OmpA: A Flexible Clamp for Bacterial Cell Wall Attachment

Firdaus Samsudin[1], Maite L. Ortiz-Suarez[1], Thomas J. Piggot[1,2], Peter J. Bond[3,4]*, and Syma Khalid[*1]

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

[2]CBR Division, Defence Science and Technology Laboratory, Porton Down, Salisbury, Wiltshire SP4 0JQ

[3]Bioinformatics Institute (A*STAR), 30 Biopolis Street, 07-01 Matrix, 138671 Singapore

[4]National University of Singapore, Department of Biological Sciences, 14 Science Drive 4, Singapore 117543

*E-mail: S.Khalid@soton.ac.uk, peterjb@bii.a-star.edu.sg; Corresponding authors

Summary

The envelope of Gram-negative bacteria is highly complex, containing separate outer and inner membranes and an intervening periplasmic space encompassing a peptidoglycan (PGN) cell wall. The PGN scaffold is anchored non-covalently to the outer membrane via globular OmpA-like domains of various proteins. We report atomically detailed simulations of PGN bound to OmpA in three different states, including the isolated C-terminal domain (CTD), the full-length monomer, or the complete full-length dimeric form. Comparative analysis of dynamics of OmpA CTD from different bacteria helped to identify a conserved PGN binding mode. The dynamics of full-length OmpA, embedded within a realistic representation of the outer membrane containing full-rough (Ra) lipopolysaccharide, phospholipids, and cardiolipin, suggested how the protein may provide flexible mechanical support to the cell wall, and the mechanism by which dimerisation occurs may help maintain the integrity of the cell wall. An accurate model of the heterogeneous bacterial cell envelope should facilitate future efforts to develop novel antibacterial agents.

Introduction

The bacterial cell wall plays a critical role in maintaining the structural integrity of the cell, providing a surrounding mechanical support that protects it from high internal osmotic pressures (Höltje 1998; Silhavy et al. 2010). The major components of the bacterial cell wall are long glycan chains, covalently cross-linked by short peptides, forming a flexible network called the peptidoglycan (PGN) (Vollmer et al. 2008). The glycan scaffold comprises alternating β-1,4-linked N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) subunits, with the latter attached to a pentapeptide of variable sequence (Schleifer & Kandler 1972). Additionally, in some species such as *Escherichia coli*, the terminal residues of the pentapeptide can be degraded to produce peptidoglycan chains with NAM linked to di, tri, or tetrapeptide sequences (Vollmer & Bertsche 2008). Apart from these various peptides, the length of each glycan strand can also differ substantially, with those in *E. coli* typically between 25-35 disaccharide units long (Harz et al. 1990). As such, PGN is a highly complex and heterogeneous mesh of polymers. Since this structure is unique to bacteria, PGN and the enzymes involved in PGN biosynthesis have become primary targets for antibiotics including, for example, the β-lactams and glycopeptides (Kohanski et al. 2010). A detailed understanding of the molecular organisation of the PGN within the cell envelope is therefore crucial for further development and refinement of novel therapeutic agents to fight bacterial infections.

In Gram-negative bacteria, PGN is found in the periplasmic space, delineated by the inner and the outer membranes. The outer membrane is covalently attached to the PGN network via the Braun's lipoprotein (Lpp) through a peptide bond between the ε-amino group of the C-terminal lysine of the protein and the carboxyl end of the diaminopimelate (DAP) residue on the peptide side chain of the PGN (Braun 1975). Apart from Lpp, PGN is also anchored non-covalently to proteins via OmpA-like domains. These include three distinct protein families: i) the integral outer membrane proteins, e.g. OmpA, ii) PGN-associated lipoproteins, e.g. PAL, and iii) flagellar motor proteins, e.g. MotB (Koebnik 1995; Parsons et al. 2006; Roujeinikova 2008). The outer membrane protein OmpA from *E. coli* (*Ec*OmpA) is perhaps the most extensively studied amongst these proteins (Smith et al. 2007). OmpA is a multi-domain protein whose N-terminus is composed of a membrane-embedded β-barrel, and whose C-terminus contains a globular periplasmic domain, which are linked together by an unstructured loop. Structures of several OmpA homologues from different bacteria have been determined using various biophysical methods (Pautsch & Schulz 1998; Arora et al. 2001; Grizot & Buchanan 2004; Døvling Kaspersen et al. 2014; Marcoux et al. 2014), although very little is currently known about how they interact with the network of PGN. Recently, the crystal structure of the C-terminal domain (CTD) of OmpA from the bacterium *Acitenobacter baumannii* (*Ab*OmpA) bound to a PGN-derived pentapeptide was solved (Park et al. 2012), providing a first glimpse into the molecular mechanism of PGN attachment.

Nevertheless, the molecular-level details concerning the relative arrangement of individual components of the cell envelope remain unclear, due to both experimental and computational difficulties associated with studyingthis complex system,. For example, since PGN is covalently linked to Lpp, it is not clear why the cell requires the non-covalent attachment to OmpA. It is therefore of immense interest to understand the structural details of this interaction in the context of the surrounding cell envelope. While OmpA is generally investigated in its monomeric form, it can also form homodimers, as has shown using SDS-PAGE analysis (Stenberg et al. 2005) and *in vivo* crosslinking studies (Zheng et al. 2011). Again, however, the functional relevance and molecular details of dimerisation vis-à-vis PGN binding are sparse.

In the present study, we have developed a new atomistic model for PGN. Based on the X-ray structure of *Ab*OmpA, we modelled CTDs from various members of the OmpA-family, and studied the structure and dynamics of their interactions with PGN. A model of full-length *E. coli* OmpA (*Ec*OmpA) in monomeric and dimeric forms was subsequently embedded within a realistic outer membrane environment, to determine the influence of oligomerization and membrane association upon PGN binding. Key conserved interactions observed between PGN and OmpA-like domains lead us to propose a universally conserved binding mechanism. Monomeric PGN-bound OmpA CTD became distorted and interacted with the outer membrane during simulation. Such dynamics were absent for the OmpA dimer; we thus

hypothesize that dimerisation may be important for maintaining the structural integrity of the PGN network.

Table 1: Summary of Simulations

System	Ionic Strength	Simulation Time (ns)	Number of Repeats
*Ab*OmpA CTD	Counter-ions only	500	3
*Se*OmpA CTD	Counter-ions only	500	3
*Nm*RmpM CTD	Counter-ions only	500	3
*Ec*OmpA CTD	Counter-ions only	500	3
*Ec*OmpA full-length monomer	Counter-ions only	500	2
*Ec*OmpA full-length homodimer	Counter-ions only	500	2
*Ec*OmpA full-length homodimer	1 M $MgCl_2$	500	2

Results

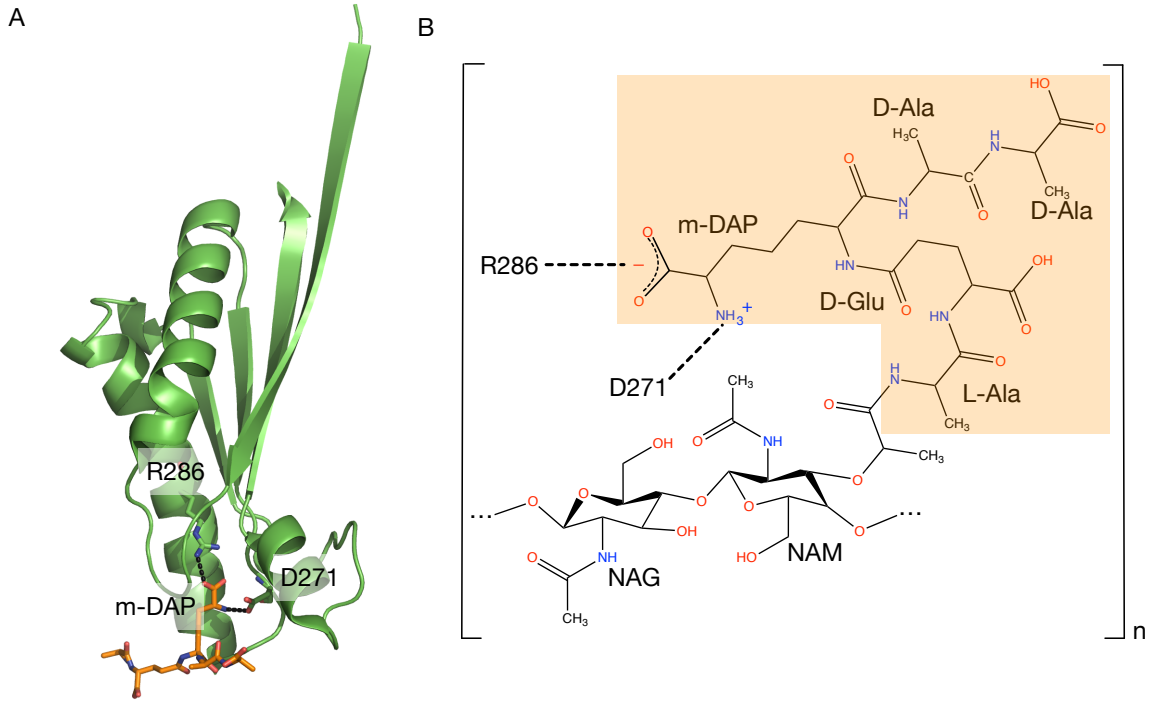
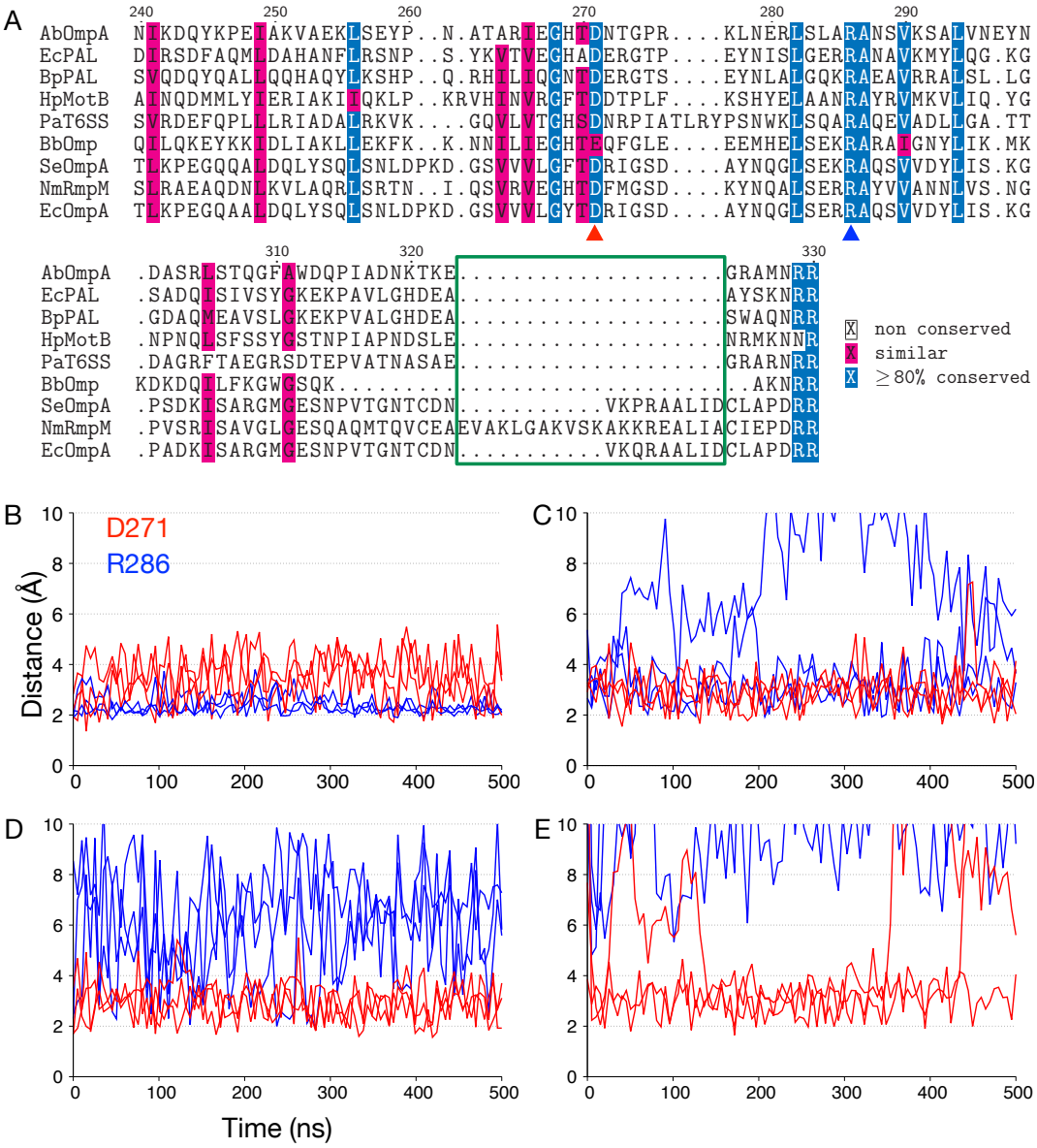


Figure 1. PGN binding to the CTD of *Ab*OmpA. (A) The crystal structure of OmpA CTD from the bacterium *A. baumanni* bound to a pentapeptide from PGN (PDB: 3TD5) (Park et al., 2012). Two residues, D271 and R286, which are key to coordinating the peptide in the binding site, are shown. (B) The chemical structure of one repeat unit of PGN with the same peptide side chain as found in the crystal structure (highlighted in orange). The contact points of PGN (via its m-DAP residue) with the CTD of *Ab*OmpA are indicated.

Development of PGN parameters. In this study, we developed a new set of atomistic parameters for the simulation of PGN. To test the reliability of these parameters to study PGN binding, we simulated the crystal structure of *Ab*OmpA CTD (PDB: 3TD5) (Park et al. 2012) in complex with a PGN molecule of two repeat units in length. The PGN molecule was docked into the

CTD binding site based on the position of the pentapeptide L-Ala-gamma-D-Glu-m-DAP-D-Ala-D-Ala found in the crystal structure (Figure 1). The crystal structure suggests that the interaction between PGN and the OmpA CTD is primarily facilitated by two key residues, D271 and R286, which respectively form salt bridges with the amino and carboxylate moieties of the m-DAP residue in PGN. To probe the stability of PGN binding, we measured the maintenance of these salt bridges throughout three 500 ns simulation replicas. It was expected that, with a reliable set of parameters, the PGN strand should bind stably to the OmpA CTD binding site. Indeed, in all three repeats, these salt bridges persisted (Figure 2B), suggesting that our PGN parameters are sufficiently accurate to study the binding of PGN to the OmpA CTD.



140

141 **Figure 2. Binding mode of PGN is conserved across species.** (A) Sequence
142 alignment of *Ab*OmpA with eight other proteins containing the OmpA-like
143 domain, i.e. PGN associated lipoprotein (PAL) from *E. coli* (*Ec*PAL) (PDB:
144 1OAP), PAL from *Burkholderia pseudomallei* (*Bp*PAL) (4B5C), flagellar
145 motor protein MotB from *Helicobacter pylori* (*Hp*MotB) (3CYP), cell wall
146 anchor from *Pseudomonas aeruginosa* (*Pa*T6SS) (4B62), outer membrane
147 protein of *Borrelia burgdorferi* (*Bb*Omp) (3OON), OmpA from *Salmonella*
148 *enterica* (*Se*OmpA) (4RHA), OmpA-like domain from *Neisseria meningitides*
149 (*Nm*RmpM) (1R1M), and OmpA from *E. coli* (*Ec*OmpA). A large insert

present only in *Se*OmpA, *Nm*RmpM and *Ec*OmpA, highlighted by a green box, contributes to the flexibility of the binding site. Two key residues in PGN binding, D271 and R286, are indicated by the red and blue triangles respectively. Conserved or similar residues are marked in light blue and magenta respectively. (B) Distance between two key residues and PGN during the 500 ns simulation. Measurement was made between: the carboxylate group of the protein aspartate residue and the amino group of the PGN m-DAP residue; and between the guanidinium group of the protein arginine residue and the carboxylate group of the PGN m-DAP residue. Equivalent measurements were performed for simulations of *Se*OmpA (C), *Nm*RmpM (D) and *Ec*OmpA (E). For each system, three independent simulation replicas were conducted. In (E) only one blue line is visible since the distance between the R286 equivalent residue and PGN in two other repeats exceeded 10 Å.

PGN binding is conserved across OmpA family members. It is not currently known whether the binding of PGN to other OmpA family members follows the same mechanism as observed in the crystal complex of *Ab*OmpA CTD, as this is the only high-resolution structural information available to date. To answer this question, we performed similar simulations with three other members of the family: i) OmpA from *S. enterica* (*Se*OmpA), ii) RmpM from *N. meningitides* (*Nm*RmpM), and iii) *Ec*OmpA. The former two proteins have *apo* crystal structures available for their CTDs (PDB: 4ERH and 1R1M respectively), whilst the latter was modelled based on the crystal structure of *Se*OmpA, which shares ~94% sequence identity. It is worth noting that while the conservation of amino acid sequences across the OmpA family is generally poor, the two key residues for PGN binding as characterised by the crystal

176 structure (Park et al. 2012) are universally conserved (Figure 2A). We therefore
177 again measured the distance between these two residues and the PGN as a
178 metric of the strength of binding, during a series of triplicate 500 ns simulation
179 studies for each family member. We found that in both *Se*OmpA and
180 *Nm*RmpM, the D271 equivalent residues formed persistent salt bridge
181 interactions with the modelled PGN, throughout each trajectory (Figure 2C and
182 2D). The R286 equivalent residues also interacted with PGN, albeit to a lesser
183 extent; continuous salt bridges were observed in two out of three simulations of
184 *Se*OmpA and in one out of three simulations of *Nm*RmpM. In the simulations
185 of *Ec*OmpA, on the other hand, only the D271 equivalent residue (D241)
186 formed a stable interaction with the PGN molecule (Figure 2E). Intermittent
187 salt bridges were observed for the R286 equivalent residue (R256) but these
188 mostly lasted for around 10-20 ns. The lack of stable interactions in these
189 simulations likely stemmed from the less robust starting structure, i.e. the use
190 of a homology model instead of a high-resolution crystal structure as in the
191 other simulations. Despite the subtle differences in these key interactions, it is
192 noteworthy that in all simulations the PGN strand remained tightly bound to
193 the protein and did not dissociate into the bulk solution.

194 One major difference found in the sequence of *Se*OmpA, *Nm*RmpM and
195 *Ec*OmpA compared to *Ab*OmpA is a large insert between residues 320 and 330
196 (Figure 2A). Interestingly, this region comprised the least stable part of the
197 protein during our simulations, as indicated by the large root mean square
198 fluctuations (RMSF) (Figure S1). This is in agreement with a recent NMR

study, which suggested that this region in *Ec*OmpA has a lower melting temperature than the rest of the protein and therefore a higher flexibility (Ishida et al. 2014). It is possible that the presence of this mobile insert in certain members of the OmpA family may result in weaker binding of PGN due to its close proximity to the binding site, which could explain the less stable salt bridge interaction observed for the R286 equivalent residues in *Se*OmpA, *Nm*RmpM and *Ec*OmpA compared to *Ab*OmpA. Taken together, our results suggest that PGN binding is conserved across species, but the strength of binding is likely to differ owing to subtle differences in protein fold.

Monomeric PGN-bound OmpA monomer binds to the outer membrane. To study PGN binding in a more realistic environment, the full-length *Ec*OmpA monomer model (Marcoux et al. 2014) was simulated within an asymmetric outer membrane containing lipopolysaccharide (LPS) in the outer leaflet, and a mixture of phospholipids (1-palmitoyl,2-cis-vaccenyl-phosphatidyl ethanolamine (PVPE), 1-palmitoyl,2-cis-vaccenyl-phosphatidyl glycerol (PVPG) and cardiolipin (DPG)) in the inner leaflet. The N-terminal β-barrel was first embedded in the outer membrane model. One strand of PGN of ten repeat units length was then docked onto the CTD following the crystal structure of *Ab*OmpA-pentapeptide complex (Park et al. 2012). This PGN strand was much longer than the one used in the CTD-only simulations, such that the two ends of the chain were in close proximity with their periodic

221 copies. To imitate the mesh of PGN in the cell wall, these two ends were linked
222 together across the periodic boundary with a harmonic bond. Sufficient Mg^{2+}
223 ions (around 0.3 M) were then added to neutralise the system, and to satisfy the
224 essential role of Mg^{2+} for maintenance of the integrity of the LPS-containing
225 membrane (Clifton et al. 2015). Two independent repeats of 500 ns simulation
226 were performed.

227 We first probed the stability of PGN binding by measuring the distance
228 between two key residues, D241 and R256, with the PGN (Figure 3A). In both
229 independent repeats, D241 formed strong, stable interactions with the PGN, as
230 previously observed in the CTD-only simulations, suggesting that the inclusion
231 of the N-terminal domain and the linker region did not noticeably degrade the
232 stability this key salt bridge. Intriguingly, compared to the CTD-only
233 simulations, the R256 residue formed more stable interactions with PGN in this
234 configuration, as indicated by the distance remaining ~5 Å for most of the 500
235 ns simulations. This result hints at possible stabilisation of the binding pocket
236 when PGN binds to OmpA in its monomeric state.

237 Previous simulations of the *apo* OmpA monomer showed that the CTD
238 makes stable contacts with the inner leaflet of the outer membrane (Ortiz-
239 Suarez et al. 2016). We observed similar behaviour in these PGN-bound
240 simulations (Figure 3C); the CTD drifted upwards to interact with the lipids,
241 concomitant with the linker region connecting the N- and C-terminal domains
242 contracting to a final length of ~5 Å. This membrane binding event only

occurred after ~100 ns in both simulation replicas (Figure 3B), compared to 20-40 ns in the previous apo simulations. This discrepancy in time scale is most likely due to the attached PGN strand imposing a resistance on the movement of the CTD, as it is infinitely linked to its periodic images.

Also of interest in this study is whether the interaction with the outer membrane is mediated by a conserved membrane-binding surface. Further inspection showed that, similar to the apo simulations, initial contact between the CTD and the inner leaflet of the membrane was primarily mediated by residues 270-300 (Figure S2), suggesting a potentially conserved binding mechanism. We note that this stretch of residues includes the mobile insert found in certain members of the OmpA family (Figure 2A). The interaction of this insert with the membrane reduced its flexibility and therefore stabilised the nearby binding pocket, which could explain the stronger salt bridge formed by R256 with PGN (Figure 3A). While these residues remained in contact with lipids for the remainder of both simulations, other residues and the PGN strand itself also began to interact with the membrane after ~200 ns in both simulation repeats. Both the peptide side chains and the backbone sugar moieties of the PGN strand facilitated membrane binding by forming hydrogen bonds and salt bridge interactions with the lipid head groups.

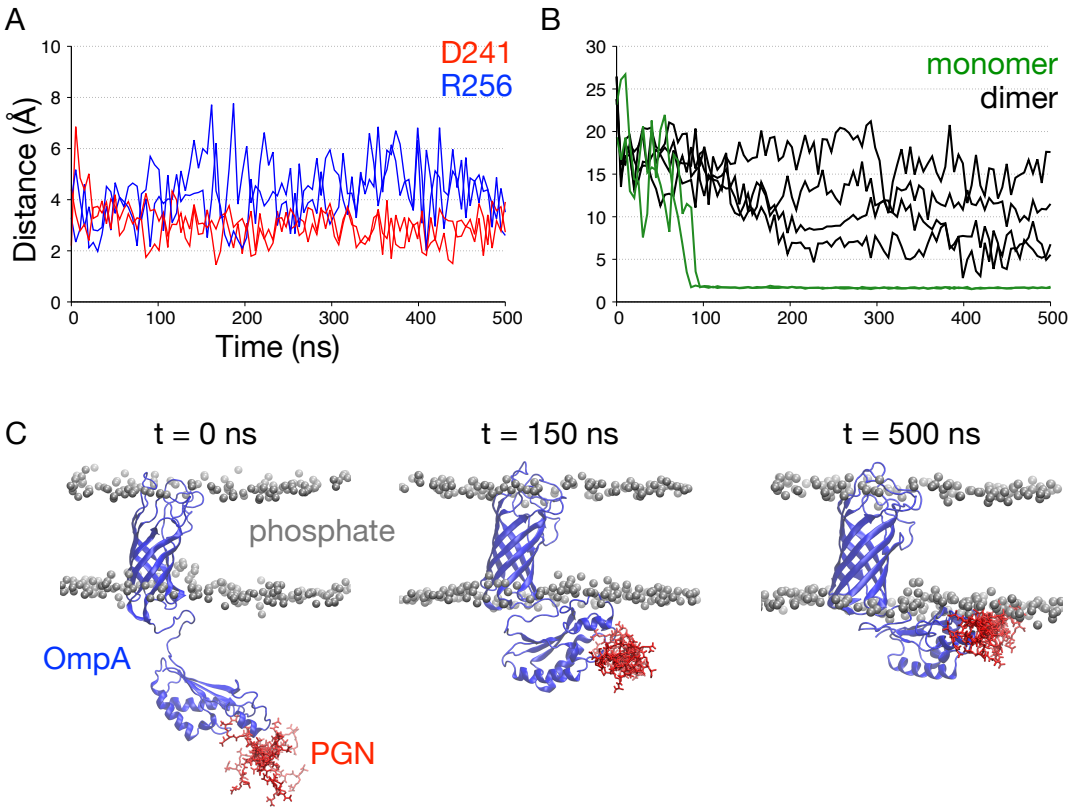


Figure 3. Interactions of PGN-bound CTD of OmpA monomer with the outer membrane. (A) Distance between key residues in the binding pocket, D241 and R256, and PGN throughout the two independent repeats of 500 ns simulations. Measurement was performed as described in Figure 2B. (B) Minimum distance between the CTD of OmpA (residue 188-316) and the inner leaflet of the outer membrane, taken from the simulations of either the monomeric (green) or the dimeric (black) form of the protein. For clarity, only the results from dimer simulations in low ionic strength are shown. (C) Snapshots taken from one of the OmpA monomer simulations highlighting the position of the protein (blue) and the bound PGN strand (red), with respect to the membrane (phosphate atoms depicted by the silver spheres).

PGN binding to OmpA homodimer. To study the effects of OmpA dimerisation on PGN binding, we simulated the full-length *Ec*OmpA

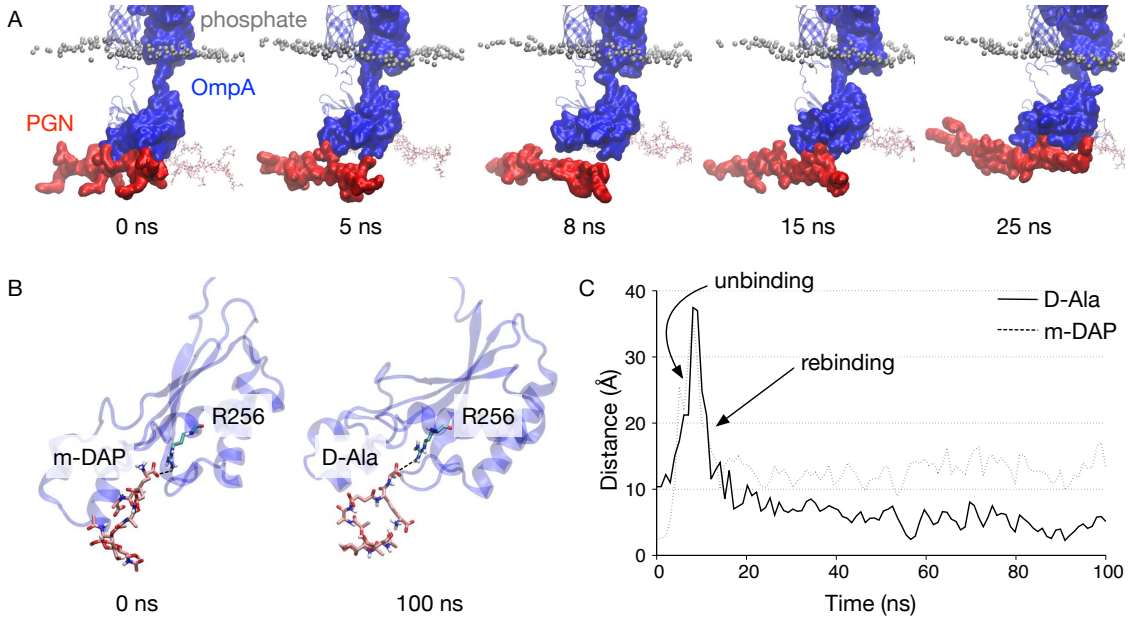
277 homodimer (Marcoux et al. 2014) with two strands of PGN bound (one for
278 each monomer), with the two N-terminal domains embedded within an
279 asymmetric outer membrane, following the same protocol described above for
280 the monomer simulations. It has previously been reported that the dynamics of
281 the OmpA homologue from *Pasteurella multicoda* are sensitive to the ionic
282 strength (Carpenter et al. 2007); we therefore also simulated this system in two
283 conditions: i) with neutralising ions only (amounting to an approximate
284 concentration of 0.3 M Mg^{2+}); and ii) at 1 M $MgCl_2$, thus spanning the range of
285 possible physiological ionic states. Again, two repeats of 500 ns simulations
286 were conducted for each condition.

287 In both salt concentrations, we found similar dynamics to those
288 observed in the CTD-only simulations, i.e. the D241 residue made a stable salt
289 bridge with the PGN for the most part of the simulations, whilst the R256
290 residue formed occasional, less stable contacts (Figure S3). While the linker
291 region connecting the NTD and CTD of each OmpA chain extended and
292 contracted during the simulations, the CTD did not make contact with the inner
293 leaflet of the outer membrane, as previously observed in the monomer
294 simulations (Figure 3B), and in agreement with previous simulations of OmpA
295 in its apo state (Ortiz-Suarez et al. 2016). The membrane-binding surface of the
296 CTD in our monomer simulations was located close to the dimerisation
297 interface. Dimerisation would therefore be expected to reduce the mobility and
298 solvent accessibility of this surface and prevent interactions with the
299 membrane. The high salt concentration in simulations with 1 M $MgCl_2$ masked

300 the polar and charged residues of the protein and lipids, further inhibiting any
301 contact. Due to the absence of membrane binding, the mobile region close to
302 R256 remained flexible throughout the entire simulations, which likely
303 explains the weaker interaction between this residue and PGN compared to that
304 of the monomer simulations.

305 The PGN strands in the dimer simulations also exhibited different
306 behaviour to that of the monomer simulations. As the CTDs of the OmpA
307 homodimer did not interact with the outer membrane, the PGN strands could
308 similarly not bind to the membrane. Instead, the peptide side chains from the
309 two adjacent PGN strands that were distant from the protein interacted with
310 each other at the end of the 500 ns simulations (Figure S4A). These
311 interactions were primarily mediated via salt bridges between the carboxyl and
312 amino groups on the m-DAP residue, and the carboxyl group on the D-Ala and
313 D-Glu residues. In the newly formed bacterial cell wall, adjacent PGN strands
314 are cross-linked by transpeptidases via the m-DAP and D-Ala terminal
315 residues. Our simulations therefore shed light on how these PGN cross-links
316 are positioned with respect to PGN-binding proteins such as OmpA (Figure
317 S4B). Some PGN peptide side chains associate with these proteins while the
318 ones that are far from the proteins are cross-linked to form a robust network
319 parallel to the membrane.

320



321

322 **Figure 4. Interactions between CTD of OmpA dimer and PGN.** (A)
323 Snapshots taken at different time points during a simulation of OmpA
324 homodimer in 1 M $MgCl_2$, highlighting the unbinding and rebinding of PGN to
325 the CTD. (B) Upon rebinding, the PGN peptide side chain reorientated itself
326 such that the R256 residue interacted with the D-Ala on PGN, instead of the m-
327 DAP residue as observed in the crystal structure. For clarity, only one repeat
328 unit of PGN is shown. (C) Distance between the R256 residue of the protein
329 with either the D-Ala or the m-DAP residue of the PGN during the first 100 ns
330 of the simulation.

331

332 Whilst the conformational dynamics of the OmpA homodimers and the
333 PGN strands in both ionic strengths studied here were generally very similar,
334 we found one key difference: a brief unbinding of PGN occurred in one of the
335 monomers during a simulation conducted at 1 M $MgCl_2$. The binding pocket
336 opened up, releasing the PGN peptide side chain into the bulk solution for a
337 few nanoseconds before the PGN reassociated with the binding site (Figure

4A). This rebinding event, intriguingly, resulted in a reorientation of the terminal residues of the PGN peptide side chain, such that the carboxyl group of the D-Ala formed a salt bridge with the R256 residue (Figure 4B and 4C). The D-Ala replaced the m-DAP residue, which is central to positioning the PGN within the binding site in the crystal structure of *Ab*OmpA. As the terminal D-Ala on the PGN does not have an amino group like the m-DAP residue, the salt bridge with D241 was lost in this new conformation. This result suggests that the PGN may adopt different configurations in the binding site due to the presence of the multiple carboxyl groups in its side chain that could form a salt bridge with the R256 residue. It is likely that the m-DAP residue forms the strongest interaction with the binding site as it can interact with both D241 and R256 via its amino and carboxyl termini respectively, and such a conformation was therefore captured within the crystallographic state (Park et al. 2012).

Discussion

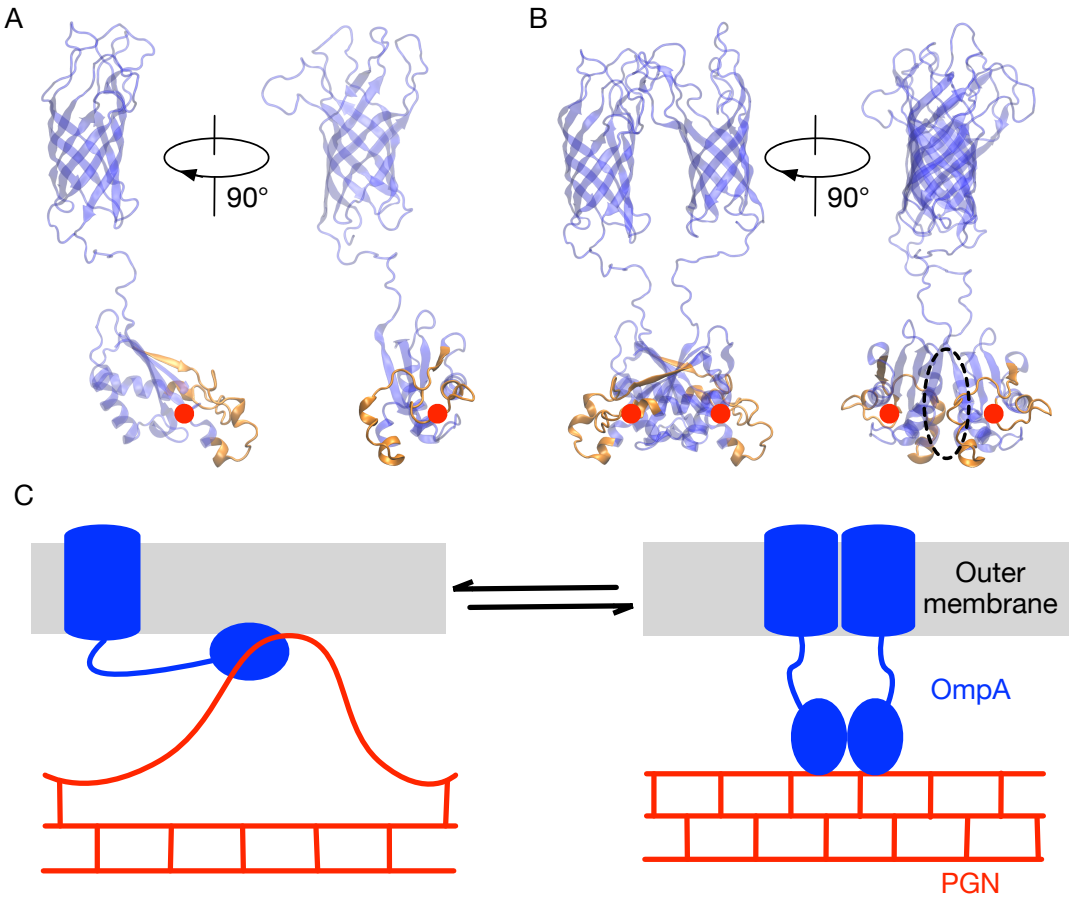
The present study utilised atomistic molecular dynamics simulations to gain insights into the interactions of PGN with members of the OmpA protein family. We first demonstrated a stable binding of PGN to OmpA CTD from four species of bacteria, suggesting a universally conserved recognition mechanism. The PGN also bound stably to a model of full-length *Ec*OmpA in both monomeric and dimeric states. When attached to OmpA monomer, the PGN strand drifted upwards and interacted with the outer membrane. Such contact

was, however, not evident in simulations of the OmpA homodimer; instead, the PGN remained within the region equivalent to the periplasmic space. This implies that dimer formation may influence the local conformation of the PGN network. PGN binding is also labile, as exemplified by the unbinding and rebinding events observed in one of the dimer simulations.

Despite very low amino acid sequence identity between members of the OmpA protein family, our simulations strongly indicated a similar PGN recognition mechanism, involving two key residues that are conserved across all species. This binding mechanism is also likely to prevail in other proteins with OmpA-like domains, such as the PAL and MotB families, in which these residues are also present (Figure 2A). NMR spectroscopy of PAL from the bacterium *Haemophilus influenzae* in complex with a PGN precursor revealed interactions between the m-DAP moiety and D71, which is equivalent to the D271 residue that coordinates the binding of PGN to *Ab*OmpA (Parsons et al. 2006). While no structure is currently available for MotB bound to PGN, a biochemical study showed that the PGN-binding domain of this protein is interchangeable with PAL (Hizukuri et al. 2009), corroborating the notion of a universal PGN binding mechanism.

Although the binding mechanism itself is conserved, our simulations suggest that the strength of binding is likely to differ for different types of proteins and for proteins from different bacteria. While we did not attempt to directly calculate binding affinities in this study, the stability of PGN binding

382 to OmpA CTDs from the four species studied appeared to vary, potentially due
383 to subtle structural differences. More specifically, certain members of the
384 OmpA CTD family have a 10-20 residue insert close to the PGN binding site
385 that is highly mobile (Ishida et al. 2014), and therefore weakens the
386 interactions. This insert was stabilised in the monomer simulations as it formed
387 part of the membrane-binding surface, which then led to stronger interactions
388 with PGN. Interestingly, such an insert is not present in the PAL and MotB
389 families, leading us to predict stronger binding affinity for these groups of
390 proteins. Also, the non-covalent interactions between the PGN network and
391 these proteins are likely to be labile *in vivo*. This is supported by the unbinding
392 and rebinding events observed in one of our dimer simulations, whereby a PGN
393 strand detached briefly from the CTD of *Ec*OmpA. We conjecture that the
394 ability to detach from the PGN network may be advantageous to allow protein
395 diffusion and reorganisation, especially for the OmpA family, whose N-
396 terminal domain can form an outer membrane pore (Arora et al. 2000) and
397 exists as a dynamic homodimer (Zheng et al. 2011). Thus, OmpA forms what
398 can be thought of as a flexible molecular clamp for attachment to the cell wall,
399 which cannot be achieved by the covalent binding exemplified by Lpp.
400 Additional support for our hypothesis comes from a recent fluorescence study
401 which demonstrated that MotB diffused freely in the membrane and exchanged
402 rapidly with the flagellar motor complex (Leake et al. 2006).



403

404 **Figure 5: OmpA dimerisation inhibits PGN association with the outer**
405 **membrane.** (A) The full-length OmpA monomer is shown, highlighting in
406 orange the position of residues that interacted significantly with the outer
407 membrane durin simulation. The approximate location of the PGN binding site
408 is marked by the red dot. (B) The position of the same residues mapped onto
409 the structure of the OmpA homodimer. The dimerisation interface is indicated
410 by the black dashed circle. (C) A schematic diagram of the possible role of
411 OmpA dimerisation in the formation of the cell wall.

412

413 Although most outer membrane porins, such as OmpF and OmpC, form
414 homotrimers, OmpA exists in equilibrium between monomeric and dimeric
415 states (Zheng et al. 2011). In this study we show that, in the monomeric state,

416 the CTD of OmpA can bind to the inner leaflet of the outer membrane,
417 resulting in interactions between the PGN strand and lipid residues. We
418 conjecture that such interactions, *in vivo*, are likely to be undesirable to the cell
419 as they may cause local distortions in the PGN network by preventing cross-
420 linking between adjacent strands, as a result of which the integrity of the cell
421 wall may be compromised. Our simulations also delineated a membrane
422 binding surface close to the PGN binding pocket (Figure 5A and Figure S2).
423 OmpA dimerisation occurs in large part through a network of salt bridges
424 between the CTDs (Marcoux et al. 2014, Ortiz-Suarez et al. 2016), positioned
425 nearby to the proposed membrane binding surface (Figure 5B). Dimerisation
426 therefore may sterically impede the interactions between the CTDs, and hence
427 the bound PGN strands, with the outer membrane, as evidenced by the absence
428 of such interactions in four independent dimer simulations presented in this
429 study. The formation of OmpA dimers could therefore have favourable
430 implications for the integrity of the cell wall, primarily by facilitating cross-
431 linking between adjacent strands and thus smoothening any local distortion in
432 the PGN network (Figure 5C). This is further supported by the prevalent
433 interactions between the peptide side chains of the two PGN strands in our
434 dimer simulations. Our results therefore shed light on the potential functional
435 role of OmpA dimerisation in the formation of the cell wall.

436 The complex architecture of the bacterial cell envelope has remained the
437 perennial topic of interest in microbiology. Computational modelling provides
438 the essential link between biochemical and biophysical data by illuminating the

439 microscopic behaviours that are otherwise currently inaccessible to direct
440 imaging techniques. While a number of other computational studies have been
441 performed involving peptidoglycan (Gumbart et al. 2014; Nguyen et al. 2015),
442 ours represent, to the best of our knowledge, the first to directly examine its
443 interaction with outer membrane proteins, specifically OmpA. Our atomic scale
444 model of peptidoglycan will thus enable future studies to answer other crucial
445 questions to further our understanding of the bacterial cell wall structure and
446 function.

447 Experimental Procedures

448 **PGN parametrisation.** Parameters for the peptidoglycan chain, consisting of
449 alternating β-1,4-linked N-acetylglucosamine (NAG) and N-acetylmuramic
450 acid (NAM) subunits with the latter attached to a tetrapeptide of L-Ala, D-Glu,
451 m-DAP, D-Ala were derived from analogous parameters available with the
452 GROMOS 45A4 carbohydrate (Lins & Hünenberger 2005) and GROMOS
453 54A7 protein force field (Schmid et al. 2011). This ensured consistency with
454 both protein and membrane parameters used in this work. We note that the
455 final D-Ala residue of the NAM attached pentapeptide (as depicted in Figure 1)
456 was not included, due to the rapid degradation of this residue in *E. coli*
457 (Vollmer & Bertsche 2008).

458 **Outer membrane construction.** The outer leaflet of the membrane was
459 composed entirely of Ra (i.e. full rough) LPS molecules containing lipid A, the
460 inner core and the outer core sugar molecules. As per our previous simulations

461 that included the inner core (Piggot et al. 2011; Piggot et al. 2013), the core
462 structure used for the Ra LPS was the R1 core type (Vinogradov et al. 1999),
463 which is the most prevalent core structure in *E. coli* species (Appelmelk et al.
464 1994). The inner leaflet of the membrane comprised 90% 1-palmitoyl,2-cis-
465 vaccenyl phosphatidyl ethanolamine (PVPE), 5% 1-paltimoyl,2-cis-vaccenyl
466 phosphatidyl glycerol (PVPG) and 5% 1-palmitoyl,2-cis-vaccenyl,3-
467 palmitoyl,4- cis-vaccenyl diphosphatidyl glycerol (DPG) (or otherwise known
468 as cardiolipin) (Lugtenberg & Peters 1976; Aibara et al. 1972; Kito et al. 1975;
469 Yokota et al. 1980). The full rough outer membrane system was constructed
470 through the manual addition of the outer core sugars to a small-equilibrated
471 outer membrane system containing 16 Rd_1 LPS molecules in the outer leaflet
472 (Piggot et al. 2011). We note here that, with longer simulations to the
473 previously reported 200 ns, this Rd_1 LPS containing membrane system
474 equilibrated to a slightly higher area per acyl chain for the LPS molecules. This
475 is presumably due to the very slow diffusing LPS molecules taking
476 substantially longer than typical phospholipid membranes to equilibrate. The
477 manual addition of the outer core was performed using the VMD program
478 (Humphrey & Dalke 1996). A 500 ns simulation of this system was performed
479 in the NVT ensemble to allow for an equilibration of the new LPS structures
480 without any artificial disruptions of the membrane packing due to the addition
481 of the outer core sugars. Subsequently two 1 μs simulations were performed in
482 the NPT ensemble to further equilibrate this system. The addition of the outer
483 core led to an LPS area per acyl chain of 0.269±0.001 nm^2 and 0.265±0.001

484 nm^2 (averaged over the final 500 ns of the 1 μs simulations). Finally, the
485 GROMACS program genconf was used to create a larger outer membrane
486 system with 64 Ra LPS molecules in the outer leaflet, which was simulated for
487 another 1 μs (LPS area per acyl chain of 0.267±0.001 nm^2 averaged over the
488 final 500 ns).

489 **Simulation systems.** CTD only simulations for *Ab*OmpA, *Se*OmpA and
490 *Nm*RmpM were performed using X-ray crystal structures (PDB: 3TD5, 4ERH
491 and 1R1M respectively) and for *Ec*OmpA, a homology model based on the
492 crystal structure of *Se*OmpA was utilised. A PGN molecule containing two
493 repeat units was docked to the binding site in the same orientation observed in
494 the crystal complex of *Ab*OmpA-pentapeptide (PDB: 3TD5) using the
495 sculpting tool in PyMOL (Schrödinger LLC, 2010). This structure was
496 solvated and counter-ions were added to neutralise the system. Short 100 ps
497 equilibration simulations (NVT followed by NPT runs) were conducted with
498 heavy atoms restrained. Each CTD was then simulated in triplicate for 500 ns
499 with no restraints using different initial velocities.

500 The full length *Ec*OmpA monomer and dimer models were built by
501 attaching the N-terminal β-barrel from X-ray crystal structure (PDB: 1G90) to
502 an 18-residue linker from an NMR structure of *K. pneumoniae* (PDB: 2K0L),
503 which in turn was connected to a homology model of the CTD generated based
504 on the crystal structure from *Se*OmpA (PDB: 4ERH) (Marcoux et al. 2014).
505 The structure was inserted into a model of *E. coli* outer membrane using

g_membed (Wolf et al. 2010). Due to the asymmetric nature of the outer membrane, during the insertion process the protein was resized by a factor of 0.3 in the plane of the membrane and the asymmetry option was applied to ensure an appropriate packing of the LPS around the protein.

A PGN chain consisting of ten repeating NAG-NAM-peptide units was constructed using the program VMD (Humphrey & Dalke 1996) to manually attach repeating units to one another in an appropriate orientation. This chain was solvated in a box constructed to have the same size in the *x* dimension as the outer membrane system. This size of system brought the two ends of the chain within close proximity across the periodic boundary in this *x* dimension. Therefore, to mimic a longer chain of PGN as would typically be found *in vivo*, the two ends of the ten repeat unit chain were linked together across the periodic boundary. This linking was performed by applying a harmonic bond (GROMACS bond type 6) with a distance of 0.3 nm and a relatively weak force constant of 222 kJ mol^{-1} nm^{-2} (i.e. 10 kcal mol^{-1} $Å^{-2}$), so as to allow flexibility within the chain structure. Subsequently, this system was simulated in the *NVT* ensemble for 150 ns, with 2 repeat simulations performed each using different starting velocities, to allow for an equilibration of the periodic chain and to ensure that the system was stable.

One of these PGN molecules containing the ten repeat units was docked onto the CTD of the membrane-inserted protein following the same method described for the CTD only simulations. The system was solvated and

528 neutralising ions were added. A similar equilibration protocol was performed
529 before running two replicates of production simulation, each for 500 ns, with
530 different initial velocities. 1 M $MgCl_2$ was added to the system for simulations
531 at a high ionic strength.

532 **Simulation protocols.** All simulations were performed using the GROMACS
533 package version 4.6.1 (Van Der Spoel et al. 2005; Hess et al. 2008), the
534 GROMOS 54A7 force field (Schmid et al. 2011) and the SPC water model
535 (Berendsen et al. 1981). The temperature was kept constant at 313 K using the
536 Nose-Hoover thermostat with a time constant of 0.5 ps (Nosé 1984). The
537 pressure was kept at 1 atm using the Parrinello-Rahman barostat with a time
538 constant of 5 ps (Parrinello & Rahman 1981). The CTD only simulation was
539 coupled to the barostat isotropically, while for the full-length OmpA simulation
540 the semi-isotropic coupling was utilised. Electrostatic interactions were
541 calculated using the Particle-Mesh Ewald (PME) method (Essmann et al. 1995)
542 with short range cut-off of 0.9 nm. Van der Waals interactions were truncated
543 at 1.4 nm with a long-range dispersion correction applied to the energy and
544 pressure. The LINCS algorithm was used to constrain all bonds to allow a 2 fs
545 time step for calculation (Hess et al. 1997).

546 # References

547 Aibara, S. et al., 1972. Changes in Positional Distribution of Fatty Acids in the
548 Phospholipids of Escherichia coli after Shift-Down in Temperature. *Biochim.*
549 *Biophys. Acta*, 270, pp.301–306.

550 Appelmelk, B.J. et al., 1994. Frequencies of lipopolysaccharide core types in
551 Escherichia coli strains from bacteraemic patients. *Microbiology*, 140(5),
552 pp.1119–1124.
553 Arora, A. et al., 2000. Refolded Outer Membrane Protein A of Escherichia
554 coliForms Ion Channels with Two Conductance States in Planar Lipid
555 Bilayers. *J. Biol. Chem.*, 275(3), pp.1594–1600. Available at:
556 http://www.jbc.org/content/275/3/1594.long.
557 Arora, A. et al., 2001. Structure of outer membrane protein A transmembrane
558 domain by NMR spectroscopy. *Nat. Struct. Biol.*, 8(4), pp.334–8. Available at:
559 http://www.ncbi.nlm.nih.gov/pubmed/11276254.
560 Berendsen, H.J. et al., 1981. Interaction models for water in relation to protein
561 hydration. In B. Pullman, ed. *Intermolecular Forces*. Boston: Reidel
562 Publishing, pp. 331–342.
563 Braun, V., 1975. Covalent lipoprotein from the outer membrane of escherichia
564 coli. *Biochim. Biophys. Acta.*, 415(3), pp.335–377. Available at:
565 http://www.sciencedirect.com/science/article/pii/0304415775900131.
566 Carpenter, T., Khalid, S. & Sansom, M.S.P., 2007. A multidomain outer membrane
567 protein from Pasteurella multocida: modelling and simulation studies of
568 PmOmpA. *Biochim. Biophys. Acta.*, 1768(11), pp.2831–40. Available at:
569 http://www.ncbi.nlm.nih.gov/pubmed/17888868 [Accessed March 23,
570 2011].
571 Clifton, L.A. et al., 2015. An Accurate In Vitro Model of the E. coli Envelope. *Angew*
572 *Chem Int Ed Engl.*, 54(41), pp.11952–11955. Available at:
573 http://doi.wiley.com/10.1002/anie.201504287.
574 Døvling Kaspersen, J. et al., 2014. Low-resolution structures of OmpA·DDM

protein-detergent complexes. *Chembiochem*, 15(14), pp.2113–2124.

Essmann, U. et al., 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.*, 103(19), pp.8577–8593.

Grizot, S. & Buchanan, S.K., 2004. Structure of the OmpA-like domain of RmpM from Neisseria meningitidis. *Mol. Microbiol.*, 51(4), pp.1027–1037.

Gumbart, J.C. et al., 2014. Escherichia coli Peptidoglycan Structure and Mechanics as Predicted by Atomic-Scale Simulations. *PLoS Comput. Biol.*, 10(2).

Harz, H., Burgdorf, K. & Holtje, J.-V., 1990. Isolation and Separation of the Glycan Strands from Murein of Escherichia coli by Reversed-Phase High-Performance Liquid Chromatography. *Anal. Biochem.*, 128, pp.120–128.

Hess, B. et al., 1997. LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.*, 18(12), pp.1463–1472. Available at: http://doi.wiley.com/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

Hess, B., Kutzner, C. & Spoel, D.V.D., 2008. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, 4(3), pp.435–447. Available at: http://pubs.acs.org/doi/abs/10.1021/ct700301q [Accessed April 12, 2012].

Hizukuri, Y. et al., 2009. The peptidoglycan-binding (PGB) domain of the Escherichia coli pal protein can also function as the PGB domain in E. coli flagellar motor protein MotB. *J. Biochem.*, 146(2), pp.219–229.

Höltje, J. V, 1998. Growth of the stress-bearing and shape-maintaining murein sacculus of Escherichia coli. *Microbiol. Mol. Biol. Rev.*, 62(1), pp.181–203. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=98910&tool=pmcentrez&rendertype=abstract.

Humphrey, W. & Dalke, A., 1996. VMD: visual molecular dynamics. *J. Mol. Graphics*, 15, pp.33–38. Available at: http://www-s.ks.uiuc.edu/Publications/Papers/PDF/HUMP96/HUMP96.pdf [Accessed April 12, 2012].

Ishida, H., Garcia-Herrero, A. & Vogel, H.J., 2014. The periplasmic domain of Escherichia coli outer membrane protein A can undergo a localized temperature dependent structural transition. *Biochim. Biophys. Acta.*, 1838(12), pp.3014–3024. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25135663.

Kito, M. et al., 1975. Metabolism of the phosphatidylglycerol molecular species in Escherichia coli. *Eur. J. Biochem.*, 54(1), pp.55–63.

Koebnik, R., 1995. Proposal for a peptidoglycan-associating alpha-helical motif in the C-terminal regions of some bacterial cell-surface proteins. *Mol. Microbiol.*, 16, pp.1269–1270.

Kohanski, M.A., Dwyer, D.J. & Collins, J.J., 2010. How antibiotics kill bacteria: from targets to networks. *Nat. Rev. Microbiol.*, 8(6), pp.423–435. Available at: http://dx.doi.org/10.1038/nrmicro2333.

Leake, M.C. et al., 2006. Stoichiometry and turnover in single, functioning membrane protein complexes. *Nature*, 443(7109), pp.355–358.

Lins, R.D. & Hünenberger, P.H., 2005. A new GROMOS force field for hexopyranose-based carbohydrates. *J. Comp. Chem.*, 26(13), pp.1400–1412.

Lugtenberg, E.J.. & Peters, R., 1976. Distribution of lipids in cytoplasmic and outer membranes of Escherichia coli K12. *Biochim. Biophys. Acta*, 441(195

6), pp.38–47.

Marcoux, J. et al., 2014. Mass spectrometry defines the C-terminal dimerization domain and enables modeling of the structure of full-length OmpA. *Structure*, 22(5), pp.781–790. Available at: http://dx.doi.org/10.1016/j.str.2014.03.004.

Nguyen, L.T. et al., 2015. Coarse-grained simulations of bacterial cell wall growth reveal that local coordination alone can be sufficient to maintain rod shape. *Proc. Natl. Acad. Sci. U.S.A.*, 112(28), pp.E3689–E3698. Available at: http://www.pnas.org/content/112/28/E3689\nhttp://www.ncbi.nlm.nih.gov/pubmed/26130803\nhttp://www.pnas.org/content/112/28/E3689.full\nhttp://www.pnas.org/content/112/28/E3689.full.pdf.

Nosé, S., 1984. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.*, 52(2), pp.255–268.

Park, J.S. et al., 2012. Mechanism of anchoring of OmpA protein to the cell wall peptidoglycan of the gram-negative bacterial outer membrane. *FASEB J.*, 26(1), pp.219–228.

Parrinello, M. & Rahman, A., 1981. Polymorphic Transitions in Single Crystals: a New Molecular Dynamics Method. *J. Appl. Phys.*, 52(12), pp.7182–7190.

Parsons, L.M., Lin, F. & Orban, J., 2006. Peptidoglycan recognition by Pal, an outer membrane lipoprotein. *Biochemistry*, 45(7), pp.2122–2128.

Pautsch, a & Schulz, G.E., 1998. Structure of the outer membrane protein A transmembrane domain. *Nat. Struct. Biol.*, 5(11), pp.1013–1017.

Piggot, T.J., Holdbrook, D.A. & Khalid, S., 2013. Conformational dynamics and membrane interactions of the E. coli outer membrane protein FecA: A molecular dynamics simulation study. *Biochim. Biophys. Acta.*, 1828(2),

650 pp.284–293. Available at:
651 http://dx.doi.org/10.1016/j.bbamem.2012.08.021.
652 Piggot, T.J., Holdbrook, D. a & Khalid, S., 2011. Electroporation of the E. coli and S.
653 Aureus membranes: molecular dynamics simulations of complex bacterial
654 membranes. *J. Phys. Chem.*, 115(45), pp.13381–8. Available at:
655 http://www.ncbi.nlm.nih.gov/pubmed/21970408.
656 Roujeinikova, A., 2008. Crystal structure of the cell wall anchor domain of MotB,
657 a stator component of the bacterial flagellar motor: implications for
658 peptidoglycan recognition. *Proc. Natl. Acad. Sci. U.S.A.*, 105(30), pp.10348–
659 53. Available at: http://www.pnas.org/content/105/30/10348.full.
660 Schleifer, K.H. & Kandler, O., 1972. Peptidoglycan types of bacterial cell walls and
661 their taxonomic implications. *Bacteriol. Rev.*, 36(4), pp.407–477.
662 Schmid, N. et al., 2011. Definition and testing of the GROMOS force-field versions
663 54A7 and 54B7. *Eur. Biophys. J.*, 40(7), pp.843–856.
664 Silhavy, T.J., Kahne, D. & Walker, S., 2010. The bacterial cell envelope. *Cold Spring*
665 *Harb. Perspect. Biol.*, 2(5), pp.1–17.
666 Smith, S.G.J. et al., 2007. A molecular Swiss army knife: OmpA structure, function
667 and expression. *FEMS Microbiol. Lett.*, 273(1), pp.1–11.
668 Van Der Spoel, D. et al., 2005. GROMACS: fast, flexible, and free. *J. Comput. Chem.*,
669 26(16), pp.1701–1718. Available at:
670 http://www.ncbi.nlm.nih.gov/pubmed/16211538 [Accessed March 5,
671 2012].
672 Stenberg, F. et al., 2005. Protein complexes of the Escherichia coli cell envelope. *J.*
673 *Biol. Chem.*, 280(41), pp.34409–34419.
674 Vinogradov, E. V. et al., 1999. The structures of the carbohydrate backbones of

the lipopolysaccharides from Escherichia coli rough mutants F470 (R1 core type) and F576 (R2 core type). *Eur. J. Biochem.*, 261(3), pp.629–639.

Vollmer, W. & Bertsche, U., 2008. Murein (peptidoglycan) structure, architecture and biosynthesis in Escherichia coli. *Biochim. Biophys. Acta.*, 1778(9), pp.1714–1734.

Vollmer, W., Blanot, D. & De Pedro, M.A., 2008. Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*, 32(2), pp.149–167.

Wolf, M.G. et al., 2010. g_membed : Efficient Insertion of a Membrane Protein into an Equilibrated Lipid Bilayer with Minimal Perturbation. *J. Comp. Chem.*, 31(11), pp.2169–2174.

Yokota, K., Kanamoto, R. & Kito, M., 1980. Composition of cardiolipin molecular species in Escherichia coli. *J. Bacteriol.*, 141(3), pp.1047–1051.

Zheng, C. et al., 2011. Cross-linking measurements of in vivo protein complex topologies. *Mol. Cell Proteomics*, 10(10), p.M110.006841.