Carbohydrate Polymers 318 (2023) 121094

Contents lists available at ScienceDirect

Carbohydrate Polymers

journal homepage: www.elsevier.com/locate/carbpol

Immune evasion through Toll-like receptor 4: The role of the core oligosaccharides from α2-Proteobacteria atypical lipopolysaccharides

Alejandra Matamoros-Recio [a,*], Javier Merino [b], Alicia Gallego-Jiménez [b], Raquel Conde-Alvarez [c], Manuel Fresno [b], Sonsoles Martín-Santamaría [a,*]

[a] Department of Structural and Chemical Biology, Centro de Investigaciones Biológicas "Margarita Salas", CSIC, C/ Ramiro de Maeztu, 9, 28040 Madrid, Spain
[b] Centro de Biología Molecular "Severo Ochoa", CSIC-Universidad Autónoma de Madrid, C/ Nicolás Cabrera, 1, 28049 Madrid, Spain
[c] Dpto. de Microbiología y Parasitología, Instituto de Salud Tropical (ISTUN), Instituto de Investigación Sanitaria de Navarra (IdiSNA), Universidad de Navarra, Campus Universitario, 31009 Pamplona, Spain

ARTICLE INFO

Keywords:
Toll-like receptor 4
Atypical lipopolysaccharides
Immuno-evasion
Molecular modeling
Computational simulations

ABSTRACT

Lipopolysaccharides (LPS) are major players in bacterial infection through the recognition by Toll-like receptor 4 (TLR4). The LPS chemical structure, including the oligosaccharide core and the lipid A moiety, can be strongly influenced by adaptation and modulated to assure bacteria protection, evade immune surveillance, or reduce host immune responses. Deep structural understanding of TLRs signaling is essential for the modulation of the innate immune system in sepsis control and inflammation, during bacterial infection. To advance this knowledge, we have employed computational techniques to characterize the TLR4 molecular recognition of atypical LPSs from different opportunistic members of α2-Proteobacteria, including *Brucella melitensis, Ochrobactrum anthropi,* and *Ochrobactrum intermedium*, with diverse immunostimulatory activities. We contribute to unraveling the role of uncommon lipid A chemical features such as bearing very long-chain fatty acid chains, whose presence has been rarely reported, on modulating the proper heterodimerization of the TLR4 receptor complex. Moreover, we further evaluated the influence of the different oligosaccharide cores, including sugar composition and net charge, on TLR4 activation. Our studies contribute to elucidating, from the molecular and biological perspectives, the impact of the α2-Proteobacteria LPS cores and the chemical structure of the atypical lipid A for immune system evasion in opportunistic bacteria.

1. Introduction

Members of α2-Proteobacteria such as *Brucella* and *Ochrobactrum*, phylogenetically closed, are likely to share similar mechanistic strategies of host evasion. *Brucella* is an opportunistic pathogen usually found in immunocompromised patients (Smith, 2018) whereas *Ochrobactrum* is a living soil organism, and has been reported to display some degree of pathogenicity (Barquero-Calvo et al., 2009). *O. anthropi*, *O. intermedium*, and *Brucella spp.* pathogens express atypical lipopolysaccharides (LPSs), comprising at least two very long acyl chain fatty acids (VLCFAs) with 19 and 28 carbons attached to a 2,3-diamino-2,3-dideoxyglucose backbone with two phosphate groups, a core oligosaccharide and an O-antigen with distinct sugars and, importantly, different net charge compared with the classical *E. coli* LPS (Fig. 1.A) (Barquero-Calvo et al., 2009; B. S. Park et al., 2009). The peculiar LPS structure confers to this class of bacteria a stealthy strategy for immune system recognition.

It is well known that the length and number of acyl chains in lipid A greatly influence the endotoxicity of the LPS and can affect Toll-like Receptor 4 (TLR4) activation. In contrast to the hexaacylated *E. coli* lipid A (PDB ID 3FXI) (B. S. Park et al., 2009), the tetraacylated precursor molecule, lipid IVa (PDB ID 2E59) (Onto et al., 2007) displays antagonistic activity in human TLR4 (hTLR4) signaling. In the X-ray crystallographic structure of lipid IVa in complex with the TLR4 co-receptor protein Myeloid Differentiation factor 2 (MD-2) (PDB ID 2E59) (Onto et al., 2007), the four acyl chains are completely buried in the MD-2 hydrophobic pocket to fill the empty space, whereas in the *E. coli* lipid A structure in complex with the activated conformation of the TLR4/MD-2 heterodimer, one acyl chain protrudes from the hydrophobic pocket of MD-2 and interacts with Phe440 and Phe463 of the partner TLR4 (B. S. Park et al., 2009). The glucosamine backbone of *E. coli* lipid A shows a shift upwards of 5 Å in comparison with lipid IVa (Onto et al., 2007). This shift of the lipid A position in MD-2 provides

* Corresponding author.
E-mail addresses: alejandra.matamoroso@cib.csic.es (A. Matamoros-Recio), smsantamaria@cib.csic.es (S. Martín-Santamaría).

https://doi.org/10.1016/j.carbpol.2023.121094
Received 28 February 2023; Received in revised form 21 May 2023; Accepted 4 June 2023
Available online 9 June 2023

additional space for acyl chains accommodation inside the MD-2 cavity and facilitates the interaction of the two phosphate groups with the positively charge residues at the MD-2 pocket rim and the TLR4 surface (B. S. Park et al., 2009). These observations indicate that the upwards displacement of the lipid A is important for the TLR4 activation, as the MD-2 cavity is not big enough to contain all six acyl chains of *E. coli* lipid A, and additional space for ligand binding is generated by this shift.

Besides the lipid A structure, the importance of the LPS core structure for inducing the dimerization of TLR4 has also been reported (Smith, 2018). *O. anthropi, O. intermedium,* and *Brucella* LPSs have similar lipid A but a markedly different core structure and charge distribution (Fig. 1. B). In particular, the *Brucella* core is composed of four glucosamines, which confer a strong positive charge (Smith, 2018). *O. anthropi* contains a similar composition, but only harbors one glucosamine on the core side branch and a negative GalA residue linked to Kdo II, which confers a lower positive net charge (Velasco et al., 1998). Interestingly *O. intermedium* core contains the same side branch than *O. anthropi*, but the two residues are phosphorylated (Barquero-Calvo et al., 2009). Removal of the glucosamines branch of the *B. melitensis* LPS core (*Bm-wadC* LPS mutant) increases the binding to MD-2, suggesting that TLR4 activation is hampered by virulent *B. melitensis* intact core (Fontana et al., 2016). Therefore, not only the lipid A structure but also the core

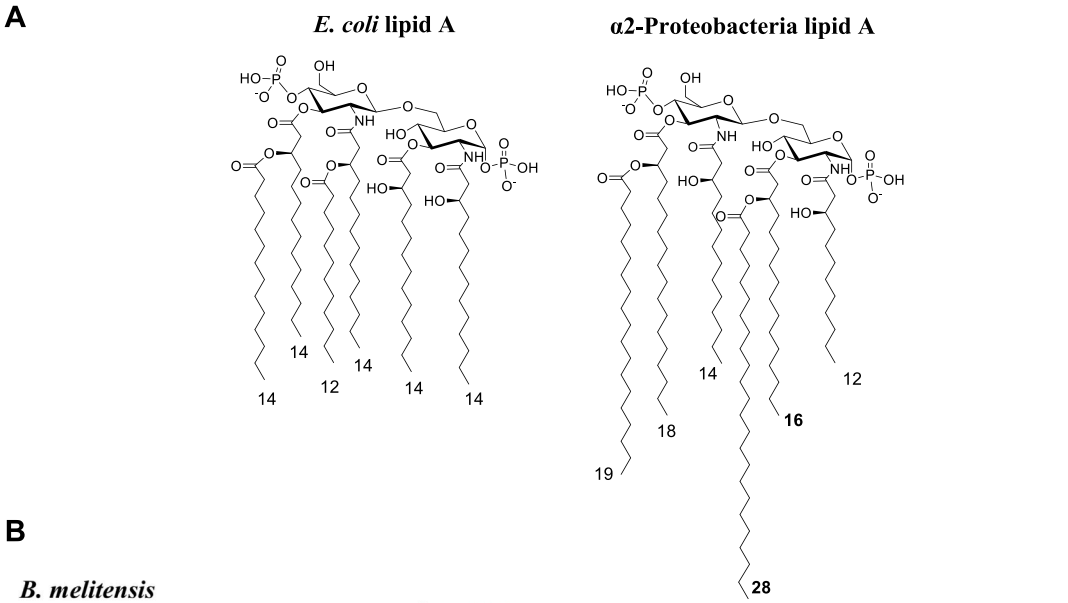

Fig. 1. Structure model of *Brucella*, *O. anthropi*, and *O. intermedium* LPS cores. The three LPSs bear a similar lipid A moiety with two very long FA chains (19 C and 28 C chains), and different cores. (A) The lipid A is composed of a linked 2,3-diamino-2,3-dideoxyglucose backbone with two phosphate groups and six fatty acid chains. (B) Core oligosaccharide structures. The core net charge is +2 for Brucella, -2 for *O. anthropi*, and -3 for *O. intermedium*. On the other hand, the *E. coli* LPS core shows an overall charge of -4 due to the presence of phosphorylated heptoses. Residues colored in red are negatively charged (charge of -1), whereas residues colored in blue are positively charged (charge of +1).

structures are responsible for the different TLR4 activation capacity, exemplified in some members of α2-Proteobacteria.

On this basis, we performed computational studies to elucidate the core-dependant different activity exhibited by several α2-Proteobacteria LPSs towards TLR4. Docking calculations followed by all-atom molecular dynamics (MD) simulations, and molecular mechanics generalized Born surface area (MM/GBSA) calculations were carried out, to propose plausible binding modes and understand the structure-activity relationship of *O. anthropi, O. intermedium*, *B. melitensis*, and *Bm-wadC* LPSs, in complex with human TLR4 in the agonist-activated conformation.

2. Results

2.1. Binding mode of α2-Proteobacteria lipid A to TLR4

We aimed to study the interaction of the full structure of the LPSs from several α2-Proteobacteria as *O. anthropi, O. intermedium*, *B. melitensis* species, plus *Bm-wadC* mutant, which lacks the branched core structure, with the human $(TLR4/MD\text{-}2)_2$ complex in the agonist-activated conformation. On this basis, the LPS ligands were divided into two moieties corresponding to their lipid A, common to all the studied LPSs, and to the core oligosaccharide (Fig. 1). We first docked the lipid A region inside the MD-2 cavity (Fig. 2). Preliminary binding poses obtained with AutoDock Vina were used as starting geometries for redocking calculations with AutoDock 4. Docking results were evaluated based on the predicted binding scores and on the apparent degree of similarity with *E. coli* lipid A (PDB ID 3FXI) (B. S. Park et al., 2009) considering both, the insertion of the fatty acid chains into the MD-2 pocket and the positioning of the glucosamine disaccharide backbone. The docking poses featured the phosphorylated saccharide backbone of the lipid A at the MD-2 pocket rim, similar to the binding mode observed for *E. coli* lipid A, and other previously reported lipids A. Nevertheless, it was placed shifted upward by about 3.5–4 Å in comparison with *E. coli* lipid A (PDB ID 3FXI) (B. S. Park et al., 2009) (Fig. 2.A). As a result, the diglucosamine backbone and the ester and amide groups connecting the saccharide moiety to the lipid chains lacked key interactions with polar residues located at the MD-2 entrance. Only one H-bond was observed between the hydroxyl group of the (*R*)-2′-hydroxymyristate chain of the

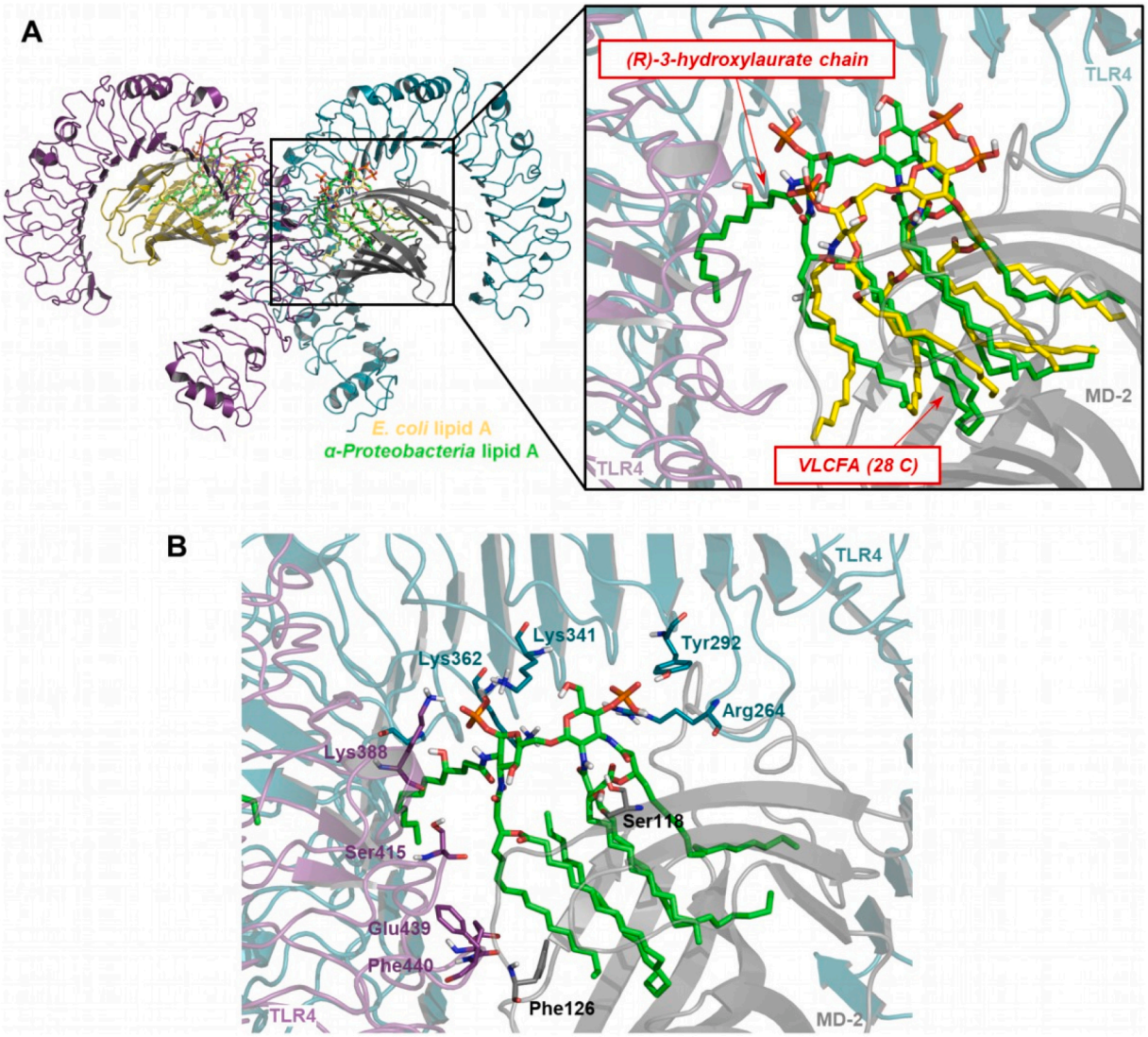

Fig. 2. Binding mode of *Ochrobactrum* and *Brucella* α2-Proteobacteria lipid A to TLR4. Superimposition of the best docking pose with the X-ray crystal structure of TLR4 in complex with *E. coli* lipid A (PDB ID 3FXI). (A) On the left: general view of the extracellular domain of TLR4, in complex with α2-Proteobacteria and *E. coli* lipids A. On the right: close-up view of the binding pocket of MD-2. (B) Close-up of the binding pocket of MD-2, showing the main interactions between α2-Proteobacteria lipid A and the TLR4 receptor complex. Ligands are represented in sticks. *E. coli* lipid A is colored in yellow, and α2-Proteobacteria lipid A in green. Proteins MD-2 (in grey and yellow), and TLR4 (in purple and aquamarine blue), are represented in semitransparent cartoon. The residues of the receptor that interacted with the ligand are also in sticks with their corresponding individual labels.

lipid A, and the backbone carbonyl group of MD-2 Ser118, in some of the poses, including the best-ranked docking pose. Regarding the lipid A phosphate groups, they established polar interactions with residues at the surface of TLR4 and the partner TLR4. Most of the poses showed the same trend; the 1-phosphate group participated in H-bond with the carbonyl group of Lys362 backbone and formed salt bridges with the amino group of the side chain of Lys341 and Ly388 from TLR4 and the counterpart TLR4, respectively, whereas the 4′-phosphate established H-bonds with the side chain of TLR4 Tyr292, and salt bridges with the guanidinium of TLR4 Arg264 (Fig. 2.B). Noteworthy, these residues are also important for the recognition of *E. coli* LPS (PDB ID 3FXI) (B. S. Park et al., 2009). In particular, Lys341, Lys362, and Lys388 participate in polar interactions with the inner core of *E. coli* LPS, whereas Arg264 interacts with the 4′-phosphate of *E. coli* lipid A (B. S. Park et al., 2009).

Five of the six lipid chains of α2-Proteobacteria lipid A were buried inside the MD-2 cavity, including the two VLCFAs. The longest fatty acyl chain (*i.e.*, the 28 carbons chain) was accommodated inside MD-2 by undergoing a fold, occupying the space of two shorter (12–14 carbons) chains found in enterobacterial LPSs (Fig. 2.B). Inside MD-2, the lipid chains established hydrophobic interactions with the alkyl side chains from many hydrophobic residues, such as Val24, Ala30, Ile32, Ile44, Ile46, Val48, Ile52, Leu54, Leu61, Ile63, Tyr65, Phe76, Leu78, Ile80, Phe104, Val113, Ile117, Phe119, Phe121, Ile124, Tyr131, Val135, Phe147, Leu149, Phe151, and Ile153, and H-π interactions with Phe104 and Phe151 side chains (Fig. 2.B). In particular, the primary acyl chain at position 3 of the disaccharide backbone of lipid A remained close to the partner TLR4 and established polar interactions with the hydroxyl group of Ser415, and hydrophobic interactions with partner TLR4 Glu439 and Phe440, and MD-2 Phe126. The sixth lipid A acyl chain, *i.e.,* the (*R*)-3-hydroxylaurate chain at position 2 of the disaccharide backbone, lied outside the MD-2 pocket protruding towards the receptor interface, stabilized by hydrophobic interactions with the non-polar parts of the side chains of residues Lys362, Gly384, Phe408, Asn409, Gly410, and His431 from TLR4, and Lys388, Val411, Thr413, and Lys435 from the partner TLR4 (Fig. 2.B). Conversely, the (*R*)-3-hydroxymyristate chain at the same position (position 2 of the diglucosamine backbone) of *E. coli* lipid A is exposed to the surface of MD-2 and interacts directly with a small hydrophobic patch on the surface of the partner TLR4, formed by residues Phe440, Leu44, and Phe463, thus completing the receptor dimerization interface (PDB ID 3FXI, Fig. 2.A) (B. S. Park et al., 2009). Remarkably, no interaction of α2-Proteobacteria lipid A was observed with Leu444 neither Phe463 (Fig. 2.B).

2.2. Binding mode of α2-Proteobacteria cores to TLR4

In a further step, we docked the inner core oligosaccharide moieties of the *O. anthropi, O. intermedium*, and *B. melitensis* LPSs, together with the *Bm-wadC* LPS mutant, using as a reference the TLR4 region where *E. coli* LPS inner core binds (PDB ID 3FXI) (B. S. Park et al., 2009). The best binding pose for each core was selected, considering the docking binding score and a suitable orientation for building the full LPSs structures. Theoretical free energies of binding predicted from the docking calculations correlated with the experimentally observed LPSs agonist activity (Table 1, see Supplementary data), suggesting that the oligosaccharide core could modulate the TLR4 response to LPS by affecting binding affinity. The lipid A (best result from AutoDock 4 redocking calculations) was then assembled to the LPSs cores, leading to full $(TLR4/MD\text{-}2/LPS)_2$ complexes (Fig. 3).

Protein-ligand interactions were further studied by MD simulations, and the stability of the complexes was predicted by MM/GBSA analysis. We performed 200 ns all-atom MD simulations and monitored the systems root-mean-square deviation (RMSD) to determine the equilibration state of all the systems during simulation time. In all cases, the complexes gradually stabilized after 40 ns, and then remained stable until 170 ns of simulation time (Fig. S.1). During this period, RMSD values were distributed around 2.7 Å for *O. anthropi* and *Bm-wadC* systems, and around 3.9 Å for *O. intermedium.* and *Brucella*, suggesting that *O. anthropi* and *Bm-wadC* complexes are the most stable structures along MD simulations. The following networks of protein-ligand interactions were observed in each system:

2.2.1. B. abortus

All the docking poses of the *Brucella* core, including the best-ranked pose, placed the four glucosamines branch of the oligosaccharide core (Fig. 1.B) in a negatively charged patch at the receptor interface, composed of residues Glu369, Asp371, and Asp395 of the partner TLR4 (Fig. 4 and Fig. 5). At this position, the glucosamines branch shields the lipid A 1-phosphate group and hampers its interaction with the TLR4 counterpart. During MD simulations, the glucosamine branch remained stable at this negative patch and established H-bonds and ionic interactions with the partner TLR4 (Fig. 3 and Fig. 4). The GlcN II formed H-bonds with the side chain carboxylate of Glu369, the GlcN III participated in H-bond with the backbone carbonyl group of Phe367 and the GlcN IV with the side chain carboxylate of Asp395. On the other hand, the GlcN I participated in salt bridge interaction with the 1-phosphate group of lipid A. As a consequence, during the simulation time, the 1-phosphate group lost all the interactions with the receptor (with both the TLR4 and the counterpart TLR4 chains). Regarding the other residues in the LPS core region, the Man moiety interacted with the carboxylate and guanidinium groups of TLR4 Glu321 and Arg322, respectively, through the formation of H-bonds. The Kdo I of the *Brucella* core was found near the position where Kdo I of *E. coli* is located (PDB ID 3FXI) (B. S. Park et al., 2009), and the Kdo II formed one H-bond with the hydroxyl group from the TLR4 Tyr296 side chain. The Glc was solvent exposed, participating in intramolecular H-bonds rather than interacting with the receptor. The position of the core did not hamper the interaction of the lipid A 4′-phosphate and the guanidinium group of TLR4 Arg264, which remained stable during the whole simulation. Moreover, one additional H-bond was formed between the 4′-phosphate and the hydroxyl group from the side chain of MD-2 Ser118 (Fig. 4).

Five of the six lipid A acyl chains remained inside the MD-2 pocket, as predicted by docking calculations, and established hydrophobic interactions with the alkyl side chains from many MD-2 hydrophobic residues. Also, one H-bond was established between the hydroxyl group of the (*R*)-2′-hydroxymyristate chain of the lipid A and the backbone carbonyl group of MD-2 Ser118. The secondary chain at position 3 of the disaccharide backbone of lipid A (*i.e.,* the VLFCA of 28C) remained folded inside MD-2, and close to the partner TLR4. It formed hydrophobic interactions with the side chain of key residues Glu439 and Phe440 from the counterpart TLR4, and with the side chain of MD-2 Phe126 (Fig. 3 and Fig. 4). The (*R*)-3-hydroxylaurate chain at position 2 of the disaccharide backbone, remained outside the MD-2 pocket, stabilized at the protein-protein dimerization interface by hydrophobic interactions with the non-polar parts of the side chains of residues Lys362, Gly384, Phe408, Asn409, Gly410, and His431 from TLR4, and Lys388, Val411, Thr413, and Lys435 from the partner TLR4. Also, the carbonyl group of this lipid chain formed one H-bond with the side chain amino group of TLR4 Lys362 (Fig. 4 and Fig. S.2).

2.2.2. Bm-wadC

In the complex of *Bm-wadC* LPS mutant, which lacks the branched residues of *Brucella* WT (Fig. 1.B), with TLR4, the Kdo I of the core was also found near the region where Kdo I of *E. coli* is located (PDB ID 3FXI) (B. S. Park et al., 2009), similar to the observed for the *Brucella* Kdo I (Fig. 3 and Fig. 6). It established H-bonds with the side chains of TLR4 residues. In particular, with the hydroxyl group of Tyr296, the guanidinium of Arg322, the carboxylate group of Glu321, and the amino group of Lys341. The Kdo II was pointed towards MD-2, as the Kdo II of *E. coli* LPS (PDB ID 3FXI) (B. S. Park et al., 2009) and participated in salt bridge interaction within its carboxylate group, with the side chain amino group of MD-2 Lys122, and in H-bond with the hydroxyl group of MD-2 Ser120. The Glc was solvent exposed, participated in

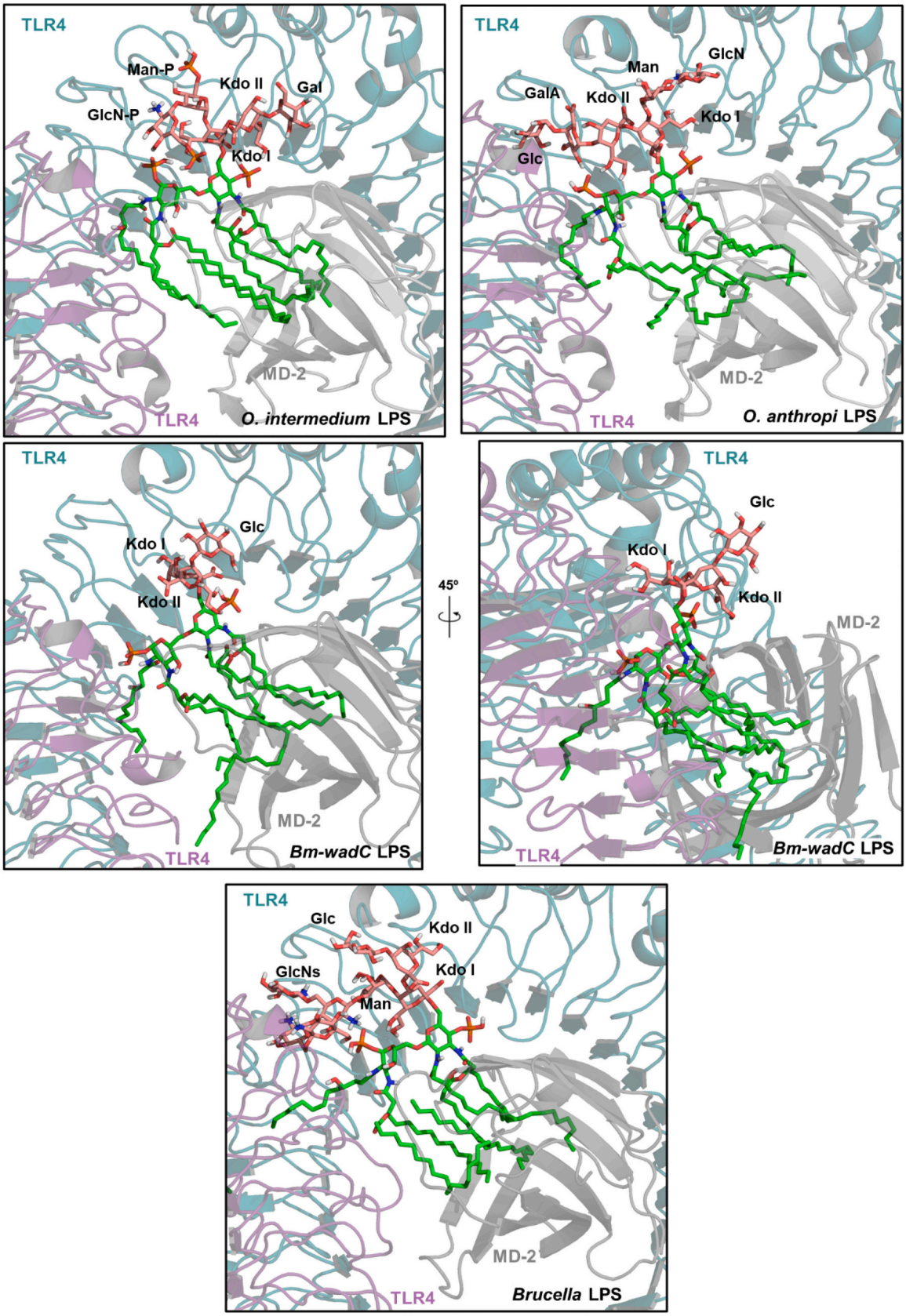

Fig. 3. Binding modes of the studied LPSs to the activated conformation of the TLR4/MD-2 receptor. Structures were selected from representative moments of the MD simulations. LPS structures are represented as sticks. The lipid A moiety is colored in green, and the residues of the cores are in pink salmon, with their corresponding individual labels. Proteins MD-2 (in grey), TLR4 (in purple), and the counterpart TLR4 (in aquamarine blue), are shown in semitransparent cartoon. The counterpart ligands (LPSs bound to the partner MD-2 molecule) are hidden for clarity.

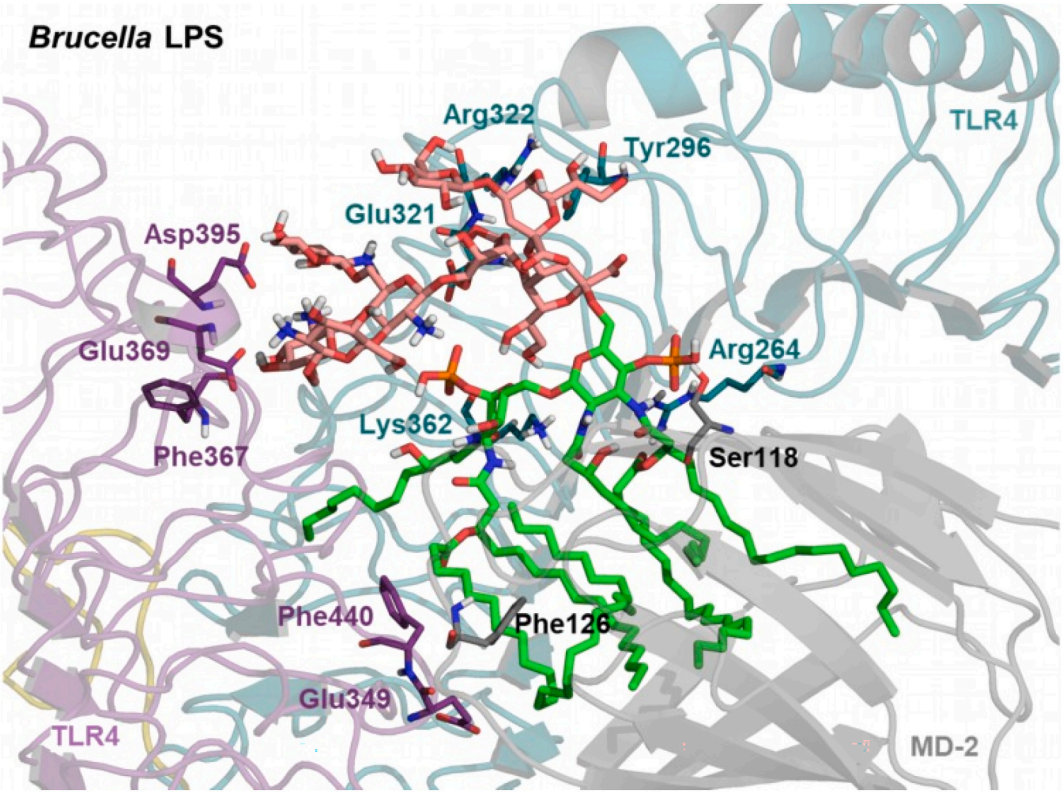

Fig. 4. Close-up of the binding pocket of MD-2, showing the major interactions between *Brucella abortus* core and the activated TLR4 dimer. LPS is represented in sticks, with the lipid A moiety colored in green, and the core in salmon pink. The residues of the receptor that interacted with the ligands are also in sticks, with their corresponding individual labels. Proteins MD-2 (in grey), TLR4 (in purple), and the counterpart TLR4 (in aquamarine blue) are shown as semitransparent cartoon. The counterpart ligand (LPSs bound to the partner MD-2 molecule) is hidden for clarity.

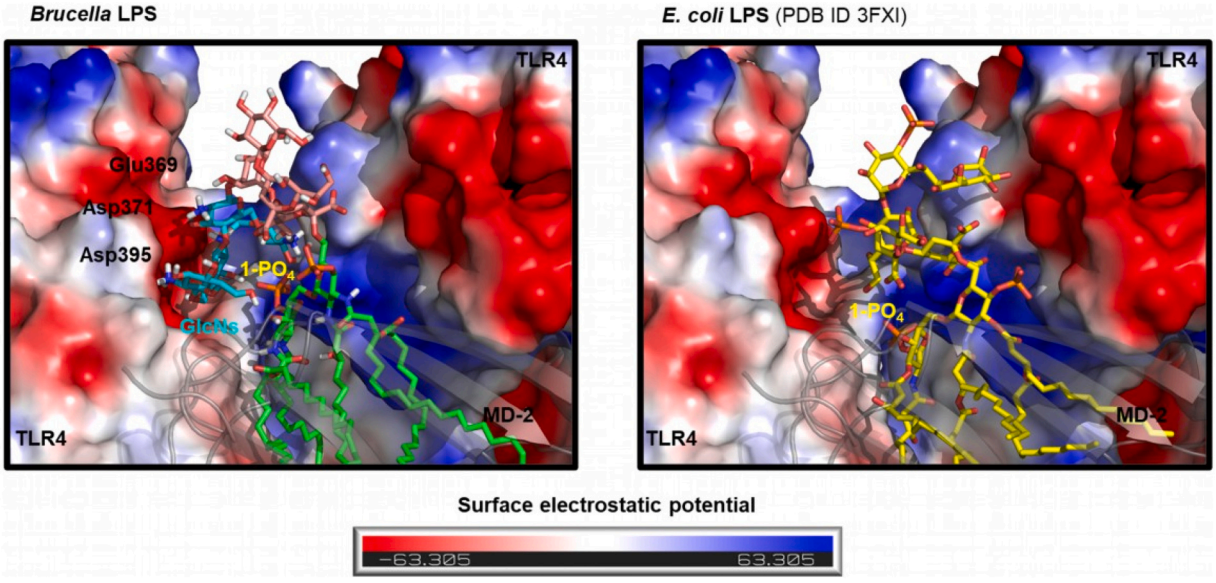

Fig. 5. Binding mode of *Brucella abortus* core to the activated conformation of the TLR4/MD-2 receptor. The four glucosamines branch of the *Brucella* core is accommodated in a negatively charged patch at the receptor interface. As a consequence, the recognition of the 1-phosphate lipid A is hampered. On the left: close-up view of the TLR4 dimerization interface, showing the glucosamines branch (in cyan) located in the above-mentioned negatively charged region. On the right: close-up view of the receptor dimerization interface in the crystal structure of TLR4 in complex with *E. coli* LPS (PDB ID 3FXI). LPSs are represented as sticks, with the *Brucella* lipid A moiety colored in green, and the core in salmon pink. The four glucosamines are colored in cyan, for the sake of clarity. The LPS of *E. coli* is colored in yellow. Residues of the counterpart TLR4 that compose the negatively charged region that interacts with the glucosamines are labeled. The surface electrostatic potential of TLR4 chains is depicted, whereas MD-2 is represented as semitransparent grey cartoon.

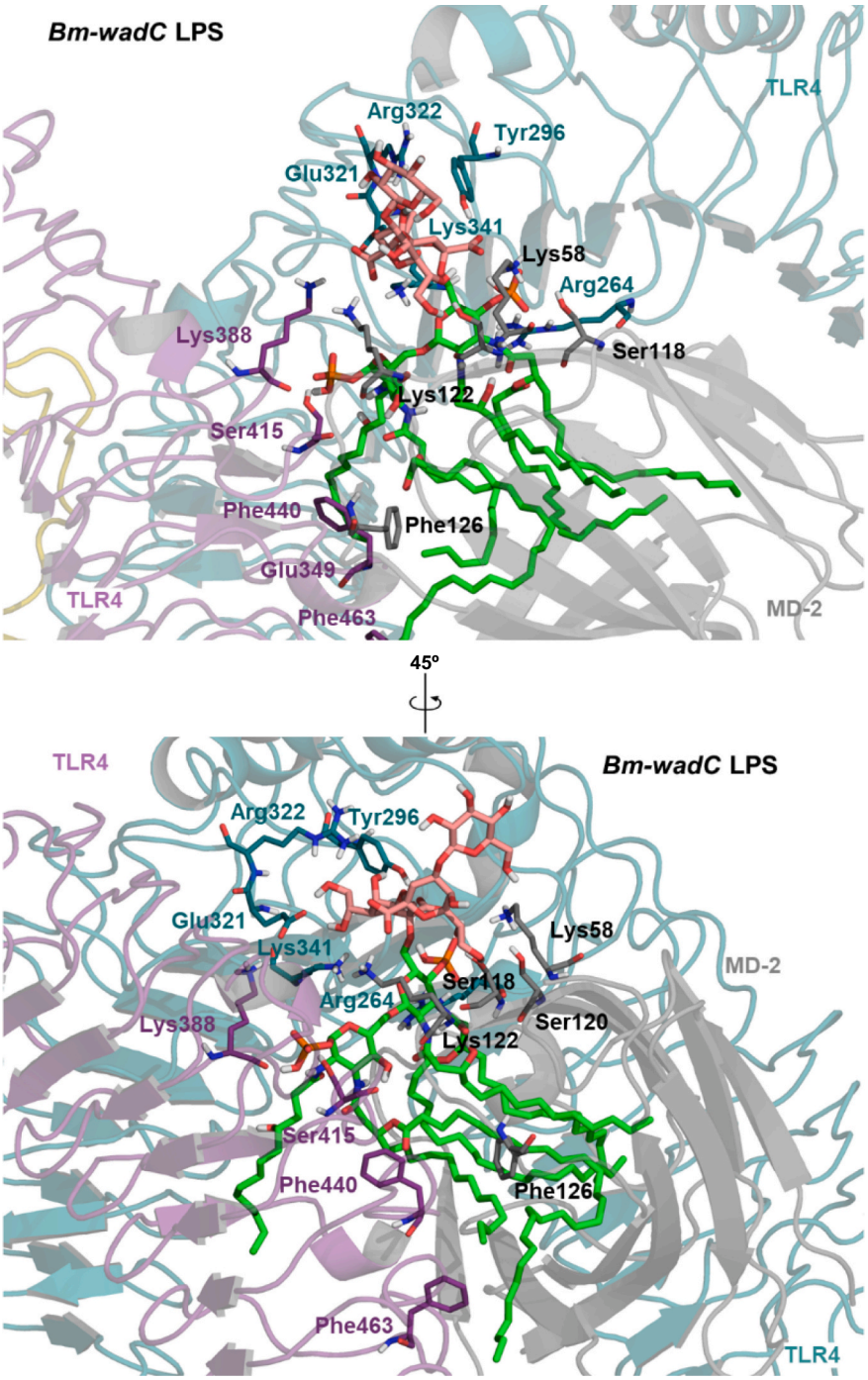

Fig. 6. Close-up of the binding pocket of MD-2, showing the major interactions between the *Bm-wadC* core and the activated TLR4 dimer. LPS is represented in sticks, with the lipid A moiety colored in green, and the core in salmon pink. The residues of the receptor that interacted with the ligands are also in sticks, with their corresponding individual labels. Proteins MD-2 (in grey), TLR4 (in purple), and the counterpart TLR4 (in aquamarine blue) are shown as semitransparent cartoon. The counterpart ligand (LPSs bound to the partner MD-2 molecule) is hidden for clarity.

intramolecular H-bonds, rather than interacting with the receptor, as previously described in the *Brucella* complex (Fig. 3 and Fig. 6). During simulation time, none of the core oligosaccharide moieties masked the lipid A diglucosamine components, thus allowing the lipid A phosphate groups to interact with the receptor. In particular, the 1-phosphate group participated in polar interactions with residues from the partner TLR4; it formed H-bonds with the side chain hydroxyl group of Ser415, and with the carbonyl group of Lys388 backbone, as well as salt bridge with the amino group of the side chain of the same residue. On the other hand, the 4′-phosphate established salt bridge interactions with the guanidinium of Arg264 and the side chain amino group of MD-2 Lys58, plus one H-bond with the hydroxyl group of MD-2 Ser118 (Fig. 3 and Fig. 6).

Five of the six lipid A acyl chains remained inside the MD-2 pocket during the entire simulation, participating in hydrophobic interactions with the side chains of MD-2 hydrophobic residues. Interestingly, in this case, part of the secondary lipid chain at position 3 of the disaccharide backbone of lipid A (*i.e.,* the VLFCA of 28 C) in one TLR4 unit, and the secondary acyl chain at the same lipid A position (*i.e.,* the (*R*)-3-hydroxypalmitate chain) in the other TLR4 unit, were accommodated in the MD-2 channel, protruding towards the TLR4 counterpart, as elucidated for *E. coli* lipid A (PDB ID 3FXI) (B. S. Park et al., 2009) (Fig. 3 and Fig. 6). Hydrophobic interactions between these chains of the lipid A and residues Phe440 and Phe463 from the counterpart TLR4 were observed, as well as with MD-2 residues such as Val82, Met85, Leu87, Ile124 and Phe126. The (*R*)-3-hydroxylaurate chain at position 2 of the

disaccharide backbone remained outside the MD-2 pocket and participated in hydrophobic interactions with residues at the receptor dimerization interface, such as Asn365, Gly384, and Phe408 from TLR4, and Lys388, Thr413, and Lys435 from the counterpart TLR4 (Fig. 6 and Fig. S.2).

2.2.3. *O. intermedium*

Along the simulation time of TLR4 in complex with *O. intermedium* LPS, the LPS molecule was observed to move away from the dimerization interface, with the LPS core and the lipid A diglucosamine backbone shifted towards TLR4, lacking interactions with the partner TLR4 chain (Fig. 3 and Fig. 7). As a consequence, the docking predicted interactions between the 1-phosphate group of lipid A and the partner TLR4 were lost. On the contrary, the interactions between lipid A phosphates groups and TLR4 were maintained, including the H-bond between the 1-phosphate and the amino group from the Lys341 side chain, and the H-bond predicted among the 4′-phosphate and the guanidinium group of Arg264. In addition, one H-bond was formed between the 4′-phosphate and the hydroxyl group from the side chain of MD-2 Ser118, as previously described for the *Brucella* complex (Fig. 3 and Fig. 7). Interaction of the core LPS with the receptor was mainly driven by the phosphate groups of the GlcN-P and Man-P residues, and the carboxylate groups of both Kdo moieties, that interact with positively charged residues located at the surface of both, the TLR4 and the MD-2 proteins. Indeed, the GlcN-P and Man-P phosphate groups of *O. intermedium* were located in the same TLR4 regions as the *E. coli* Hep I phosphate and Kdo II carboxylate,

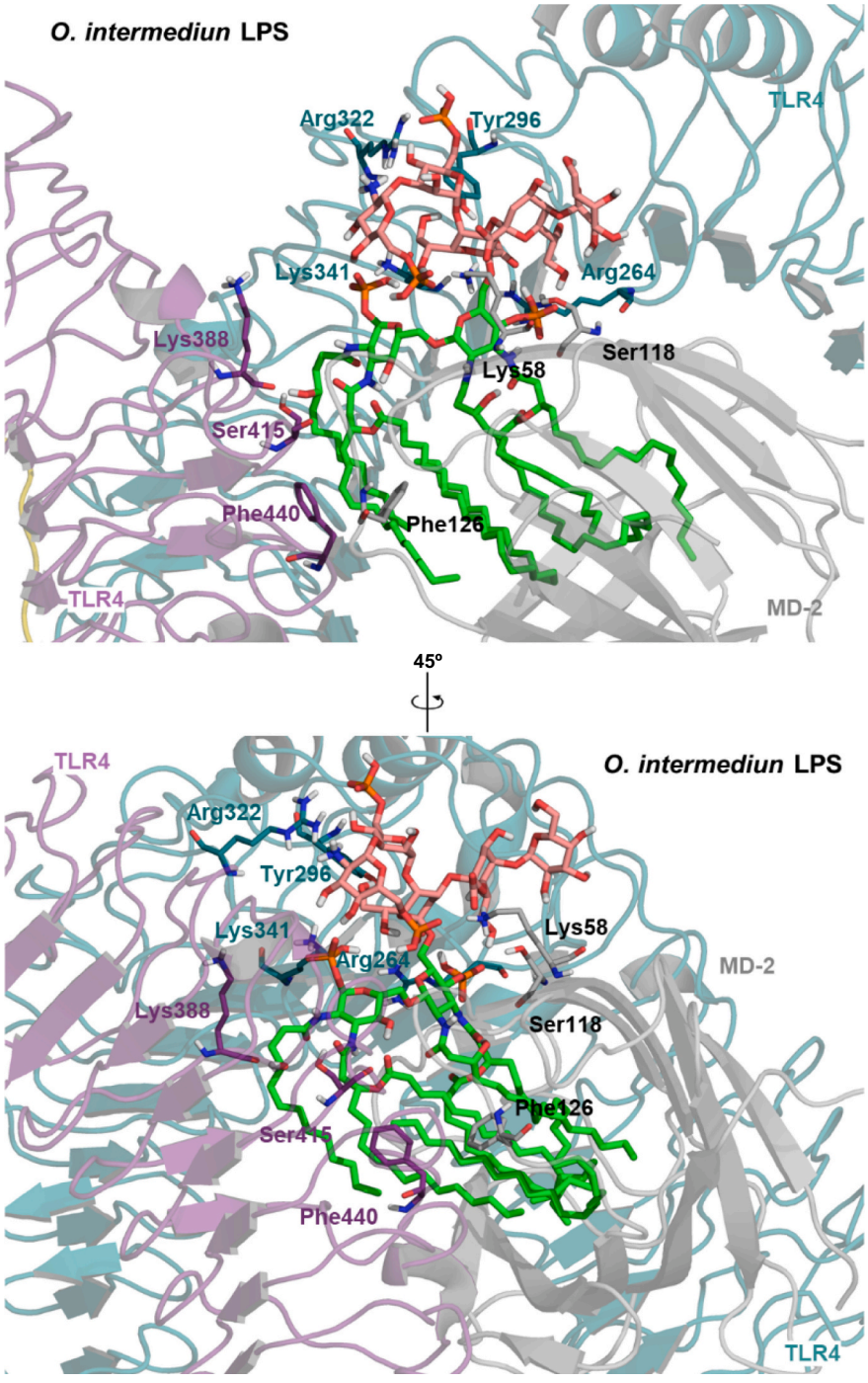

Fig. 7. Close-up of the binding pocket of MD-2, showing the major interactions between *O. intermedium* core and the activated TLR4 dimer. LPS is represented in sticks, with the lipid A moiety colored in green, and the core in salmon pink. The residues of the receptor that interacted with the ligands are also in sticks, with their corresponding individual labels. Proteins MD-2 (in grey), TLR4 (in purple), and the counterpart TLR4 (in aquamarine blue) are shown as semitransparent cartoon. The counterpart ligand (LPSs bound to the partner MD-2 molecule) is hidden for clarity.

respectively. In particular, the GlcN-P phosphate group formed one H-bond with the hydroxyl group from the side chain of MD-2 Ser120 and participated in salt bridge interaction with the amino group from the side chain of MD-2 Lys58. On the other hand, the Man-P phosphate established salt bridge interaction with the guanidinium group of TLR4 Arg322. Regarding the carboxylate groups of the Kdo moieties, the carboxylate of Kdo I also participated in salt bridge interaction with the guanidinium of TLR4 Arg264, and in H-bond with the hydroxyl group of TLR4 Tyr296, whereas the carboxylate of Kdo II established salt bridge interaction with the amino group from the side chain of MD-2 Lys58. The remaining saccharide residues participated in intramolecular H-bonds. Interestingly, the positively charged amino group of the Man-P does not seem to play an important role in the TLR4 molecular recognition of the *O. intermedium* core, as this group remained exposed to the solvent, without participating in any interaction with the receptor (Fig. 3 and Fig. 7).

As in the case of the TLR4 in complex with *Brucella*, five of the six acyl chains of the lipid A remained inside the MD-2 pocket, participating in hydrophobic interactions with side chains from MD-2 hydrophobic residues, such as Phe126 residue. Although the LPS molecule was displaced from the receptor interface over simulation time, the primary acyl chain at position 3 of the disaccharide backbone of lipid A remained close to the partner TLR4 and formed one H-bond with the hydroxyl group of Ser415, and hydrophobic interactions with TLR4 Phe440 and MD-2 Phe126 (Fig. 3 and Fig. 7). The (*R*)-3-hydroxylaurate chain at position 2 of the disaccharide backbone, remained outside the MD-2

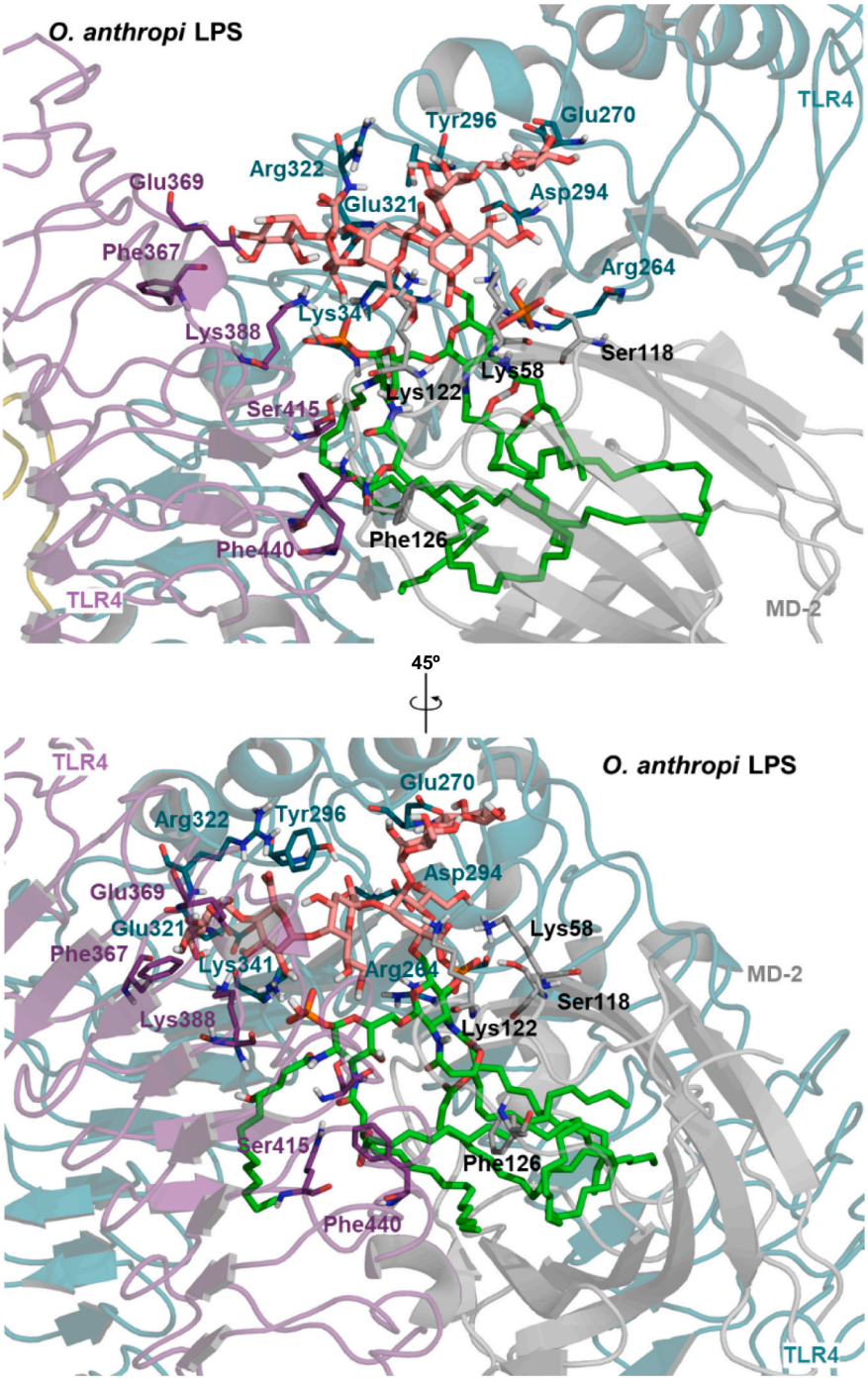

Fig. 8. Close-up of the binding pocket of MD-2, showing the major interactions between *O. anthropi* core and the activated TLR4 dimer. LPS is represented in sticks, with the lipid A moiety colored in green, and the core in salmon pink. The residues of the receptor that interacted with the ligands are also in sticks, with their corresponding individual labels. Proteins MD-2 (in grey), TLR4 (in purple), and the counterpart TLR4 (in aquamarine blue) are shown as semitransparent cartoon. The counterpart ligand (LPSs bound to the partner MD-2 molecule) is hidden for clarity.

pocket. Nevertheless, this chain was displaced from its initial position at the TLR4 interface and bent towards MD-2. As a result, the chain remained in the TLR4-interacting dimerization, but not as deep as the same chain in the *Brucella* complex. The chain formed hydrophobic interactions with the side chains of counterpart TLR4 residues Thr413 and Gln436, whereas the hydroxyl group of the acyl chain participated in H-bond with the backbone carbonyl group of the partner TLR4 Lys388 (Fig. 7 and Fig. S.2).

2.2.4. *O. anthropi*

During the simulation trajectory, residues of the *O. anthropi* core extended along the entire TLR4, also encompassing part of the receptor dimerization interface, and formed an extensive network of protein-ligand interactions. In this extended conformation, the core masked part of the lipid A diglucosamine backbone, including the 1-phosphate group (Fig. 3 and Fig. 8). Consequently, only the salt bridge between the 4′-phosphate of the lipid A and the guanidinium group of TLR4 Arg264 predicted by docking was maintained, whereas interactions between the 1-phosphate group of lipid A and the partner TLR4 were lost. Instead, both the 1-phosphate group and the corresponding residues of the counterpart TLR4 (*i.e.,* Lys341 and Ly388) established interactions with the LPS core region (Fig. 3 and Fig. 8). The core Glc residue was placed at the TLR4 interface, where participated in H-bonds with the carboxylate group of the TLR4 Glu321 side chain, and with residues from the partner TLR4, in particular, with the carboxylate and amino groups from the side chains of Glu369 and Lys388 residues, respectively, and with the carbonyl group of the Phe367 backbone. The GalA moiety also formed a H-bond with the amino group from the side chain of the partner TLR4 Lys388 residue, plus two more H-bonds with the carboxylate and amino groups from the TLR4 Glu321 and Lys341 side chains, respectively. Moreover, salt bridge interaction was observed between the carboxylate group of the GalA saccharide and the guanidinium group of TLR4 Arg322. The core Kdo II mainly formed intramolecular hydrogen bonds, and only one H-bond with the protein, within the saccharide carboxylate group. In particular, with the hydroxyl group from the TLR4 Tyr296 side chain. The Kdo I also participated in polar interactions with TLR4 through the formation of two H-bonds with the carboxylate group from the Asp294 side chain. Interestingly, the negatively charged carboxylate group of Kdo I was oriented towards MD-2 and formed two H-bonds with the amino groups from the side chains of Lys58 and Lys122. These two interactions were the only polar interactions observed between the ligand and MD-2 protein, together with the H-bond established among the hydroxyl group of the (*R*)-2′-hydroxymyristate chain of the lipid A and the backbone carbonyl group of MD-2 Ser118 (Fig. 3 and Fig. 8). The Man residue only participated in intramolecular H-bonds, whereas the GlcN amino group formed salt bridge interaction with the carboxylate group from the TLR4 Glu270 side chain. Again, five of the six lipid A acyl chains remained inside the MD-2 pocket, as predicted by docking calculations, and established hydrophobic interactions with the alkyl side chains from many MD-2 hydrophobic residues, including Phe126 residue.

Polar residues from both primary and secondary acyl chains at position 3 of the disaccharide backbone of lipid A remained close to the partner TLR4 and participated in H-bonds with the amide NH2 from the Gln436 side chain, and with the hydroxyl group of Ser415, and in hydrophobic interactions with the side chain of MD-2 Phe126 (Fig. 3 and Fig. 8). All these residues play a key role in the recognition of *E. coli* lipid A (B. S. Park et al., 2009). However, no hydrophobic interactions were observed between the lipid A chains and the partner TLR4 residues such as Phe440, that are also key players in the *E. coli* lipid A-driven TLR4 assembly (B. S. Park et al., 2009). The (*R*)-3-hydroxylaurate chain at position 2 of the disaccharide backbone remained outside the MD-2 pocket and interacted with residues from the TLR4 dimerization interface, but in a lesser extent to *Brucella* lipid A, and in a more similar manner to the observed for the *O. intermedium* complex. In this case, the chain was stabilized by hydrophobic interactions with the side chains of TLR4 residues Phe408 and His431, and *via* the formation of one H-bond between the hydroxyl group of the lipid chain and the backbone carbonyl group of TLR4 Gly363 (Fig. 8 and Fig. S.2).

2.3. TLR4 conformational changes

We also monitored the motion of the Phe126 side chain, known as the switch on/off of the receptor complex (Fig. S.3) (Paramo et al., 2013), over simulation time. In all the complexes, the MD-2 loop from residues 123 to 129, which contains the Phe126 residue, was stabilized by hydrophobic interactions with either the primary (*i.e.,* the (*R*)-3-hydroxypalmitate chain), or the secondary (*i.e.,* the VLFCA of 28C) acyl chains at position 3 of the lipid A disaccharide. Therefore, no conformational changes of this loop occurred during MD simulations, and the Phe126 remained in an agonist-like conformation in all the systems (Paramo et al., 2013).

Nevertheless, although the MD-2 Phe126 agonist conformation (Paramo et al., 2013) was retained over simulation time (Fig. S.3), the *O. intermedium* and *B. abortus* complexes presented global conformational changes compared to the crystal structure of the activated TLR4 (PDB ID 3FXI) (B. S. Park et al., 2009), that affected the correct assembly of the receptor. During the MD trajectories of both the *O. intermedium* and *Brucella* complexes, TLR4 chains displayed a tendency to dissociate one from each other, disfavoring the activated agonist geometry of the receptor. To evaluate the dissociation degree of the TLR4 complex interface, we calculated the angle θ, defined between the centers of mass of TLR4 residues from Glu286 to Ala291 of one chain, and residues from Phe533 to Phe538 of the counterpart chain. The angle was plotted over time for all the studied complexes (Fig. 9.A-B). The *B. abortus* system exhibited the largest θ values which range from θ ~75–80°, followed by the *O. intermedium* complex, with values from θ ~72–77°, whereas *O. anthropi* and *Bm-wadC* complexes showed no TLR4 interface dissociation, with stable θ values around 70° (Fig. 9.B). As a consequence, the network of protein-protein interactions at the receptor binding interface was weakened in the *O. intermedium* and *B. abortus* systems (Fig. S.4). This observation backs up the hypothesis that *O. intermedium* and *Brucella* LPSs act as poor agonists by impairing the proper formation of the TLR4 receptorial complex. On the other hand, the protein-protein interface from the TLR4 activated complex (PDB ID 3FXI) (B. S. Park et al., 2009) was well maintained in the *O. anthropi* and *Bm-wadC* complexes, as deduced from the contact maps between pair of residues belonging to the leucine-rich repeat modules 12–18 of the TLR4 extracellular domain (residues from 400 to 565) (B. S. Park et al., 2009) (Fig. 9.B). In the *O. anthropi* complex, new interactions were established at the TLR4 region comprising residues from 550 to 590, indicating that the LPS brings the C-terminus of the TLR4 ectodomain closer together. After the dissociation of the TLR4 subunits in the *O. intermedium* and *Brucella* TLR4 complexes, the interface of both binding TLR4 chains gained solvent accessibility (Fig. S.5), which is a direct consequence of the loss of protein-protein interactions.

Finally, to investigate the stabilization role of the studied lipopolysaccharides towards the TLR4 receptor, we also predicted the interaction energy at the dimerization interface of the $(TLR4/MD\text{-}2/LPS)_2$ complexes and correlated with angle θ. We calculated the Gibbs free energy (ΔG) for protein-protein interactions between both TLR4/MD-2 heterodimers during MD simulations. As observed in Fig. 9.C the ΔG values were -104.6344 ± 9.4979 kcal·mol^{-1} for *O. anthropi*, −99.8927 ± 11.5952 kcal·mol^{-1} for *Bm-wadC*, −96.5931 ± 12.3484 kcal·mol^{-1} for *O. intermedium*, and -69.082 ± 10.4565 kcal·mol^{-1} for *Brucella*. It is clear that the most energetic conformation (*i.e.,* the less stable) was displayed in the *Brucella* system. This suggests that larger θ values impair the stabilization of the TLR4 receptor, in agreement with the observed loss of protein-protein interactions during MD simulations of the *Brucella* and *O. intermedium* complexes (Fig. 9). Altogether, results point out that the LPSs of *O. anthropi* and *Bm-wadC* stabilize the $(TLR4/MD\text{-}2)_2$ system more efficiently than *O. intermedium* and *Brucella* LPSs.

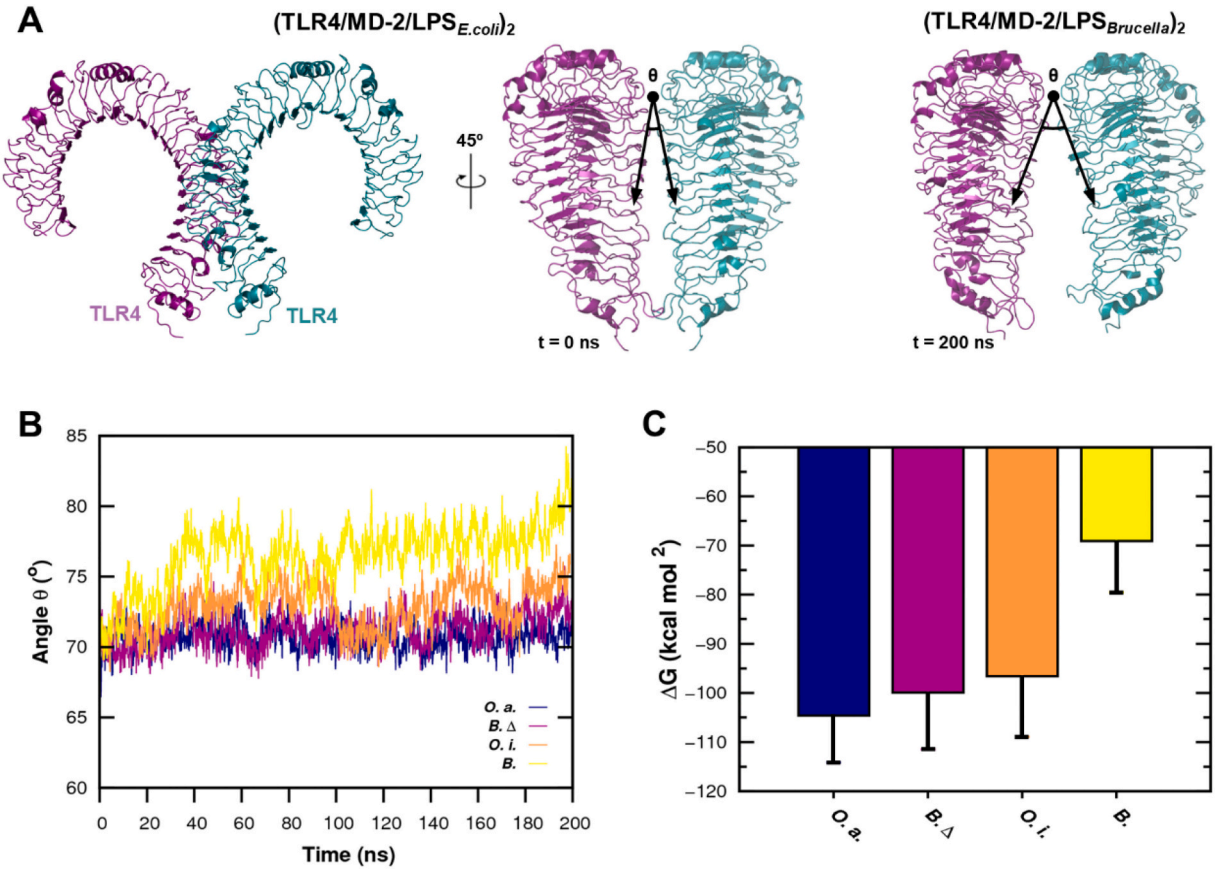

Fig. 9. Evolution of the TLR4 in complex with the α2-Proteobacteria LPSs during MD simulations. (A) Representation of angle θ, defined as the angle between both TLR4 chains, used to follow the TLR4 conformation changes associated with the presence of the ligands. (B) Monitorization of angle θ plotted over simulation time. (C) Gibbs free energy, ΔG, predicted for the TLR4 protein-protein interface (between chains A and B). O.a: *O. anthropi*; B.Δ: *Bm-wadC*; O.i. *O. intermedium*; and B.: *Brucella abortus*.

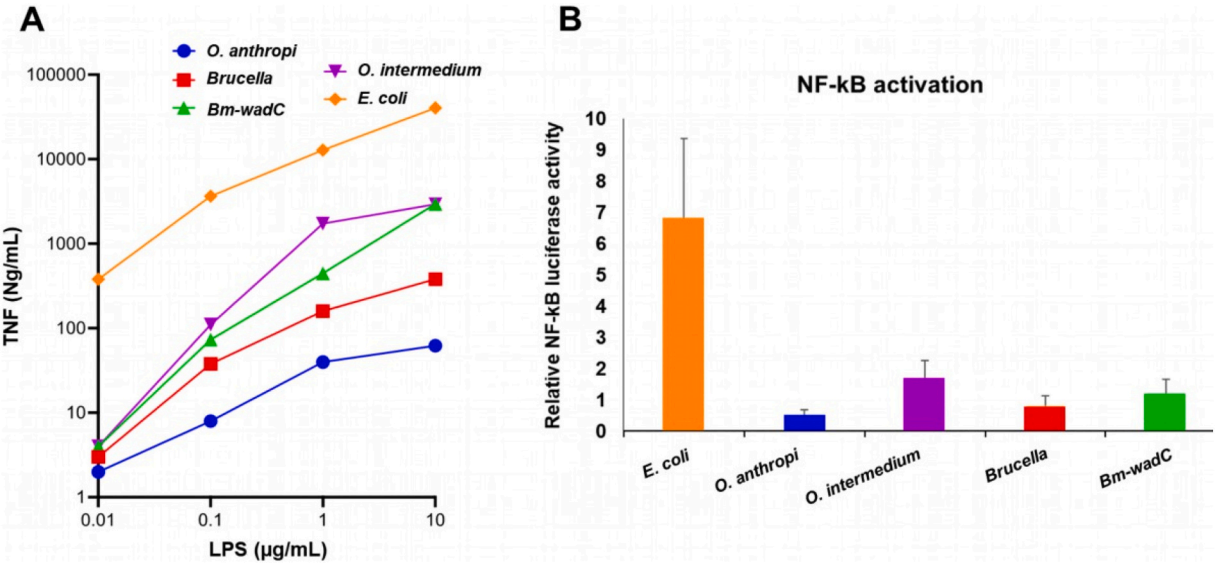

Fig. 10. LPS from α2-Proteobacteria induce lower proinflammatory TNF levels and NF-kB activation, compared to the *E. coli* LPS. (A) Macrophage cell line Raw was stimulated with α2-Proteobacteria LPS (*O. intermedium*, *O. anthropi*, *B. melitensis* or *Bm-wadC* LPS) or *E. coli* LPS at a concentration of 0,01, 0,1, 1 and 10 μg/mL for 24 h. TNF secretion was assayed from macrophage supernatants by ELISA and data are expressed as the mean of two independent experiments in duplicate. (B) HEK 293-hTLR2/1 and HEK 293/hTLR4A-MD2-CD14 cells were cotransfected with the pNF3ConA Luc (NF-kB) firefly reporter construct and the thymidine kinase promoter-Renilla reporter plasmid, and stimulated with LPS purified from *Escherichia coli* (100 ng/mL) and LPS (10 mg/mL) from the other indicated bacteria. The activity of Firefly and Renilla luciferases were measured 24 h after, normalized among them, and the ratio of values obtained in TLR4 to TLR2 (non-specific-one) was represented.

2.4. Proinflammatory activities of α2-Proteobacteria LPSs

In an attempt to translate those structural results to biological activity, we tested over a wide dose-response range, the effect of *O. anthropi, O. intermedium, B. melitensis* or *Bm-wadC* LPS in the induction of TNF in murine Raw macrophages. As seen in Fig. 10, all LPS were at least 100-fold less potent than the standard pro-inflammatory LPS from *E. coli*. Interestingly, the elimination of the branched side of the core of *Brucella* LPS increased its inflammatory potency as described previously (Fontana et al., 2016).

3. Discussion

Our observations, together with previously reported results,(Francisco et al., 2022) provide key insights into the structure-activity relationship of some α2-Proteobacteria LPSs. First, we demonstrate that the length of the acyl chains is a critical factor in determining the defective TLR4 activation by lipids A that bear VLFCAs. It is clear that the bulky lipid A is not able to be fully accommodated in the MD-2 hydrophobic pocket. Consequently, the lipid A disaccharide moiety is displaced from the MD-2 pocket, lacking key interactions with residues from the MD-2 pocket rim and the partner TLR4. Additionally, the sixth acyl chain of lipid A (the (*R*)-3-hydroxylaurate chain) is placed at the protein-protein interface, impairing the proper dimerization of the receptor complex.

Second, from the MD simulations analysis, a common pattern can be established for *O. intermedium*, *O. anthropi*, and *Brucella* LPSs; the presence of the core hinders the correct recognition of the lipid A 1-phosphate by the TLR4. In the *O. anthropi* and *Brucella* complexes, part of the core oligosaccharide was accommodated at the receptor dimerization interface, masking determinant interactions between the lipid A 1-phosphate and TLR4. Interactions among *O. intermedium* core and the receptor led to the displacement of the lipid A diglucosamine backbone from the dimerization interface of the complex, also lacking interactions between the 1-phospate and the partner TLR4 chain. On the other hand, MD simulations of the TLR4 in complex with the *Bm-wadC* LPS mutant demonstrated that the absence of the positively charged glucosamine branch in the core (as in *Brucella* WT) allows the interaction of the lipid A 1-phosphate with key residues of the receptor, which may explain its relatively increased pro-inflammatory activity in absence of core (Smith, 2018).

The importance of the lipid A 1-phosphate group for the TLR4 activation is well known, as the *E. coli* lipid A lacking the 1-phosphate (*i.e.,* the monophosphoryl lipid A, MPLA) is at least 100-fold less toxic than the bisphosphorylated lipid A (Molinaro et al., 2015). Hence, our results point to the impairment of the correct recognition of the lipid A 1-phosphate in the *O. intermedium*, *O. anthropi*, and *Brucella* complexes as one of the causes of their immune evasion. These observations are in agreement with previous experimental results that linked the specific removal of the glucosamines side branch with the increased inflammatory response of *Bm-wadC* LPS compared to the *Brucella* WT LPS (Fontana et al., 2016).

In the case of *O. anthropi*, the loss of interactions between the *O. anthropi* 1-phosphate and the TLR4 counterpart was also observed, although it may establish many interactions with key receptor residues, including those residues of the TLR4 partner that interact with the 1-phosphate group of *E. coli* lipid A (B. S. Park et al., 2009). Therefore, the loss of interactions between the *O. anthropi* 1-phosphate and the TLR4 counterpart is somehow compensated through the interaction of these TLR4 residues with the LPS core.

Finally, MD simulations of the LPSs complexes reinforced the hypothesis that the MD-2 pocket is not big enough to accommodate the six lipid A chains of α2-Proteobacteria LPS since the lipid A disaccharide moiety was shifted upwards from MD-2 cavity, and the (*R*)-3-hydroxylaurate was protruding towards the protein-protein interface, impairing the proper dimerization of the receptor complex.

Together, it could be suggested that some atypical α-2 Proteobacteria LPSs may evade immune system activation by either preventing TLR4 complex formation or by disturbing the complex stability. Nonetheless, we cannot discard that its peculiar lipid A-core structures, different from typical LPS of Gram-negative bacteria, may also result in some cases, in the heterodimer formation with TLR2 as we described in the LPS of *O. intermedium.*(Francisco et al., 2022) Thus, despite that the LPSs of *O. anthropi.* and *Bm-wadC may* stabilize the $(TLR4/MD\text{-}2)_2$ system more efficiently than *O. intermedium* LPS, this LPS may compensate it through heterodimerization with TLR2, which could explain the biological results in Fig. 10.

In conclusion, our results clearly indicate that the biological activity of LPSs and their relationship with TLRs is more complex than generally thought. Thus, not only the well know interaction of lipid A with TLR4, which is affected by the number and length of acyl chains, but also subtle differences in LPS core structure affect TLR4/MD2 interaction, and homo and heterodimerization (with TLR2) of TLR4. A deep structural understanding of TLRs activation (and/or evasion) could help to modulate the innate immunity system in sepsis control, inflammation-based diseases, and cancer vaccines, among other relevant therapeutic applications.

4. Methods

4.1. Macromolecule preparation

3D coordinates from the X-ray structure of the human $(TLR4/MD\text{-}2/E.\ coli\ LPS)_2$ ectodomain (PDB ID 3FXI)(B. S. Park et al., 2009) were retrieved from the Protein Data Bank (www.rcsb.org). Solvent, ligands, and ions were removed. Hydrogen atoms were added to the X-ray structure using the preprocessing tool of the Protein Preparation Wizard of the Maestro package (Maestro | Schrödinger, 2020) The protein structure went through a restrained minimization under the OPLS3 force field (Harder et al., 2016) with a convergence parameter to RMSD for heavy atoms kept default at 0.3 Å.

4.2. Construction and optimization of the ligands

The 3D structure of α2-Proteobacteria lipid A was built with PyMOL molecular graphics and modeling package (Schrödinger & DeLano, 2020) using as a template the *E. coli* lipid A (PDB ID 3FXI) (B. S. Park et al., 2009). The oligosaccharide cores of α2-Proteobacteria LPSs were constructed using the Glycan Modeler implemented in CHARMM-GUI (S. J. Park et al., 2019) and attached to the lipid A moiety, using the builder tool in PyMOL (Schrödinger and DeLano, 202). The resulting structures were first refined at the AM1 level of theory and then optimized at the Hartree-Fock (HF) level (HF/6-311G**) with Gaussian09 (M. J. Frisch et al., 2016).

4.3. All-atom parametrization of the ligands

α2-Proteobacteria LPSs structures were split into residues to facilitate and homogenize the parametrization process. The partial charges and atom types of monosaccharides composing the oligosaccharide backbone were retrieved from the GLYCAM06 force field (Kirschner et al., 2008), by means of the Glycam Carbohydrate Builder web server (www.glycam.org). The parameters needed for α2-Proteobacteria lipid and hexa-AfLAs acyl chains, and for those monosaccharides not available in the GLYCAM06 force field (Kirschner et al., 2008) were obtained using the standard Antechamber procedure implemented in Amber16 (Case et al., 2016). The partial charges were derived from the HF calculations and formatted for AmberTools15 and Amber16 (Case et al., 2016) with Antechamber (J. Wang et al., 2006) using RESP charges (Bayly et al., 1993) and assigning the general Amber force field (GAFF) atom types (J. Wang et al., 2004). For the monosaccharides, the B3LYP/6-31G++(2d, 2p) level was used, consistent with the approach used in the GLYCAM force field development (Kirschner et al., 2008). Later, the atom types of the saccharide atoms were changed to the GLYCAM06

force field (Kirschner et al., 2008) atom types, and the atoms constituting the lipid chains to the Lipid14 force field (Dickson et al., 2014) atom types.

4.4. Refinement of α2-Proteobacteria cores

Due to the number and inherent flexibility of glycosidic linkages in the core, α2-Proteobacteria LPSs structures were subjected to MD refinement prior to docking. The accuracy and the computational cost of ligand conformational search during docking calculations depend on the number of rotatable bonds in the ligand and the size of the rotational angle increment (Woods, 2018). Indeed, conformational search of molecules with several rotatable bonds usually reduces docking accuracy and involves longer computing times. For that reason, the LPSs structures were submitted to all-atom MD simulations during 100 ns in the Amber16 suite (Case et al., 2016). The simulation box was designed such as the edges were distant of at least 10 Å of any atom. The system was solvated with the TIP3P water molecules model. Na^+ and Cl^- ions were added to counterbalance the eventual charges of the LPS molecules. All the simulations were performed with the same equilibration and production protocol. First, the system was submitted to 1000 steps of the steepest descent algorithm followed by 7000 steps of the conjugate gradient algorithm. A 100 kcal·mol^{-1}·A^{-2} harmonic potential constraint was applied to the ligand. In the subsequent steps, the harmonic potential was progressively lowered (respectively to 10, 5, and 2.5 kcal·mol^{-1}·A^{-2}) for 600 steps of the conjugate gradient algorithm each time, and then the whole system was minimized uniformly. Next, the system was heated from 0 to 100 K using the Langevin thermostat in the canonical ensemble (NVT) while applying a 20 kcal·mol^{-1}·A^{-2} harmonic potential restraint on the protein and the ligand. Finally, the system was heated up from 100 to 300 K in the isothermal-isobaric ensemble (NPT) under the same restraint condition as the previous step, followed by simulation for 100 ps with no harmonic restraint applied. At this point, the system was ready for the production run, which was performed using the Langevin thermostat under the NPT ensemble, at a 2 fs time step. Long-range electrostatics were calculated using the Particle Mesh Ewald (PME) method (Essmann et al., 1995).

4.5. Docking calculations

α2-Proteobacteria LPS ligands were divided into two moieties corresponding to their lipid A, common to all the studied LPSs, and to the core oligosaccharides. Carbohydrate conformational search was performed as a separate step before molecular docking, by means of MD simulations (Woods, 2018). To avoid the limitation of the use of only one scoring function, AutoDock Vina 1.1.2 (Oleg & Arthur, 2010) and AutoDock 4.2 (Morris et al., 2009) were used for the docking of α2-Proteobacteria lipid A and the α2-Proteobacteria cores in the human TLR4 agonist model. Preliminary docking poses were obtained with AutoDock Vina and the best-predicted docking poses were redocked with AutoDock 4. AutoDockTools 1.5.6 program (Morris et al., 2009) was used to assign the Gasteiger-Marsili empirical atomic partial charges to the atoms of both, the ligands, and the receptor. Non-polar hydrogens were merged for the ligands. The structure of the receptor and the ligands were set rigid and flexible, respectively. In AutoDock 4.2, the Lamarckian evolutionary algorithm was selected, and all parameters were kept default except for the number of genetic algorithm runs that was set to 100 to enhance the sampling. The box spacing was set to the default value of 1 Å in AutoDock Vina, and to 0.375 Å in AutoDock. The size of the box was set to 33.00, 40.50, and 35.25 Å in the x, y, z-axes respectively, with the box center located equidistant to the mass center of residues Arg90 (MD-2), Lys122 (MD-2), and Arg264 (TLR4), in both docking programs. The structure of the receptor was always kept rigid, whereas the structure of the ligands was set partially flexible by providing freedom to some carefully selected rotatable bonds. Since a prior MD conformational search was performed for the α2-Proteobacteria LPSs, those glycosidic bonds in the cores oligosaccharides that reached stability during the MD simulations were kept rigid (Woods, 2018).

4.6. All-atom MD simulations

Selected docking complexes were submitted to all-atom MD simulations during 200 ns in the Amber16 suite (Case et al., 2016). The protein was described by the ff14SB all-atom force field (Maier et al., 2015), the saccharide moieties of the ligands by the GLYCAM06 force field (Kirschner et al., 2008), and the lipid chains with the Lipid14 force field (Dickson et al., 2014). The simulation box was designed such as the edges were distant of at least 10 Å of any atom. The systems were solvated with the TIP3P water molecules model. Na^+ ions were added to counterbalance the eventual charges of the protein-ligand systems, when needed. All the simulations were performed with the same equilibration and production protocol. First, the system was submitted to 1000 steps of the steepest descent algorithm followed by 7000 steps of the conjugate gradient algorithm. A 100 kcal·mol^{-1}·A^{-2} harmonic potential constraint was applied to both the proteins and the ligand. In the subsequent steps, the harmonic potential was progressively lowered (respectively to 10, 5, and 2.5 kcal·mol^{-1}·A^{-2}) for 600 steps of the conjugate gradient algorithm each time, and then the whole system was minimized uniformly. Next, the system was heated from 0 to 100 K using the Langevin thermostat in the canonical ensemble (NVT) while applying a 20 kcal·mol^{-1}·A^{-2} harmonic potential restraint on the protein and the ligand. Finally, the system was heated up from 100 to 300 K in the isothermal-isobaric ensemble (NPT) under the same restraint condition as the previous step, followed by simulation for 100 ps with no harmonic restraint applied. At this point, the system was ready for the production run, which was performed using the Langevin thermostat under the NPT ensemble, at a 2 fs time step. Long-range electrostatics were calculated using the PME method (Essmann et al., 1995).

4.7. All-atom MD analysis

Trajectory analysis was performed using the cpptraj module (Roe & Cheatham, 2013) of AmberTools15 (Case et al., 2016) and a combination of in-built Gromacs(Abraham et al., 2015) utilities and in-house written scripts. Trajectories were converted to Gromacs (Abraham et al., 2015) suite formats by using ParmEd tools. Molecular visualization and graphics were generated using visual molecular dynamics (VMD) software (Humphrey et al., 1996), and PyMOL molecular graphics and modeling package (Schrödinger & DeLano, 2020).

Root-mean-square deviation (RMSD): we calculated the RMSD values for the α-carbons of each protein conformation in a given trajectory with respect to the α-carbons in the initial conformation of the same trajectory using the standard *gmx rms* function. RMSD analysis was used as parameter to determine the equilibrated state of all systems.

Number of contacts: the number of contacts between both TLR4 and MD-2 were calculated by means of the *gmx mindist* tool. A contact was determined within a distance $\leq$4.5 Å.

Contact maps: the *gmx mdmat* tool and a python script written in-house (see Supplementary data) were used to determine the contacts between all atoms from residues involved in a given protein-protein interface, in each frame. A contact was determined within a range of $0 \leq$ distance $\leq$4.5 Å. Calculations were performed during the time interval from 120 to 160 ns of MD simulations, as the RMSD values for all the complexes show the lowest fluctuations during this period.

Solvent-accessible surface area: the SASA of the extracellular domain of TLR4 was calculated using the *gmx sasa* tool, as implemented in Gromacs (Abraham et al., 2015).

Phe126 angle: cpptraj was used to predict the angle over simulation time between two arbitrarily selected vectors starting both from the α-carbon (CA) of residue MD-2 Phe126 to, respectively, the zeta-carbon (CZ) of the same residue and the α-carbon of residue MD-2 Ser31.

Angle θ: angle between both TLR4 chains was predicted using a tcl script written in-house (see Supplementary data), and the visual molecular dynamics (VMD) software (Humphrey et al., 1996).

Gibbs free energy (ΔG): Molecular Mechanics Generalized Born Surface Area approach (MM/GBSA) (E. Wang et al., 2019), implemented in Amber16 (Case et al., 2016), was applied with the igb = 2 implicit solvent model to calculate the enthalpic contribution to the protein-protein binding affinity between both TLR4/MD-2 heterodimers in complex with α2-Proteobacteria LPSs. Calculations were performed during the time interval from 120 to 160 ns of MD simulations, as the RMSD values for all the complexes show the lowest fluctuations during this period.

4.8. LPS stimulation of macrophages

LPS from *O. anthropi* LMG 3331 and from *E. coli* were obtained from Julian Velasco, and those from *B. melitensis* and *Bm-wadC,* were obtained as reported somewhere else (Francisco et al., 2022). *O. intermedium* LGM 3306 was purified and characterized as previously described. The purity of those compounds was assessed by mass-spectrometry, with a purity level higher than 98 %.

The murine macrophage cell line Raaw was cultured in RPMI 1640 medium (Gibco) (2 mM L-glutamine, antibiotics), supplemented with 5 % FBS (Merck). Cells were cultured in 12-well plates and serum deprived for 16 h prior LPS stimulation for 24 h. TNF-α concentration was determined using DuoSet ELISA kit from R&D systems, according to the manufacturers protocol.

4.9. HEK-TLR-expressing cells NF-kB reporter assays

Stable immortalized HEK 293T, HEK 293-hTLR2/6, and HEK 293/hTLR4A-MD-2-CD14 cells (InvivoGen, San Diego, CA, USA) were plated at 3×10^6 cells in 6-well plates growing at 37 °C in DMEM culture medium supplemented with 5 % FBS, 2 mM L-glutamine, 100 U/mL gentamycin, 0.01 % pyruvate, and 0.4 mM non-essential amino acids. Then, 24 h later, cells were cotransfected with the pNF3ConA Luc [nuclear factor (NF)-kB] Firefly reporter construct and the thymidine kinase promoter Renilla reporter plasmid (100:1 ratio) using Metafectene PRO (Biontex, Plannegg, Germany). Transfection medium was changed after 24 h, and cells were seeded at 1.3×10^4 cells per well in 96-well plates. Then, 24 h later, ligands were added. Activities of Firefly and Renilla luciferases were measured 24 h after using TwinLite Firefly and Renilla Luciferase Reporter Gene Assay System (PerkinElmer) in Fluo Star Optima (BMG) plate reader (three replicates per condition). All ratios were compared with the control condition (non-stimulated cells). We used FSL-1 (InvivoGen), TNF-a (InvivoGen), and LPS purified from *E. coli* (Sigma), *O. intermedium*, *O. anthropi*, *B. melitensis* and *Bm-wadC*.

Credit authorship contribution statement

M.F., and S.M.-S. carried out the conception of the work. A.M.-R., and S.M—S. designed the computational studies. A.M.-R. performed the docking calculations, simulations, and analysis. J.M., A.G.-J., and R. C-A. performed the biological assays. S.M.-S. performed the supervision of the computational studies. M.F. performed the supervision of the biological assays. A.M.-R., M.F., and S.M.-S. wrote and reviewed the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was financially supported by the Spanish Ministry for Science and Innovation (grants CTQ2017-88353-R, and PID2020-113588RB-I00 to S.M.S.; grant PRE2018-086249 to A.M.R., and SAF2016-75988 to M.F.), Instituto de Salud Carlos III (RD16/0027/0006), European Commission – NextGenerationEU (Regulation EU 2020/2094) through CSIC's Global Health Platform (PTI Salud Global), Comunidad de Madrid (S2017/BMD-3671. INFLAMUNE-CM) and from Laboratorios Esteve to M.F. and institutional grants from "Fundación Ramón Areces" and "Banco de Santander" to CBMSO. Endless support from S.H.J. is sincerely acknowledged.

Abbreviations

3D	Three-dimensional
B. melitensis	*Brucella melitensis*
C	Carbon
CA	α-Carbon
CD14	Cluster of Differentiation 14
CZ	Zeta-carbon
E. coli	*Escherichia coli*
FA	Fatty acid
GAFF	General Amber force field
Gal	Galactose
GalA	Galacturonic acid
GLA	Glucopyranosyl lipid A
GlcN	Glucosamine
GlcN-P	Glucosamine 6-phosphate
Glu	Glucose
Hep	Heptose
HF	Hartree Fock
hTLR2	Human Toll-like receptor 2
hTLR4	Human Toll-like receptor 4
Kdo	Ketodeoxyoctonic acid
LPS	Lipopolysaccharide
Man	Mannose
Man-P	Mannose 6-phosphate
MD	Molecular Dynamics
MD-2	Myeloid Differentiation factor 2
MM/GBSA	Molecular Mechanics Generalized Born Surface Area
MPLA	Monophosphoryl lipid A
NF-Kb	Early nuclear factor kappa-light-chain enhancer of activated B cells
O. antrhopi	*Ochrobactrum anthropi*
O. intermedium	*Ochrobactrum intermedium*
PDB	Protein Data Bank
PME	Particle Mesh Ewald
RESP	Restrained electrostatic potential
RMSD	Root-mean-square deviation
SASA	Solvent-accessible surface area
t	Time
T	Temperature
TLR	Toll-like receptor
TLR2	Toll-like receptor 2
TLR4	Toll-like receptor 4
TNF	Tumor necrosis factor receptor
VLFA	Very long-chain fatty acid
VMD	Visual molecular dynamics
WT	Wild type
ΔG	Free energy of Gibbs

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.carbpol.2023.121094.

References

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindah, E. (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX, 1–2*, 19–25.

Barquero-Calvo, E., Conde-Alvarez, R., Chacón-Díaz, C., Quesada-Lobo, L., Martirosyan, A., Guzmán-Verri, C., … Chaves-Olarte, E. (2009). The differential interaction of Brucella and ochrobactrum with innate immunity reveals traits related to the evolution of stealthy pathogens. *PLoS One, 4*(6).

Bayly, C. I., Cieplak, P., Cornell, W. D., & Kollman, P. A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *Journal of Physical Chemistry, 97*(40), 10269–10280.

Case, D. A., Betz, R. M., Cerutti, D. S., Cheatham, T. E., III, Darden, T. A., Duke, R. E., … Homeyer, N. (2016). *AMBER 2016 reference manual* (pp. 1–923). San Francisco, CA, USA, April: University of California.

Dickson, C. J., Madej, B. D., Skjevik, Å. A., Betz, R. M., Teigen, K., Gould, I. R., & Walker, R. C. (2014). Lipid14: The amber lipid force field. *Journal of Chemical Theory and Computation, 10*(2), 865–879.

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *The Journal of Chemical Physics, 103*(19), 8577–8593.

Fontana, C., Conde-Álvarez, R., Ståhle, J., Holst, O., Iriarte, M., Zhao, Y., … Widmalm, G. (2016). Structural studies of lipopolysaccharide-defective mutants from Brucella melitensis identify a core oligosaccharide critical in virulence. *The Journal of Biological Chemistry, 291*(14), 7727–7741.

Francisco, S., Billod, J. M., Merino, J., Punzón, C., Gallego, A., Arranz, A., … Fresno, M. (2022). Induction of TLR4/TLR2 interaction and heterodimer formation by low Endotoxic atypical LPS. *Frontiers in Immunology, 12*.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., … D., J. F. (2016). Expanding the limits of computational chemistry. In *Gaussian 09, Revision A.02*. Wallingford CT: Gaussian, Inc.

Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J. Y., … Friesner, R. A. (2016). OPLS3: A force field providing broad coverage of drug-like small molecules and proteins. *Journal of Chemical Theory and Computation, 12*(1), 281–296.

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics, 14*(1), 33–38.

Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L., & Woods, R. J. (2008). *GLYCAM06: A generalizable biomolecular*.

Maestro | Schrödinger. (2020). Available at: https://www.schrodinger.com/maestro.

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of Chemical Theory and Computation, 11*(8), 3696–3713.

Molinaro, A., Holst, O., Di Lorenzo, F., Callaghan, M., Nurisso, A., D'Errico, G., Zamyatina, A., Peri, F., Berisio, R., Jerala, R., Jiménez-Barbero, J., Silipo, A., & Martín-Santamarí, S. (2015). Chemistry of lipid a: At the heart of innate immunity. In , *Vol. 21, Issue 2. Chemistry - A European journal* (pp. 500–519). Wiley-VCH Verlag.

Morris, G. M., Ruth, H., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry, 30*(16), 2785–2791.

Oleg, T., & Arthur, J., .O. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry, 31*, 455–461.

Onto, U., Fukase, K., Miyake, K., & Satow, Y. (2007). Crystal structures of human MD-2 and its complex with antiendotoxic lipid IVa. *Science., 316*(5831), 1632–1636.

Paramo, T., Piggot, T. J., Bryant, C. E., & Bond, P. J. (2013). The structural basis for endotoxin-induced allosteric regulation of the Toll-like receptor 4 (TLR4) innate immune receptor. *The Journal of Biological Chemistry, 288*(51), 36215–36225.

Park, B. S., Song, D. H., Kim, H. M., Choi, B. S., Lee, H., & Lee, J. O. (2009). The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature, 458*(7242), 1191–1195.

Park, S. J., Lee, J., Qi, Y., Kern, N. R., Lee, H. S., Jo, S., … Im, W. (2019). CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology, 29*(4), 320–331 (h).

Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation, 9*(7), 3084–3095.

Schrödinger, L., & DeLano, W. (2020). PyMOL. Available at: http://www.pymol.org/pymol.

Smith, J. A. (2018). Brucella lipopolysaccharide and pathogenicity: The core of the matter. *Virulence, 9*(1), 379–382.

Velasco, J., Moll, H., Knirel, Y. A., Sinnwell, V., Moriyón, I., & Zähringer, U. (1998). Structural studies on the lipopolysaccharide from a rough strain of Ochrobactrum anthropi containing a 2,3-diamino-2,3-dideoxy-D-glucose disaccharide lipid a backbone. *Carbohydrate Research, 306*(1–2), 283–290.

Wang, E., Sun, H., Wang, J., Wang, Z., Liu, H., Zhang, J. Z. H., & Hou, T. (2019). End-point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design. *Chemical Reviews, 119*(16), 9478–9508.

Wang, J., Wang, W., Kollman, P. A., & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling, 25*(2), 247–260.

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general Amber force field. *Journal of Computational Chemistry, 25*(9), 1157–1174.

Woods, R. J. (2018). Predicting the structures of glycans, glycoproteins, and their complexes. *Chemical Reviews, 118*(17), 8005–8024.