The outer membrane proteins OmpA, FhuA, OmpF, EstA, BtuB and OmpX have unique lipopolysaccharide fingerprints.

Jonathan Shearer, Damien Jefferies and Syma Khalid[*1]

[1]School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ

*E-mail: S.Khalid@soton.ac.uk, Corresponding author

Abstract

The outer membrane of Gram-negative bacteria has a highly complex asymmetrical architecture, containing a mixture of phospholipids in the inner leaflet, and almost exclusively lipopolysaccharide (LPS) molecules in the outer leaflet. In *E. coli,* the outer membrane contains a wide range of proteins with a beta barrel architecture, that vary in size from the smallest having eight strands to larger barrels composed of twenty-two strands. Here we report coarse-grained molecular dynamics simulations of six proteins from the *E. coli* outer membrane OmpA, OmpX, BtuB, FhuA, OmpF and EstA in a range of membrane environments, which are representative of the *in vivo* conditions for different strains of *E. coli*. We show that each protein has a unique pattern of interaction with the surrounding membrane, which is influenced by the composition of the protein, the level of LPS in the outer leaflet and the differing mobilities of the lipids in the two leaflets of the membrane. Overall we present analyses from over 200 microseconds of simulation for each protein.

Introduction

The outer membrane proteins (Omps) of Gram-negative bacteria perform a wide range of functions including signalling, host cell adhesion, catalysis of crucial reactions, and transport (active and passive) of solutes/nutrients into and out of the cell. Given the range of functions they perform, it is not surprising that Omps have a wide range of sizes and oligomeric states. Currently the structures of > 70 Omps from Gram-negative bacteria have been determined; all of these with the exception of Wza from *E. coli*, have beta barrel architectures (see http://blanco.biomol.uci.edu/mpstruc). In *E. coli*, the sizes of Omps range from barrels composed of eight beta strands to twenty-two strands for a monomer[1]. The structure-function relationships of a number of Omps have been the focus of experimental and simulation studies over the last twenty years or so, and consequently many details have emerged, some of these are reviewed in [2]. More recently it has become evident that the interactions of Omps with lipids within their local membrane environment can be of a specific nature, these interactions are therefore likely to have functional consequences. For example, experimental studies have shown that OmpF has a specific LPS-binding region on its outer surface. In addition to their interaction with lipids, Omps also interact with each other as well as with peripheral membrane-binding proteins, which is perhaps not surprising given the crowded nature of the outer membrane. However, as with lipid-Omp interactions, Omp-protein interactions are also often specific and have functional

consequences that may impact upon the life-cycle of the bacterium. For example, it is known that the translocation of vitamin B12 across the outer membrane is facilitated by binding to BtuB, and then the subsequent formation of a translocon complex with OmpF. To understand these interactions and how they are stabilised, it is important to characterise the orientation and dynamics of Omps in their native membrane environment, as well as the impact they have on the local membrane.

Molecular dynamics simulations at fine-grain and coarse-grain levels of resolution have been employed to study the orientations of many membrane proteins. The most comprehensive study is perhaps that of Stansfeld *et al* in which coarse-grained simulations of membrane proteins embedded in symmetric phospholipid bilayers are performed and the results deposited in an online database[3]. The study covers all known structures of membrane protein and is updated as more structures become available. For Omps, however the analyses is somewhat complicated by the biochemistry of the outer membrane; the two leaflets differ in their lipid composition. The outer leaflet contains lipopolysaccharide (LPS) and the inner leaflet contains a mixture of zwitterionic and anionic phospholipids. Given that the length of the LPS molecule can vary, and it differs from phospholipids, it is important to ascertain if this impacts upon the orientation of Omps. We and others have previously shown that inclusion of this biochemical heterogeneity is important for understanding the conformational dynamics of Omps[4]. However a thorough study of the interplay between these proteins and their local membrane environment

and how this may differ in a protein-dependent manner, such as the study by Tieleman and co-workers in which the 'fingerprint' of each protein is identified, has so far not been reported for Omps in their natural lipid environment[5]. One particular technical issue that arises when simulating LPS-containing membranes is that of conversion, given the slowly diffusing nature of the lipids. It has been shown previously that atomistic LPS-containing membranes require simulation on the order of 500 ns for adequate sampling[6].

To identify whether the patterns of Omp orientations, and their interactions with the local membrane environment are uniform or dependent upon the protein and/or membrane type, we have performed coarse-grained simulations of six Omps of varying sizes, namely OmpX, OmpA (N-terminal domain), BtuB, FhuA, OmpF and EstA in outer membrane models that vary in the level of LPS in the outer leaflet. This set of proteins has been chosen as it includes small (OmpX and OmpA), large (BtuB and FhuA), multimeric (OmpF) and multidomain (EstA) Omps. We find that the level of LPS, but also the differences in mobility of the two leaflets of the outer membrane impact the protein-membrane interactions of the six Omps in different ways. Thus to use the terminology of Tieleman and co-workers, each one of these proteins has a unique LPS 'fingerprint'. In undertaking this study we have also been able to determine the timescales that are required for convergence of various properties of LPS-containing systems; these are significantly longer than those needed for simpler phospholipid membranes.

Results.

For clarity, the results are divided into two broad categories; Omps in varying levels of LPS and Specific-Omp interactions, each of which is subdivided into smaller sections. The four different membrane environments have the same inner leaflet composition (90% 16:0–18:1 phosphoethanolamine (POPE), 5% 16:0–18:1 phosphatidylglycerol (POPG) and 5% cardiolipin), whereas the outer leaflet is composed of one of the following: either 100 % rough lipopolysaccharide at (i) Re level, termed ReLPS (ii) or Ra level, termed RaLPS (iii) or smooth LPS which incorporates 5 sugar units of O-antigen, termed OANT (iv) or a mixture of smooth LPS and POPE in a 4:1 ratio, termed OANT_PE, further details are given in the methods section (Figure S1).

Table 1 provides a summary of the simulations that have been performed for this study. For each of the six proteins, at least 2 x 30 microseconds simulations have been performed of the protein in 4 different environments. Giving a simulation time of 240 microseconds per protein. Inevitably with this type of study, see for example Stansfeld et al, Scott et al, Corradi et al, an enormous amount of data is generated, and it is not possible to present or discuss comparative analyses for all proteins across all metrics[3,5,7]. Therefore we have presented the analyses for all protein in the supplementary information and have compared interesting cases in the main text of the paper.

Membrane protein	Membrane	Time (μs)	Box dimensions (nm)
OmpA	ReLPS	2x40	13.0x13.0x16.2
	RaLPS	2x60	13.0x13.0x18.7
	OANT	2x30	13.0x13.0x27.2
	OANT_PE	2x30	13.1x13.1x26.8
	DPPC	2x20	13.5x13.5x15.8
OmpX	ReLPS	2x30	13.5x13.5x16.4
	RaLPS	2x30	13.5x13.5x18.7
	OANT	2x30	13.5x13.5x26.8
	OANT_PE	2x30	13.5x13.5x26.9
	DPPC	2x20	13.5x13.5x16.0
OmpF	ReLPS	2x30	16.6x16.6x16.1
	RaLPS	2x30	16.9x16.9x18.4
	OANT	2x30	16.2x16.2x26.8
	OANT_PE	2x30	16.2x16.2x26.8
FhuA	ReLPS	2x30	13.6x13.6x17.8
	RaLPS	2x30	13.1x13.1x18.7
	OANT	2x30	13.1x13.1x27.1
	OANT_PE	2x30	13.1x13.1x27.1
EstA	ReLPS	2x30	15.8x15.8x20.1
	RaLPS	2x30	15.8x15.8x20.1
	OANT	2x30	15.8x15.8x26.5
	OANT_PE	2x30	15.8x15.8x26.9
BtuB	ReLPS	2x40	13.0x13.0x16.7
	RaLPS	2x40	13.0x13.0x18.5
	OANT	2x40	13.0x13.0x30.6
	OANT_PE	2x40	13.1x13.1x26.7

Table 1. Details for all of the simulations carried out for this paper. The box dimensions refer to the initial size of each system generated by CHARMM-GUI.

Orientation of Omps in varying levels of LPS.

Molecular Dynamics is an established technique for prediction of protein orientation within membranes. In particular the smoother energy landscapes and longer timescales accessible to coarse-methods make them especially useful for this

purpose[8]. As has been demonstrated previously, for symmetrical bilayers they can be used to self-assemble a bilayer around the protein, thus removing any bias in terms of human intervention in determining the orientation[7]. For asymmetric bilayers, self-assembly is not straightforward as one cannot control the composition of each leaflet. But the faster dynamics afforded by coarse-grained models do provide a route to determining equilibrium protein orientation that is independent of the initial orientation[3]. We have calculated the tilt angle with respect to the membrane plane of all six proteins in the four model membranes summarised in Table 1 (Figures 1 and S2).

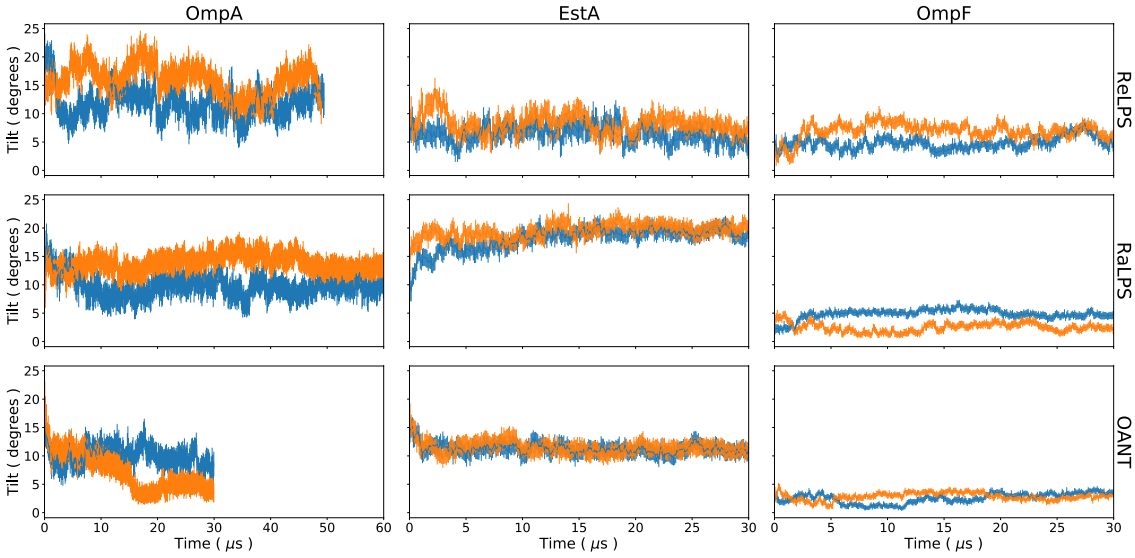


Figure 1. The protein tilt of the transmembrane region of each protein backbone with respect to the z axis as a function of time for each repeat. In the case of OmpF the tilt was determined using the entire protein backbone. Further details of the protein tilt calculation can be found in the Methods section. Each column corresponds to a single

membrane protein, while each row contains a certain level of LPS: ReLPS, RaLPS and OANT in descending order. The errors were generated from the standard deviation of block averages taken over 100 ns intervals.

Across two independent repeats of each simulation, convergence of the tilt angles is largely achieved after 30-40 μs for most of the proteins, with the O-antigen containing membranes showing the greatest variation between independent repeat simulations. Our results indicate that the proteins differ in the way they are orientated depending on the level of LPS in the membrane, although we don't see a simple pattern that can describe the dependency. Furthermore, compared to simple phospholipid membranes, the Omps can access a wider range of angles in LPS-containing membranes. Interestingly, the magnitude of differences in tilt angles, and pattern of tilt versus LPS type varies from protein to protein. For example, if we consider the two eight stranded barrels, OmpA and OmpX (Figures 1 and S2B), then we note that OmpA is tilted at an angle of ~ 12 ° (where 0 ° and 90 ° indicate the protein is perpendicular and parallel to the plane of the membrane respectively), in ReLPS and O-antigen level of LPS. In RaLPS there is less overlap of the two repeat simulations, but the trends indicate a tilt of ~10-13 °. Interestingly in the OANT_PE system OmpA is tilted at an angle of ~ 8 °, so there is no simple correlation between tilt angle and length of the LPS molecule. OmpX, is tilted at about 5-7 ° in ReLPS, ~ 12-13 ° in RaLPS, 8-9 ° in OANT and in OANT_PE the two independent

simulations vary without overlapping. Similarly, for the other four proteins, there is no simple correlation between LPS length and protein tilt. The protein EstA is the only Omp in our test set that has an extracellular domain large enough to interact with all sugar sections in each LPS type. The EstA tilt angles are about 5 °, 20 ° and 10 ° in ReLPS, RaLPS and OANT systems, respectively (Figure 1). The increased tilting in RaLPS systems seems to be the result of the extracellular bulb-like domain tilting to form interactions with the core oligosaccharides. In ReLPS systems the bulk domain is too far from the sugars for tilting to occur; whereas in O-Antigen systems the level of tilting observed in RaLPS is unnecessary (Figure S3) as stabilising interactions form with the O-Antigen chains. Thus, from our results it seems that for Omps with substantial extracellular domains, the level of LPS can play an important role in the membrane alignment of the protein.

While it is clear that the level of the LPS can alter protein orientation, it is pertinent to ask do the same residues tilt towards the membrane? To answer this question a 'relative tilt' was required that was invariant to the rotation of the protein in the xy plane, but indicated which direction the protein tilted in the outer membrane (i.e. which residues tilt towards the membrane). We picked a reference residue that these angles would be relative to for OmpA, EstA and OmpF (Thr 144, Ser 469 and Gly 232, respectively). Then the relative tilt was defined as the angles between the shortest vector, **R**, from the backbone of the reference residue to the principal axis used for the protein tilt (See Methods section for more details) and the z-axis (Figure S4). Thus, if

175 the relative tilt is greater than 90° this indicates that the vector **R** points towards the
176 membrane. The results show that for proteins with small transmembrane radii,
177 OmpA and EstA, in both independent simulations, the protein tilts in the same
178 general direction (Figure S5). While for OmpF relative tilt values calculated from
179 the repeat simulations of the same system, indicate the tilt directions differ between
180 the simulations.

181 Here it is important to consider an aspect of the outer membrane asymmetry that is
182 usually neglected in the context of Omp orientation; mobility of the lipids within
183 the two leaflets. The LPS molecules in the outer leaflet are less mobile than the
184 phospholipids of the inner leaflet; this is now well-established[2b,9]. To understand how
185 this impacts the motion of smaller Omps in each leaflet, we considered the proteins
186 OmpX and OmpA. We divided the barrels of each protein into the region that is
187 contained within the outer leaflet and the region that is found in the inner leaflet.
188 We calculated the centre of geometry motion of each portion of the barrel for time
189 period 0-30 μs (Figures 2 and S6). It can be clearly seen that in general, the portion
190 of the protein in the inner leaflet samples more of the membrane than the portion
191 within the outer leaflet, but the trend is more marked in the smaller proteins. If we
192 consider OmpX in detail (Figure 2); during the first 5 microseconds of simulation
193 there is greater movement of the protein within the inner leaflet compared to the
194 outer leaflet, while during the last 5 microseconds as an equilibrium orientation is
195 approached, the amount of movement within the two leaflets is comparable. For

reference we calculated the centre of geometry motion of OmpX and OmpA in simple phospholipid bilayers (Figure S7); where the differences in the amount of the membrane traversed by the protein in each leaflet is far smaller. Taken together the results of the tilt and motion analyses indicate that although it has been suggested for many years in the literature that Omp location and orientation within the outer membrane are governed by aromatic residues on their outer surfaces which are positioned at the lipid headgroup/tail interface, matching the hydrophobic surface of the Omp with the hydrophobic region of the surrounding bilayer respectively, certainly the latter is rather an oversimplification for the Omps with smaller barrels.

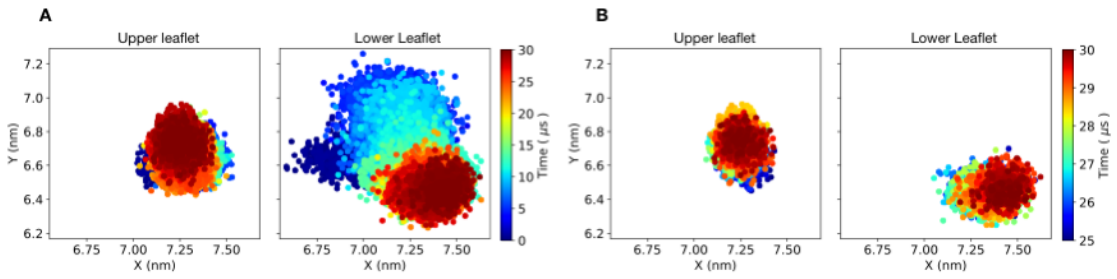


Figure 2. The barrel of OmpX was divided into regions that were in either the lower (see right) or upper (see left) leaflet of the ReLPS outer membrane based on the center of geometry of the entire transmembrane region of the protein backbone. The center of geometry of the protein backbone of each region was then plotted as a function of time for A) 0-30 μs and B) 25-30 μs. Before this analysis was carried out each trajectory was centered on the protein center of mass.

Specific Omp-lipid interactions:

The hydrophobic regions of all our model outer membranes have the same width, as the lipid A component is unchanged, therefore given we see differences in their behaviour in the different membranes, interactions with the sugar moieties of LPS clearly have an influence on the protein orientation and location. To probe specific protein residues that form stabilising interactions with the lipids of the outer membrane, we have calculated the number of interactions of all protein residues with the LPS headgroups and sugars. Given the headgroup of LPS is the same in all membrane models, we analysed only the ReLPS containing OM model for these interactions. In general we observe that aromatic and basic residues have greatest propensity to interact with the ReLPS headgroups (Figures 3 and S8A). The interactions with aromatic residues are in agreement with many previous studies of Omps in phospholipid bilayers, and are posited to anchor the proteins in the outer membrane. Given that ReLPS does not have an extensive sugar region, this phospholipid-like pattern of interactions makes sense. In LPS we observe these interactions to be dominated by tyrosine residues, although in OmpA the numbers of interactions with tyrosine and tryptophan are comparable. The headgroup interactions with basic residues are largely with the phosphate particles of the LPS headgroups (Figure S8B) as would be expected on the basis of electrostatics. Each protein has its own distinct pattern of interactions; in OmpX and OmpF interactions of the basic residues with LPS headgroups are largely though Lysine residues

whereas for BtuB and FhuA the number of interactions are comparable between arginines and lysines and for EstA there are more interactions with arginine than lysine. Intriguingly for OmpA, interactions with asparagine and histidine dominate. The extracellular loops of OmpA and OmpX both form contacts with the sugars and the lipid A headgroups (Figure S9). Closer inspection of the composition of the extracellular loops in OmpA reveals a reduced number of acidic amino acids compared to OmpX. Previous atomistic studies have reported that loops rich in positively charged amino acids in OprH from *P. aeruginosa* mediate interactions to the lipid A headgroup[4b]. OmpA dimers have recently been characterised, although their biological significance is still debated. We note here that the loops of OmpA are known to define the interface of the OmpA dimer through polar and non-polar interactions, therefore having basic residues in these loops would likely alter the propensity of the protein to dimerise/destabilise the dimerization interface[10]. The dimer was not studied here as OmpA is thought to be present mostly in the monomeric form.

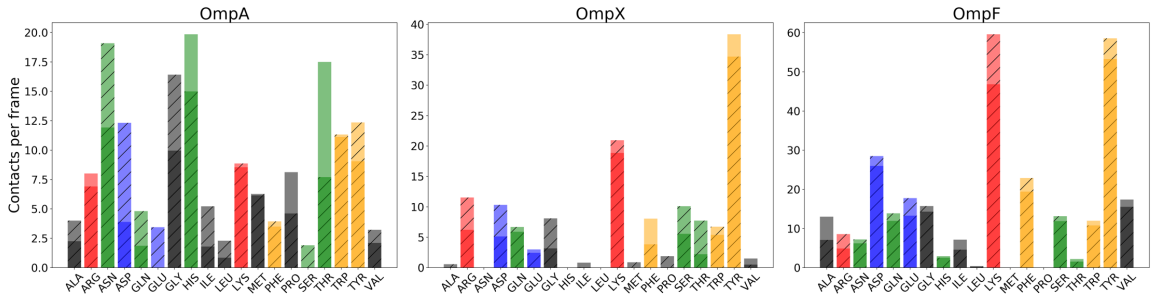


Figure 3. Average number of contacts per frame between each residue in a protein and the Lipid A headgroup for 25-30 μs across repeats (one repeat has hash markings). Each bar was coloured with respect to the polarity of each residue: red=acidic, blue=basic, orange=aromatic, green=polar and black=non-polar.

We next sought to quantify the interactions of the six proteins with the sugars of LPS (Figures 4 and S10). Once again these interactions are largely dominated by basic residues in ReLPS, with the exception of BtuB, however they are more variable in RaLPS and O-antigen. For example if we compare BtuB and FhuA given their size and functional similarities we may expect their interactions with LPS to be consistent with each other. We find in BtuB the highest frequency of contacts are with tyrosine residues in all membrane types. The highest number of contacts with charges residues are with lysines in ReLPS, aspartates in RaLPS and comparable between lysines and aspartates in OANT. On the other hand, in FhuA the highest frequency of contacts at the ReLPS level are with histidines closely followed by

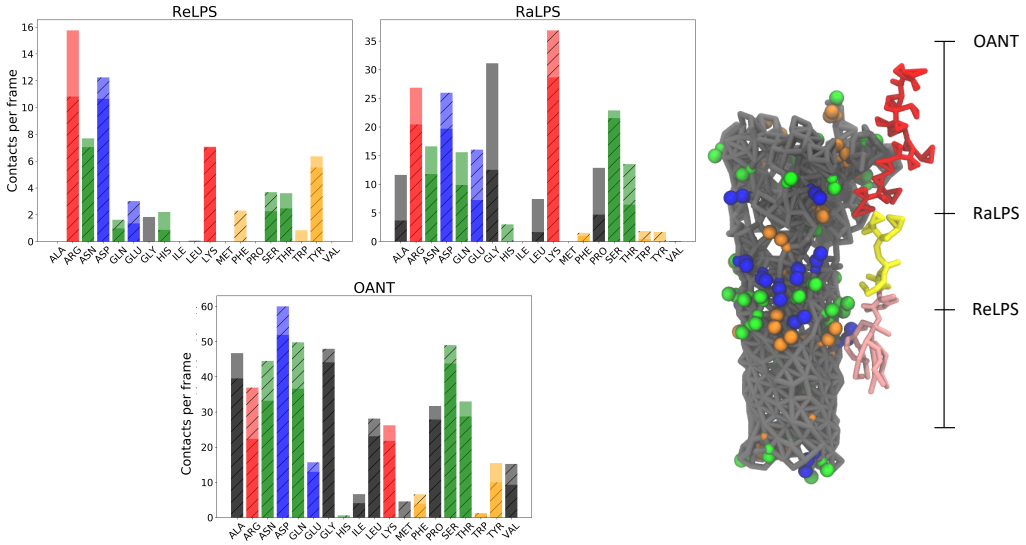


Figure 4. (left) Average number of contacts per frame between each residue in EstA and the lipopolysaccharide (LPS) core oligosaccharides and O-Antigen (if present) from 25-30 μs across repeats (one repeat has hash markings)**.** Every bar was coloured with respect to the polarity of each residue: red=acidic, blue=basic, orange=aromatic, green=polar and black=non-polar. Note that multiple contacts between a single residue and another Lipid A headgroup were only counted once. (right) Model of EstA backbone (grey) with all Arginine (orange), Aspartate (green) and Lysine (blue) residues shown next to a smooth LPS lipid. The height of LPS in ReLPS (pink), RaLPS (yellow) and OANT (red) systems is marked next to the LPS model.

lysines whereas in RaLPS and OANT contacts with lysines dominate. The interactions with sugars of the multidomain protein, EstA, also show variations in the polarity of the charged interactions (Figure 4); they are largely through arginine and aspartate residues in ReLPS, but with lysine in RaLPS and then again with

282 aspartate in OANT. This can be explained by considering the distribution of
283 residues. There is a cluster of five lysine resides in the ~ middle section of the
284 extracellular domain (Figure 4). This cluster of lysines can form contacts with the
285 sugars in RaLPS, but is not accessible to ReLPS. Likewise, many of the aspartate
286 residues that form interactions with sugars in the OANT membrane are not
287 accessible to RaLPS. Thus, the pattern of contacts with charged residues observed
288 for EstA can be explained by a comparative deficiency of lysine in the upper and
289 lower regions of the extracellular domain of EstA. Overall, the protein-LPS contact
290 data show that the pattern of contacts is not general (for the six proteins we have
291 studied here), the interactions may be dominated by charged residues which can be
292 acidic or basic, or they can be dominated by polar or aromatic residues, precisely
293 which, is dependent on the Omp but also the level of LPS present in the outer leaflet
294 of the outer membrane. As a validation of our methods it is important to compare to
295 experimental data, where this is available. Experimental details regarding the
296 interactions of FhuA and OmpF are available in the literature; one of the X-ray
297 structures of FhuA is resolved with a lipid A molecule bound to the protein[11] and
298 OmpF has also been shown to have specific LPS binding sites[12]. Our simulations
299 identify the same region of FhuA as making contacts with LPS, as that bound to
300 Lipid A in the X-ray structure (Figure 5). Lakey and co-workers have shown that
301 OmpF requires LPS to form a timer *in vivo* through the binding of LPS to acidic
302 amino acids[12]. Our simulations are consistent with these observations, with OmpF
303 interactions with LPS sugars being dominated by lysine residues (Figure 4). The

interactions with LPS headgroups are dominated by lysine and tyrosine residues (Figure 3). This is consistent with the X-ray structure of the porin OmpE36 in complex LPS, determined by the same authors which shows interactions between LPS headgroups and Lysine and tyrosine residues [12].

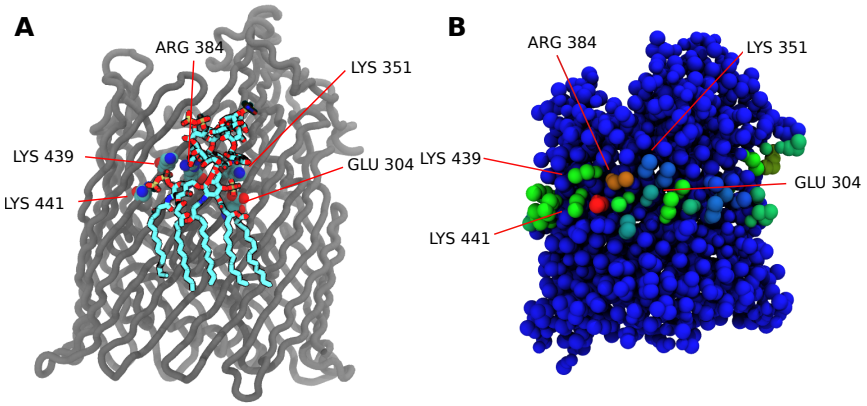


Figure 5. A) Image of crystal structure of FhuA (pdb code = 1QFG) which contains a bound lipopolysaccharide (LPS). The marked residues where within 0.4 nm of the Lipid A phosphate groups. B) Contact map of the coarse-grained (CG) structure of FhuA, between each FhuA residue and the phosphate beads of Lipid A in a ReLPS membrane. The contact analysis was carried out from 25-30 μs using a cutoff of 0.6 nm (as is standard for CG models). Each residue was coloured based on their respective number of contacts from low (blue) to high (red). The residues identified in Figure 5A are marked for the sake of comparison.

Thus far the analyses have been protein-focussed; but it is just as important from the perspective of the structure-dynamics-function relationship, to ask how the

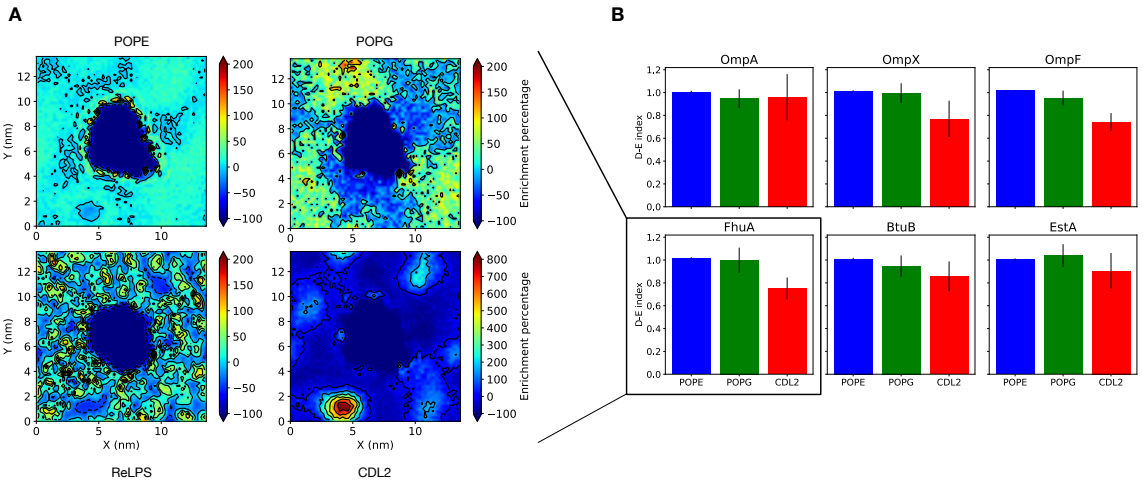
membrane responds to the proteins. To this effect we have analysed the enrichment/depletion of lipids around each Omp, the membrane thickness, and the effect of the proteins on the orientation of the O-antigen polysaccharide chains.

Membrane thickness and lipid enrichment/depletion.

Previous simulation studies have shown that Omps have a tendency to cause local thinning of phospholipid bilayers[7, 13]. Given the lipid A tails are shorter than the tails of the phospholipids used in both leaflets of the bilayer in previous studies, it is of interest to determine whether Omps also induce thinning of the more biologically relevant models of the outer membrane used here. Indeed, we find the membrane to be thinner within ~ 2.5 nm of all six proteins, in all 4 outer membrane models, beyond which can be considered 'bulk' membrane (Figure S11). This is the same distance from the protein that membrane thinning and thickening effects were reported for a range of beta barrel and alpha helical membrane proteins by Scott *et al* in phospholipid bilayers and for helical membrane proteins by Corradi et al. In our simulations we find the effect is always thinning of the membrane around the protein, regardless of the level of LPS. It is of functional relevance to explore the lipidic origins of the thinning effect; is it a general feature of membranes around Omps, or a function of specific lipid types? In order to explore this, we have calculated the lipid enrichment/depletion around the proteins in the inner leaflet of the ReLPS outer membrane using the methods described by Corradi et al [5]. Overall, in line with our findings from the analyses presented already, we find each protein

has its own characteristic pattern of lipids around it (Figure 6B). In general there is no overall enrichment of PE nor PG lipids around the Omps. The greatest variations are seen with cardiolipin; while overall there is depletion of cardiolipin around the Omps (within a 1.4 nm radius from the centre of geometry of the protein). On the basis of electrostatics alone, one might expect PG and cardiolipin to behave similarly around the proteins; from our results it is evident that this is not the case for the six Omps studied here. The D-E index provides a view of the lipid levels within a given radius around the protein rather than details of specific regions within that radius. To get a more detailed characterisation of specific regions of lipid depletion or enrichment, contact maps of PG with each protein were determined for the 25-30 μs interval of each system (Figures 6A and S12). We then cross-referenced to determine persistent binding site between repeat simulations of the same protein. We found that in particular FhuA and BtuB had prominent PG binding sites. The three persistent bindings sites of FhuA mainly consisted of charged or aromatic residues. Interestingly, one of these binding sites is located directly below the LPS binding site resolved in the X-ray structure (PDB code = 1QFG)[11]. Similarly two persistent PG binding sites were observed in the lower leaflet of BtuB. For both BtuB and FhuA there were also additional patches of PG enrichment, but these were not consistent between repeat simulations in terms of size and location. Interestingly clustering of cardiolipins is seen beyond the 1.4 nm radius used to calculate the D/E index in some of the simulations. This effect is seen in a number of simulations and appears to be independent of protein type. Perhaps the most pronounced effect is

363 seen in one of the simulations of FhuA (Figure 6); while cardiolipin is depleted
364 within 1.4 nm of the protein, molecules of cardiolipin are clustered together in one
365 region of the membrane beyond this distance. This cluster of cardiolipin molecules
366 in the lower leaflet corresponds to a region of LPS depletion above it, in the upper
367 leaflet. This compensation effect is also seen in simulations of OmpX and BtuB.
368 We previously reported this phenomenon in simulations of OmpA and OmpF.
369 Extension to four additional Omps in the present study reveals that the effect is
370 common, but likely not specific. Given this effect is not seen in both independent
371 repeats of the Omp simulations, other than FhuA, and it occurs outside of the annular
372 lipids of the protein, this suggests that the compensation for LPS depletion, by
373 cardiolipin clustering is a general feature of the outer membrane and is not
374 dependent upon the protein type. This is consistent with our previous work which
375 found such cardiolipin clustering even in systems which did not contain any
376 membrane proteins[14]. To probe density gradients in more detail we calculated the 2D
377 enrichment maps for the entire membrane, all tail beads, lower leaflet phosphates
378 and lower leaflet tails for the FhuA system shown in Figure 6 (Figure S13). It can
379 be seen that in areas of cardiolipin clustering there is an increase of lipid tail
380 enrichment in these regions. For the entire membrane and all lipid tails in regions
381 of cardiolipin clustering local depletion and enrichment were observed, respectively
382 (Figure S13A and B). Given that the lower leaflet headgroups show significant
383 enrichment in such areas (Figure S13C) it can be concluded that the LPS headgroups
384 and sugars groups are depleted in these regions.



385

386 **Figure 6. A) 2D lipid density analysis for the ReLPS outer membrane systems**.

387 Density maps were done for, bottom left) entire Re lipopolysaccharide (ReLPS) lipid, top

388 left) 16:0-18:1 phosphoethanolamine (POPE), top right) phosphatidylglycerol (POPG)

389 and bottom right) cardiolipin (CDL2) lipid phosphate beads and were averaged over 25-

390 30 μs. The density maps are coloured by enrichment (>0) or depletion (<0) of each lipid

391 type with respect to the average density of a given lipid type. B) Depletion-Enrichment

392 (D-E) indices for POPE, POPG and CDL2 for each protein in the ReLPS membranes.

393 The D-E index was obtained by dividing the lipid composition the 1.4 nm shell around

394 the protein by the bulk membrane composition; therefore a D-E index > 1 indicates

395 enrichment, while a D-E index < 1 indicates depletion. The D-E index was calculated

396 from 10-30 μs in 5 μs blocks for both repeats to get eight values for each system; these

397 eight values were then averaged and the error determined from their standard deviation.

398 For further details see the Methods section.

399

400 It is well known that LPS has diffuses at rates of more than order of magnitude slower than

401 most phospholipid systems[9]. The rate of lipid sorting was determined through the lifetime

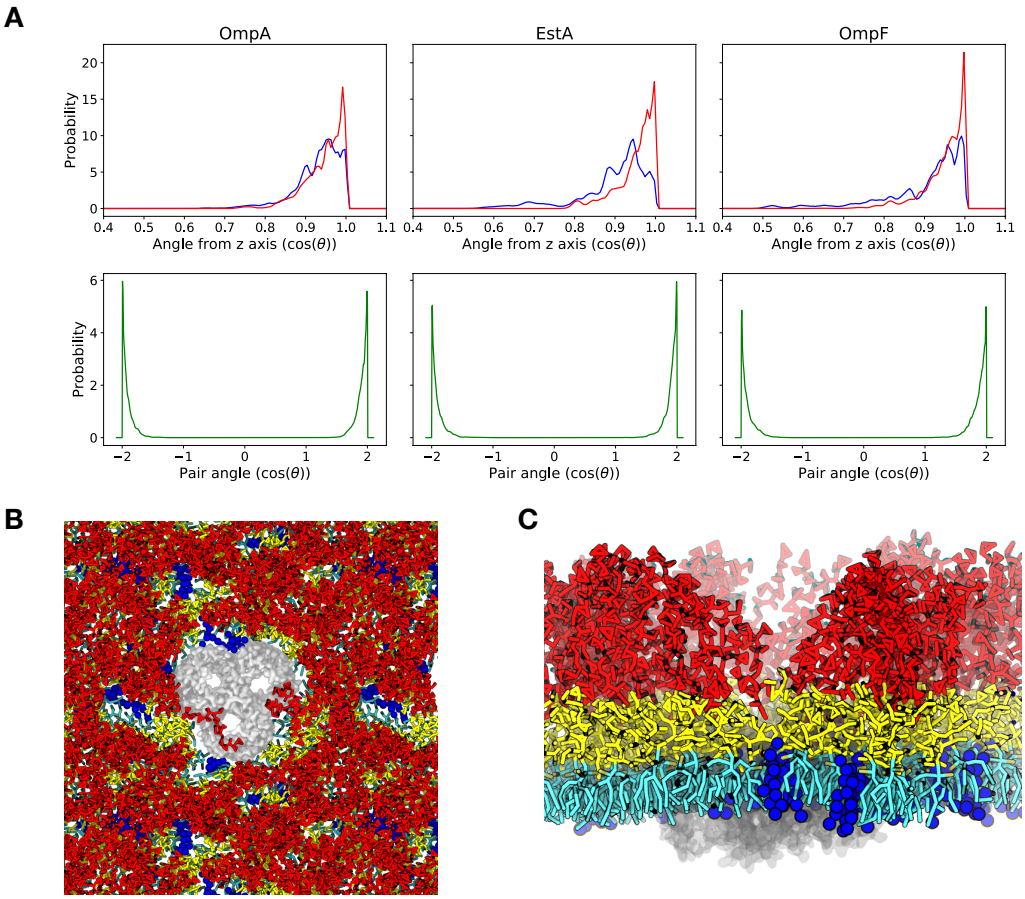
of the lipids at the surface of a protein through the decay of the probability P(t+Δt) as a function of Δt, as described by Horn et.al[15], where P(t+Δt) describes the probability that a lipid is bound at t as well as Δt (Figure S14). The lipid survival probability, P(t+Δt), was then fitted to a double exponential to determine a fast and slow decay constant of lipid turnover (Table S1). To determine these constants P(t+Δt) was determined from 20 – 40 μs and 40 – 60 μs for both repeats of the OmpA RaLPS systems. The probability values were then averaged and fit to a double exponential, obtaining the decay constants of each lipid. The minimum value of the double exponential was shifted by the a constant, C, which was to referred to as the decay limit (Table S1). The decay limit is the minimum value of the exponential in the range of the provided data. The fast decay constants of POPE and POPG were both on the order of magnitude of the resolution of the trajectory (1 ns). The slow decay constants of POPE and POPG (96.4 ± 2.5 and 83.5 ± 3.9 ns) compare well to previous measured slow decay constants of phospholipids in Martini systems[16]. Cardiolipin has a slow decay constant that is around 3 times larger than the other lipids in the lower leaflet. Cardiolipin's slow diffusion is likely a result of its comparatively (to POPE and POPG) large size, combined with the formation of nanodomains of cardiolipin further reducing diffusion[17]. For the lipids in the lower leaflet the decay limits are all close to 0 (Table S1). Whereas for RaLPS the decay limit was 0.96 ± 0.21, indicating that the sorting of LPS does not occur on the timescales of these simulations.

O-antigen orientation within the outer leaflet of the outer membrane.

The O-Antigen chains attached to the RaLPS outer core act as a chemical and

physical barrier to the permeation of molecule across the outer membrane. In previous sections of this paper, we have shown that the patterns of contacts of the six Omps with sugar moieties are protein dependant, here we ask if the Omp-sugar interactions impact upon the behaviour of the O-Antigen chain. To this end we calculated the end-to-end tilt of each O-Antigen chain with respect to the z-axis and distance from the protein. The distances are divided into 'local' and 'bulk', where local is defined as the lipid A section of each smooth LPS molecule being within 2.5 nm of the protein surface, as this is the approximate radius of the annular regions around most membrane proteins, see for example [5,14]. The bulk region is defined as the region beyond the local region. The relative orientation of O-Antigen chains was determined by the calculation of the angle between end-to-end vectors of neighbouring O-Antigen chains. The angles took values in the interval of [2, 0] if the chains tilt toward each other and [-2, 0) if they did not. In both cases the results were presented as probability distributions with respect to the angles. For the O-Antigen systems, bulk region sugar chains rarely have tilts angles greater than 45° ($< 0.71 \cos(\theta)$), whereas in the local region, O-Antigen tilt angles can be as large as 65° (as small as $0.42 \cos(\theta)$) (Figures 7 and S15A). Tilting of O-Antigen chains with respect to the z axis in the local region in general has a positive correlation with the size of the Omps; this observation was particularly pronounced for the OmpA and OmpF, respectively, where the O-antigen chains in the vicinity of OmpF had a considerably larger tilt angle than the chains nearer to the smaller protein, OmpA

445 (Figure 7). It is interesting to note that EstA also showed a greater propensity of O-
446 Antigen tilt at short-ranges even though the transmembrane region of EstA is not
447 much larger than that of OmpA; this observation reflects the extensive interactions
448 of the large extracellular region of EstA with the O-Antigen chains. The pairwise
449 angles were all distributed from -50° to 50° with the most probable relative tilt being
450 ~ ± 25° (± 1.91). Once the most probable angle is reached the distributions undergo
451 a gradual decay to 0, highlighting the packed nature of these system. For every Omp,
452 pairs of neighbouring O-Antigens are more likely to tilt towards each other (tenting),
453 than they are away from each other; this leads to O-Antigen chains clustering
454 together. Similarly, shorter, atomistic simulations reported by Im and co-workers
455 shown tenting of O-Antigen chains around OmpF[18].



456

457 **Figure 7. A) Probability distributions of (top) O-Antigen chain tilt with respect to**
458 **the z axis and tilt between neighbouring pairs of O-Antigen chains for 25-30 μs of**
459 **the OANT systems.** The tilt of a single O-Antigen chain was calculated using its end to
460 end vector. For the chain tilt with respect to the z axis, two probability distributions were
461 generated for smooth lipopolysaccharides (LPS) closer than 2.5 nm (blue line) or further
462 than 2.5 nm away from a given protein's surface (red line). For the pairwise O-Antigen
463 distributions if the chains tilted towards each other the angle was positive; in all other
464 cases the angles would be negative. Angles are presented in units of cos(θ) where θ is an
465 angle. For the pairwise O-Antigen distributions if the chains tilted towards each other
466 cos(θ) takes the range [0, 2]; in all other cases the angles would be between (-2, 0). B)

Snapshot of OANT_PE system with an embedded OmpF taken at 30 μs from above (extracellular side) and C) the same system viewed from the side of the membrane. Colour key: protein = grey, O-Antigen chain = red, LPS core oligosaccharides = yellow, lipid A = cyan and gold and 16:0-18:1 phosphoethanolamine (POPE) lipids in the upper leaflet = blue.

The addition of POPE to the upper leaflet of the outer membrane results in an overall rise in the probability of tilt angles above 45° (< 0.71), especially within the local region of the protein (Figure S15B). Whereas the pair angles seem largely unchanged from the OANT membrane simulations, which maybe suggest that extreme tilts are mirrored by their local environment. Here we focus on the OmpF OANT_PE simulations, as these systems exhibited the largest differences in O-Antigen tilting behaviour when POPE was added to the outer leaflet. Snapshots taken at the 30 μs point from one of the simulations show a few O-antigen chains curved over the extracellular loops of OmpF, but only partially obstructing one of the pores of the trimer (Figure 7); an effect that was seen throughout the trajectory. We determined the average O-antigen contacts per frame with all residue numbers within each repeating unit (1-340) of OmpF of the system in Figure 5 from 25-30 μs (Figure S16). As in the atomistic study of OmpF in an O-antigen containing model the OM reported by Im and co-workers, we observed significant contacts with parts of the loops L4 (residues 160-172), L5 (residues 197-209), L6 (residues

237-253) and L8 (residues 318-330). Compared to the study of Im and co-workers, we see less extreme tenting of O-Antigen chains near OmpF; which is a likely consequence of the O-Antigen chains in this study being constructed of 4 monomer sugar units compared to 5 units in the work of Im, and thus it would not be unreasonable to observe increasing O-Antigen flexibility with an increase in chain length. In simulations of the other Omps presented here we did not observe O-antigen chains to block access to the pore of the Omps and rarely did they tilt over the protein. Furthermore, in simulations in which POPE was present in the outer leaflet, we observed O-antigen chains tilt away from small clusters or even single POPE lipids to preferentially interact with neighbouring sugars. Thus, despite the presence of a thick O-antigen layer in the outer leaflet of the outer membrane, on the basis of our simulations, this layer does not impede access to the proteins embedded within the membrane.

Discussion

We have shown that the six outer membrane proteins we have studied here, OmpA, OmpX, FhuA, BtuB, OmpF and EstA have unique patterns of interaction with the outer membrane. Furthermore, the patterns of interactions are not particularly consistent in terms of protein function or size. In the inner leaflet we find that cardiolipin is depleted within the vicinity of all six Omps, in contrast to findings by ourselves and others that cardiolipin is enriched near some bacterial inner membrane proteins. We observe protein-independent clustering of cardiolipin molecules in the

509 inner leaflet, directly beneath regions of lipopolysaccharide enrichment in the upper
510 leaflet. We previously reported such behaviour close to OmpA and OmpF, the
511 present study suggests this phenomenon is perhaps a more general characteristic of
512 the outer membrane. In the outer leaflet we do not see any trends in protein
513 interactions with LPS even in proteins that are similar in size or functionality. The
514 Omp-LPS interactions vary across proteins, but also for a single protein they vary
515 as a function of the level of lipopolysaccharide in the outer leaflet. For example
516 interactions with basic residues dominate at the ReLPS level for most Omps, but at
517 the RaLPS, in OmpX there are more interactions with acidic residues. These
518 findings can in part be rationalised by the variations in the distribution of amino
519 acids on the outer surface of these proteins. For example, the extracellular loops of
520 OmpA contain fewer lysine residues that point out towards the LPS than OmpX,
521 despite the similar size of the proteins; both beta barrels contain just eight
522 antiparallel strands. These compositional differences impact upon the orientations
523 of the proteins within the different membrane environments as they dictate the
524 pattern of interactions with LPS molecules. In the past it had been thought that the
525 aromatics residues that are in general present in two bands, one near each mouth of
526 the barrel of Omps anchor the protein into the headgroups of the lipids and then the
527 barrels tilts to match the hydrophobic regions on their surface to the hydrophobic
528 width of the surrounding membrane. We show that this is rather too simplistic a
529 view, as the large extracellular loops and their interactions with the sugar moieties
530 of the LPS molecules also play a role in Omp orientation. Furthermore the very

different dynamics of the lipids in the two leaflets also impact upon the tilt angles of these proteins within the outer membranes. In the context of the *in vivo* outer membrane, which is known to be crowded by proteins, it is likely that the orientations of the proteins will also vary as a function of which other proteins are nearby, for the reasons already discussed; namely compositional differences. Interestingly we find that O-antigen chains preferentially tilt towards each other compared to towards Omps. We note that this may be important for recognition of Omps by other molecular species as they navigate through the 'hairy' layer of O-antigen chains of the outer membrane. It is important to place the current work in the appropriate biological context and discuss any limitations. For this study these are associated with protein sample size, and almost inevitably for molecular dynamics simulation, with simulation time and length scales. Of course, there are very many more outer membrane proteins in *E. coli* than those studied here, thus it is possible that there may be some that are very similar in terms of their patterns of LPS interactions. The proteins we have chosen do cover a range of sizes and functions and while we are confident that the hypothesis of unique interactions will hold based on our findings here, it would be interesting to see similar studies of other Omps in future and indeed of Omps in other bacterial species too. The timescales we have achieved here (30-40 μs per simulation) are state-of-the-art, and as the data on tilt angles reveals, they are absolutely necessary to achieve convergence for the slow-moving LPS molecules. It is important to note that even these timescales many not be sufficient to achieve full convergence, indeed, our data

553 on OmpA, in which the simulations were extended to 60 μs suggest longer
554 timescales may still be needed, and in future enhanced sampling methods are likely
555 to be necessary for simulations of LPS-containing membranes. Our simulations
556 show some interesting features of the lipidic components of the outer membrane,
557 namely the clustering of cardiolipins above regions of low LPS density in the outer
558 leaflet and also the presence of clusters of PG lipids which appear to be independent
559 of the protein they surround. In summary, the present study reveals that the six
560 proteins studied here, have unique lipid fingerprints particularly with respect to
561 lipopolysaccharide, and are not in general occluded or engulfed by O-antigen chains
562 of the surrounding membranes. The presence of shells of lipids or annular lipids
563 surrounding proteins has been known for some time, along with the idea that
564 specific protein-interactions often have functional roles[19]. More recently however,
565 for phospholipid membranes a picture has been emerging of these lipid shells having
566 compositions that are unique to the proteins they surround[5, 20]. Here we show
567 evidence, within the caveat of the timescales accessible to us, for unique patterns of
568 interactions for proteins in LPS-containing membranes. We find that these patterns
569 cannot be easily generalised and thus while the composition of the outer leaflet of
570 the outer membrane is almost exclusively composed of LPS, the extent to which
571 native outer membrane proteins interact with the different moieties of LPS seems to
572 be unique to the protein.

573 ## Experimental Procedures

O-antigen parameterisation.

The GROMOS 53A6 force field was used to perform simulations of duration 100 ns of *E. coli* O42 O-antigen chain units (with terminal glucose sugar), solvated with SPC water[21]. Pressure was maintained at 1.0 bar using the Parrinello-Rahman barostat[22], while the temperature was held constant at 313 K using the Nosé-Hoover thermostat[23]. Electrostatic interactions were treated with the smooth particle mesh Ewald algorithm with a short-range cut-off of 0.9 nm, while the van der Waals interactions were truncated at 1.4 nm with long-range dispersion corrections for energy and pressure[24]. All bonds were constrained with the LINCS algorithm, enabling a timestep of 2 fs[25].

Subsequently, the united-atom particles were mapped into coarse-grained pseudo-atoms, and equilibrium bond lengths and bond angles (for the coarse-grained O-antigen model) were determined from the relative conformations of these pseudo-atoms, while the corresponding force constants were resolved from their dynamic movement. The coarse-grained system was then simulated in water for 100 ns and the resulting simulation data was compared with the united atom simulations of equivalent systems, enabling further fine-tuning of the coarse-grained parameter set. Thereafter the coarse-grained parameters (bond lengths, angles and dihedrals) were iteratively optimized as additional coarse-grained solvation simulations were performed with adjusted parameter sets until agreement was obtained between the similar coarse-grained and fine-grain solvation simulations within the range used

595 for our original coarse-grained model of LPS (additional details are provided in the
596 supplementary information, Figure S17). In other words, we manually adjusted the
597 parameters slightly to achieve a better agreement with the united-atom model. The
598 resulting O-antigen model and parameter set were combined with our pre-existing
599 coarse-grained model and parameter set for *E. coli* RaLPS[26]. The parameters
600 developed for the O-antigen model are available from the Martini-maker module of
601 the Charmmgui website and are also provided in the supplementary information of
602 this paper[27].

603 **Membrane construction.** Five different membrane models were simulated: 4 outer
604 membrane systems and a 16:0 dipalmitoylphosphocholine (DPPC) system. The
605 inner leaflet in all four of the outer membrane models was composed of 90% 16:0–
606 18:1 palmitoylphosphoethanolamine (POPE), 5% 16:0–18:1
607 palmitoylphosphatidylglycerol (POPG) and 5% cardiolipin. The outer leaflet was
608 comprised of either 100 % rough lipopolysaccharide (LPS) molecules; at either Re
609 or Ra levels or smooth LPS which incorporates O-antigen, or a mixture of smooth
610 LPS and POPE in a 4:1 ratio. The membranes were constructed using the Martini-
611 maker module of the charmmgui website[27].

612 **Simulation setup.** Initial protein structures were obtained from the RCSB protein
613 databank, they are OmpA: 1QJP[28], OmpX: 1QJ8[29], OmpF: 3POX[30], BtuB: 1NQE[31],
614 FhuA: 1QFG [32] and EstA: 3KVN [33]. Any missing residues or broken loops where
615 added with Modeller[34] and then the protein was given a transmembrane orientation

616 determined by MemEmbed[35]. All initial membrane protein structures were generated
617 with the Martini-maker module of CHARMM-GUI [27]. Each protein model was
618 coarse-grained to the ElNeDyn model with an elastic network cut-off of of 0.9 nm
619 and force constant of 500 $kJmol^{-1}nm^{-2}$ [36]. For each system the box height was set such
620 that the distance above and below the membrane was 5 nm. For the smaller proteins
621 the initial box dimensions in the xy plane were around 13x13 nm, while for larger
622 proteins (OmpF and EstA) initial box dimensions were as large as 17x17 nm (see
623 Table 1). The core of LPS was neutralised with Ca^{2+} ions, while the rest of the charge
624 deficit was removed via the addition of Na^{+} ions. Following this 0.15 M of NaCl
625 was added to each system to give biologically relevant salt concentrations.

626

627 **Simulation protocols.** All simulations were performed using the GROMACS
628 package version 2016 and the Martini version 2.2 force field[37,38]. All simulations were
629 carried out at 323 K with a stochastic velocity rescale thermostat with a coupling
630 constant of 1.0 ps. Initial minimisations followed the CHARMM-GUI protocol[39],
631 which involved an initial steepest decent minimisation followed by a series of
632 equilibrations with 5 fs to 10 fs timesteps for a total of 20 ns with a Berendsen
633 barostat (4.0 ps coupling constant)[40]. During these equilibrations the position of the
634 protein was restrained with 1000 $kJmol^{-1}nm^{-2}$ harmonic restraints on the backbone
635 atoms. Following this production runs of up to 40 μs (see Table 1) were generated
636 with a 10 fs timestep and analysed as described elsewhere. For all production runs,

637 the Parrinello-Rahman semi-isotropic barostat was used with a 12.0 ps coupling
638 constant [22]. Repeat simulations were performed for all systems by assigning random
639 velocities from the stage after which the initial system had been generated from
640 CHARMM-GUI[27].

641

642 The Lennard-Jones potential was cut off using a Potential shift Verlet scheme at
643 long ranges. The reaction field method was used for the electrostatics, with
644 dielectric constants of 15 and infinity for charge screening in the short- and long-
645 range regimes, respectively (as recommended for use with Martini2). The short-
646 range cut-off for both the non-bonded and electrostatic interactions was 1.2 nm. The
647 non-bonded Neighbour lists were updated every 10 steps.

648

649 **Analysis.** All analysis was carried over 25-30μs, unless stated otherwise, using a
650 combination of GROMACS tools and in-house scripts which utilised the
651 MDAnalysis python module[41 42]. To calculate the tilt of each protein the principal axis
652 of inertia of the transmembrane region of the protein backbone, or for OmpF the
653 backbone entire timer, was determined. The protein tilt was then obtained by the
654 angle between the previously obtained principal axis and the z axis. The relative tilt
655 angle, θ, was defined as the angle between the vector **R** and the z axis (Figure S4);
656 Where **R** is the shortest vector from the center of geometry of a given residue

backbone to the principal axis for the protein used for the protein tilt. As shown in Figure S4, if $\theta < 90°$ then **R** is tilting away from the membrane, whereas the opposite is true if $\theta > 90°$. It is worth noting that this method is invariant to the rotation of the beta barrel in the xy plane. The residues used for the relative angles were THR 144, SER 469 and GLY232 for OmpA, EstA and OmpF, respectively.

Contact histograms and maps between a single protein and a range of target atoms were determined with in-house scripts that used a cut-off of 0.6 nm. If multiple contacts were present between a protein residue and a single target atom only one contact was recorded. The membrane thickness was determined with an in-house script that utilised Voronoi diagrams to get the membrane thickness, in the same manner as the APL@Voro analysis tool [43]. The thickness was measured using only the phosphate beads in each leaflet. The depletion/enrichment indices were determined by first counting the number of lipids with a center of geometry within 1.4 nm and then comparing this number to the number expected in the bulk of the membrane using the procedure described by Corradi et al [5]. The enrichment maps were generated by first determining the 2D density map of the membrane using tools developed by Castillo et al[44]. Following this enrichment was determined using the procedure described by Corradi et al [5]. The D-E index was obtained by dividing the lipid composition the 1.4 nm shell around the protein by the bulk membrane composition; therefore a D-E index > 1 indicates enrichment, while a D-E index < 1 indicates depletion. The D-E index was calculated from 10-30 μs in 5 μs blocks

679 for both repeats to get eight values for each system; these eight values were then
680 averaged and the error determined from their standard deviation (as in [5]). Any
681 images of systems seen here were generated with VMD[45].
682
683 The tilt of each O-Antigen chain was determined by first generating an end to end
684 vector for each chain that spanned from the sugar connected to the LPS core to the
685 end of the chain (n.b this is a normalised vector). The tilt angle of each chain was
686 the angle between the end to end vector of the chain and the z axis. Each chain was
687 then split into two groups based on whether the center of geometry Lipid A
688 headgroup of each LPS was closer than 2.5 nm to the protein surface. The values in
689 a single group were then used to generate a probability distribution that was
690 normalised to have a total integral of one. The pairwise O-Antigen angle probability
691 distributions were determined by first finding all pairs of O-Antigen chains within
692 a cut-off of 0.6 nm from each other. Then the angle between the end to end vectors
693 of each pair of chains was determined in units of cos(θ). If the O-Antigen chains
694 were tilting away from each other, the tilt angle was between [-2, 0); otherwise the
695 tilt angle was between [2, 0]. The scripts developed for this paper are available on
696 GitHub (https://github.com/js1710/LPS_fingerprints_paper).

697 ## Associated Content

698 **Supporting information**

699 The supporting information is available free of charge on the ACS publication
700 website at doi: xxx.

701 Contains further plots, tables and images considered too long for the main text. In
702 most cases figures in the supporting information are the rest of the plots for analysis
703 included in the main text for only a subset of the systems in this paper (Supporting
704 Information.pdf).

705 The parameters of the O-Antigen model developed here(martini_lps_oant.itp).

706

707

708 ## Funding Sources

709 Jonathan Shearer is funded by EPSRC grant number EP/L015722/1.

710 ## Acknowledgment

711 We acknowledge the use of the University of Southampton high performance
712 computing resources, Iridis5. We are grateful to the UK Materials and Molecular
713 Modelling Hub for computational resources, which is partially funded by EPSRC
714 (EP/P020194/1).

715 ## References

1. (a) Koebnik, R.; Locher, K. P.; Van Gelder, P., Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* **2000,** *37* (2), 239-253; (b) Delcour, A. H., Structure and function of pore-forming beta-barrels from bacteria. *Journal of Molecular Microbiology and Biotechnology* **2002,** *4* (1), 1-10.
2. (a) Khalid, S.; Berglund, N. A.; Holdbrook, D. A.; Leung, Y. M.; Parkin, J., The membranes of Gram-negative bacteria: progress in molecular modelling and simulation. *Biochemical Society transactions* **2015,** *43* (2), 162-7; (b) Boags, A.; Hsu, P. C.; Samsudin, F.; Bond, P. J.; Khalid, S., Progress in Molecular Dynamics Simulations of Gram-Negative Bacterial Cell Envelopes. *J Phys Chem Lett* **2017,** *8* (11), 2513-2518.
3. Stansfeld, P. J.; Goose, J. E.; Caffrey, M.; Carpenter, E. P.; Parker, J. L.; Newstead, S.; Sansom, M. S., MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. *Structure* **2015,** *23* (7), 1350-61.
4. (a) Piggot, T. J.; Holdbrook, D. A.; Khalid, S., Conformational dynamics and membrane interactions of the E. coli outer membrane protein FecA: a molecular dynamics simulation study. *Biochimica et biophysica acta* **2013,** *1828* (2), 284-93; (b) Lee, J.; Patel, D. S.; Kucharska, I.; Tamm, L. K.; Im, W., Refinement of OprH-LPS Interactions by Molecular Simulations. *Biophys J* **2017,** *112* (2), 346-355; (c) Holdbrook, D. A.; Piggot, T. J.; Sansom, M. S.; Khalid, S., Stability and membrane interactions of an autotransport protein: MD simulations of the Hia translocator domain in a complex membrane environment. *Biochimica et biophysica acta* **2013,** *1828* (2), 715-23; (d) Balusek, C.; Gumbart, J. C., Role of the Native Outer-Membrane Environment on the Transporter BtuB. *Biophys J* **2016,** *111* (7), 1409-1417.
5. Corradi, V.; Mendez-Villuendas, E.; Ingolfsson, H. I.; Gu, R. X.; Siuda, I.; Melo, M. N.; Moussatova, A.; DeGagne, L. J.; Sejdiu, B. I.; Singh, G.; Wassenaar, T. A.; Magnero, K. D.; Marrink, S. J.; Tieleman, D. P., Lipid-Protein Interactions Are Unique Fingerprints for Membrane Proteins. *Acs Central Sci* **2018,** *4* (6), 709-717.
6. Kirschner, K. N.; Lins, R. D.; Maass, A.; Soares, T. A., A Glycam-Based Force Field for Simulations of Lipopolysaccharide Membranes: Parametrization and Validation. *Journal of Chemical Theory and Computation* **2012,** *8* (11), 4719-4731.
7. Scott, K. A.; Bond, P. J.; Ivetac, A.; Chetwynd, A. P.; Khalid, S.; Sansom, M. S., Coarse-grained MD simulations of membrane protein-bilayer self-assembly. *Structure* **2008,** *16* (4), 621-30.
8. Marrink, S. J.; Tieleman, D. P., Perspective on the Martini model. *Chemical Society Reviews* **2013,** *42* (16), 6801-6822.
9. Piggot, T. J.; Holdbrook, D. A.; Khalid, S., Electroporation of the E. coli and S. Aureus membranes: molecular dynamics simulations of complex bacterial membranes. *The journal of physical chemistry. B* **2011,** *115* (45), 13381-8.
10. (a) Ortiz-Suarez, M. L.; Samsudin, F.; Piggot, T. J.; Bond, P. J.; Khalid, S., Full-Length OmpA: Structure, Function, and Membrane Interactions Predicted by Molecular Dynamics Simulations. *Biophys J* **2016,** *111* (8), 1692-1702; (b) Marcoux, J.; Politis, A.; Rinehart, D.; Marshall, D. P.; Wallace, M. I.; Tamm, L. K.; Robinson, C. V., Mass spectrometry defines the C-terminal dimerization domain and enables modeling of the structure of full-length OmpA. *Structure* **2014,** *22* (5), 781-90.

11. Ferguson, A. D.; Hofmann, E.; Coulton, J. W.; Diederichs, K.; Welte, W., Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. *Science* **1998,** *282*, 2215-2220.
12. Arunmanee, W.; Pathania, M.; Solovyova, A. S.; Le Brun, A. P.; Ridley, H.; Basle, A.; van den Berg, B.; Lakey, J. H., Gram-negative trimeric porins have specific LPS binding sites that are essential for porin biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **2016,** *113* (34), E5034-E5043.
13. Holdbrook, D. A.; Huber, R. G.; Piggot, T. J.; Bond, P. J.; Khalid, S., Dynamics of Crowded Vesicles: Local and Global Responses to Membrane Composition. *PloS one* **2016,** *11* (6), e0156963.
14. Shearer, J.; Khalid, S., Communication between the leaflets of asymmetric membranes revealed from coarse-grain molecular dynamics simulations. *Sci Rep* **2018,** *8* (1), 1805.
15. Horn, J. N.; Kao, T. C.; Grossfield, A., Coarse-Grained Molecular Dynamics Provides Insight into the Interactions of Lipids and Cholesterol with Rhodopsin. *Adv Exp Med Biol* **2014,** *796*, 75-94.
16. Horn, J. N., Exploring Lipid-Peptide Interactions Through Computer Simulation of Antimicrobial Lipopeptides and a G Protein-Coupled Receptor. *PhD thesis* **2013,** *University of Rochester, New York*.
17. Koldso, H.; Shorthouse, D.; Helie, J.; Sansom, M. S. P., Lipid Clustering Correlates with Membrane Curvature as Revealed by Molecular Simulations of Complex Lipid Bilayers. *PLoS computational biology* **2014,** *10* (10).
18. Patel, D. S.; Re, S.; Wu, E. L.; Qi, Y.; Klebba, P. E.; Widmalm, G.; Yeom, M. S.; Sugita, Y.; Im, W., Dynamics and Interactions of OmpF and LPS: Influence on Pore Accessibility and Ion Permeability. *Biophys J* **2016,** *110* (4), 930-8.
19. (a) Lee, A. G., How lipids affect the activities of integral membrane proteins. *Bba-Biomembranes* **2004,** *1666* (1-2), 62-87; (b) Lee, A. G., Lipid-protein interactions in biological membranes: a structural perspective. *Biochim. Biophys. Acta* **2003,** *1612*, 1-40.
20. Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V., Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **2014,** *510* (7503), 172-+.
21. (a) Oostenbrink, C.; Soares, T. A.; van der Vegt, N. F.; van Gunsteren, W. F., Validation of the 53A6 GROMOS force field. *European biophysics journal : EBJ* **2005,** *34* (4), 273-84; (b) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J., *Intermolecular Forces*. Reidel: Dordrecht, 1981.
22. Parrinello, M.; Rahman, A., Polymorphic transitions in single-crystals - a new molecular dynamics method. *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.
23. (a) Nose, S.; Klein, M. L., Constant Pressure Molecular-Dynamics for Molecular-Systems. *Mol. Phys.* **1983,** *50* (5), 1055-1076; (b) Hoover, W. G., Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev.* **1985,** *A31*, 1695-1697.
24. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995,** *103* (19), 8577-8593.
25. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.* **1997,** *18*, 1463-1472.

26. Hsu, P. C.; Jefferies, D.; Khalid, S., Molecular Dynamics Simulations Predict the Pathways via Which Pristine Fullerenes Penetrate Bacterial Membranes. *The journal of physical chemistry. B* **2016,** *120* (43), 11170-11179.
27. Hsu, P. C.; Bruininks, B. M. H.; Jefferies, D.; de Souza, P. C. T.; Lee, J.; Patel, D. S.; Marrink, S. J.; Qi, Y. F.; Khalid, S.; Im, W., CHARMM-GUI Martini Maker for modeling and simulation of complex bacterial membranes with lipopolysaccharides. *Journal of Computational Chemistry* **2017,** *38* (27), 2354-2363.
28. Pautsch, A.; Schulz, G. E., High-resolution structure of the OmpA membrane domain. *J. Mol. Biol.* **2000,** *298*, 273-282.
29. Vogt, J.; Schulz, G. E., The structure of the outer membrane protein OmpX from *Escherichia coli* reveals possible mechanisms of virulence. *Structure Fold. Des.* **1999,** *7* (10), 1301-1309.
30. Efremov, R. G.; Sazanov, L. A., Structure of Escherichia coli OmpF porin from lipidic mesophase. *J Struct Biol* **2012,** *178* (3), 311-8.
31. Chimento, D. P.; Kadner, R. J.; Wiener, M. C., The Escherichia coli outer membrane cobalamin transporter BtuB: structural analysis of calcium and substrate binding, and identification of orthologous transporters by sequence/structure conservation. *J Mol Biol* **2003,** *332* (5), 999-1014.
32. Ferguson, A. D.; Welte, W.; Hofmann, E.; Lindner, B.; Holst, O.; Coulton, J. W.; Diederichs, K., A conserved structural motif for lipopolysaccharide recognition by procaryotic and eucaryotic proteins. *Struct. Fold. Des.* **2000,** *8*, 585-592.
33. van den Berg, B., Crystal structure of a full-length autotransporter. *J Mol Biol* **2010,** *396* (3), 627-33.
34. Sanchez, R.; Sali, A., Comparative protein structure modeling. Introduction and practical examples with modeller. *Methods in molecular biology* **2000,** *143*, 97-129.
35. Nugent, T.; Jones, D. T., Membrane protein orientation and refinement using a knowledge-based statistical potential. *BMC Bioinformatics* **2013,** *14*, 276.
36. Periole, X.; Cavalli, M.; Marrink, S. J.; Ceruso, M. A., Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *J Chem Theory Comput* **2009,** *5* (9), 2531-43.
37. Abraham, M. J.; van der Spoel, D.; Lindahl, E.; Hess, B.; team., a. t. G. d., GROMACS User Manual version 2016. *www.gromacs.org* **2018**.
38. de Jong, D. H.; Singh, G.; Bennett, W. F.; Arnarez, C.; Wassenaar, T. A.; Schafer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J., Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J Chem Theory Comput* **2013,** *9* (1), 687-97.
39. Kumar, R.; Iyer, V. G.; Im, W., CHARMM-GUI: A graphical user interface for the CHARMM users. *Abstr Pap Am Chem S* **2007,** *233*, 273-273.
40. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R., Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984,** *81* (8), 3684-3690.
41. Abraham, M. J.; Murtola, T.; Schulz, R.; Pall, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *Software X* **2015,** *1-2*, 19-25.

42. Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O., MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* **2011,** *32* (10), 2319-27.
43. Lukat, G.; Kruger, J.; Sommer, B., APL@Voro: a Voronoi-based membrane analysis tool for GROMACS trajectories. *J Chem Inf Model* **2013,** *53* (11), 2908-25.
44. Castillo, N.; Monticelli, L.; Barnoud, J.; Tieleman, D. P., Free energy of WALP23 dimer association in DMPC, DPPC, and DOPC bilayers. *Chemistry and physics of lipids* **2013,** *169*, 95-105.
45. Humphrey, W.; Dalke, A.; Schulten, K., VMD - Visual Molecular Dynamics. *J. Molec. Graph.* **1996,** *14*, 33-38.