

# Prediction of the Closed Conformation and Insights into the Mechanism of the Membrane Enzyme LpxR

Graham M. Smeddle,<sup>1</sup> Hannah E. Bruce Macdonald,<sup>1</sup> Jonathan W. Essex,<sup>1</sup> and Syma Khalid<sup>1,\*</sup>

<sup>1</sup>Department of Chemistry, University of Southampton, Highfield, Southampton, United Kingdom

**ABSTRACT** Covalent modification of outer membrane lipids of Gram-negative bacteria can impact the ability of the bacterium to develop resistance to antibiotics as well as modulating the immune response of the host. The enzyme LpxR from *Salmonella typhimurium* is known to deacylate lipopolysaccharide molecules of the outer membrane; however, the mechanism of action is unknown. Here, we employ molecular dynamics and Monte Carlo simulations to study the conformational dynamics and substrate binding of LpxR in representative outer membrane models as well as detergent micelles. We examine the roles of conserved residues and provide an understanding of how LpxR binds its substrate. Our simulations predict that the catalytic H122 must be  $N\epsilon$ -protonated for a single water molecule to occupy the space between it and the scissile bond, with a free binding energy of  $-8.5 \text{ kcal mol}^{-1}$ . Furthermore, simulations of the protein within a micelle enable us to predict the structure of the putative “closed” protein. Our results highlight the need for including dynamics, a representative environment, and the consideration of multiple tautomeric and rotameric states of key residues in mechanistic studies; static structures alone do not tell the full story.

## INTRODUCTION

The outer membrane (OM) of Gram-negative bacteria is a formidable barrier to the permeation of molecular species seeking to enter the bacterial cell (1). It is only selectively permeable, enabling molecules essential for the survival of the bacterium such as nutrients to get across the membrane but excluding those that are harmful, such as antibacterial agents. The chemical natures of the lipids of the membrane are thought to play a key role in achieving this selective permeability. The membrane is a lipid bilayer, with the leaflet facing the external environment, known as the outer leaflet, almost exclusively composed of lipopolysaccharide (LPS) molecules, whereas the inner leaflet contains a mixture of zwitterionic and anionic lipids. To date, LPS is perhaps the most chemically complex natural lipid known. The internal membrane component is known as lipid A, which is covalently linked to a polymer of sugars, some of which are phosphorylated. The chemical structure of the LPS molecules varies across bacterial species and sometimes even within one species; for example, in *Escherichia coli* and *Salmonella typhimurium*, the lipid A component has six acyl tails, whereas in *Pseudomonas aeruginosa*, it only has five tails (2). In *E. coli* and *S. typhimurium*, these are laurate and myr-

istate tails in a ratio of 5:1. LPS may be referred to as smooth, rough LPS (Ra LPS), or deep-rough LPS (Re LPS), referring to the number of sugars attached to the lipid A segment. Smooth LPS includes the O antigen and the full core segment, whereas rough LPS contains just the six core sugars and deep rough only two keto-deoxyoctulosonate (Kdo) sugars.

A number of pathogenic bacteria have been shown to synthesize LPS molecules with modified lipid A. Some of these modifications to lipid A can facilitate the development of resistance to drugs; for example, addition of the 4-amino-4-deoxy-L-arabinopyranose group, which is positively charged at pH 7, neutralizes the negative charge of a lipid A phosphate group in *E. coli*, *S. typhimurium*, and *P. aeruginosa*, which reduces the susceptibility of these bacteria to antimicrobial peptides. A number of different bacterial enzymes that catalyze the covalent modification of LPS have been identified. Three of these enzymes, all of which modify lipid A tails, PagP, PagL, and LpxR, are embedded in the OM (3–6). LpxR catalyzes the removal of two acyl chains of lipid A, in the form of 3-(tetradecanoyloxy)tetradecanoic acid. The X-ray structure of the protein has been resolved to a 1.9-Å resolution, and although the structure of the lipid substrate bound to the protein remains elusive, mutational studies have identified residues that are essential for catalytic activity thus providing clues to the location of the active site (5). The X-ray structure of LpxR was obtained

Submitted April 26, 2018, and accepted for publication September 6, 2018.

\*Correspondence: s.khalid@soton.ac.uk

Editor: Michael Brown.

<https://doi.org/10.1016/j.bpj.2018.09.002>

© 2018 The Authors.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



by cocrystallizing the protein with  $\text{Zn}^{2+}$ , and although there is some evidence of  $\text{Zn}^{2+}$  binding to LpxR in the X-ray structure, the density is low and likely indicative of partial occupation. Interestingly, it has been shown that whereas  $\text{Ca}^{2+}$  is essential for LpxR activity, some other divalent cations, such as  $\text{Sr}^{2+}$  and  $\text{Cd}^{2+}$ , but not  $\text{Zn}^{2+}$  can replace  $\text{Ca}^{2+}$  without a total loss of catalytic activity (6). Thus, although we know  $\text{Ca}^{2+}$  is essential, it is unclear from the X-ray structure precisely where it is bound for catalytic activity to occur. Based on their structures and mechanistic predictions, the mechanism of deacylation in LpxR is speculated to be similar to that displayed by a number of phospholipase A2 enzymes (7,8). The proposed mechanism of deacylation is thought to occur via a histidine, with H122 acting as a base by activating a nearby water molecule, so the latter can react with the substrate lipid molecule to hydrolyze the ester bond. In the X-ray structure, there is a water molecule resolved near H122, which may well be the crucial mechanistic water. Although the structure of LpxR, the location of the water, and the docked lipid substrate provide a plausible static model for the reaction mechanism of the enzyme, the dynamic stability of the model protein-substrate complex, the location of any additional key water and ion binding sites (especially given the very low electron density for  $\text{Zn}^{2+}$  in the X-ray structure), and the effect of the complex on the local membrane are still unexplored, and thus the mechanism of action has not been proven.

To this end, here we present a molecular simulation study in which we use a combination of molecular dynamics and Monte Carlo simulations to investigate the molecular interactions between LpxR and the Re LPS substrate, with the specific aim of uncovering mechanistic insights into the process of deacylation.

Details of the calculations are presented at the end of the article, but for ease of reading, we present the tables summarizing the molecular dynamics (Table 1) and Monte Carlo (Table 2) simulations below.

## METHODS

### Simulation systems

The coordinates of the LpxR X-ray structure (Protein Data Bank (PDB): 3FID), with Re LPS,  $\text{Ca}^{2+}$  ion, and a single water molecule modeled as described by Rutten et al. (5), were obtained from Piet Gros. The  $\text{Ca}^{2+}$  ion is positioned such that it replaces the zinc ion that was resolved in the X-ray structure, given that the former is known to be essential for catalytic activity. These initial coordinates were manipulated to set up a series of simulations that are summarized in Table 1. We truncated the LPS molecule at the Re LPS level, because in this study, we are focused on the lipid A binding site of the protein. This is thought to be located at the lipid headgroup interface; therefore, omission of other sugars and the O antigen is a reasonable approximation; additionally, this allows for longer simulations given the smaller system size compared to the inclusion of the full LPS.

The protein,  $\text{Ca}^{2+}$  ion, and single water molecule modeled by Rutten et al. (5), were retained in all “Model” simulations. The Re LPS molecule was removed from the putative binding site in “apo” simulations, and the

$\text{Ca}^{2+}$  ion and single water molecule were removed in “Unbiased” simulations. In “Bond Break” systems, the scissile covalent bond of Re LPS was removed, and the termini were protonated, thereby removing two acyl chains from the molecule. The two resultant molecules were parameterized in a manner consistent with the Groningen molecular simulation (GROMOS) 54a7 force field (9). Both molecules were retained in the simulations. Once the LPS molecule had been converted into two smaller molecules, these systems were then subjected to an additional 500 ns of production simulation. Table 1 summarizes the four mutant proteins also studied, which were constructed by mutating the residues of interest using the PyMOL code (10). In “H122\_p” simulations, H122 was *Ne*-protonated, rather than *N $\delta$* -protonated as in all other simulations.

The model membrane used in each simulation was a mix of 90% phosphatidylethanolamine, 5% phosphatidylglycerol, and 5% cardiolipin in the inner leaflet and Ra LPS in the outer leaflet. This is the same membrane composition previously used in the study of OmpA (11). This membrane was created by Piggot et al. (12) to reproduce the most common OM composition of *E. coli*. As in the previous studies,  $\text{Mg}^{2+}$  ions were used to neutralize the system.

### Coarse-grained simulation protocols

Coarse-grained simulations of the protein in an Ra LPS-phospholipid bilayer were performed to determine the preferred orientation of the protein within the membrane. Coarse-grained simulations used the MARTINI force field, with the LPS parameters of Hsu et al. (13). The Parrinello-Rahman barostat was used for semi-isotropic pressure coupling with a time constant of 1 ps. The velocity rescale thermostat was used, with a time constant of 1 ps. The time step for integrations was 10 fs. Coulombic interactions were cut off at 1.4 nm, and van der Waals forces were reduced to zero between 0.9 and 1.4 nm. Groningen Machine for Chemical Simulations (GROMACS) molecular dynamics software package (version 5.1.4) was used (14,15).

### Atomistic membrane simulation protocols

Simulations were set up and performed using the GROMACS molecular dynamics software package (version 5.1.4) with the GROMOS 54a7 force field (9,14,15). The parameters for Re LPS and Ra LPS are identical to those described by Piggot et al. (12) and Samsudin et al. (11), respectively, and the equilibrated protein-free membranes were taken from these earlier studies. The protein was embedded in the membrane using the GROMACS membed tool (12,15,16). The simple point-charge water model was used throughout the simulations (17). Systems were maintained at temperatures of either the biological 310 K or slightly higher at 323 K to improve sampling, using the Nosé-Hoover thermostat with a time constant of 0.5 ps (18,19). The pressure of the system was maintained at 1 atm, with a time constant of 5 ps, using semi-isotropic pressure coupling with the Parrinello-Rahman barostat (20,21). The same pressure coupling and barostat were used for NPT equilibration as well as production runs. All van der Waals interactions were cut off at 1.4 nm, and a smooth particle mesh Ewald algorithm was used to treat electrostatic interactions with a short-range cutoff of 1.4 nm. Simulation parameters were chosen based on similar published studies of OmpA (11). Each system was subjected to 500 ps of NVT simulation, followed by 20 ns of NPT for equilibration purposes. Positional restraints ( $1000 \text{ kJ mol}^{-1} \text{ nm}^2$ ) were placed on the  $\alpha$  atoms of the protein and modeled  $\text{Ca}^{2+}$  ion and water molecules during NVT and NPT equilibration. For some of the simulations, which are highlighted in Table 1, the restraints were kept in place for the first 100 ns of the production runs. Production runs of 300 or 500 ns were then performed for simulations at 310 K or 1  $\mu\text{s}$  for simulations at 323 K. The results were analyzed using GROMACS (14,15) tools and in-house scripts. Visualization was performed using the visual molecular dynamics software package (22).

**TABLE 1 Summary of All Molecular Dynamics Simulations Reported in This Work**

System	Notes	Length of Each Simulation	Temperature and Number of Repeats
CG_LpxR	Coarse-grained system	2 $\mu$ s	310 K ( $\times 2$ ) 323 K ( $\times 1$ )
Apo Model	X-ray structure with water and Ca <sup>2+</sup> as reported by Rutten et al. (5)	300 ns ( $\times 2$ ) 1 $\mu$ s ( $\times 1$ )	310 K 323 K
Apo Model <sup>a</sup>	As above but with positional restraints for the first 100 ns	300 ns ( $\times 2$ ) 1 $\mu$ s ( $\times 1$ )	310 K 323 K
Apo Unbiased	X-ray structure with key water and Ca <sup>2+</sup> removed from the starting structure	300 ns ( $\times 2$ ) 1 $\mu$ s ( $\times 1$ )	310 K 323 K
Ligand-Bound Model	X-ray structure with water and Ca <sup>2+</sup> and modeled in substrate as reported by Rutten et al. (5)	300 ns ( $\times 2$ ) 1 $\mu$ s ( $\times 1$ )	310 K 323 K
Ligand-Bound Model <sup>a</sup>	As above but with positional restraints for the first 100 ns	300 ns ( $\times 2$ ) 1 $\mu$ s ( $\times 1$ )	310 K 323 K
Ligand-Bound Unbiased	X-ray structure with modeled in substrate but without water and Ca <sup>2+</sup>	300 ns ( $\times 2$ ) 1 $\mu$ s ( $\times 1$ )	310 K 323 K
H122_p <sup>b</sup>	H122 is N $\epsilon$ protonated	300 ns ( $\times 2$ ) 1 $\mu$ s ( $\times 1$ )	310 K 323 K
D10A	Single point mutation in protein	300 ns ( $\times 2$ ) 1 $\mu$ s ( $\times 1$ )	310 K 323 K
D11A	Single point mutation in protein	300 ns ( $\times 2$ ) 1 $\mu$ s ( $\times 1$ )	310 K 323 K
T34A	Single point mutation in protein	300 ns ( $\times 2$ ) 1 $\mu$ s ( $\times 1$ )	310 K 323 K
H122A	Single point mutation in protein	300 ns ( $\times 2$ ) 1 $\mu$ s ( $\times 1$ )	310 K 323 K
Bond Break Model	Scissile bond removed in substrate	500 ns	310 K ( $\times 2$ ) 323 K ( $\times 1$ )
Bond Break Unbiased	As above but with key water and Ca <sup>2+</sup> removed from the starting structure	500 ns	310 K ( $\times 2$ ) 323 K ( $\times 1$ )
Micelle	Protein in self-assembled DPC micelle	500 ns	323 K ( $\times 3$ )
Apo_DPPC	X-ray structure inserted into simple DPPC membrane	500 ns	310 K ( $\times 2$ ) 323 K ( $\times 1$ )

The total simulation time was 29.6  $\mu$ s.

<sup>a</sup>In these systems, the protein, Ca<sup>2+</sup> ion, and modeled water molecule were subjected to positional restraints for the first 100 ns of simulation; see [Methods](#) for more details.

<sup>b</sup>H122 is N $\epsilon$  protonated in this simulation, whereas it is N $\delta$  protonated in all other simulations.

## Atomistic micelle simulation protocols

To study protein structure in a more labile environment, LpxR was placed in a box with 100 dodecylphosphocholine (DPC) molecules, Na<sup>+</sup> ions to neutralize charge, and simple point-charge water. Using a detergent micelle to study protein structure is a well-established method both in silico and in vivo (23–25). Positional restraints (1000 kJ mol<sup>-1</sup> nm<sup>2</sup>) were placed on the C $\alpha$  atoms of the protein during NPT equilibration, during which a DPC detergent micelle formed around the protein. DPC parameters were downloaded from <http://wcm.ucalgary.ca/tieleman/downloads>. Equilibration lasted for 50 ns at 350 K to ensure the micelle remained localized around the protein. Produc-

tion runs of 500 ns were then implemented at 323 K. Principal component and cluster analyses were performed on the resultant trajectories.

## GCMC simulation protocols

Grand canonical Monte Carlo (GCMC) simulations allow for the determination of both the location of water molecules within a defined region and their binding free energies. GCMC involves simulating in the grand canonical ensemble, that is the  $\mu$ VT ensemble, where  $\mu$  is the chemical potential of the system. The chemical potential of the simulation controls the water

**TABLE 2 Summary of All GCMC Simulations Reported in This Work**

Simulation	Repeats	B-Value(s)	Production Steps (Millions)	GCMC Box
H122 $\epsilon$	3	-40.79 to -10.79	20	0.2 $\times$ 0.2 $\times$ 0.2 nm <sup>3</sup>
H122 $\delta$	3	-40.79 to -10.79	20	0.2 $\times$ 0.2 $\times$ 0.2 nm <sup>3</sup>
H122 $\epsilon^a$	3	-7.94	40	0.73 $\times$ 0.49 $\times$ 1.16 nm <sup>3</sup>
H122 $\delta^a$	3	-7.94	40	0.73 $\times$ 0.49 $\times$ 1.16 nm <sup>3</sup>
H122 $\epsilon^b$	3	-40.79 to -10.79	20	0.2 $\times$ 0.2 $\times$ 0.2 nm <sup>3</sup>
H122 $\delta^b$	3	-40.79 to -10.79	20	0.2 $\times$ 0.2 $\times$ 0.2 nm <sup>3</sup>

<sup>a</sup>These systems refer to the water location identification simulations.

<sup>b</sup>In these systems, the imidazole ring of H122 was flipped 180°.

occupancy of a GCMC region. A  $B$  value, proportional to the chemical potential will be used from this point forward, because it encapsulates both the chemical potential and additional constant parameters. A detailed explanation of these simulations, including a relation of  $B$  value and chemical potential have been published by Ross et al. (26,27). Two types of GCMC simulations were performed; one was performed to calculate the locations of water molecules around the site of esterification using a larger GCMC box, and the other simulations were performed to calculate the binding free energy of the water molecule identified as being likely to be catalytic using a smaller GCMC box.

Within one of the ligand-bound molecular dynamics simulations at 100 ns, H122 was seen to move closer to the scissile bond, from  $\sim 0.9$  to  $0.6$  nm. A snapshot of the system was taken from this and used as the starting point for GCMC simulations using Prototype Molecular Simulation (28). The membrane was discarded, and the protein-ligand complex solvated in a  $4.5$ -nm sphere of the transferrable intermolecular potential four point of water (29). The membrane was discarded to simplify calculations; because the GCMC box was to be placed over an area accessible by bulk solvent, it was unnecessary to include additional lipids. The protein and ligand were modeled using the Amber 14SB and gaff16 force fields respectively (30–32). A  $\text{Ca}^{2+}$  ion (parameters taken from (32)) was included in the simulation, where the location of the ion was taken from our atomistic (AT) trajectories. Given H122 is key for the catalytic mechanism, to identify the correct protonation state and provide clues as to its mechanistic role, two sets of simulations were performed (one in which the  $\text{N}\delta$  was protonated and the other in which  $\text{N}\epsilon$  was protonated). Because of the similar electron density of carbon and nitrogen, they are difficult to differentiate between in crystallographic electron density. This means that it can be difficult to resolve the orientation of histidine residues experimentally. For this reason, the alternative rotameric states (referred to as the flipped forms) of both the  $\text{N}\delta$  and  $\text{N}\epsilon$  will be considered in the binding free-energy simulations. Two equilibration stages were performed; each stage involved one million MC moves. The first equilibration stage involved only GCMC sampling moves, while the second equilibration stage involved both GCMC sampling and sampling of the protein, ligand and bulk water. The protein was sampled as fully flexible with both backbone and side-chain moves. For the larger GCMC box, 40 million production steps were performed, whereas 20 million production steps were performed for the binding free-energy calculations. For the binding free-energy simulations, in which multiple  $B$  values are simulated, replica exchange in  $B$  was attempted every 100,000 steps (27). A list of all GCMC simulations can be found in Table 2.

For location simulations, a large GCMC box of dimensions  $0.73 \times 0.49 \times 1.16 \text{ nm}^3$  was placed over the region of mechanistic interest, i.e., the H122 and the Re LPS ester group. This simulation was performed to determine the hydration sites over the region. GCMC was performed at the equilibrium  $B$  value,  $-7.94$ , following the Prototype Molecular Simulation default methodology (28). Three independent repeats were performed for both of the H122 tautomeric states.

In addition, for binding free-energy simulations, GCMC calculations were performed to calculate the binding free energy of the specifically identified likely catalytic water molecule. The GCMC titration calculations were performed over a range of chemical potentials (16 equally spaced  $B$  values over an inclusive range of  $-40.79$  to  $-10.79$ ) for  $\text{N}\delta$ ,  $\text{N}\epsilon$ ,  $\text{N}\delta$ -flipped, and  $\text{N}\epsilon$ -flipped H122 conformations. The water site of catalytic interest was defined as the water position found for the  $\epsilon$ -protonated H122 of the larger GCMC box simulation. A small cubic GCMC box of length  $0.2 \text{ nm}$  was used, with the center of the box taken as the centroid of the water molecule. Three repeats were performed, and the binding free energy is calculated by performing 100 bootstrapping calculations of the data.

## RESULTS AND DISCUSSION

The protein and simulation setup is presented in graphical form in Fig. 1. Positioning of coarse-grain and united-

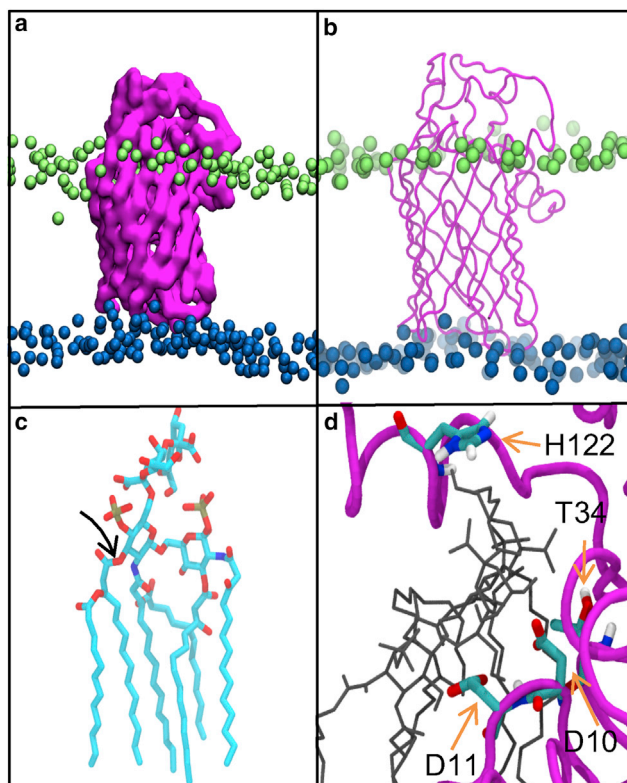


FIGURE 1 (a) Coarse-grained and (b) united-atom models of LpxR from the X-ray structure (PDB: 3FID) embedded within a model of the OM. Ra LPS headgroups are shown in lime, phospholipid headgroups are shown in blue, and lipids tails are omitted for clarity. (c) Re LPS with scissile bond are indicated and hydrogen atoms are removed for clarity. (d) A close-up of the docked Re LPS molecule (black) with residues thought to be key for catalysis through mutagenesis studies is highlighted (5). To see this figure in color, go online.

atom LpxR models with respect to phosphate headgroups can be seen in Fig. 1, a and b. The covalent bond broken when LpxR catalyses the removal of two acyl chains of lipid A is highlighted in Fig. 1 c, and a summary of the mutational residues that are key for catalytic activity are shown in active site (Fig. 1 d).

### Protein orientation and stability

Coarse-grained simulations were performed to provide an unbiased prediction of the localization and orientation of LpxR with respect to the OM. The densities of molecular components as well as the tilt angle of the protein with respect to the plane of the membrane calculated from the coarse-grained simulations were used to set up the AT systems. Comparisons of these metrics between AT and coarse-grained simulations are provided in the Supporting Materials and Methods. The principal axis of the protein barrel remains tilted by  $\sim 13^\circ$  relative to the bilayer normal during simulations at both resolutions (this is in close agreement with the tilt angle of  $\sim 16^\circ$  calculated from the



snapshot deposited in the membrane protein simulation database, MemProtMD (33); the differences arise from the latter being simulated in a symmetric phospholipid membrane and thus missing some of the key LPS-protein interactions included in our simulations). As has been reported for other outer membrane proteins (OMPs), the orientation of the protein is maintained in part by the aromatic residues on the outer surface of the barrel being positioned at the lipid headgroup-tail interface of the membrane (34). In the [Supporting Materials and Methods](#), we provide a heatmap of protein interactions with Ra LPS molecules, in which a contact is defined as an interatomic distance of  $\leq 0.4$  nm (Fig. S1).

Next, we calculated the area per lipid for both leaflets in our apo-protein systems to ensure well-equilibrated systems. Taking the area per lipid in the inner leaflet of our asymmetric membrane before protein insertion and an average after two independent apo-protein simulations (1- $\mu$ s simulation time for each) reveals an increase of only  $0.01$  nm<sup>2</sup>, suggesting little influence from the protein on inner-leaflet packing. For the outer leaflet, we calculated the area per lipid acyl chain, for which the experimental value is  $0.26$  nm<sup>2</sup> for LPS tails (35). We calculate an area per LPS acyl chain of  $0.28$  nm<sup>2</sup>, which is an increase of  $0.01$  nm<sup>2</sup> from the membrane before protein insertion and in excellent agreement with the experimentally determined value.

Interestingly we observe some local deformations of the membrane; specifically, the membrane is  $\sim 0.8$  nm thinner within a radius of  $\sim 1.2$  nm around the protein, compared to further away in what can be considered the bulk lipid region (Fig. S2 a). Previous simulation studies of bacterial OMPs have reported thinning of the local membrane, driven by the need to match the width of the hydrophobic region of the membrane with the hydrophobic surface of the protein (36).

A density profile of the molecular components of the simulation systems (Fig. S2 b) does not suggest any particular phospholipids are involved in the membrane thinning process. Indeed, comparison of a membrane-only system to our data from apo LpxR in an Ra LPS-containing membrane shows that phosphate headgroups of phosphatidylethanolamine, phosphatidylglycerol, and cardiolipin remain at a distance of 2 nm from the bilayer core despite the presence of the membrane-thinning protein.

Comparison of the order parameters of LPS acyl tails within 0.5 nm of the protein and in the bulk lipid region provides further evidence of the extent of membrane distortion caused by the protein (Fig. S3). There is greater ordering of the acyl tails away from the protein, and those closest to the protein are more disordered. Taken together, these results reveal a significant effect of the protein on the structure of the local membrane. Given LPS has a slow rate of diffusion compared to phospholipids, and often penetration of molecules into the monolayer is rarely observed in MD

simulation (37–39), the structural effects observed here are notable and imply nuanced analysis of these membranes is needed to tease out the effect of local lipids.

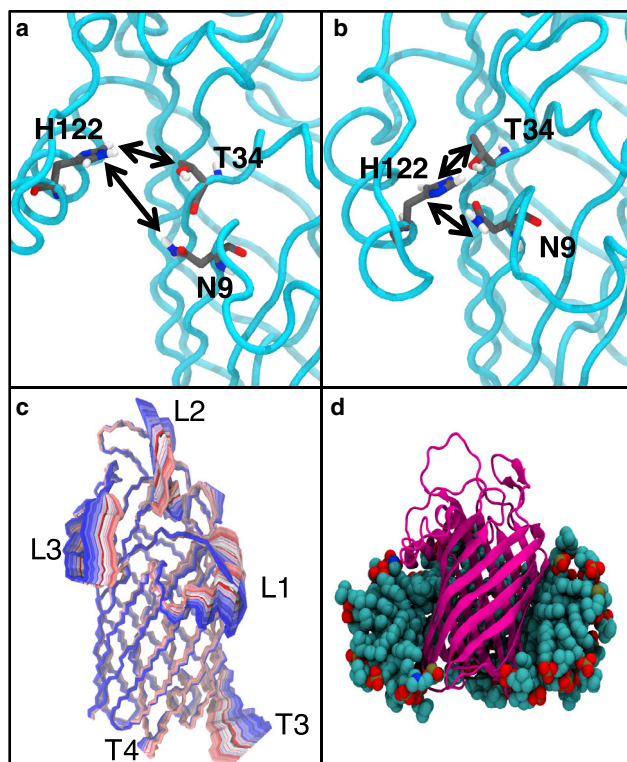
Root mean-square deviation (RMSD) and root mean-square fluctuations (RMSF) revealed LpxR to be stable within the OM, both in the apo form and when in complex with Re LPS. The RMSD of the barrel was  $0.1$ – $0.15$  nm for both bound and apo forms of the protein; these values are similar to those reported from other simulation studies of OMPs (40,41) (Fig. S4). There was a marked decrease in RMSD of the  $\alpha$ -helical regions of the protein from  $0.3$ – $0.35$  nm to  $\sim 0.25$  nm when comparing apo to bound forms of the protein. Given the RMSD fluctuates at about the same plateau value at 1  $\mu$ s as it is at 300 ns, we can be confident that our 300-ns simulations are as valid as the 1- $\mu$ s simulations for calculating equilibrium properties of the protein.

The extracellular loops showed greater flexibility than the barrel in terms of RMSF, which agrees with other previously reported OMPs (16,42); this is expected because the extracellular residues are not afforded the same scaffold-like support as the barrel by bulk lipid tails. Details of consistency in tilt angle and system partial density between coarse-grained and united-atom simulations along with further information on RMSD and RMSF can be found in Fig. S4.

### Conformational dynamics of the apo protein

Given the slow diffusion rate of LPS (39), to study the conformational states of the protein, we used two alternative environments (a detergent micelle environment and a symmetrical phospholipid bilayer, with the former considered first). Detergent micelles are often used in NMR studies of OMPs and are known to retain the stability of the protein barrels while showing the conformational flexibility of the extracellular loops in the ensemble of structures that is generated (3,42). Previously reported simulation studies have shown that the labile nature of detergent micelles allows for faster conformational dynamics (38,41,43). Over the course of three independent 500-ns simulations at 323 K (labeled “protein in self-assembled DPC micelle” in Table 1), we observed substantial conformational rearrangements of the protein. Whereas the  $\beta$ -barrel retained its original conformation as expected, the extracellular loops rearranged such that access to the LPS-binding pocket (Fig. 2, a and b) was occluded. Furthermore, residues previously identified as key for the catalytic process, N9, T34, and H122, were observed to form intermolecular hydrogen bonds. Thus, a putative closed conformation of the protein has been identified from our simulations.

To further investigate the conformational space sampled by LpxR in the micelle, we performed principal component analysis on the protein backbone by analyzing a concatenated trajectory of all the micelle simulations. The first principal component accounted for 31.5% of the total variance



**FIGURE 2** Conformational dynamics of LpxR within a detergent micelle at 323 K. (a) The position of catalytic residues in the binding pocket in the energy-minimized LpxR structure (PDB: 3FID). The protein backbone is shown in cyan, with carbon, oxygen, nitrogen, and hydrogen colored gray, red, blue, and white, respectively. (b) Closed conformation of the protein revealed by simulations in DPC micelles, with catalytic residues within hydrogen-bonding distance of each other. (c) The motion of the protein is depicted by extrapolating between the two extreme projections described by eigenvector 1 and then overlaying the conformation of the protein after every 20 ps. The image is colored on a blue, white, and red scale (blue at the start of the simulation, through white to red at the end of the simulation). (d) A representative snapshot of the protein in detergent micelle, some of the detergent molecules have been removed for clarity. To see this figure in color, go online.

in the backbone. The motion represented by this principal component was movement of extracellular loops L1, L2, and L3, which constitute the walls of the LPS-binding pocket to the “closed” conformation of the protein (Fig. 2 c). Furthermore, secondary structure analysis confirmed the stability and conformational integrity of the  $\beta$ -barrel segment of LpxR (Fig. S5).

The protein was also simulated within a symmetric DPPC bilayer (three independent simulations of 500 ns duration each) to allow for conformational changes that may lead to the closed state, given these lipids are faster diffusing than LPS. The putative closed conformation identified from the micelle simulations was reproduced in the phospholipid bilayer (Fig. S6 a). Interestingly, in simulations of the apo-protein-within DPPC bilayers, L2 retained the  $\alpha$ -helical conformation observed in the X-ray structure, in contrast to in the micelle simulations. In the DPPC bilayer, very little

movement of this loop was observed (Fig. S6 b), with residue backbone atoms having an RMSD of 0.34 nm, compared to 0.23 nm in Ra LPS. This is most likely due to the lateral packing of LPS sugar moieties around the extracellular regions of the protein. Principal component analysis was also performed on the protein backbone in DPPC and asymmetric membranes, and two-dimensional projections across the first two eigenvectors can be seen in Fig. S7 a. As expected, a smaller area of the conformational space was sampled when the protein was in the more structured environment of the bilayer. That is to say, the DPC micelle enabled the greatest conformational flexibility, followed by the DPPC membrane and finally, the Ra LPS membrane allowed the least. Images of conformational change in the apo-protein backbone in DPPC and Ra LPS membrane are provided in Fig. S7, b and c.

### Cation binding sites in the apo protein

Having established the structural stability of the protein, we then characterized the behavior of the apo protein with respect to the putative Re LPS and cation-binding sites. We looked for evidence of a hydrogen bond between the carboxylate side chain of E128 and  $N\epsilon$  of H122, as is suggested in the proposed mechanism. However, across the 4.8- $\mu$ s simulations of the apo protein in bilayer molecular dynamics simulation, there was little evidence of a stable hydrogen bond here; because of steric hindrance produced by the backbone of extracellular loop L3 (residues 110–139, with the  $\alpha$ -helix within this loop defined by residues 110–126), it was easier for a hydrogen bond to form between the backbone carbonyl of E128 and  $N\epsilon$  of H122. The distance between the center of mass of H122 and E128 over 300 ns is shown in Fig. S9 a. We return to this later when discussing the mechanism of deacylation.

In each of the “model” simulations, the  $Ca^{2+}$  ion moved away from its initial binding site as soon as positional restraints were removed, indicating that either 1) a ligand is required to stabilize the ion within the binding site or 2) this is not the cation binding site. When subject to positional restraints, the  $Ca^{2+}$  remained  $\sim 0.25$  nm from T34; when restraints were removed, the ion moved away to  $\sim 1.2$  nm from T34. The positioning of  $Ca^{2+}$  in the model corresponds to electron density attributed to a  $Zn^{2+}$  ion in the X-ray structure (PDB: 3FID) (5).

### Cation binding sites in the ligand-bound complex

The protein-ligand membrane system was simulated in several states (Table 1); each one was initiated with the protein-lipid complex as predicted by docking calculations reported by Rutten et al. (5). In “Model” simulations, the  $Ca^{2+}$  and water molecule proposed to be essential for the catalytic mechanism were retained, whereas in some simulations these were removed. We consider the former first. The Re LPS substrate remained noncovalently bound to

the protein through electrostatic interactions between the Kdo-sugar-hydroxyl groups and the basic residues K67 and R68 for the duration of all of our simulations. Interestingly, Reynolds et al. (6) noted that lipid A, which lacks Kdo sugars, is a poor substrate for LpxR. Our results suggest that this is likely due to the absence of the salt bridges between the Kdo sugars and the protein. Positioning of Re LPS with respect to the binding site on LpxR for “Model” simulations may be seen in Fig. 3, *a* and *c* and for “Unbiased” simulations in Fig. 3, *b* and *d*; each image in Fig. 3 was produced from a representative simulation at 323 K.

H122 sampled two major rotamers throughout the simulations (Fig. 3 *e*). The distance between the center of mass of H122 and the scissile bond varied between 0.3 and 0.9 nm, although for most simulations the imidazole of H122 remained in a conformation angled toward the scissile ester bond. As expected, the H122A mutation did not affect protein interaction with Re LPS, nor did it affect the conformation of the L3  $\alpha$ -helix (residues 110–126), as shown in Fig. S8 *b*.

Interestingly simulations of the wild-type protein predict a second cation-binding site in which an  $Mg^{2+}$  ion is coor-

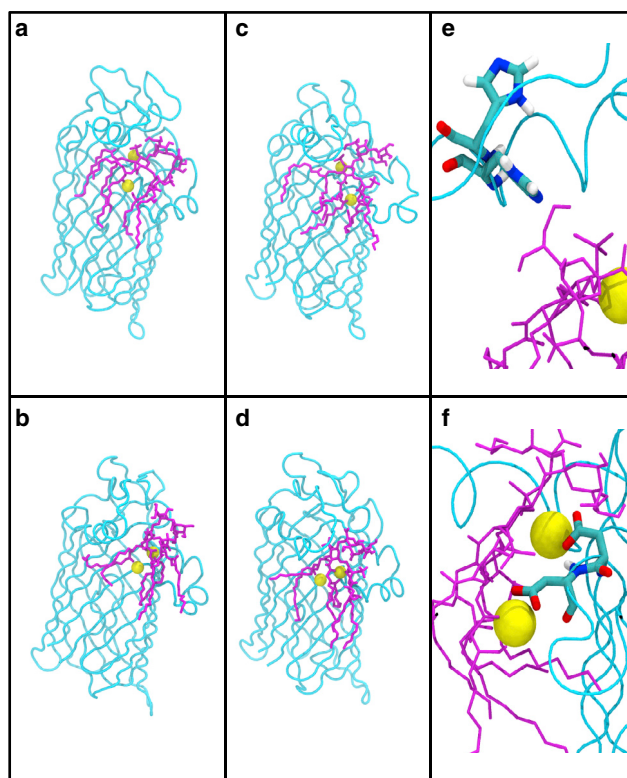


FIGURE 3 Snapshots of the protein-lipid-ion complex after 300 ns from two independent simulations of the ion-biased (*a* and *c*) and ion-unbiased (*b* and *d*) simulations. The two rotameric states of H122 are shown in (*e*), and (*f*) shows coordination of cations by residues D10 and D11. The protein is colored cyan, Re LPS is magenta, and ions are yellow. The membrane, water, and other ions are omitted for clarity. To see this figure in color, go online.

inated by the glycerol oxygen adjacent to the scissile bond, the carboxylic acid moiety of D11, and water molecules found persistently in this region throughout the simulations. Coordination of cations by D10 and D11 is shown in Fig. 3 *f*. When the  $Ca^{2+}$  ion is removed before equilibration, the conformational behavior of the putative active site is rather different. Although two  $Mg^{2+}$  ions are observed to enter the active site region during equilibration and remain within this region throughout the simulations, neither one is located precisely in the same spot as the  $Ca^{2+}$  ion from the Rutten model (5). Furthermore, H122 is observed to flip out of the active site such that the side chain is pointing toward the extracellular loops. This movement of H122 is accompanied by snorkeling of K67 toward the phosphate group of the lipid A sugar (Fig. S9) rather than interacting with the hydroxyl groups of the Kdo sugars as observed with the simulations of the Rutten model when  $Ca^{2+}$  is located within the proposed binding site (5). Residues T34 and D10 are not involved in interactions with the two cations in these simulations; instead one  $Mg^{2+}$  ion is observed to interact with D11 and two glycerol oxygens from adjacent acyl tails. The second cation-binding site identified in the wild-type simulations is again observed here with D11 involved in coordination along with a glycerol oxygen from one of the substrate lipid tails. The D11A mutation leads to the cation moving out of this binding site, indicating the importance of residue D11 for coordination of the second ion. In Fig. S10, we provide heat maps showing protein contacts with the  $Ca^{2+}$  ion and  $Mg^{2+}$  ions, where contact is defined as an interatomic distance of  $\leq 0.4$  nm.

The scissile bond is oriented through interaction of the carbonyl oxygen with the  $Ca^{2+}$  ion; the latter is also coordinated by N9, D10, T34, and three water molecules throughout the simulations of the wild-type protein. In simulations of the protein with the single-point mutation D10A, the  $Ca^{2+}$  moved out of the binding site such that after 300 ns it was no longer interacting with the protein, providing strong evidence that residue D10 plays a key role in coordination of the  $Ca^{2+}$  ion. In simulations of the T34A mutant however, the ion  $Ca^{2+}$  remains in the pocket, still coordinated by D10. Thus, our simulations suggest that for  $Ca^{2+}$  ion coordination, residue D10 is essential but T34 is not. Fig. 4 shows the effect of specific residue mutation on the conformation of the protein-ion-lipid complex. Again, all images in Fig. 4 were produced from simulations at 323 K.

Our simulations therefore predict that both the aspartate residues within the putative binding site, D10 and D11, play a role in binding cations. Interestingly, when  $Ca^{2+}$  is not coordinated by D10, the protein remains stable and the protein-substrate complex does not dissociate over the timeframe of our simulations; however, the local conformational rearrangements of the nearby residues are such that it is difficult to envisage a deacylation process occurring. As such, our simulations support the hypothesis of Rutten et al. (5) in which D10 must coordinate a cation. However,



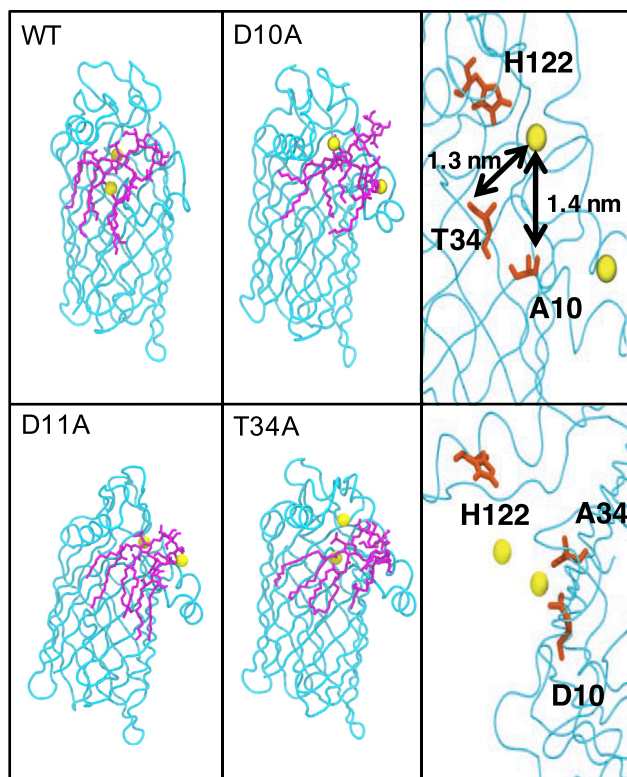


FIGURE 4 Snapshots of the protein-lipid-ion complex with specific residue mutations after 300 ns, compared to the wild-type. The membrane and solvent have been omitted for clarity. The top right panel shows the cation moving away from residue 10 when it is mutated from D to A. The bottom right panel shows the cation is still near its original location when residue 34 is mutated from T to A. The color scheme is the same as Fig. 3. To see this figure in color, go online.

this does leave open the question of why does D11 also bind a cation? Cation binding in this region is observed in all our simulations (other than those of the D11A mutant); therefore, cation binding by D11 is likely to be important for ion recruitment to the active site.

### Catalytic mechanism

Having identified the conditions under which cations bind to the enzyme, we next sought to gain some insights into the catalytic mechanism of deacylation by LpxR by combining the results of molecular dynamics and Monte Carlo simulations, with clues from what is known about the mechanisms of phospholipase A2 enzymes (8) and a previously hypothesized mechanism for deacylation from the static X-ray structure, docking studies, and mutagenesis studies (5,6). It is important to note here that initial predictions do not specify whether the mechanism requires one mechanistic water or a cascade of two water molecules; the latter is known to be the case for the mechanism of a secreted pancreatic phospholipase A2 enzyme (7,8). Attention was paid to the number of water molecules found between

H122 and the scissile ester during GCMC calculations. The putative catalytic protein residues,  $\text{Ca}^{2+}$  ion, known to be essential for activity, and the ester moiety of the lipid substrate are shown in Fig. 5.

To study aspects of the mechanism of deacylation and to predict the number of water molecules likely to be involved in the mechanism, we performed GCMC simulations to determine the positions of water molecules near the protein-ligand complex. Simulations were performed at  $B_{\text{eq}}$ , which corresponds to equilibrium with bulk water. The two protonation states of H122 were once again studied. H122 was more mobile in simulations in which it was N $\delta$  protonated. In two of the three independent simulations, the H122 shifted its position by rotating away from the Re LPS ester group, leaving a predominantly dry vacancy between H122 and the ligand. In the one independent repeat simulation in which the directional interaction between the H122 and the ester group was maintained, two water molecules were located by GCMC in between the two moieties as part of a chain of water molecules. Although the molecules were suitably positioned, the orientation of the water was such that formation of a hydrogen bond to the ester group of the substrate was not possible. However, the water was able to form a compensating hydrogen bond with another nearby water molecule. A snapshot of this water network and hydrogen-bonding interactions is shown in Fig. 6.

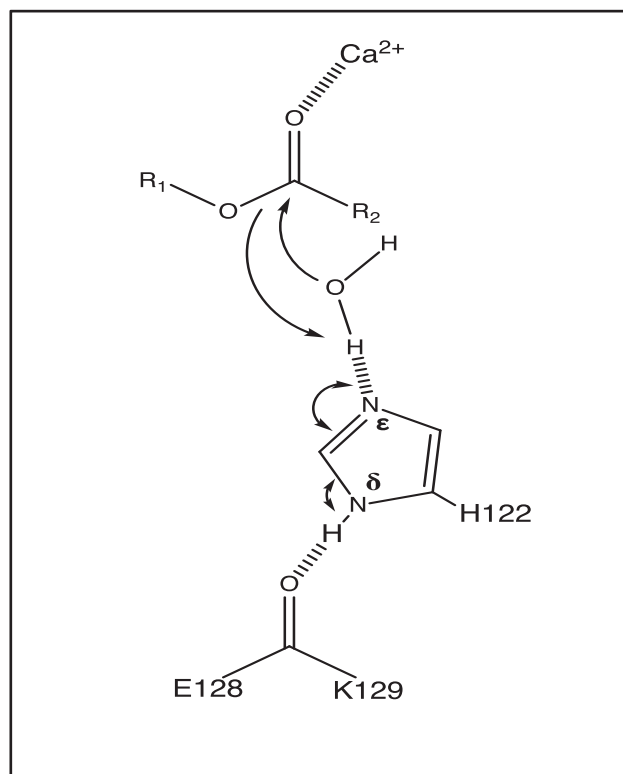


FIGURE 5 Catalytic mechanism including the residues E128, K129, H122, the portion of Re LPS to be cleaved, and one  $\text{Ca}^{2+}$  ion.



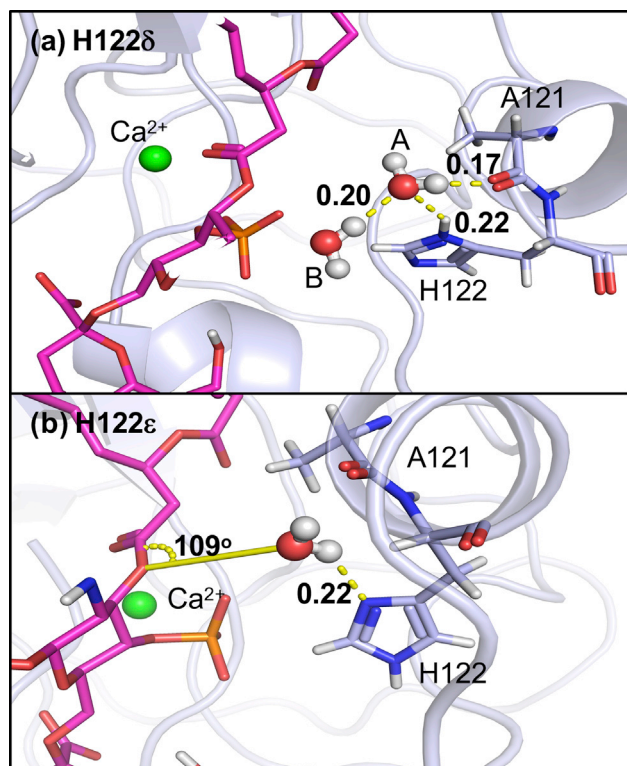


FIGURE 6 (a) Water locations from the GCMC simulation where H122 is  $N\delta$  protonated. Two water molecules found between the H122 and the ester. Hydrogen-bonding interactions are illustrated with a dashed yellow line, with length shown in nanometers. Water molecule B is forming hydrogen bonds with the H122  $H\delta$ , water A, and A121. The water A molecule is  $\sim 0.3$  nm from the ester group throughout the simulation and is oriented with the oxygen atom pointed toward the group, which means that it is unable to hydrogen bond to the ester group and would therefore not result in hydrolysis. The water is stabilized in this orientation by hydrogen bonding to another water close by. (b) A snapshot of the  $N\epsilon$ -protonated H122. A121, H122, and ligand are shown in stick representation, and the protein backbone is shown in light blue. The calcium ion is shown in green. The relevant water(s) in each case are shown in red and white. Part of the ligand and other GCMC water molecules have been removed from the image for clarity. To see this figure in color, go online.

When the H122 was  $N\epsilon$ -protonated, the hydrogen-bonding interaction was flipped, with a hydrogen bond accepting  $N\delta$  oriented toward the ester bond. In all three repeats, the GCMC method located a single water molecule interacting with both the H122 and the ester bond. The orientation of this water molecule, shown (Fig. 6) was such that it would enable it to act as a nucleophile in the catalysis. This water molecule was  $\sim 0.2$  nm away from His122 ( $N\delta$  to water O) and  $\sim 0.3$  nm away from the ester (water O to carbonyl C) throughout the simulation. The oxygen atom of the water molecule was at an angle of  $109^\circ$  to the plane of the ester group, close to the Bürgi-Dunitz angle of  $107^\circ$  (44). Based on the clustering of the GCMC-inserted water molecules, the water was present in this position 82.9% of the time when H122 was  $N\epsilon$ -protonated. Analysis of our MD simulation over 1  $\mu$ s revealed that water was present

in this region for  $\sim 95\%$  of the simulation, therefore clearly showing the two methods converging to this water site. The binding free energy of the catalytic water was determined using GCMC titration simulations for all combinations of rotameric and tautomeric states of H122, conformations of which are shown in Fig. S11. The water molecule in the catalytic position was found to have binding free energies of  $N\epsilon$   $-8.5$  (0.2),  $N\delta$   $-5.9$  (0.3),  $N\epsilon$  flipped  $-6.0$  (0.2), and  $N\delta$  flipped  $-7.0$  (0.2) kcal mol $^{-1}$  respectively, with SD shown. The corresponding GCMC titration curves are provided in Fig. S12. In all cases, the water molecule is tightly bound, but the binding free energy is most favorable for the  $N\epsilon$  conformation (the proposed catalytic conformation). The orientation of this water molecule relative to H122 and the ester bond, as well as its energetic stability, provide compelling support for it being the catalytic water molecule. We note here that Bahnsen (8) suggests that the catalytic mechanism of phospholipase A2 involves water activation by a histidine that is  $N\epsilon$  protonated. Based on the GCMC results favoring the  $N\epsilon$  protonation of H122 and the results from our molecular dynamics simulations, we can predict a mechanism for deacylation. We hypothesize here that the role of  $Ca^{2+}$  in the deacylation is to further polarize the  $sn_2$  carbonyl oxygen of the ester bond. The mechanism shown in Fig. 7 has H122 directly increasing the nucleophilicity of the catalytic water via hydrogen bonding from  $N\delta$ . As previously mentioned, although E128 has been hypothesized to stabilize the orientation of H122 through hydrogen bonding, we do not observe a persistent E128-H122 hydrogen bond in any of our simulations, regardless of the protonation state of H122. The lack of hydrogen bonding between these two residues could in fact account for the relatively low activity of LpxR in *S. typhimurium* (6,45); it remains to be seen whether a modified membrane composition could alter LpxR orientation and conformation, thereby stabilizing the proposed E128-H122 hydrogen bond to increase the basicity of H122 therefore activating the catalytic residue. Further detail of the rotamers sampled in GCMC simulation can be seen in Fig. S13.

### Protein and substrate after deacylation

To predict the diffusion of the products that occurs postcatalysis, the Re LPS ligand was manually deacylated by removing the covalent bond that would be enzymatically cleaved by LpxR. Simulations initiated with the two postcatalysis molecules placed within the hypothesized LPS binding region reveal that the cleaved acyl tails rapidly move away from the protein, toward the lower leaflet of the OM. Meanwhile, the modified LPS molecule remains within the putative binding site for the duration of our simulations. Each one of our six postcatalysis systems show the acyl tails moving toward the inner leaflet within 500 ns, as shown in Fig. 8. Images in Fig. 8 are produced from the bond-break unbiased simulation at 323 K.

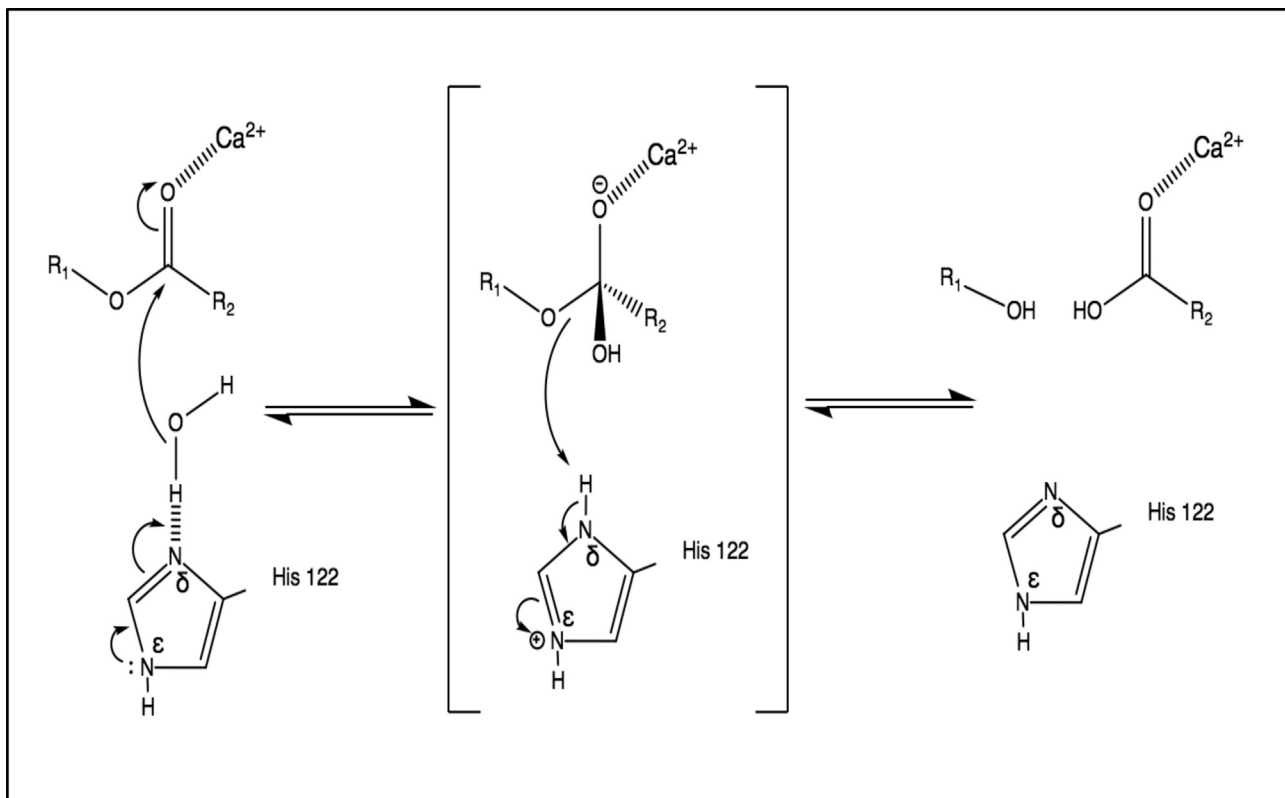


FIGURE 7 The predicted mechanism of deacylation based on our molecular dynamics and GCMC simulations. The catalytic histidine, H122, is  $N_{\epsilon}$  protonated. To establish whether the GCMC and molecular dynamics simulations were identifying similar water sites, we defined a  $1\text{-nm}^3$  box around H122 and the ester moiety of the ligand in the ligand-bound unbiased trajectory. Using visual molecular dynamics, we found that water molecules occupied this space for 95% of our  $1\text{-}\mu\text{s}$  molecular dynamics simulation at 323 K. More than one water molecule was found when the distance between H122 and the ester moiety increased.

In the case of OmpLA, the movement of lipid tails toward the inner leaflet leads to increased permeability and fluidity of the inner leaflet. This plays a role in bacteriocin release; it remains unclear as to whether LpxR could be implicated in the same way or whether the acyl chains are promptly repurposed in lipid synthesis.

It is useful to consider the main limitation of the current study. This arises from the initial enzyme substrate originating from a docked model. Although the simulations show that lipid A is stable in the docked conformation within the putative binding site, structural data would confirm the structure of the enzyme-substrate complex. It

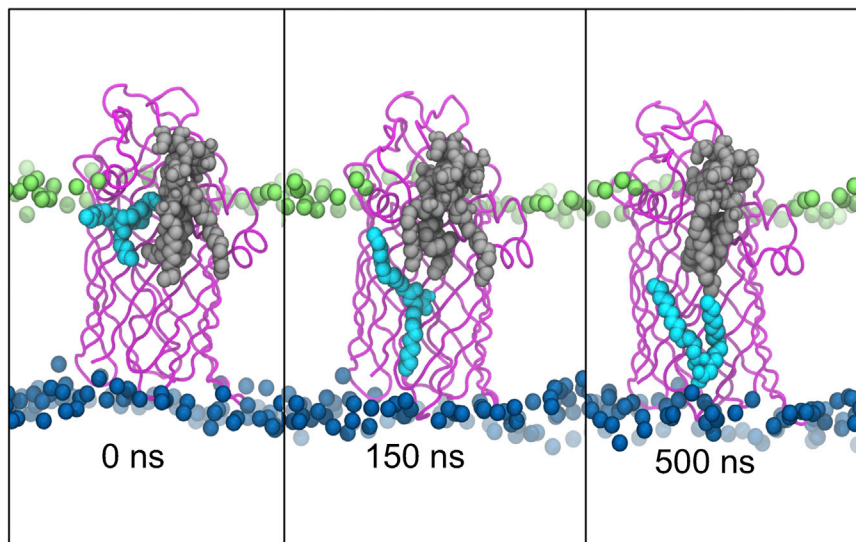


FIGURE 8 Movement of catalytically cleaved acyl tails (cyan) away from the protein (magenta) and the truncated Re LPS substrate (gray). To see this figure in color, go online.

is also worth noting that our study does not include polarization and other quantum effects, which would allow for a more detailed study of the chemical reaction mechanism. Having said that, quantum mechanics calculations of membrane-embedded proteins are immensely challenging, with very little available in the literature in terms of best practice. The approach we have used here of studying multiple protonation states of the catalytically active residue and using GCMC to identify water-binding sites provides a sound classical alternative to using quantum methods such as transition-state searching. However, it is clear that such mechanistic studies would benefit from representation at the quantum level, and we are moving toward developing protocols to enable this in the future. We note here that representation of ions with a classical molecular mechanics force field is notoriously difficult given we rely on Lennard-Jones parameters and a fixed charge. Here, we have employed the GROMOS 54a7 force field, in which the Lennard-Jones parameters for  $\text{Ca}^{2+}$  are unchanged from those in the 53A6 force field. These parameters have been successfully used in a number of studies in which specific ion binding plays a key role in the system dynamics (9,46,47). Overall, the models and methods used here come with their caveats, as do all scientific methods, but by using a combination of modeling, molecular dynamics, GCMC, and performing multiple simulations of each system we have ensured that the limitations are addressed as far as possible.

The mechanism of deacylation by LpxR is of immense interest from an immunological and microbiological viewpoint, given that modification of LPS can lead to antimicrobial resistance. Here, we have employed molecular dynamics and Monte Carlo simulations to study the conformational behavior of the protein, which has enabled us to predict the structure of the putative “closed” state of the protein in the absence of a substrate. The closed state of the protein exhibits hydrogen bonds between residues known to be key for the catalytic process. Thus, we hypothesize that in the absence of substrate, these residues play a role in stabilizing the protein in the closed conformation. For the substrate-bound state of the protein, we identify key protein-substrate interactions that hold the substrate within the active site. For the catalysis itself, we show that residue D10 plays a key role in coordinating a divalent cation. Our simulations predict the catalytic histidine to be  $\text{N}\epsilon$  protonated, which differs from the previously proposed mechanism from structural data and docking studies in which the histidine would have to be  $\text{N}\delta$  protonated. Furthermore, we provide quantitative evidence that one water molecule occupies the space between the catalytic histidine and the scissile bond in a tightly bound conformation and thus seems suitably placed to participate in the hydrolysis reaction. We show that the tails that are removed from the LPS molecule are able to rapidly diffuse toward the inner leaflet and presumably insert into this leaflet, whereas movement of the remainder of the LPS molecule out of the

active site is slower. The combined data from the different types of simulations enable us to hypothesize the full mechanism of catalysis and also the movement of the newly formed chemical species, postcatalysis. Indeed we show the importance of considering the conformational dynamics of membrane enzymes and ligands in a suitable local environment when attempting to decipher their mechanism of action.

## SUPPORTING MATERIAL

Thirteen figures are available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(18\)31026-9](http://www.biophysj.org/biophysj/supplemental/S0006-3495(18)31026-9).

## AUTHOR CONTRIBUTIONS

G.M.S. and H.E.B.M. performed the simulations. G.M.S., H.E.B.M., J.W.E., and S.K. analyzed the data. G.M.S., H.E.B.M., and S.K. wrote the manuscript.

## ACKNOWLEDGMENTS

H.E.B.M. is funded by the Engineering and Physical Sciences Research Council through grant number EP/L015722/1 (<http://gow.epsrc.ac.uk/NGBOViewGrant.aspx?GrantRef=EP/L015722/1>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## REFERENCES

1. Nikaïdo, H. 2003. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* 67:593–656.
2. Kim, S., D. S. Patel, ..., W. Im. 2016. Bilayer properties of lipid A from various Gram-negative bacteria. *Biophys. J.* 111:1750–1760.
3. Hwang, P. M., and L. E. Kay. 2005. Solution structure and dynamics of integral membrane proteins by NMR: a case study involving the enzyme PagP. *Methods Enzymol.* 394:335–350.
4. Rutten, L., J. Geurtsen, ..., J. Tommassen. 2006. Crystal structure and catalytic mechanism of the LPS 3-O-deacylase PagL from *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA.* 103:7071–7076.
5. Rutten, L., J. P. Mannie, ..., P. Gros. 2009. Active-site architecture and catalytic mechanism of the lipid A deacylase LpxR of *Salmonella typhimurium*. *Proc. Natl. Acad. Sci. USA.* 106:1960–1964.
6. Reynolds, C. M., A. A. Ribeiro, ..., M. S. Trent. 2006. An outer membrane enzyme encoded by *Salmonella typhimurium* lpxR that removes the 3'-acyloxyacyl moiety of lipid A. *J. Biol. Chem.* 281:21974–21987.
7. Berg, O. G., M. H. Gelb, ..., M. K. Jain. 2001. Interfacial enzymology: the secreted phospholipase A(2)-paradigm. *Chem. Rev.* 101:2613–2654.
8. Bahnson, B. J. 2005. Structure, function and interfacial allostereism in phospholipase A2: insight from the anion-assisted dimer. *Arch. Biochem. Biophys.* 433:96–106.
9. Schmid, N., A. P. Eichenberger, ..., W. F. van Gunsteren. 2011. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* 40:843–856.
10. DeLano, W. L. 2014. The PyMOL Molecular Graphics System, Version 1.8. Schrödinger LLC, New York <http://www.pymol.org>.
11. Samsudin, F., M. L. Ortiz-Suarez, ..., S. Khalid. 2016. OmpA: a flexible clamp for bacterial cell wall attachment. *Structure.* 24:2227–2235.



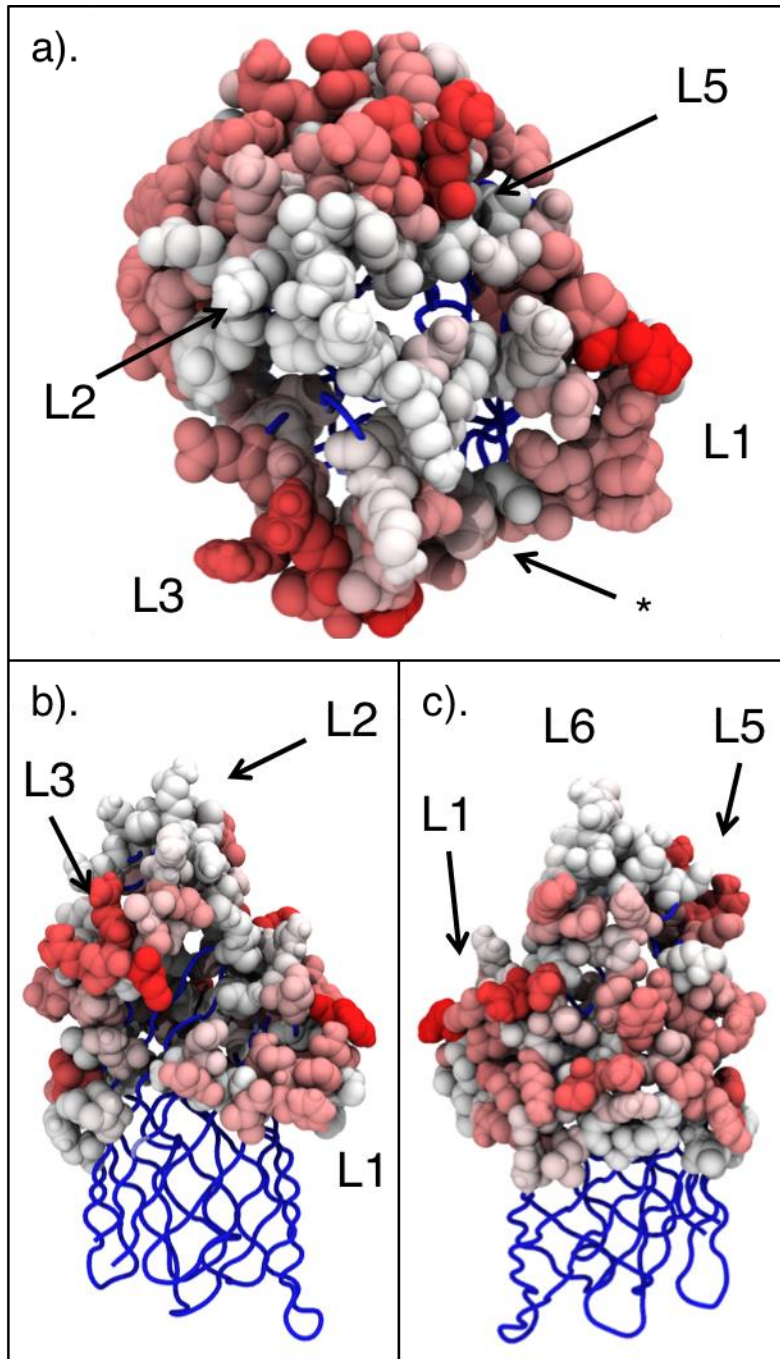
12. Piggot, T. J., D. A. Holdbrook, and S. Khalid. 2011. Electroporation of the *E. coli* and *S. Aureus* membranes: molecular dynamics simulations of complex bacterial membranes. *J. Phys. Chem. B.* 115:13381–13388.
13. Hsu, P. C., D. Jefferies, and S. Khalid. 2016. Molecular dynamics simulations predict the pathways via which pristine fullerenes penetrate bacterial membranes. *J. Phys. Chem. B.* 120:11170–11179.
14. Van Der Spoel, D., E. Lindahl, ..., H. J. Berendsen. 2005. GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26:1701–1718.
15. Martins-Costa, M. T., and M. F. Ruiz-López. 2017. Reaching multi-nanosecond timescales in combined QM/MM molecular dynamics simulations through parallel horsetail sampling. *J. Comput. Chem.* 38:659–668.
16. Ortiz-Suarez, M. L., F. Samsudin, ..., S. Khalid. 2016. Full-length OmpA: structure, function, and membrane interactions predicted by molecular dynamics simulations. *Biophys. J.* 111:1692–1702.
17. Berendsen, H. J. C., J. P. M. Postma, ..., J. Hermans. 1981. Interaction models for water in relation to protein hydration. In *Intermolecular Forces*. B. Pullman, ed. Springer, pp. 331–342.
18. Nosé, S. 1984. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* 52:255–268.
19. Hoover, W. G. 1985. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A Gen. Phys.* 31:1695–1697.
20. Parrinello, M., and A. Rahman. 1982. Strain fluctuations and elastic constants. *J. Chem. Phys.* 76:2662–2666.
21. Parrinello, M., and A. Rahman. 1980. Crystal structure and pair potentials: a molecular-dynamics study. *Phys. Rev. Lett.* 45:1196–1199.
22. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28.
23. Bond, P. J., J. D. Faraldo-Gómez, ..., M. S. Sansom. 2006. Membrane protein dynamics and detergent interactions within a crystal: a simulation study of OmpA. *Proc. Natl. Acad. Sci. USA.* 103:9518–9523.
24. Marrink, S. J., D. P. Tieleman, and A. E. Mark. 2000. Molecular dynamics simulation of the kinetics of spontaneous micelle formation. *J. Phys. Chem. B.* 104:12165–12173.
25. Zhang, L., R. N. DeHaven, and M. Goodman. 2002. NMR and modeling studies of a synthetic extracellular loop II of the  $\kappa$  opioid receptor in a DPC micelle. *Biochemistry.* 41:61–68.
26. Ross, G. A., M. S. Bodnarchuk, and J. W. Essex. 2015. Water sites, networks, and free energies with grand canonical Monte Carlo. *J. Am. Chem. Soc.* 137:14930–14943.
27. Ross, G. A., H. E. Bruce Macdonald, ..., J. W. Essex. 2017. Replica-exchange and standard state binding free energies with grand canonical Monte Carlo. *J. Chem. Theory Comput.* 13:6373–6381.
28. Bodnarchuk, M. S., R. Bradshaw, ..., C. Woods. 2017. ProtoMS. University of Southampton, Southampton, UK.
29. Neumann, M. 1986. Dielectric relaxation in water. Computer simulations with the TIP4P potential. *J. Chem. Phys.* 85:1567–1580.
30. Dickson, C. J., L. Rosso, ..., I. R. Gould. 2012. GAFFlipid: a general Amber force field for the accurate molecular dynamics simulation of phospholipid. *Soft Matter.* 8:9617.
31. Maier, J. A., C. Martinez, ..., C. Simmerling. 2015. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* 11:3696–3713.
32. Åqvist, J. 1990. Ion-water interaction potentials derived from free energy perturbation simulations. *J. Phys. Chem.* 94:8021–8024.
33. Stansfeld, P. J., J. E. Goose, ..., M. S. Sansom. 2015. MemProtMD: automated insertion of membrane protein structures into explicit lipid membranes. *Structure.* 23:1350–1361.
34. Koebnik, R., K. P. Locher, and P. Van Gelder. 2000. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* 37:239–253.
35. Snyder, S., D. Kim, and T. J. McIntosh. 1999. Lipopolysaccharide bilayer structure: effect of chemotype, core mutations, divalent cations, and temperature. *Biochemistry.* 38:10758–10767.
36. Schiffrin, B., A. N. Calabrese, ..., S. E. Radford. 2017. Effects of periplasmic chaperones and membrane thickness on BamA-catalyzed outer-membrane protein folding. *J. Mol. Biol.* 429:3776–3792.
37. Dias, R. P., G. C. da Hora, ..., T. A. Soares. 2014. Outer membrane remodeling: the structural dynamics and electrostatics of rough lipopolysaccharide chemotypes. *J. Chem. Theory Comput.* 10:2488–2497.
38. Edrington, T. C., E. Kintz, ..., L. K. Tamm. 2011. Structural basis for the interaction of lipopolysaccharide with outer membrane protein H (OprH) from *Pseudomonas aeruginosa*. *J. Biol. Chem.* 286:39211–39223.
39. Soares, T. A., and T. P. Straatsma. 2008. Assessment of the convergence of molecular dynamics simulations of lipopolysaccharide membranes. *Mol. Simul.* 34:295–307.
40. Baaden, M., C. Meier, and M. S. Sansom. 2003. A molecular dynamics investigation of mono and dimeric states of the outer membrane enzyme OMPLA. *J. Mol. Biol.* 331:177–189.
41. Bond, P. J., and M. S. Sansom. 2003. Membrane protein dynamics versus environment: simulations of OmpA in a micelle and in a bilayer. *J. Mol. Biol.* 329:1035–1053.
42. Cox, K., P. J. Bond, ..., M. S. Sansom. 2008. Outer membrane proteins: comparing X-ray and NMR structures by MD simulations in lipid bilayers. *Eur. Biophys. J.* 37:131–141.
43. Abel, S., F. Y. Dupradeau, and M. Marchi. 2012. Molecular dynamics simulations of a characteristic DPC micelle in water. *J. Chem. Theory Comput.* 8:4610–4623.
44. Bürgi, H. B., J. D. Dunitz, ..., G. Wipff. 1974. Stereochemistry of reaction paths at carbonyl centres. *Tetrahedron.* 30:1563–1572.
45. Kawasaki, K., M. Teramoto, ..., S. Amamoto. 2012. Lipid A 3'-O-deacylation by *Salmonella* outer membrane enzyme LpxR modulates the ability of lipid A to stimulate Toll-like receptor 4. *Biochem. Biophys. Res. Commun.* 428:343–347.
46. Hofer, T. S., and W. F. van Gunsteren. 2012. Exploring the properties of small molecule protein binding via molecular simulations: the TRSH-p53 core domain complex. *Mol. Biosyst.* 8:2891–2900.
47. Huang, W., Z. Lin, and W. F. van Gunsteren. 2011. Validation of the GROMOS 54A7 force field with respect to  $\beta$ -peptide folding. *J. Chem. Theory Comput.* 7:1237–1243.

**Biophysical Journal, Volume 115**

**Supplemental Information**

**Prediction of the Closed Conformation and Insights into the Mechanism  
of the Membrane Enzyme LpxR**

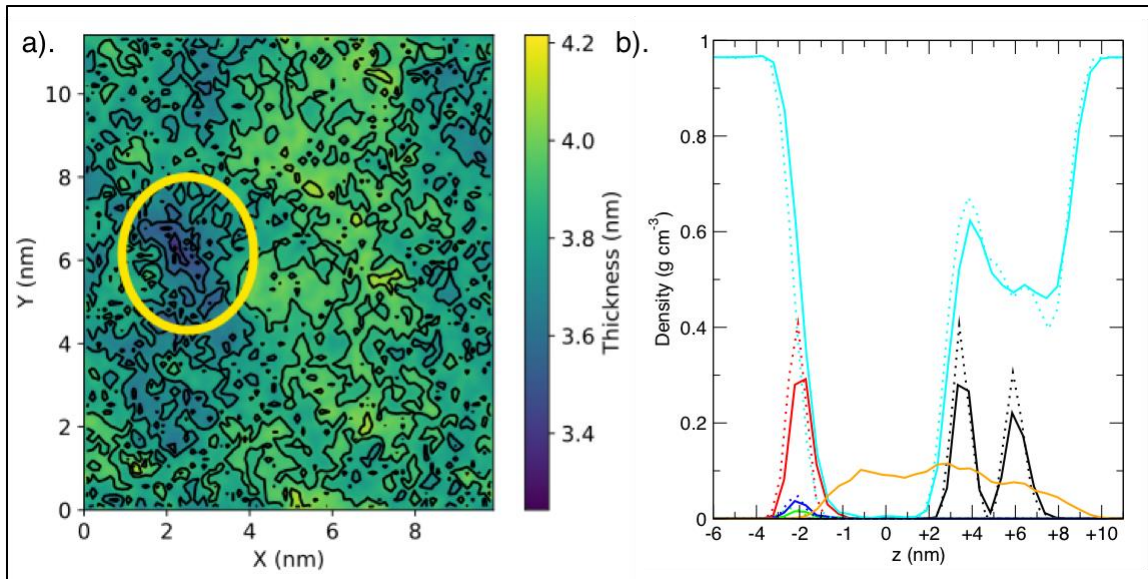
**Graham M. Smeddle, Hannah E. Bruce Macdonald, Jonathan W. Essex, and Syma Khalid**



**Figure S1.** a) Heat-map of protein residue contacts with Ra LPS molecules, with BWR colour scale and extracellular loops annotated. Contact defined as an interatomic distance of  $\leq 0.4$  nm, and images generated from a representative  $1\mu\text{s}$  *apo* protein trajectory at 323 K. a) A top-down view of the protein, with an asterisk denoting the LPS binding site. b) viewing LpxR from the side, with direct

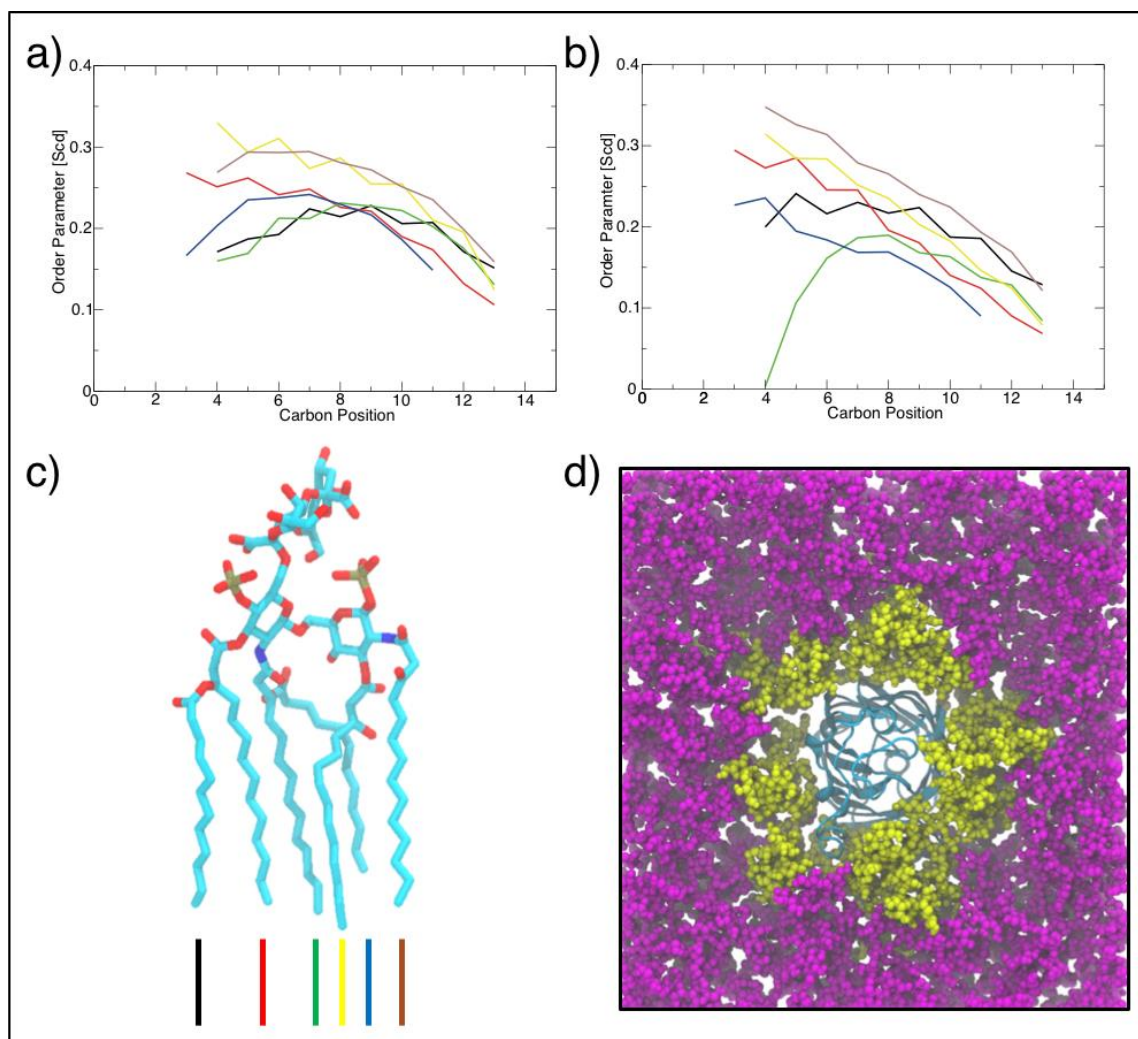


focus on the LPS binding site. **c)** viewing LpxR from the opposite side to **b)**, showing area typically immersed in LPS.



**Figure S2. a)** Depicting the thickness of the bilayer as a 2D landscape for the membrane. LpxR is shown to decrease the distance between phosphate residues of opposing leaflets in close proximity to the protein. Yellow circle denotes position of LpxR averaged across the simulation. **b).** Density of Re LPS (black), PE (red), PG (green) and DCG (blue) phosphates, as well as water (cyan) and protein (orange) across the z axis, comparing between membrane only and protein in Ra LPS membrane systems. Solid lines denote protein in membrane system and dashed line membrane only, with vertical black dashed lines highlighting the transmembrane (TM) region

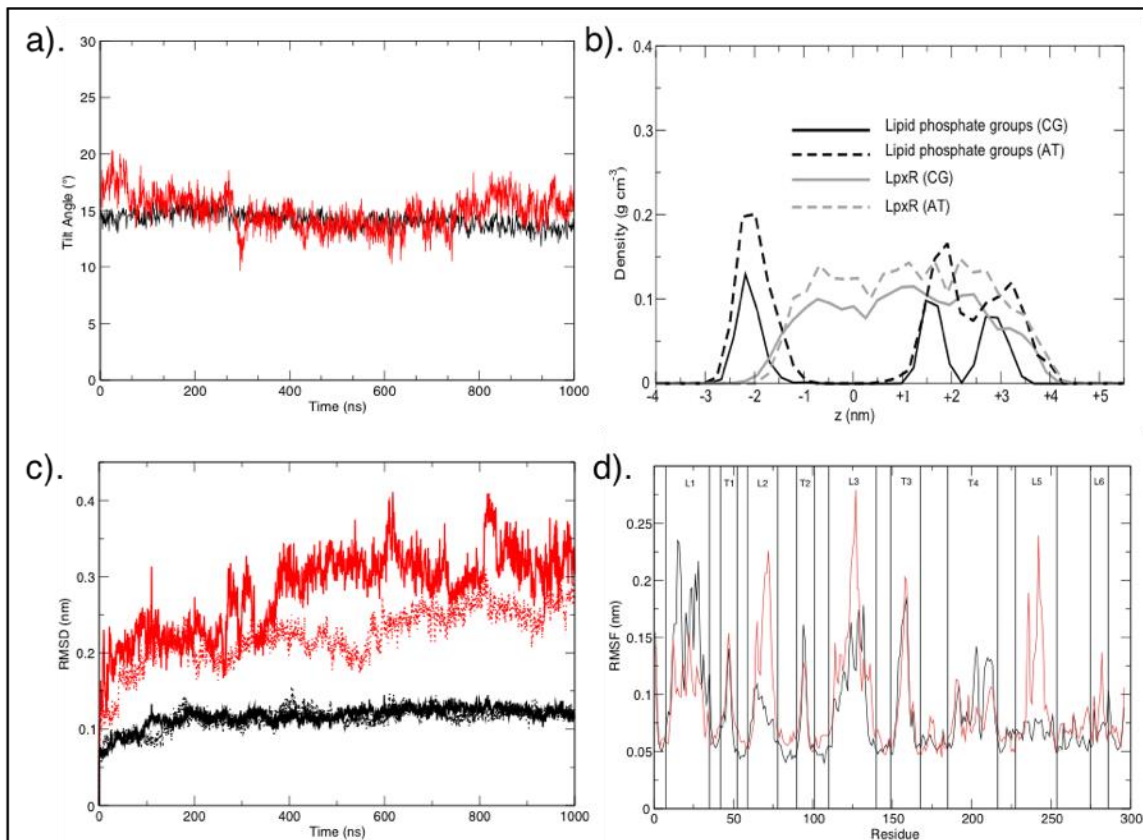
To assess any local membrane deformations caused by LpxR, membrane thickness analysis was carried out on the 2  $\mu$ s CG\_LpxR system at 323 K. Membrane thickness is here denoted as the distance between phosphate positions of opposing bilayers.



**Figure S3. a).** Deuterium order parameters of bulk LPS acyl tails during 1  $\mu$ s *apo* simulation at 323 K. **b).** Deuterium order parameters of acyl tails of LPS within 0.5 nm of LpxR. **c).** LPS molecule with tails colour coordinated to correspond with order parameter plots. **d).** Highlighting LPS residues selected for deuterium order parameter calculation. LPS included in **S2a** in magenta, **S2b** in yellow and protein in cyan.

**Figures S3a** and **S3b** show that lipid tails are more disordered in the region directly surrounding LpxR, compared with the bulk lipid region. **Figure S3c** acts as a graphical legend for **S3a** and **S3b**. This greater level of disorder supports our observation that the membrane is thinner in the 1.2 nm surrounding the

protein; LpxR disrupts the membrane, causing localised thinning and disorder of lipid acyl tails.

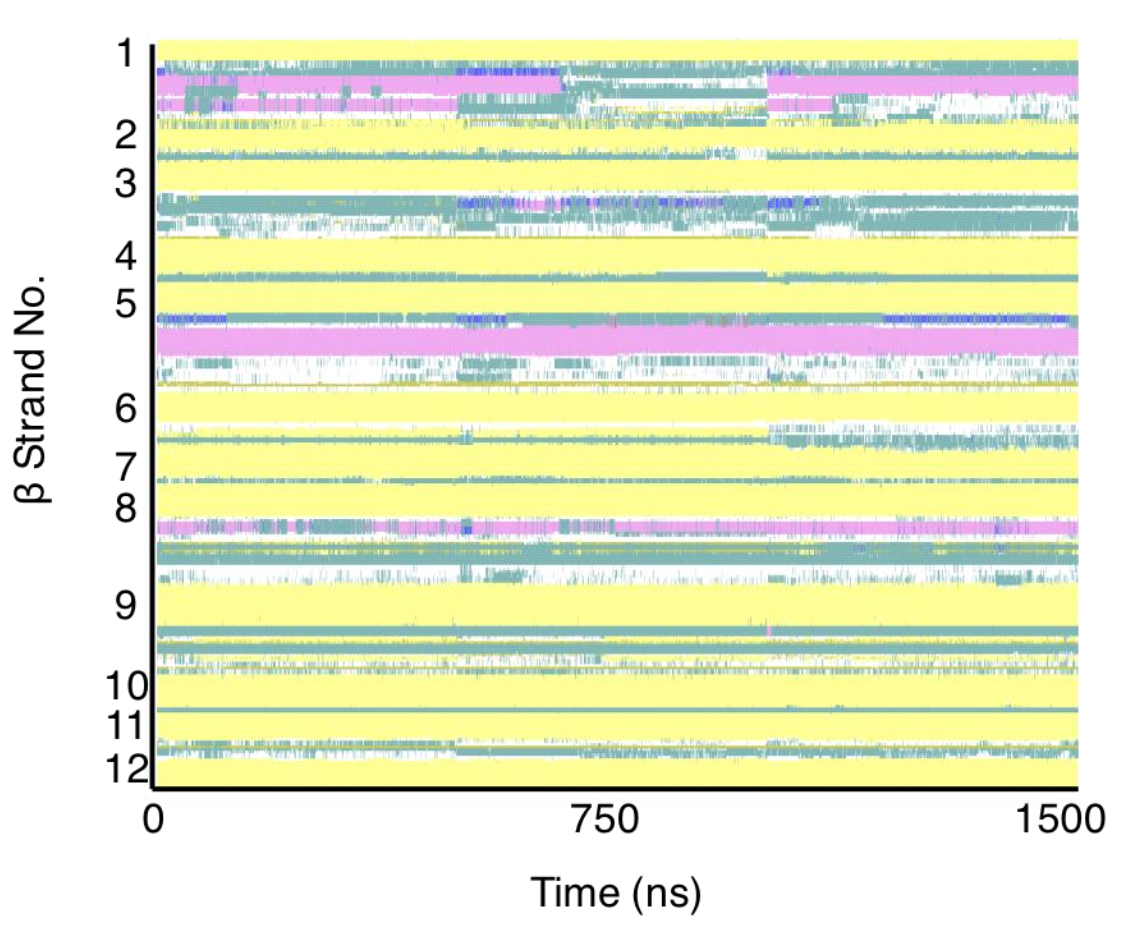


**Figure S4. a).** Protein barrel tilt angle relative to bilayer normal over 1  $\mu$ s of coarse-grain (black) or united-atom (red) simulation **b).** Relative densities of protein and lipid headgroups, coarse-grain (CG) and atomistic (AT), along the z axis **c).** RMSD of protein beta barrel (black) and alpha helices (red) in apo (solid line) and ligand-bound (broken line) systems **d).** RMSF of protein in apo (black) and ligand-bound (red) systems.

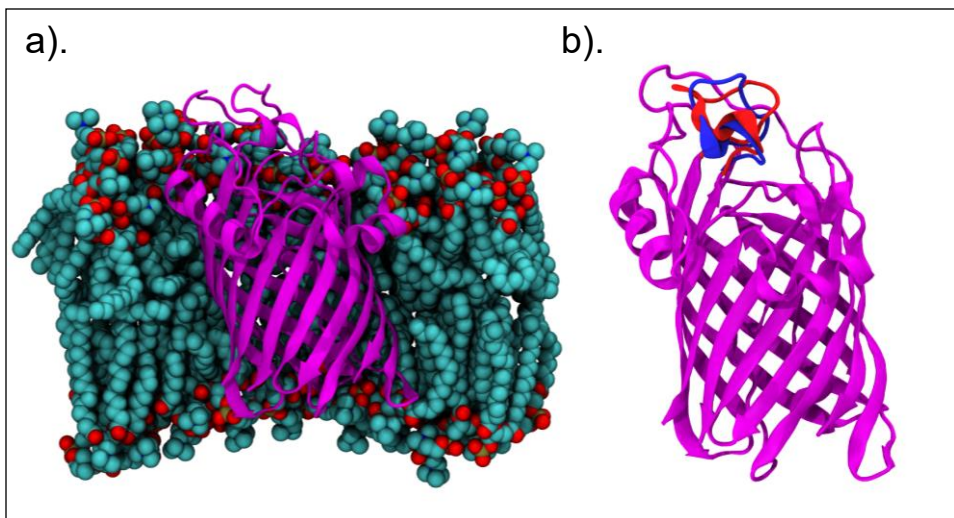
To ensure the orientation of protein in AT systems corresponded with their CG counterparts, both the tilt angle of the barrel principal axis and system partial densities were analysed and compared. The tilt angle of the principal axis of the



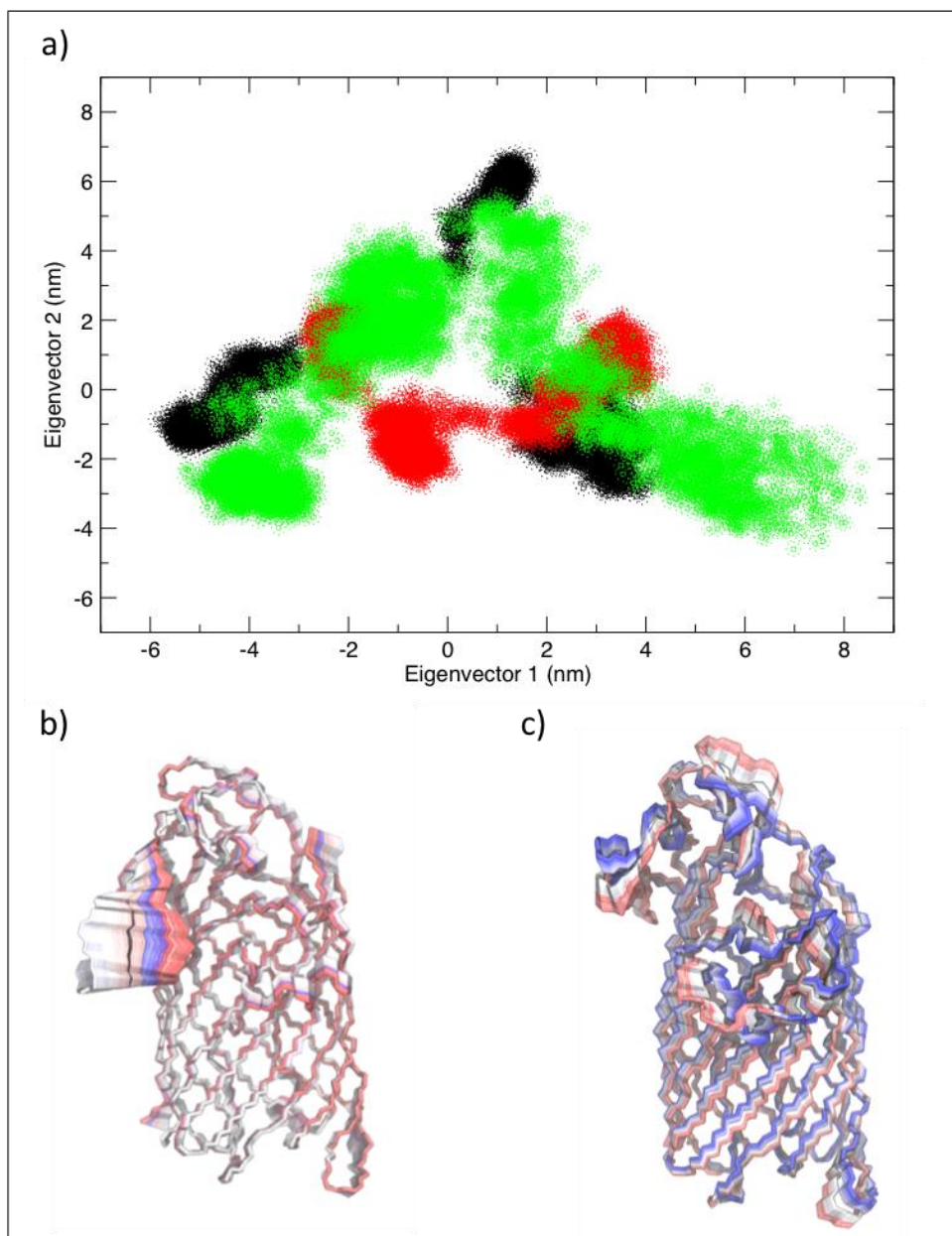
beta barrel was derived from the centre of mass of residues at the top and bottom of the barrel, relative to the bilayer normal.



**Figure S5.** Depicting the secondary structure of residues of LpxR across 1500 ns of simulation.  $\beta$  strand numbers labelled for clarity. Colouring as follows; yellow  $\beta$  sheet, pink  $\alpha$  helix, green turn, blue  $3_{10}$  helix, white coil, red  $\pi$  helix and brown isolated bridge. Secondary structure analysis was performed using the VMD script on the full length concatenated 1500 ns of protein from micelle simulations.



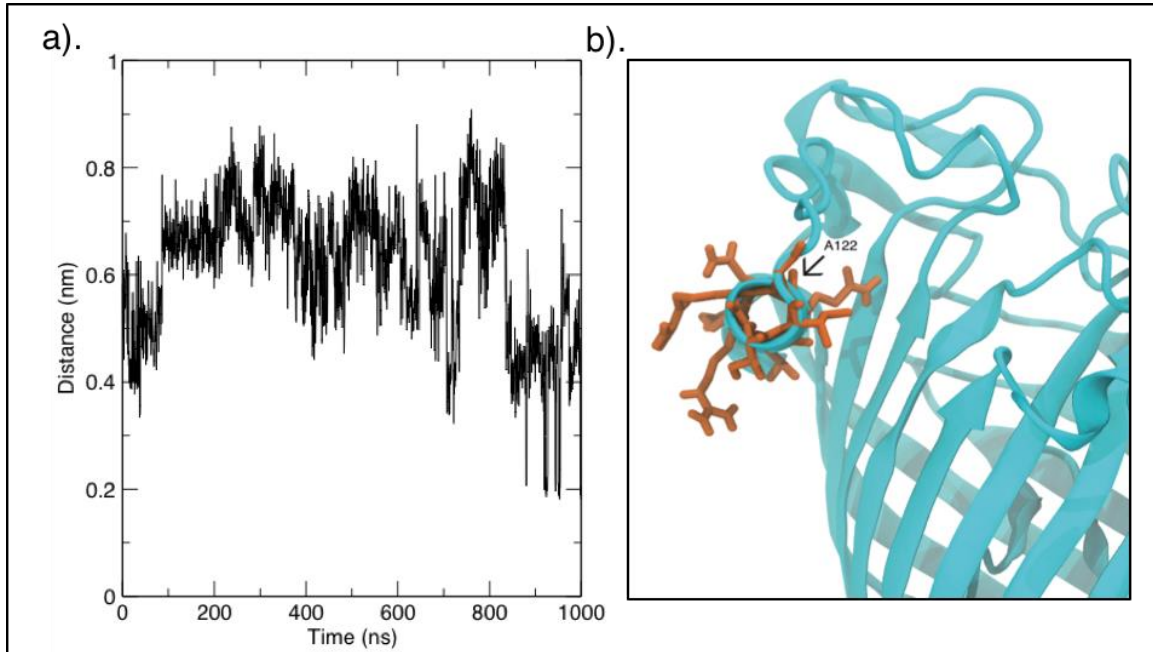
**Figure S6. a).** A representative snapshot of the closed conformation of LpxR in DPPC bilayer, with water and some lipids are removed for clarity **b).** An overlay of extracellular loop L2 throughout DPPC bilayer simulations. Blue denotes position of L2 at 0  $\mu$ s and red at 1  $\mu$ s.



**Figure S7. a).** A 2d projection of the first two principal components of protein backbone movement. Black points denote LpxR in asymmetric Ra LPS membrane, red in DPPC membrane and green in DPC micelle **b).** The motion of the protein backbone between two extreme projections described by the first principal component with protein conformations overlaid every 20 ps. Protein backbone trajectory taken from simulations with protein in DPPC membrane. The image is coloured on a BWR scale (blue at the start of the simulation, through white, to red at the end of the simulation) **c).** Identical to **b)**, but protein backbone



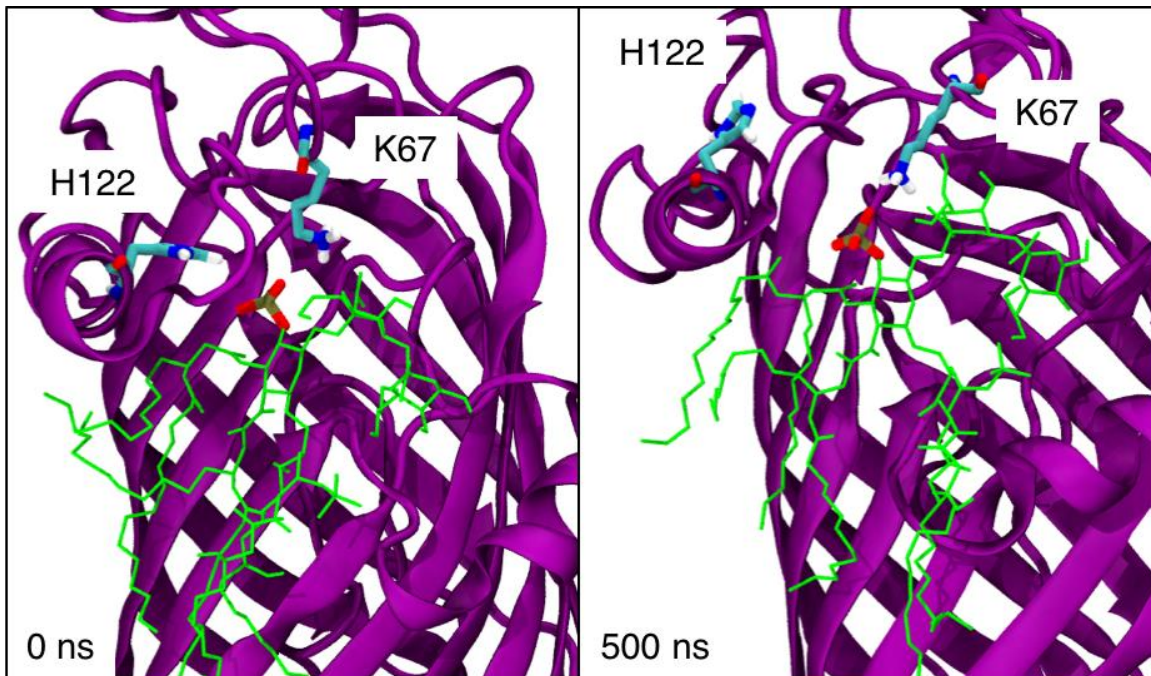
trajectory taken from simulations with protein in Ra LPS membrane.



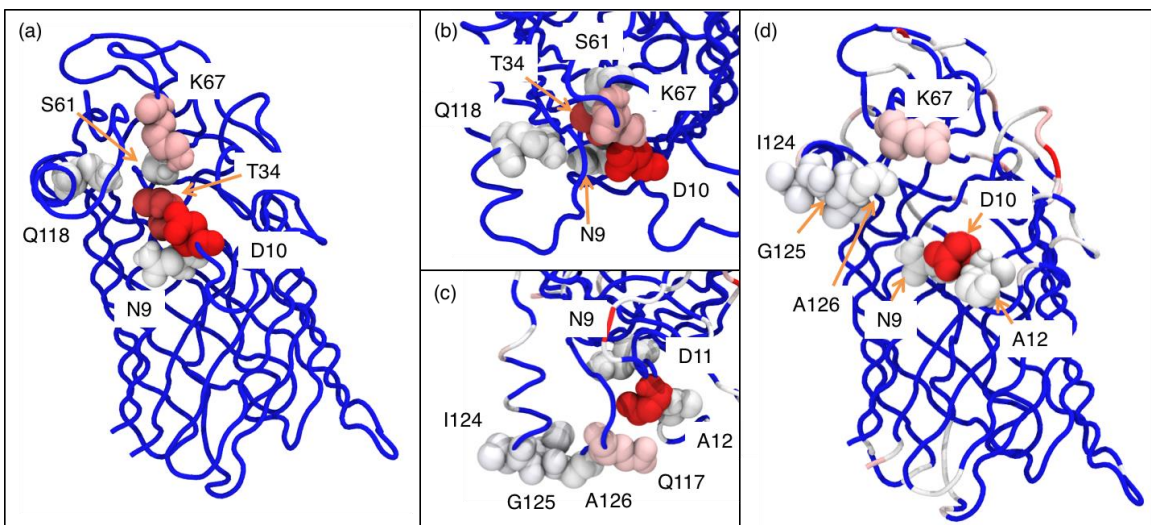
**Figure S8. a).** Distance between H122 and E128 centre of mass to indicate possibility of hydrogen bond formation between the two residues. **b).** Helix of extracellular loop L3 with H122A mutation after 1  $\mu$ s. Protein backbone coloured in cyan, residues 110-126 in orange.

As mentioned in the main text, we saw did not see evidence of a stable hydrogen bond between H122 and E128 during or simulations. We show in **Figure S8a** the distance between the centre of mass of H122 and E128 during the “model” simulation at 323 K, and how the two residues do not spend large periods of the simulation close enough to form a hydrogen bond.

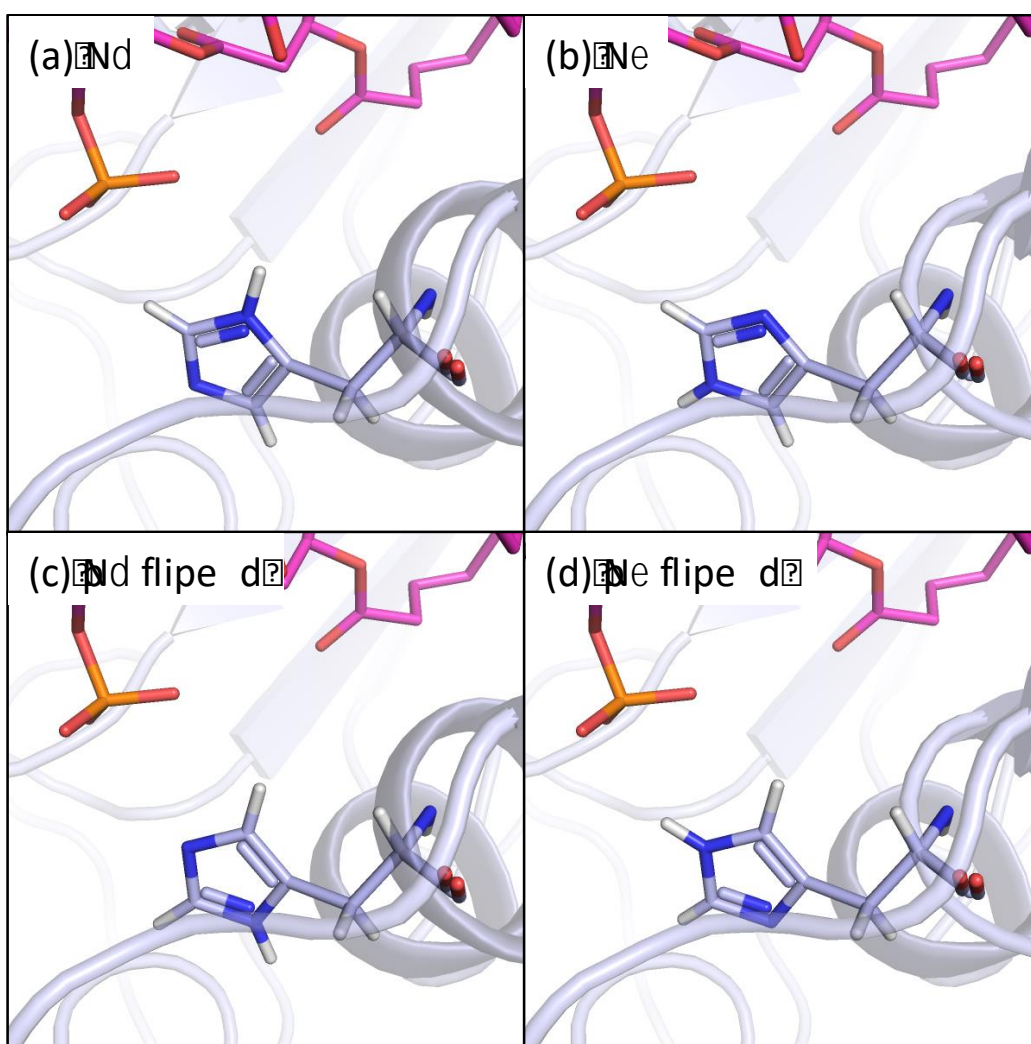
Also in the main text is the claim that the H122A mutation does not affect the conformational stability of the alpha helix of extracellular loop L3. This is demonstrated in **Figure S8b**.



**Figure S9.** Movement of H122 away from the scissile bond accompanied by 'snorkelling' of K67 towards the 1' phosphate of Re LPS bound to LpxR. Protein in purple and components of Re LPS other than the 1' phosphate in lime green. The two images represent time points of 0 ns and 500 ns. Image generated from representative model ligand-bound system at 323 K.

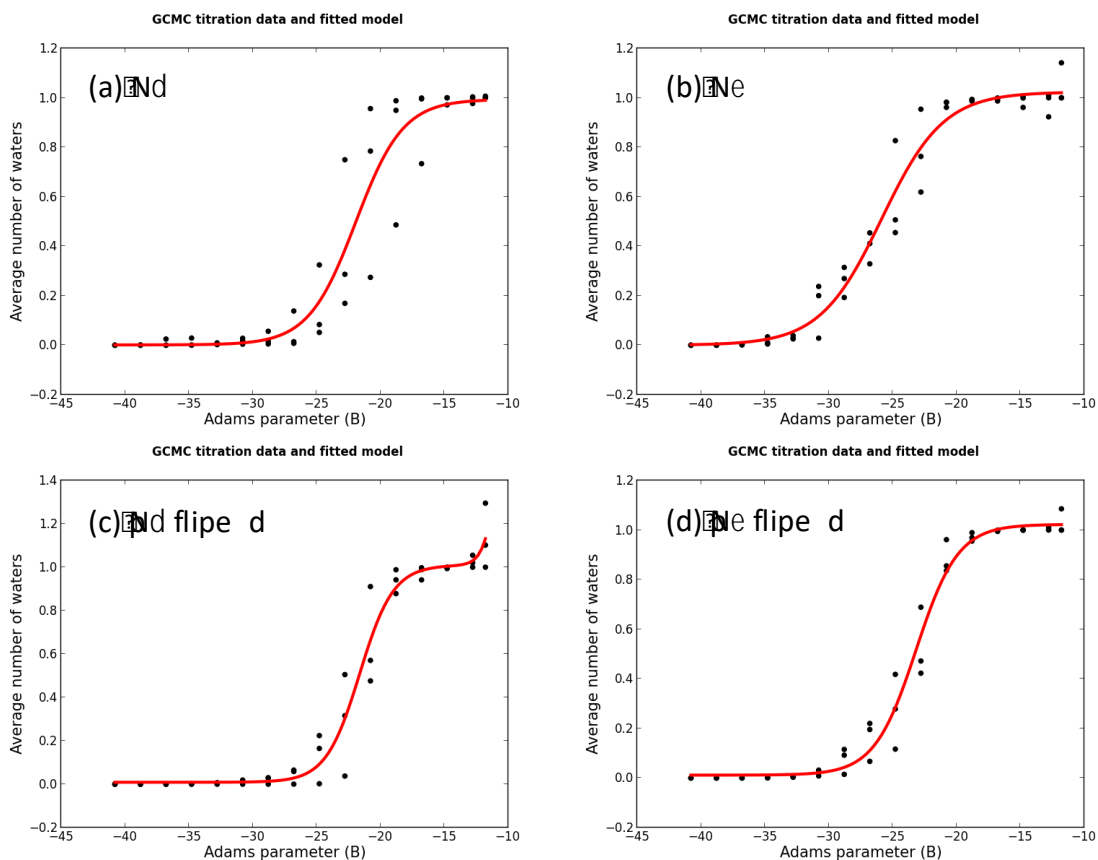


**Figure S10.** Heat-map of protein residue contacts with cations in the LPS binding site, with BWR colour scale. Contact defined as an interatomic distance of  $\leq 0.4$  nm **a).** Residues in contact with  $\text{Ca}^{2+}$  ions over  $1\mu\text{s}$  simulation at 323 K. **b)** same as **a)** but from a top-down viewpoint. **c)** Residues in contact with  $\text{Mg}^{2+}$  ions over  $1\mu\text{s}$  simulation at 323 K, with a top-down view of the binding site. **d)** same as **c)** but viewing the binding site of the protein from the side.



**Figure S11.** All starting orientation of H122 for GCMC titration simulations; **a).**  $\text{N}\delta$  **b).**  $\text{N}\epsilon$  **c).**  $\text{N}\delta$  flipped and **d).**  $\text{N}\epsilon$  protonated. Protein is shown in light blue, with histidine and REMP shown in stick representation, coloured light blue and pink

respectively.

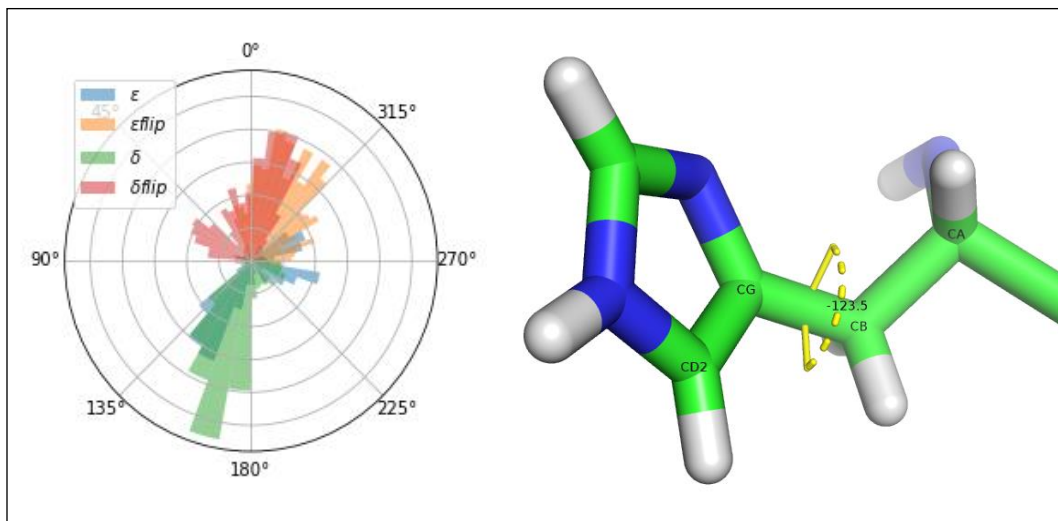


**Figure S12.** GCMC titration curves for the binding free energy calculations for water molecules with **a).**  $N_{\delta}$  **b).**  $N_{\epsilon}$  **c).**  $N_{\delta}$  flipped and **d).**  $N_{\epsilon}$  protonated H122 protein.

Titration curves show the average water occupancy for a simulation at each B value. The higher the B value, the higher the average water occupancy. A logistic function has been fit to the data (shown in red), from which the integral of this can be used to calculate the binding free energy of the number of water



molecules (a single water molecule in this case). This can be calculated using the GCI equation. The GCI equation relates the above results and their titration curves to the binding free energies presented in the manuscript.



**Figure S13.** Sampling of H122 rotamers observed in GCMC simulations with dihedral angle illustrated. Sampling of epsilon and delta is mostly observed between 135-180°, while the flipped simulations are observed to sample mostly between 315-360°.