LETTERS

The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex

Beom Seok Park[1], Dong Hyun Song[1], Ho Min Kim[1], Byong-Seok Choi[1], Hayyoung Lee[3] & Jie-Oh Lee[1,2]

The lipopolysaccharide (LPS) of Gram negative bacteria is a well-known inducer of the innate immune response[1]. Toll-like receptor (TLR) 4 and myeloid differentiation factor 2 (MD-2) form a heterodimer that recognizes a common 'pattern' in structurally diverse LPS molecules. To understand the ligand specificity and receptor activation mechanism of the TLR4–MD-2–LPS complex we determined its crystal structure. LPS binding induced the formation of an m-shaped receptor multimer composed of two copies of the TLR4–MD-2–LPS complex arranged symmetrically. LPS interacts with a large hydrophobic pocket in MD-2 and directly bridges the two components of the multimer. Five of the six lipid chains of LPS are buried deep inside the pocket and the remaining chain is exposed to the surface of MD-2, forming a hydrophobic interaction with the conserved phenylalanines of TLR4. The F126 loop of MD-2 undergoes localized structural change and supports this core hydrophobic interface by making hydrophilic interactions with TLR4. Comparison with the structures of tetra-acylated antagonists bound to MD-2 indicates that two other lipid chains in LPS displace the phosphorylated glucosamine backbone by ~5 Å towards the solvent area[2,3]. This structural shift allows phosphate groups of LPS to contribute to receptor multimerization by forming ionic interactions with a cluster of positively charged residues in TLR4 and MD-2. The TLR4–MD-2–LPS structure illustrates the remarkable versatility of the ligand recognition mechanisms employed by the TLR family[4,5], which is essential for defence against diverse microbial infection.

Minute amounts of LPS released from invading bacteria are an early sign of infection and prepare the immune system to counteract further infection[1]. They can also lead to fatal septic shock syndrome if the inflammatory response is amplified and uncontrolled. LPS is a glycolipid located in the outer membrane of Gram-negative bacteria. It is composed of an amphipathic lipid A component and hydrophilic polysaccharides of the core and O-antigen[6,7]. Lipid A represents the conserved molecular pattern of LPS and is the main inducer of immunological responses to LPS. TLR4 in association with MD-2 is responsible for the physiological recognition of LPS[8,9]. So far, ten members of the TLR family, recognizing a wide variety of microbial products, have been identified in humans[10]. The extracellular domains of TLRs consist of leucine-rich repeats (LRRs) with a horseshoe-like shape[2,4,11,12]. Binding of agonistic ligands causes dimerization of the extracellular domains and is believed to trigger the recruitment of specific adaptor proteins to the intracellular domains, thus initiating a signalling cascade[4,5] (Supplementary Fig. 1). LPS is extracted from the bacterial membrane and transferred to TLR4–MD-2 by two accessory proteins, LPS-binding protein and CD14[13]. The TLR4–MD-2 heterodimer has complex ligand specificity. It can be activated by structurally diverse LPS molecules, and apparently minor changes in synthetic derivatives of LPS can abolish their endotoxic potency[7,14,15]. The lack of a high-resolution structure is in part responsible for incomplete understanding of the basis of receptor specificity and of the activation mechanism. We have therefore determined the crystal structure of the TLR4–MD-2–LPS complex at 3.1 Å resolution.

The receptor multimer is composed of two copies of the TLR4–MD-2–LPS complex arranged in a symmetrical fashion (Fig. 1a).

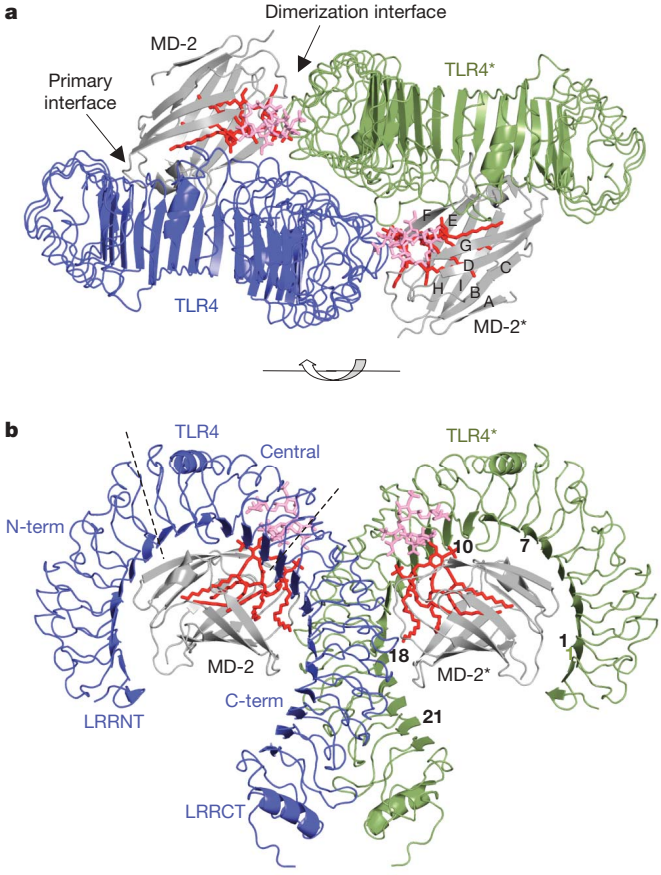

Figure 1 | Overall structure of the TLR4–MD-2–LPS complex. a, Top view of the symmetrical dimer of the TLR4–MD-2–LPS complex. The primary interface between TLR4 and MD-2 is formed before binding LPS, and the dimerization interface is induced by binding LPS. **b**, Side view of the complex. The lipid A component of LPS is coloured red, and the inner core carbohydrates of LPS are coloured pink. The module numbers of the LRRs in TLR4 and the names of the β strands in MD-2 are written in black. TLR4 is divided into N-, central and C-terminal domains[2]. The LRRNT and LRRCT modules cover the amino and carboxy termini of the LRR modules.

[1]Department of Chemistry and [2]Institute for the BioCentury, KAIST, Daejeon, 305-701, Korea. [3]Department of Biology, School of Bioscience & Biotechnology, Chungnam National University, Daejeon, 305-764, Korea.

Comparison with the recently published structure of the monomeric TLR4 and MD-2 complex shows that the overall folding of TLR4 and MD-2 is not disturbed by LPS binding or dimerization[2]. TLR4 adopts the characteristic horseshoe-like shape of the LRR superfamily[16] (Fig. 1b). MD-2 has a β-cup fold structure composed of two anti-parallel β sheets forming a large hydrophobic pocket for ligand binding[2,3]. LPS binds to this pocket and directly mediates dimerization of the two TLR4–MD-2 complexes (Fig. 1a). As shown previously[2], the primary contact interface between TLR4 and MD-2 that is formed before LPS binding involves two chemically distinct regions, the A and B patches provided by the N-terminal and the central domains of TLR4, respectively. The main dimerization interface of MD-2 is located on the opposite side of the primary interface and interacts with LRR modules 15–17 in the C-terminal domain of TLR4 (Fig. 1a and Supplementary Fig. 2a). For dimerization, TLR4 forms hydrophobic and hydrophilic bonds directly with LPS and the surrounding F126 and L87 loops of MD-2. The F126 and L87 loops connect the βG–βH and βE–βF strands of MD-2, respectively (Supplementary Fig. 2b). Throughout this paper, we have marked the second TLR4 and MD-2 molecules and their amino acid residues in the hetero-tetrameric complex with an asterisk to distinguish them from those of the primary TLR4 and MD-2 complex (Fig. 1).

The lipid A component of *Escherichia coli* LPS contains two phosphorylated glucosamines that are connected by a β(1–6) linkage and acylated by six lipid chains (Fig. 2a). The lipid chain attached to the C2′ carbon of the glucosamine is labelled R2′. The R2″ chain is connected to the R2′ lipid by an ester link. The other lipid chains are named in a similar manner. In the crystal structure these lipid chains interact with the hydrophobic pocket in MD-2 (Fig. 2b). The carbon chains of lipids R3, R2′, R3′, R2″ and R3″ are completely buried inside the pocket, but the R2 chain is partially exposed to the MD-2 surface composing the core hydrophobic interface for interaction with TLR4*. The ester and amide groups connecting the lipids to the glucosamine backbone or to the other lipid chains are exposed to the surface of MD-2. They interact with hydrophilic side chains located on the βG strand of the MD-2 pocket and on the surface of TLR4 and TLR4* (Fig. 2c). The two phosphate groups of the lipid A bind to the TLR4–MD-2 complex by interacting with positively charged residues in TLR4, TLR4* and MD-2 and making a hydrogen bond to S118 of MD-2 (Fig. 2d).

Both hydrophobic and hydrophilic interactions contribute to the main dimerization interaction between MD-2/LPS and TLR4* (Fig. 3a). The hydrophobic R2 lipid chain of LPS interacts directly with a small hydrophobic patch on the surface of TLR4* composed of two phenylalanines, F440* and F463* at the core, and L444* at the periphery (Fig. 3b, c). The hydrophobic residues V82, M85, L87, I124 and F126 of MD-2 supplement this core hydrophobic interface. Hydrophilic residues in the F126 loop and the R90 residue of MD-2 form hydrogen bonds and ionic interactions with TLR4* that surround and support the hydrophobic core of the dimerization interface. The 3-hydroxyl group of the R2 chain of LPS contributes to the interaction by forming a hydrogen bond with Q436* of TLR4*. Two phosphate groups, the 1-phosphate and 4′-phosphate of lipid A, also

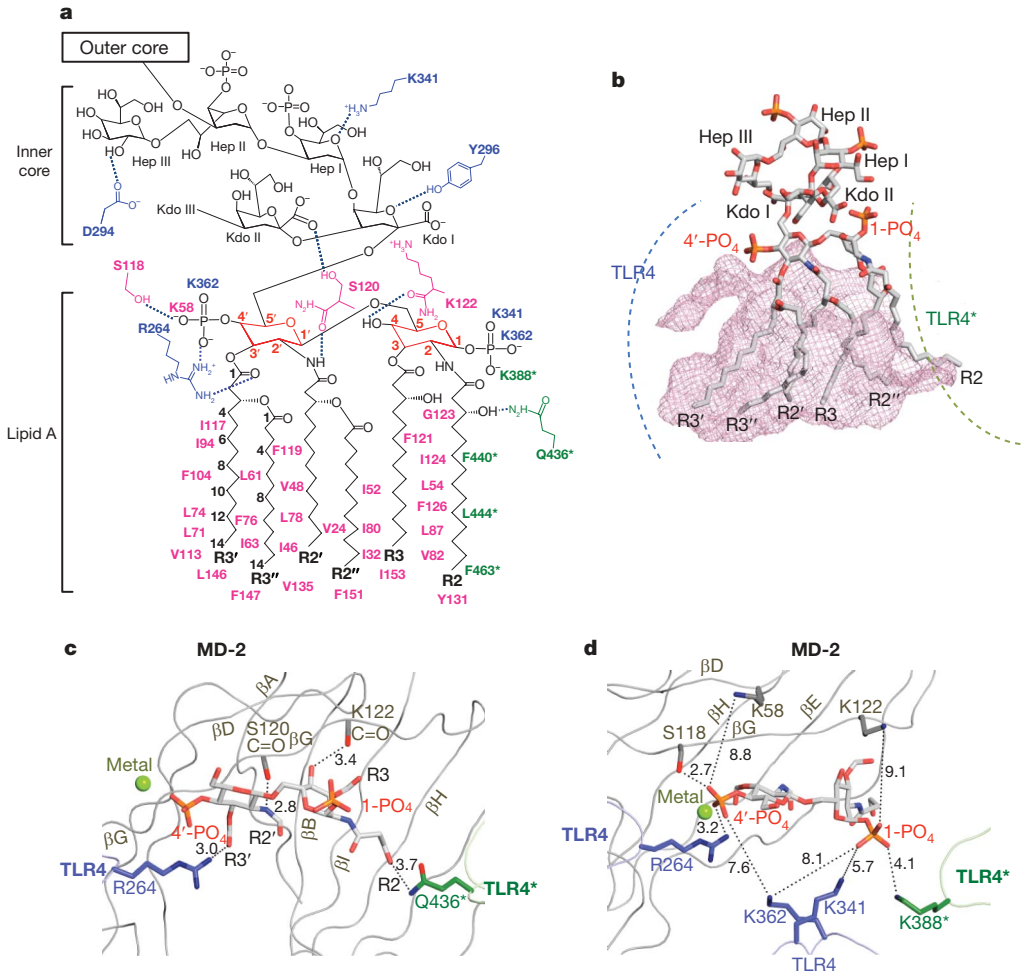

Figure 2 | Binding of LPS to TLR4 and MD-2. a, Chemical structure of the Ra chemotype of *E. coli* LPS[24]. The lipid chains are labelled. The carbons of the glucosamines and lipid chains of lipid A are numbered. Hydrogen bonds are shown by broken blue lines. **b**, The molecular surface of the MD-2 pocket is drawn in mesh. **c**, Hydrogen bonds between lipid A and TLR4–MD-2. **d**, Ionic and hydrogen bond interactions of the two phosphate groups of lipid A. Interaction distances in ångströms are written. Inner core carbohydrates and carbon chains of the lipids are omitted for clarity.

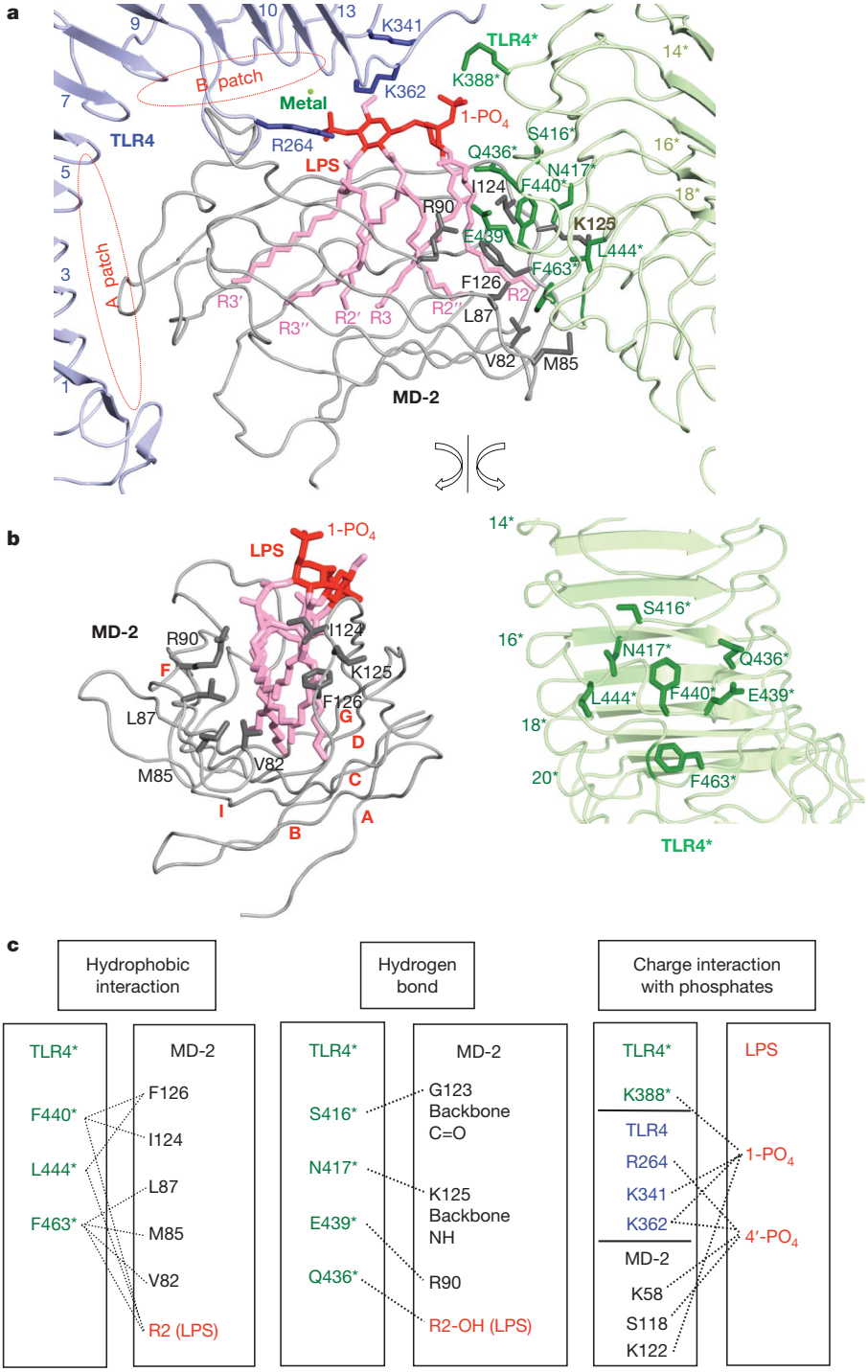

Figure 3 | The main dimerization interface of the TLR4–MD-2–LPS complex. a, Overall shape of the main dimerization interface. The inner core of LPS is omitted for clarity. The primary interface is classified into A and B patches, which are marked in red[2]. K58, S118 and K122 of MD-2 are not shown for clarity. **b**, The dimerization interface has been split and rotated to show the residues involved. The LRR module numbers of TLR4* and the β strands of MD-2 are labelled in green and red, respectively. **c**, Residues involved in the core dimerization interface. The interaction partners are connected by broken lines.

play an important role in dimerization by binding to a positively charged cluster of lysines and an arginine from TLR4, TLR4* and MD-2 (Figs 2d and 3a, c). LPS binding and dimerization induce localized changes in the structure of the F126 loop of MD-2, and in the radius and bending angle of TLR4 (Supplementary Figs 3 and 4).

In addition to the main interaction described above, several secondary interactions contribute to the dimerization. A metal ion found near the 4′-phosphate group appears to connect the MD-2–LPS complex to TLR4 either directly, or indirectly through water molecules (Supplementary Fig. 5). This is probably a magnesium ion supplied by the crystallization buffer and appears to be dispensable for LPS binding and receptor dimerization because neither are blocked by EDTA treatment (data not shown). TLR4 makes an additional contribution to dimerization by directly contacting TLR4* (Supplementary Fig. 6). Carbohydrates in the inner core of LPS form several bonds with the metal ion and the MD-2 and TLR4 molecules (Supplementary Fig. 7). The importance of these carbohydrate interactions in receptor dimerization is, however, not obvious, because

they are not essential for the endotoxic activity of LPS[17]. The Kdo III group of the inner core and all carbohydrates in the outer core are not clearly visible in the electron density map, presumably because their structures and positions are not fixed by interaction with the proteins.

Our structural observations on the TLR4–MD-2–LPS complex are supported by previous biochemical and mutagenesis studies. We have shown that truncation of the C-terminal domain of TLR4 blocks LPS-induced dimerization without compromising LPS binding activity[2]. In the crystal structure, the main dimerization interface of TLR4 is located at LRR modules 15–17 in the C-terminal domain. Therefore, truncation of all these LRR modules should completely block dimerization. Mutation of F126 and the surrounding residues in MD-2 also blocked LPS-induced dimerization[2,18]. Residue F126 is located in the core of the dimerization interface and initiates structural changes in MD-2 (Fig. 3 and Supplementary Fig. 3). In the monomeric structures of MD-2 bound to non-agonistic ligands[2,3], F126 is exposed to the solvent area without interacting with any ligand or protein residues (Supplementary Fig. 3b). On the contrary, in the dimerized structure of the TLR4–MD-2–LPS complex, F126 together with L54, Y131 and I124 of MD-2 forms extensive hydrophobic bonds with lipid chains R2 and R3 of LPS, and F440* of TLR4* (Supplementary Fig. 3c). These interactions are important for positioning the R2 lipid chain correctly and inducing a ~5 Å structural shift in the F126 loop, which moves the critical residues G123, I124 and K125 to positions suitable for the dimerization interaction with TLR4*. Hence, changing F126 to alanine, with its smaller side chain, should disrupt this core interaction and prohibit receptor dimerization. Mutations of other residues in the F126 loop also interfere with LPS binding and signalling[19–22]. Interestingly, a mutant MD-2 with K125 changed to alanine shows normal LPS binding and receptor activation activity[22]. This is because the backbone atoms, but not the side chain atoms, are involved in receptor dimerization.

Recently the structures of MD-2 bound by two antagonists, Eritoran and lipid IVa, have been reported[2,3]. In these crystal structures, the four lipid chains of the antagonists completely fill the available space in the pocket. *E. coli* LPS has two more lipid chains than these antagonists, so it has been proposed that global structural changes may take place in the MD-2 pocket to accommodate the extra lipid chains[2,3,23]. Unexpectedly, the crystal structure of the TLR4–MD-2–LPS complex demonstrates that the size of the MD-2 pocket is unchanged and that additional space for lipid binding is generated by displacing the glucosamine backbone upwards by ~5 Å (Fig. 4). This shift of the glucosamines repositions the phosphate groups such that they can interact with positively charged residues of TLR4 and TLR4*, thus promoting the dimerization and activation of the receptor complex (Figs 2d and 3a, c). Interestingly, the glucosamine backbones of the antagonists are not only translated but also rotated by ~180 degrees, so interchanging the two phosphate groups. More research is needed to establish whether all LPS derivatives with four lipid chains cause a similar rotation of the phospho-glucosamine backbones.

The structure-activity relationship of the lipid A of LPS has been extensively studied and several factors governing the immunological activity of LPS have been identified[15,24]. Of these, the total number of lipid chains is the most important factor. Lipid A with six lipid chains has optimal inflammatory activity, while lipid As with five lipid chains are ~100-fold less active, and those with four lipid chains, such as Eritoran, completely lack agonistic activity[25,26]. The lipid chains of LPS interact hydrophobically with MD-2, and hydrophobic interactions are not sensitive to distance and angle, so if the chemical structure of the chains is changed their positions can be shifted to maximize hydrophobic contact. Therefore, in LPS with five or less lipid chains, all the lipid chains probably move further into the pocket to fill the empty space, and there should be substantial energetic penalties when they move back to the surface of MD-2 for dimerization with TLR4*.

The two phosphate groups in the lipid A region also greatly affect the endotoxic activity of LPS[15,24]. Deletion of either of these phosphate groups reduces endotoxic activity ~100-fold and the resulting monophosphoryl LPS (MPL) is only a weak activator of the human innate immune response. Furthermore, an MPL based on lipid A from *Salmonella minnesota* selectively activates the TLR4–TRAM–TRIF signalling pathway but not the TLR4–Mal–MyD88 pathway[27] (Supplementary Fig. 1). This data suggest that deletion of the 1-phosphate

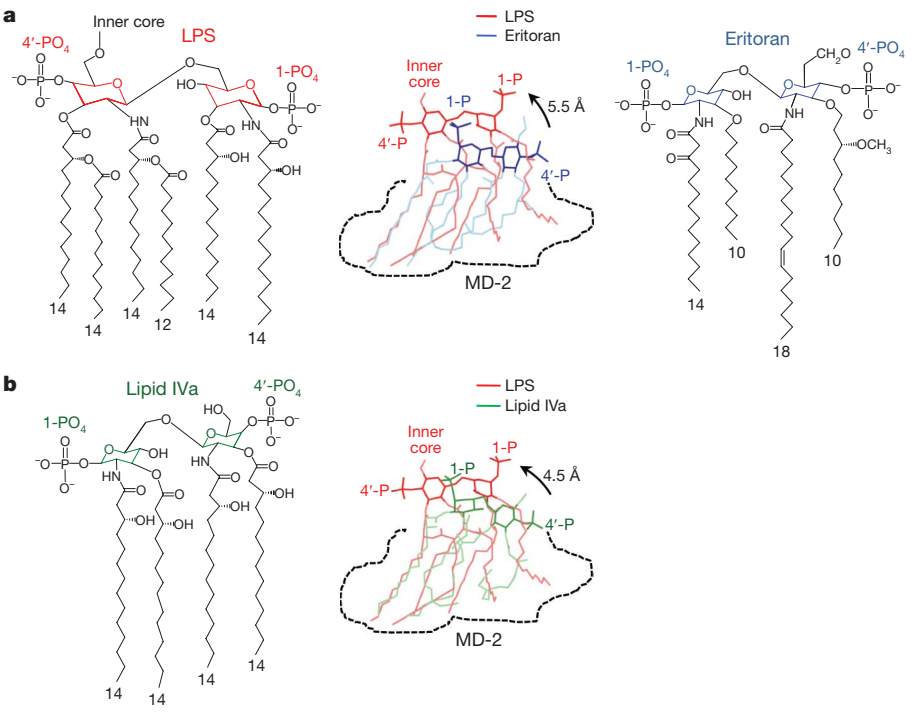

Figure 4 | Structural comparison of LPS with antagonists. **a**, Comparison of LPS and Eritoran after superimposition of MD-2. LPS and Eritoran are coloured red and blue, respectively. **b**, Comparison of LPS and lipid IVa. The structures of LPS and lipid IVa are shown after superimposition of MD-2. The shape of the MD-2 pocket is drawn schematically with broken lines. The glucosamine rings and the phosphate groups are represented in darker colours.

not only weakens the ligand affinity but also induces structural rearrangement of the TLR4–MD-2–adaptor multimer. In the crystal structure, the 1- and 4′- phosphate groups interact with a cluster of positively charged residues from TLR4, TLR4* and MD-2 (Fig. 2d and 3a, c). Therefore their deletion should have a profound impact on the receptor–ligand interaction. In natural LPSs, chemical groups such as 2-aminoethylphosphate or L-Ara4N are occasionally attached to the phosphate groups of lipid A[24]. Previous studies have also shown that removal of the negative charges of the phosphate groups by chemical modification dramatically reduces the potency of LPS, whereas replacing the phosphates with other negatively charged groups such as phosphonoxyethyls has only minor effects[15,24,28]. This is because the 1- and 4′-phosphates in the TLR4–MD-2–LPS complex form mostly medium-range ionic bonds, and there is space for extra atoms in the binding sites (Fig. 2d). Therefore, substituting other negatively charged groups for the phosphates should not always disrupt binding, although it may modulate the activity of the resulting complexes. Domain swapping data involving horse and human TLR4–MD-2s further support the importance of the phosphate groups in LPS recognition[29] (Supplementary Discussion). The structure of the TLR4–MD-2–LPS complex illustrates the remarkable variety of the ligand recognition mechanisms of the TLR family[4,5] (Supplementary Discussion).

In summary, the crystal structure shows that multiple structural components of the TLR4–MD-2 receptor are involved in LPS recognition. Further work is required to establish how these components respond to the structurally diverse LPS variants from different bacterial species. Such studies may lead to the discovery of novel LPS derivatives with desirable therapeutic properties.

METHODS SUMMARY

The ectodomain of human TLR4 (amino acids 27–631) and the human MD-2 gene (amino acids 19–160) fused to the Protein A tag were co-expressed in High Five insect cells (Invitrogen) and purified by IgG Sepharose affinity chromatography (GE Healthcare Life Sciences). A list of the polymerase chain reaction (PCR) primers used for cloning is given in Supplementary Table 1. The Protein A tag was removed with thrombin. The TLR4–MD-2 complex was further purified by SP Sepharose ion exchange and Superdex 200 gel filtration chromatography (GE Healthcare Life Sciences). The purified TLR4–MD-2 complex was concentrated to $\sim$1 mg ml^{-1} and treated with PNGase F. The Ra chemotype of *E. coli* LPS (Sigma, L9641) was incubated with the partially deglycosylated TLR4–MD-2 complex. After the reaction, the TLR4–MD-2–LPS complex was purified by Superdex 200 gel filtration chromatography to remove unbound LPS Ra. Crystals of the complex appeared after three days at 5 °C, when we employed the hanging-drop vapour diffusion method, mixing 1 μl of protein solution and 1 μl of crystallization buffer containing 50 mM $MgCl_2$, 0.1 M Na-HEPES pH 7.5 and 30% PEG MME 550. Initial phases were calculated by molecular replacement using the program PHASER[30] (Supplementary Table 2). The atomic models were refined by iterative modelling and refinement cycles. LPS was built into a strong and continuous electron density found in the MD-2 pocket (Supplementary Fig. 8).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 November 2008; accepted 27 January 2009.
Published online 1 March 2009.

1. Beutler, B. & Rietschel, E. T. Innate immune sensing and its roots: the story of endotoxin. *Nature Rev. Immunol.* **3,** 169–176 (2003).
2. Kim, H. M. *et al.* Crystal structure of the TLR4–MD-2 complex with bound endotoxin antagonist Eritoran. *Cell* **130,** 906–917 (2007).
3. Ohto, U., Fukase, K., Miyake, K. & Satow, Y. Crystal structures of human MD-2 and its complex with antiendotoxic lipid IVa. *Science* **316,** 1632–1634 (2007).
4. Jin, M. S. *et al.* Crystal structure of the TLR1–TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* **130,** 1071–1082 (2007).
5. Liu, L. *et al.* Structural basis of toll-like receptor 3 signaling with double-stranded RNA. *Science* **320,** 379–381 (2008).
6. Raetz, C. R. Biochemistry of endotoxins. *Annu. Rev. Biochem.* **59,** 129–170 (1990).
7. Raetz, C. R. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71,** 635–700 (2002).
8. Medzhitov, R., Preston-Hurlburt, P. & Janeway, C. A. Jr. A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. *Nature* **388,** 394–397 (1997).
9. Shimazu, R. *et al.* MD-2, a molecule that confers lipopolysaccharide responsiveness on Toll-like receptor 4. *J. Exp. Med.* **189,** 1777–1782 (1999).
10. Matsushima, N. *et al.* Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors. *BMC Genom.* **8,** 124 (2007).
11. Bell, J. K. *et al.* The molecular structure of the Toll-like receptor 3 ligand-binding domain. *Proc. Natl Acad. Sci. USA* **102,** 10976–10980 (2005).
12. Choe, J., Kelker, M. S. & Wilson, I. A. Crystal structure of human toll-like receptor 3 (TLR3) ectodomain. *Science* **309,** 581–585 (2005).
13. Miyake, K. Roles for accessory molecules in microbial recognition by Toll-like receptors. *J. Endotoxin Res.* **12,** 195–204 (2006).
14. Erridge, C., Bennett-Guerrero, E. & Poxton, I. R. Structure and function of lipopolysaccharides. *Microbes Infect.* **4,** 837–851 (2002).
15. Rietschel, E. T. *et al.* Bacterial endotoxin: molecular relationships of structure to activity and function. *FASEB J.* **8,** 217–225 (1994).
16. Bella, J., Hindle, K. L., McEwan, P. A. & Lovell, S. C. The leucine-rich repeat structure. *Cell. Mol. Life Sci.* **65,** 2307–2333 (2008).
17. Galanos, C. *et al.* Synthetic and natural *Escherichia coli* free lipid A express identical endotoxic activities. *Eur. J. Biochem.* **148,** 1–5 (1985).
18. Kobayashi, M. *et al.* Regulatory roles for MD-2 and TLR4 in ligand-induced receptor clustering. *J. Immunol.* **176,** 6211–6218 (2006).
19. Kawasaki, K., Nogawa, H. & Nishijima, M. Identification of mouse MD-2 residues important for forming the cell surface TLR4–MD-2 complex recognized by anti-TLR4–MD-2 antibodies, and for conferring LPS and taxol responsiveness on mouse TLR4 by alanine-scanning mutagenesis. *J. Immunol.* **170,** 413–420 (2003).
20. Re, F. & Strominger, J. L. Separate functional domains of human MD-2 mediate Toll-like receptor 4-binding and lipopolysaccharide responsiveness. *J. Immunol.* **171,** 5272–5276 (2003).
21. Visintin, A., Latz, E., Monks, B. G., Espevik, T. & Golenbock, D. T. Lysines 128 and 132 enable lipopolysaccharide binding to MD-2, leading to Toll-like receptor-4 aggregation and signal transduction. *J. Biol. Chem.* **278,** 48313–48320 (2003).
22. Teghanemt, A. *et al.* Novel roles in human MD-2 of phenylalanines 121 and 126 and tyrosine 131 in activation of Toll-like receptor 4 by endotoxin. *J. Biol. Chem.* **283,** 1257–1266 (2008).
23. Jin, M. S. & Lee, J. O. Structures of the Toll-like receptor family and its ligand complexes. *Immunity* **29,** 182–191 (2008).
24. Rietschel, E. T. *et al.* The chemical structure of bacterial endotoxin in relation to bioactivity. *Immunobiology* **187,** 169–190 (1993).
25. Teghanemt, A., Zhang, D., Levis, E. N., Weiss, J. P. & Gioannini, T. L. Molecular basis of reduced potency of underacylated endotoxins. *J. Immunol.* **175,** 4669–4676 (2005).
26. Rossignol, D. P. & Lynn, M. TLR4 antagonists for endotoxemia and beyond. *Curr. Opin. Investig. Drugs* **6,** 496–502 (2005).
27. Mata-Haro, V. *et al.* The vaccine adjuvant monophosphoryl lipid A as a TRIF-biased agonist of TLR4. *Science* **316,** 1628–1632 (2007).
28. Ulmer, A. J. *et al.* Biological activity of synthetic phosphonooxyethyl analogs of lipid A and lipid A partial structures. *Infect. Immun.* **60,** 3309–3314 (1992).
29. Walsh, C. *et al.* Elucidation of the MD-2/TLR4 interface required for signaling by lipid IVa. *J. Immunol.* **181,** 1245–1254 (2008).
30. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D* **61,** 458–464 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the staff of beamline 4A at the Pohang Accelerator Laboratory and beamline ID23-2 at ESRF for help with data collection. We thank J. Gross for critical reading of the manuscript. J.-O.L and co-workers are funded by the Creative Research Initiative (Center for Membrane Receptor Research) from the Ministry of Education, Science and Technology of Korea.

Author Contributions B.S.P., D.H.S. and H.M.K. performed the experiments. B.S.P. and J.-O.L. designed the experiments and analysed the data. B.S.P., H.L. and J.-O.L. wrote the paper. J.-O.L. managed the project and had overall responsibility for data interpretation and writing the manuscript. All authors discussed and commented on the manuscript.

Author Information Atomic coordinates and the structure factor files have been deposited in the Protein Data Bank (http://www.rcsb.org) under accession number 3FXI. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.-O.L. (jieoh.lee@kaist.ac.kr).

METHODS

Protein expression and purification. The ectodomain of human TLR4 (amino acids 27–631) was cloned into the BamHI and NotI sites of pAcGP67 vector (BD Biosciences). The human MD-2 gene (amino acids 19–160) was fused to the Protein A gene derived from pEZZ18 (GE Healthcare Life Sciences) and cloned into the BamHI and NotI sites of pAcGP67 vector (BD Biosciences). A thrombin cleavage site was introduced between the MD-2 and the Protein A genes for removal of the tag. A list of the PCR primers used for cloning is given in Supplementary Table 1. The amino acid sequences of the cloned genes were confirmed by DNA sequencing. TLR4 and MD-2 were co-expressed in High Five insect cells (Invitrogen) and purified by IgG Sepharose affinity chromatography (GE Healthcare Life Sciences). The Protein A tag was removed with thrombin. For the digestion, the concentration of the TLR4–MD-2 complex was set at 1–2 mg ml^{-1}, and 0.1% (w/w) thrombin was used at 4 °C overnight. The TLR4–MD-2 complex was further purified by SP Sepharose ion exchange and Superdex 200 gel filtration chromatography (GE Healthcare Life Sciences). The purified TLR4–MD-2 complex was concentrated to ~1 mg ml^{-1} and treated with PNGase F at 37 °C for 3 h in a buffer containing 20 mM Tris pH 8.0 and 200 mM NaCl. The molar ratio of the protein to PNGase F was 3:1. The partially deglycosylated TLR4–MD-2 complex was purified by SP Sepharose ion exchange and Superdex 200 gel filtration chromatography.

LPS binding and crystallization. To screen LPS chemotypes suitable for crystallization experiments, we tested the binding and dimerizing activities of the Ra, Rc and Re forms of LPS and lipid A. A buffer solution containing Ra or Rc, the LPS chemotypes with longer carbohydrate chains, became transparent after sonication, suggesting that these LPS molecules form small and homogeneous micelles. After incubation with these LPS types in our reaction condition, ~80% of TLR4–MD-2 protein bound to the LPS and was dimerized. On the other hand, a buffer solution containing lipid A or LPS Re with shorter carbohydrate chains, remained turbid even after sonication, suggesting that these LPS molecules form larger and highly heterogeneous aggregates owing to their greater hydrophobicity versus hydrophilicity ratio. As a result, binding of these LPS types by TLR4–MD-2 is significantly less efficient, probably because of the lower efficiency with which they are extracted from the large aggregates. We did not examine this problem systematically, but simple addition of mouse CD14 and LPS-binding protein did not make any significant difference to the binding efficiency of LPS under our reaction conditions. For crystallization, the Ra chemotype of *E. coli* LPS (Sigma, L9641) was sonicated for 10 min and incubated with the deglycosylated TLR4–MD-2 complexes at 37 °C for 3 h. The reaction buffer contained 20 mM Tris-HCl, pH 8.0 and 200 mM NaCl. The concentration of TLR4–MD-2 was ~1 mg ml^{-1} and the molar ratio of LPS to the protein was 10:1. In the Ra chemotype of LPS, the O antigen carbohydrates of LPS are absent but the lipid A and inner and outer cores remain intact[6]. After the reaction, the TLR4–MD-2–LPS complex was purified by Superdex 200 gel filtration chromatography to remove unbound LPS Ra. Binding of LPS Ra to the protein complex was monitored by 8% native polyacrylamide gel electrophoresis (PAGE) and gel filtration chromatography. The TLR4–MD-2–LPS complex was concentrated to 2 mg ml^{-1} in a buffer containing 20 mM Tris pH 8.0 and 200 mM NaCl. Crystals of the complex appeared after three days at 5 °C, when we employed the hanging-drop vapour diffusion method mixing 1 μl of protein solution and 1 μl of crystallization buffer containing 50 mM $MgCl_2$, 0.1 M Na-HEPES pH 7.5 and 30% PEG MME 550.

Data collection. For data collection, crystals were flash-frozen at −170 °C in the crystallization buffer with a 3% higher concentration of PEG MME 550. Diffraction data were collected at the 4A beamline of the Pohang Accelerator Laboratory and the ID23-2 beamline of the European Synchrotron Radiation Facility. The XDS package was used to index, integrate and scale the diffraction data (Supplementary Table 2)[31].

Model building and refinement. Initial phases were calculated by molecular replacement using the program PHASER[30]. The published structures of TV3-MD-2-Eritoran and the VT3 hybrid were used together as search probes[2]. Model building was aided by the structures of mouse TLR4–MD-2 and the TV8 hybrid. The atomic models were refined by iterative modelling and refinement cycles using programs O and CNS[32,33]. The atomic coordinates of the model were restrained by two-fold non-crystallographic symmetry throughout the refinement. The atomic model of LPS was built into a strong and continuous electron density found in the MD-2 pocket (Supplementary Fig. 8). The final structure was refined to R and R_{free} factors of 24% and 28%, respectively (Supplementary Table 2). 87% of protein residues are present in the favoured regions and no non-glycine residues are in the disallowed regions of the Ramachandran plots[34].

31. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26,** 795–800 (1993).
32. Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
33. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).
34. Lovell, S. C. *et al.* Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. *Proteins* **50,** 437–450 (2003).