxH was then extracted with *n*-dodecyl- β -D-maltopyranoside (1.5% wt/vol) for 2 h at 4 °C and subjected to high-speed centrifugation at 100,000g for 1 h. After centrifugation, the supernatant was purified by the HisPur Ni-NTA gravity-flow chromatography (ThermoFisher Scientific) at 4 °C. Approximately 2 ml of the Ni-NTA resin was used per litre of cell culture. After sample loading, the column was washed with 15–20 column volumes of buffer containing 200 mM NaCl and 20 mM HEPES (pH 8.0) and 1.8 mM *n*-decyl- β -D-thiomaltopyranoside (DTM) (2 \times critical micelle concentration, CMC). HiLpxH was eluted using ten column volumes of buffer that also contained 250 mM imidazole. For the crystallization sample, the C-terminal His₁₀ tag was not removed.

The eluted fractions from the Ni-NTA chromatography were combined, concentrated and further purified using size-exclusion chromatography (HiPrep 26/60

Sephacryl S-200 HR, GE Healthcare) at a flow rate of 1 ml min⁻¹ in the presence of 200 mM NaCl, 20 mM HEPES pH 8.0 and 1.08 mM DTM (1.2× CMC). Three fractions of 2 ml each with the highest concentrations were pooled and concentrated to 10 mg ml⁻¹ with an estimated final DTM concentration of 3.6–4.5 mM (4–5× CMC). The selenomethionine labelled LpxH sample was prepared using SelenoMet Medium (Molecular Dimensions).

LpxH crystals were obtained by mixing freshly prepared protein solution with mother liquor and additives at a drop ratio of 1:1. The optimal crystallization condition contained 5 mg ml⁻¹ LpxH in 100 mM NaCl, 10 mM HEPES, 35 mM sodium acetate pH 4.1, 2.2% polyethylene glycol (PEG) 4000, 0.3 mM sulfonylethyl piperazine, 0.7% DMSO, 22% glycerol and ~3–4 mM DTM. Crystals started to appear after 7–10 days at 15 °C. Suitable-sized crystals were cryoprotected using the corresponding crystallization buffer containing 40% glycerol before flash-freezing for data collection.

Data collection and structural analysis. Diffraction data were collected at the Northeastern Collaborative Access Team (NECAT) beamlines (24-ID-C and 24-ID-E). The selenium single anomalous dispersion (Se-SAD) data set was obtained by merging diffraction data from two crystals collected at the wavelength of the selenium absorption peak (0.97912 Å). The native data set was collected at a wavelength of 1.0000 Å. The diffraction data were processed using XDS and AIMLESS modules in the CCP4 suite and were truncated and scaled with the University of California, Los Angeles, anisotropy diffraction server^{18,19}. The identification of heavy-atom sites, selenomethionine single-wavelength anomalous dispersion (Se-SAD) phasing, and initial model building were done using AutoSol and AutoBuild modules within the PHENIX suite²⁰. All four predicted Se heavy atom sites were unambiguously identified. The resulting structure was used as the search model for molecular replacement against the native data set. Restraints of lipid X were generated using the eLBOW module within the PHENIX suite²⁰ and edited manually. Iterative model building and refinement was carried out using COOT²¹ and PHENIX²⁰. The 2mFo-DFc omit map was generated using PHENIX²⁰.

Lipid extraction and mass spectrometry. LpxH-bound lipids were extracted using the single-phase Blish–Dyer method¹³ by adding chloroform and methanol to the aqueous solution at a ratio of 1:2:0.8 (vol/vol/vol). The supernatant of the mixture was dried and re-suspended in methanol and submitted to reversed-phase liquid chromatography electron spray ionization mass spectrometry (RPLC-ESI/MS) analysis using a Shimadzu LC system (comprising a solvent degasser, two LC-10A pumps and an SCL-10A system controller) coupled to a TripleTOF5600 mass spectrometer (Sciex). The LC was operated at a flow rate of 200 µl min⁻¹ with the following linear gradients: 100% of mobile phase A was held isocratically for 2 min and then linearly increased to 100% mobile phase B over 3 min and held at 100% B for 2.5 min. At the end, the gradient was returned to 100% mobile phase A and held for 2 min. Mobile phase A consisted of methanol/acetonitrile/aqueous 1 mM ammonium acetate (60/20/20, vol/vol/vol). Mobile phase B consisted of 100% ethanol containing 1 mM ammonium acetate. A Zorbax SB-C8 reversed-phase column (5 µm, 2.1 × 50 mm) was obtained from Agilent. The LC eluent was introduced into the ESI source of the mass spectrometer. The instrument settings for the negative ion ESI/MS and MS/MS analysis of lipid species were as follows: ion spray voltage (IS) = -4,500 V; curtain gas (CUR) = 20 p.s.i.; ion source gas 1 (GS1) = 20 p.s.i.; de-clustering potential (DP) = -55 V; focusing potential (FP) = -150 V. Data acquisition and analysis were performed using the Analyst TF1.5 software (Sciex).

Accession codes. The coordinate of the LpxH–lipid X complex has been deposited at the Protein Data Bank (PDB) under accession code 5K8K.

Received 9 June 2016; accepted 22 July 2016;
published 15 August 2016

References

1. Raetz, C. R. H. & Whitfield, C. *Annu. Rev. Biochem.* **71**, 635–700 (2002).
2. Whitfield, C. & Trent, M. S. *Annu. Rev. Biochem.* **83**, 99–128 (2014).
3. Needham, B. D. & Trent, M. S. *Nat. Rev. Microbiol.* **11**, 467–481 (2013).
4. Babinski, K. J., Ribeiro, A. A. & Raetz, C. R. *J. Biol. Chem.* **277**, 25937–25946 (2002).
5. Metzger, L. E. t. & Raetz, C. R. *Biochemistry* **49**, 6715–6726 (2010).
6. Young, H. E. *et al. mBio* **7**, e00090-16 (2016).
7. Metzger, L. E. t., Lee, J. K., Finer-Moore, J. S., Raetz, C. R. & Stroud, R. M. *Nat. Struct. Mol. Biol.* **19**, 1132–1138 (2012).
8. Metzger, L. E. t. & Raetz, C. R. *Biochemistry* **48**, 11559–11571 (2009).
9. Babinski, K. J., Kanjilal, S. J. & Raetz, C. R. *J. Biol. Chem.* **277**, 25947–25956 (2002).
10. Nayar, A. S. *et al. J. Bacteriol.* **197**, 1726–1734 (2015).
11. Goldberg, J. *et al. Nature* **376**, 745–753 (1995).
12. Barb, A. W., Donohue, M. P., Smirnova, T. I., Smirnov, A. I. & Zhou, P. *J. Biol. Chem.* **288**, 26987–27001 (2013).
13. Blish, E. G. & Dyer, W. J. *Can. J. Biochem. Physiol.* **37**, 911–917 (1959).
14. Bulawa, C. E. & Raetz, C. R. *J. Biol. Chem.* **259**, 4846–4851 (1984).
15. Barb, A. W. & Zhou, P. *Curr. Pharm. Biotechnol.* **9**, 9–15 (2008).
16. Lee, C. J. *et al. Nat. Commun.* **7**, 10638 (2016).
17. Erwin, A. L. *Cold Spring Harb. Perspect. Med.* **6**, 1–14 (2016).
18. Collaborative Computational Project, N. *Acta Crystallogr. D* **50**, 760–763 (1994).
19. Strong, M. *et al. Proc. Natl Acad. Sci. USA* **103**, 8060–8065 (2006).
20. Adams, P. D. *et al. Acta Crystallogr. D* **66**, 213–221 (2010).
21. Emsley, P. & Cowtan, K. *Acta Crystallogr. D* **60**, 2126–2132 (2004).

Acknowledgements

This work was supported by grants from the National Institute of General Medical Sciences (NIGMS) (to P.Z., GM115355 and GM51310). Diffraction data were collected at the Northeastern Collaborative Access Team beamlines (24-ID-C and 24-ID-E) funded by NIGMS (GM103403). The Pilatus 6M detector on the 24-ID-C beam line is funded by a NIH-ORIP HEI grant (RR029205). This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory (contract no. DE-AC02-06CH11357). The authors thank Ziqiang Guan for assistance with mass spectrometry analysis and Brian Coggins for critical reading of this manuscript.

Author contributions

P.Z. conceptualized the study. J.C. and H.E.Y. purified and crystallized LpxH. J.C., C.-J.L. and P.Z. determined the structure. J.Z. carried out mass spectrometry analysis. P.Z. wrote the manuscript with critical inputs from all authors.

Additional information

Supplementary information is available online. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.Z.

Competing interests

The authors declare no competing financial interests.