

## REVIEW

# Recognition of lipopolysaccharide pattern by TLR4 complexes

Beom Seok Park<sup>1</sup> and Jie-Oh Lee<sup>2</sup>

Lipopolysaccharide (LPS) is a major component of the outer membrane of Gram-negative bacteria. Minute amounts of LPS released from infecting pathogens can initiate potent innate immune responses that prime the immune system against further infection. However, when the LPS response is not properly controlled it can lead to fatal septic shock syndrome. The common structural pattern of LPS in diverse bacterial species is recognized by a cascade of LPS receptors and accessory proteins, LPS binding protein (LBP), CD14 and the Toll-like receptor4 (TLR4)–MD-2 complex. The structures of these proteins account for how our immune system differentiates LPS molecules from structurally similar host molecules. They also provide insights useful for discovery of anti-sepsis drugs. In this review, we summarize these structures and describe the structural basis of LPS recognition by LPS receptors and accessory proteins.

*Experimental & Molecular Medicine* (2013) 45, e66; doi:10.1038/emm.2013.97; published online 6 December 2013

**Keywords:** lipopolysaccharide (LPS); Toll-like receptor 4 (TLR4); MD-2; CD14; LBP

## INTRODUCTION

In the initial phase of infection, the innate immune system generates a rapid inflammatory response that blocks the growth and dissemination of the infectious agent. This response is followed, in vertebrates, by an adaptive immune response that leads to its elimination. The adaptive immune receptors are highly specific for the particular pathogen that induces them. Unlike the adaptive response, the innate immune response, mediated mainly by macrophages and dendritic cells, involves a limited number of germline-encoded pattern recognition receptors that recognize common structural patterns among diverse molecules produced by microbes. Among these receptors, the representative key players are proteins of the Toll-like receptor (TLR) family.<sup>1,2</sup>

Lipopolysaccharide (LPS) is a well-characterized pathogen-associated molecular pattern found in the outer leaflet of the outer membrane of most of the Gram-negative bacteria. It can initiate a strong immune response and serves as an early warning signal of bacterial infection. LPS is initially extracted from bacterial membranes and vesicles released from them by LPS binding protein (LBP) in serum. LBP then transfers LPS to CD14, which can be found either in soluble form or linked to

the cell surface by a glycosylphosphatidylinositol anchor. CD14 splits LPS aggregates into monomeric molecules and presents them to the TLR4–MD-2 complex. Aggregation of the TLR4–MD-2 complex after binding LPS leads to activation of multiple signaling components, including NF- $\kappa$ B and IRF3, and the subsequent production of pro-inflammatory cytokines.<sup>3–7</sup>

The progress of genome sequencing projects has led to the identification of 13 TLRs in mammalian genomes, 10 in humans and 12 in mice. The TLRs are type I integral membrane receptors composed of an extracellular domain, a single transmembrane helix and an intracellular signaling domain. The extracellular domain has a horseshoe-shaped solenoid structure and is responsible for the recognition of common structural patterns in various microbial molecules. For example, lipoproteins or lipopeptides are recognized by TLR2 complexed with TLR1 or TLR6, viral double-stranded RNAs by TLR3, LPSs by the TLR4–MD-2 complex, bacterial flagellins by TLR5, viral and bacterial single-stranded RNAs by TLR7 or TLR8 and CpG-rich under-methylated DNAs by TLR9. Binding of these ligands results in the formation of large receptor–adaptor complexes that lead to production and secretion of cytokines.<sup>7–9</sup>

<sup>1</sup>Department of Biomedical Laboratory Science, College of Health Science, Eulji University, Seongnam, Korea and <sup>2</sup>Department of Chemistry, Korea Advanced Institute of Science and Technology, Daejeon, Korea

Correspondence: Professor BS Park, Department of Biomedical Laboratory Science, College of Health Science, Eulji University, Seongnam 461 713, Korea. E-mail: bspark74@eulji.ac.kr

or Professor J-O Lee, Department of Chemistry, Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon 305 701, Korea.

E-mail: jieoh@kaist.ac.kr

Received 11 July 2013; accepted 22 July 2013

In this review, we discuss the reported structures of the TLR4–MD-2 complex, LBP and CD14, and consider their implications for our understanding of TLR4–MD-2 activation in response to recognition of LPS.

### STRUCTURE OF LPS

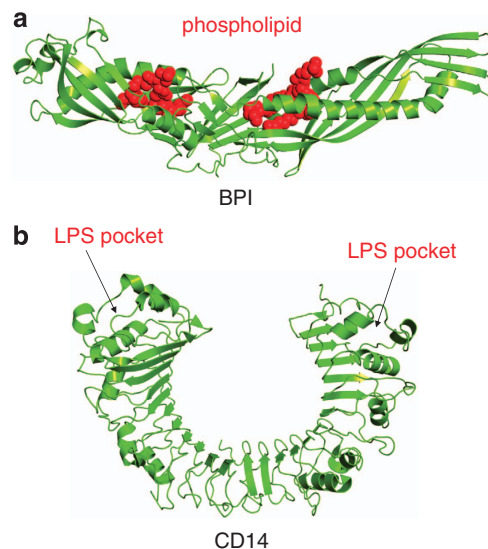
LPS is a macromolecular glycolipid composed of a hydrophobic lipid A region attached to a long and branched carbohydrate chain. The lipid A portion, which is responsible for most of the immunologic activity of LPS, is composed of a phosphorylated diglucosamine backbone with four to seven acyl chains attached to it. Four of the acyl groups are directly linked to the 2, 3, 2' and 3' positions of the glucosamine backbone, and the remaining two are attached to the hydroxyl groups of the lipid chains. Lipid A from different bacterial species shows substantial structural diversity.<sup>3,10</sup> The number and length of the acyl chains can vary, and the phosphate groups can be modified by other chemical groups.

The carbohydrate region of LPS can be divided into two areas, the core and the O-specific chain. The core region is relatively conserved among bacterial species and contains unusual carbohydrate residues such as heptose and KDO that are not usually found in host cells. The O-specific region is composed of many copies of carbohydrate repeating units. Bacterial cells produce a highly heterogeneous set of repeating units with different structures. Removal of the entire carbohydrate chain by acid hydrolysis has only a minimal effect on the inflammatory activity of LPS, demonstrating that the core and O-specific carbohydrates have only a minor role in recognition by host immune receptors.<sup>3,11</sup>

### STRUCTURES OF LBP AND CD14

LPS molecules, due to their amphipathic nature, form large aggregates in aqueous environments above a critical micellar concentration. The accessory proteins, LBP and CD14, enhance the detection of LPS by the TLR4–MD-2 complex by extracting and monomerizing LPS before its presentation to TLR4–MD-2. LBP is an acutely induced plasma protein that binds avidly to LPS aggregates and delivers them to CD14.<sup>12,13</sup> It belongs to the lipid transfer or LBP family. Other members of the family are bacterial and permeability-increasing protein (BPI), cholesterol ester transfer protein, phospholipid transfer protein and a few poorly characterized proteins.<sup>14–19</sup> Of these proteins, the structures of BPI and cholesterol ester transfer protein have been determined;<sup>16,17</sup> the structure of LBP has not been reported but it is expected to share the general features of BPI because the two proteins have 48% sequence homology. BPI is a boomerang-shaped molecule composed of a central  $\beta$  sheet with barrel-shaped domains at its termini (Figure 1a).<sup>16,20</sup> However, the two proteins differ functionally: LBP transfers LPS to TLR4–MD-2, whereas BPI does not. Structural studies are required to account for the functional difference between the two proteins.

CD14 is expressed on the surface of myelomonocytic cells in the form of a glycosylphosphatidylinositol-linked glycoprotein, and as a soluble protein in the serum.<sup>21</sup> Its crucial role in LPS



**Figure 1** The structures of accessory proteins involved in LPS recognition. (a) The crystal structure of BPI, with two phospholipid binding sites. LBP is expected to have a similar structure. (b) CD14 forms homodimers. The monomeric subunit of CD14 contains 11 LRR modules and a single LRRNT module.

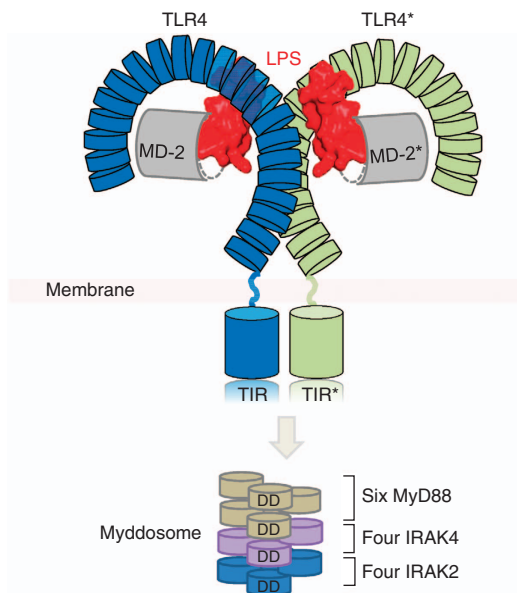
signaling has been confirmed using knock-out mice: CD14-deficient mice are highly resistant to septic shock initiated by injection of LPS or live bacteria.<sup>22</sup> CD14 binds to LPS delivered by LBP and transfers the bound LPS to the TLR4–MD-2 complex. Since the presumed LPS binding pocket of CD14 is too small for large LPS aggregates, it is likely to bind the monomeric form of LPS.

CD14 belongs to the leucine-rich repeat (LRR) family, and has the characteristic curved solenoid structure (Figure 1b).<sup>23,24</sup> LRR family proteins are composed of multiple copies of LRR modules. The individual LRR modules consist of 20–30 amino-acid residues with highly conserved ‘LxxLxLxxN’ motifs. The central LxL part of the module forms the core of a  $\beta$  strand; the two leucines point toward the interior of the protein, forming the hydrophobic core, whereas the variable x residues within the motif are exposed to solvent and some are involved in interactions with ligands. Asparagines in the motif make stable hydrogen bonds with the backbone carbonyls of neighboring  $\beta$  strands throughout the entire protein, forming an extended hydrogen bonding network called an ‘asparagine ladder’. As a result, the  $\beta$  strands are more closely packed, and assemble into a large  $\beta$  sheet making up the entire concave surface of the horseshoe. Variable amino acids outside the conserved  $\beta$  strands of each LRR module are surface exposed and some of them have important roles in ligand interactions. To prevent exposure of the hydrophobic core of the LRR modules, there are two special modules named LRRNT and LRRCT at the N- and C-termini of the proteins. These modules do not show the sequence conservation pattern of the LRR modules and often contain an anti-parallel  $\beta$  hairpin stabilized by disulfide bridges.<sup>25</sup>

CD14 exists as a homodimer.<sup>23</sup> Unlike other LRR family members, it does not contain an LRRCT module protecting its hydrophobic core. Instead, the C-terminus of the LRR modules of one CD14 molecule interacts with the C-terminus of another, forming a dimer. The LPS interaction pocket of CD14 is located at the boundary of the LRRNT and the first LRR module (Figure 1b). In addition to LPS, CD14 can bind other microbial products, such as peptidoglycan, lipoteichoic acid, lipoarabinomannan and lipoproteins. Therefore, it has broad ligand specificity and functions as a pattern recognition receptor by recognizing structural motifs in diverse microbial products.<sup>23,26</sup>

### STRUCTURE OF THE TLR4–MD-2 COMPLEX

The extracellular domain of TLR4 also belongs to the LRR family (Figure 2) and is responsible for ligand binding and receptor dimerization.<sup>27,28</sup> The structure of TLR4, as well as those of TLR1, TLR2 and TLR6, deviates substantially from the consensus LRR structures.<sup>27,29</sup> LRR family proteins can be classified into seven subfamilies. Each subfamily is defined by its conformation, which is a function of the twist, tilt, rotation angles and radii of the central  $\beta$  sheet of the LRR motifs. Most of the LRR proteins have uniform radii and  $\beta$  sheet angles throughout the protein: for example, typical subfamilies have 24 amino acids per module and 3<sub>10</sub> helices and loops in the convex area. However, TLRs 1, 2, 4 and 6 are divided by a structural transition into three subdomains: N-terminal, central and C-terminal. The structural discontinuities seem to be caused by irregular LRR sequences in the central domain.



**Figure 2** Overview of LPS recognition by TLR4–MD-2. LPS binding induces dimerization of the TLR4–MD-2 complex, which is proposed to enable dimerization of the intracellular TIR domains and recruitment of adaptor molecules such as MyD88. Aggregation of the death domains (DD) of MyD88 brings four IRAK4 and four IRAK2 molecules together forming a large tower-like structure called the ‘Myddosome’.

Interestingly, the domain boundaries in TLRs 1, 2 and 6 fulfill important roles in ligand binding.<sup>27–29</sup> In the case of TLR4, its N-terminal and central domains clearly provide charge complementarity for binding of its surface to its co-receptor MD-2, forming a stable 1:1 heterodimer.

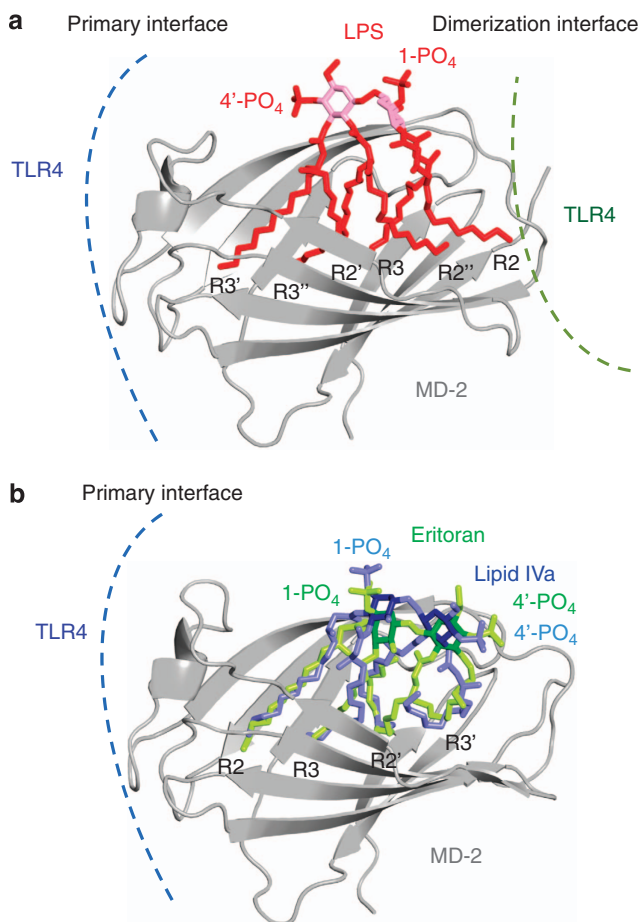
MD-2 is smaller than TLR4 and is the main LPS binding module of the TLR4–MD-2 receptor complex. MD-2 has a  $\beta$ -cup fold structure, composed of two antiparallel  $\beta$  sheets.<sup>30,31</sup> The two sheets are separated from each other on one side of the sandwich-like structure so that the hydrophobic interior is exposed for interaction with ligands. This large internal pocket is ideally shaped for binding flat hydrophobic ligands like LPS. The  $\beta$ -cup fold of MD-2 is shared by a small number of proteins mostly involved in lipid storage or transport.<sup>32</sup> This fold is very similar to the more famous immunoglobulin fold. However, it does not have the conserved disulfide bridge connecting the two  $\beta$  sheets of the immunoglobulin fold. Separation of the two  $\beta$  sheets and exposure of the internal pocket in the  $\beta$ -cup fold is possible because the connecting disulfide bridge is missing.<sup>30,31</sup>

MD-2 binds to TLR4 primarily via hydrophilic interactions such as hydrogen bonds and charge interactions, and a few hydrophobic residues are found in the binding interface.<sup>28,30</sup> The long and narrow binding surface of TLR4 can be divided into two areas, the A and B patches. The negatively charged A patch interacts with the positively charged surface of MD-2, and the positively charged B patch with the negatively charged surface of MD-2. Mutation of residues involved in this interaction blocks TLR4 and MD-2 binding, and LPS signaling.

### LPS BINDING TO THE TLR4–MD-2 COMPLEX

The crystal structure of the TLR4–MD-2 complex bound to a truncated form of *E. coli* LPS has been determined.<sup>28</sup> LPS binding induces the formation of the ‘m’ shaped 2:2:2 TLR4, MD-2 and LPS complex, with the two C-termini of the extracellular domains of TLR4 converging in the center and the N-termini stretching outwards (Figure 2). The acyl chains of the lipid A are inserted deep into the MD-2 pocket and the two phosphate groups of lipid A form charge and hydrogen bond interactions with charged residues in MD-2 and TLR4. Before it binds to LPS, MD-2 forms a stable heterodimer with the extracellular domain of TLR4. LPS binding causes dimerization of the TLR4–MD-2 complex because the LPS creates an additional binding interface between TLR4 and MD-2. The former interaction interface is named the primary interface, which is formed between one edge of MD-2 and part of its concave surface by the extensive network of charge-enhanced hydrogen bonds of TLR4, while the latter, dimerization interface of MD-2 is located opposite the primary interface and interacts with a convex surface provided by a small hydrophobic patch in the C-terminal domain of the other TLR4 in the TLR4–MD-2–LPS complex (Figure 3a).

The structure–activity relationships of LPS have been studied for decades using natural and chemically modified LPS.<sup>28</sup> The crystal structure of TLR4–MD-2–LPS provides an



**Figure 3** Binding of LPS and antagonistic ligands to the TLR4-MD-2 complex. **(a)** Structure of the primary and dimerization interfaces of the TLR4-MD-2-LPS complex. The lipid chains of LPS are labeled. MD-2 is colored grey. The lipid chains and phosphate groups of LPS are shown in red. The glucosamine backbone is pink. **(b)** Structures of Eritoran and lipid IVa bound to MD-2.

explanation for why LPS with six lipid chains is optimal for activation of TLR4 signaling. In the crystal structure, five of the six lipid chains of *E. coli* LPS are completely buried inside the pocket, but the remaining chain is partially exposed to the MD-2 surface and forms the hydrophobic interaction interface together with hydrophobic surface residues of MD-2.

Two phosphate groups attached to the glucosamine of LPS further support formation of the stable TLR4-MD-2 complex by making charge and hydrogen bond interactions simultaneously with the two TLRs in the complex. Removal of one of the phosphate groups greatly reduces the inflammatory activity of LPS. This modified form of LPS is named monophosphorylated lipid A; it has recently been approved as a vaccine adjuvant because it retains strong immune stimulatory activity but has lost most of the inflammatory toxicity of LPS.<sup>33-35</sup> Since the core carbohydrate part of *E. coli* LPS makes only a limited number of weak ionic and hydrogen bonds with TLR4 and MD-2, the carbohydrate chain should have only a minor role in the immunological activity of LPS.<sup>28</sup>

Two single-nucleotide polymorphisms of human TLR4, D299G and T399I have been identified and are thought to be associated with LPS hyporesponsiveness. Recently, Ohto *et al.*<sup>36</sup> determined the structure of polymorphic TLR4 in complex with MD-2 and LPS. They found that the alterations in their structure are quite localized, and the relative orientation and distance between the TLR4 subunits is not changed. They therefore proposed that the conformational changes induced by the two polymorphisms do not directly influence conformation of the TLR4-MD-2 complex, but rather reduce the response to LPS by affecting binding affinity, the folding efficiency, cell surface expression or stability of TLR4.

### BINDING OF ANTAGONIST TO THE TLR4-MD-2 COMPLEX

Some lipid A derivatives with four lipid chains have antagonistic activity for the TLR4-MD-2 complex.<sup>37,38</sup> The structures of two of these antagonists, Eritoran (or E5564) and lipid IVa, in complexes with TLR4-MD-2 and MD-2, respectively, have been determined.<sup>30,31</sup> Eritoran is a synthetic molecule derived from the lipid A component of the non-pathogenic LPS of *Rhodobacter sphaeroides*. It is a high affinity antagonist and has useful therapeutic activity against severe sepsis.<sup>39,40</sup> Lipid IVa, a precursor form of LPS, acts as an antagonist of human TLR4-MD-2. When bound to MD-2, Eritoran and lipid IVa adopt remarkably similar structures (Figure 3b). Their glucosamine backbones are nearly superimposable, despite structural differences between their lipid chains.<sup>28</sup> The lipid chains of both are packed inside the hydrophobic pocket in MD-2. Neither Eritoran nor lipid IVa induces dimerization of human TLR4-MD-2 because all four lipid chains are completely submerged inside the pocket and cannot provide a hydrophobic dimerization surface that can be used for interaction with TLR4.

Although lipid IVa is an antagonist to human TLR4-MD-2, it has weak agonistic activity for the mouse TLR4-MD-2 complex.<sup>41</sup> To explain this peculiar species-specific response, the structure of the lipid IVa bound to mouse TLR4-MD-2 complex has been determined.<sup>42</sup> In this structure, the positions of the glucosamine backbone and the phosphate groups are nearly identical with those of LPS bound to TLR4-MD-2, and therefore it can induce dimerization of TLR4-MD-2. However, because lipid IVa has only four lipid chains, the hydrophobic pocket of MD-2 has a large unoccupied space in this structure. This arrangement of acyl chains is energetically unfavorable, which can explain why lipid IVa has only weak agonistic activity.

### ACTIVATION OF THE TLR4-MD-2 COMPLEX

The TLR components of all TLR-ligand complexes have similar dimeric conformations, with the two C-termini of the extracellular domains of the TLRs converging in the middle.<sup>27,43-46</sup> This structural arrangement supports the hypothesis that dimerization of the extracellular domains leads to juxtaposition of the intracellular domains, signaling

adaptors being then recruited to these dimerized intracellular domains. The intracellular domains of TLRs have substantial sequence and structural homology to the interleukin-1 receptor (IL-1R) family, and are named Toll/IL-1R homology (TIR) domains. TIR domains are not only found in the receptors but also in the signaling adaptors involved in TLR and IL-1R signaling.<sup>9</sup> After binding LPS, the TIR domain of TLR4 interacts with the TIR domain of myeloid differentiation factor 88 (MyD88), in conjunction with another TIR-containing adaptor protein, referred to as MyD88 adaptor-like (Mal). Mutations of the TIR domains abrogate this interaction, demonstrating that the latter are responsible for this interaction.<sup>47–49</sup>

MyD88 has been reported to be involved in the signaling pathways initiated by all TLRs with the exception of TLR3.<sup>50</sup> Of the MyD88-dependent pathways, only those involving TLR2 and TLR4 require Mal for efficient signaling. The MyD88 aggregation signal is transmitted to IL-1 receptor kinase (IRAK) through an interaction between the death domain of MyD88 and IRAK. Phosphorylation of the signaling kinases eventually activates the transcription factors, NF- $\kappa$ B and activator protein 1 via a signaling cascade. Instead of the MyD88-dependent pathway, the TRIF (TIR-domain containing adaptor inducing interferon- $\beta$ ) pathway of TLR4 activates interferon response factors to produce and secrete type-I interferons. TIR domains are found not only in mammalian proteins but also in plant and bacterial proteins. Sequence conservation in the TIR family is relatively low, in the 20–30% range, and domain lengths vary between 135 and 160 residues because of large deletions and insertions in the loop regions.<sup>51</sup>

Crystal structures of TIR domains have been reported for TLRs 1, 2 and 10, IL-1RAPL, MAL and MyD88, and for TIR domains of several plant and bacterial proteins (Figure 4a).<sup>52–57</sup> In addition, an NMR structure has been reported for the MyD88 TIR domain.<sup>56</sup> In all these structures, alternating  $\beta$  strands and  $\alpha$  helices are arranged in the form of five central parallel  $\beta$  strands surrounded by five  $\alpha$  helices. Homomultimeric TIR interfaces have been observed in crystals of TLR2 and TLR10, but the functional relevance of these interactions remains uncertain because dimerization of these TIR domains has not been observed in solution.

On the basis of the crystal structures of these monomeric TIR domains, mutational studies and computational modeling have been carried out to predict the homo- and heteromeric structures of TIR complexes.<sup>51,58,59</sup> The results suggest that the BB and the DD loops of the TLR TIR domains have key roles in TIR domain aggregation and signal activation. The BB loop that joins strands  $\beta$ B and helix  $\alpha$ B extends away from the rest of the TIR domain, forming a protrusion from the surface (Figure 4b). The DD loop connects the  $\beta$ D strand and the  $\alpha$ D. The structures of the BB and DD loop regions vary considerably in different TIR structures.

MyD88 contains not only a TIR domain but also a death domain. Its death domain binds to downstream IRAKs after the activation of TLRs. Recently, Lin *et al.*<sup>60</sup> determined the

structure of the death domain complex of MyD88, IRAK4 and IRAK2. In this tower-like structure, the death domains form a large signaling complex, called a Myddosome, which consists of a four-layered left-handed helical arrangement of six MyD88, four IRAK4 and four IRAK2 death domains. On the basis of this structure, they proposed that formation of the Myddosome brings the kinase domains of the IRAKs into proximity so that they can be phosphorylated and activated.

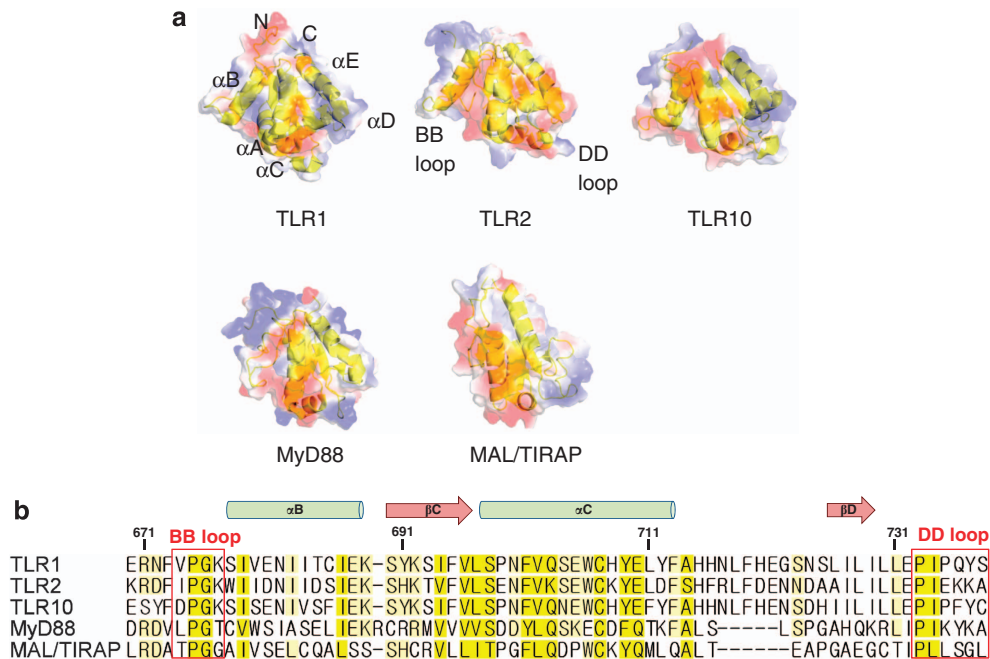
## STRUCTURES OF OTHER TLRs

To date, the structures of six of the ten human TLRs in complex with their physiological or synthetic ligands have been reported (Figure 5). TLR2 is unique among human TLRs because it can form heterodimers with other TLRs, TLR1 and TLR6. The principal ligands of the TLR1–TLR2 complex are triacyl lipopeptides, the interaction with diacyl lipopeptides being substantially weaker.<sup>61,62</sup> By contrast, the TLR2–TLR6 complex is able to bind to diacyl lipopeptides. These lipoproteins and lipopeptides are functionally and structurally diverse bacterial proteins anchored to the membrane by two or three covalently attached lipid chains.<sup>63</sup>

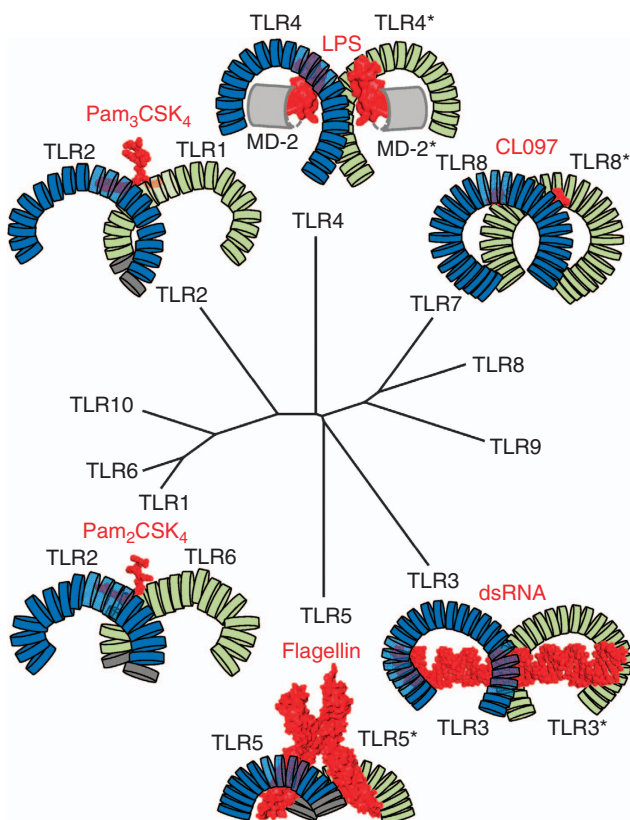
The structures of TLR1–TLR2 and TLR2–TLR6 with the synthetic lipopeptide ligands, Pam<sub>3</sub>CSK<sub>4</sub> or Pam<sub>2</sub>CSK<sub>4</sub>, have been determined.<sup>27,29</sup> The overall shapes of the complexes are similar to that of TLR4 dimers: the two N-terminal regions stretch outward and the C-termini of the extracellular domains converge in the middle. The Pam<sub>3</sub>CSK<sub>4</sub> binding pockets of TLR1 and TLR2 are formed at the boundaries between the central and C-terminal domains. The three lipid chains of Pam<sub>3</sub>CSK<sub>4</sub> mediate TLR2–TLR1 heterodimer formation: two are inserted into the TLR2 pocket, and the remaining, amide-bound, lipid chain is inserted into the narrower TLR1 channel. Heterodimers of TLR1 and TLR2 are further stabilized by protein–protein interactions near the opening of the ligand binding pocket.

Whereas the TLR1–TLR2 complex recognizes triacylated lipopeptides, the TLR2–TLR6 complex recognizes the diacylated ligand, Pam<sub>2</sub>CSK<sub>4</sub>, which lacks the lipid chain that binds in the hydrophobic pocket of TLR1. Nevertheless, the structure of the TLR2–TLR6–Pam<sub>2</sub>CSK<sub>4</sub> complex indicates that TLR6 and TLR1 form m-shaped structures very similar to the TLR1–TLR2–Pam<sub>3</sub>CSK<sub>4</sub> complex. However, there are important structural differences between TLR1 and TLR6 in their ligand binding and dimerization regions. In TLR6, the side chains of two phenylalanine residues block the lipid binding channel, creating a channel that is less than half the length of that of TLR1.<sup>29</sup> This structural feature explains the selectivity of the TLR2–TLR6 complex for diacylated over triacylated lipopeptides. Despite the absence of the amide-bound lipid chain in the diacylated lipopeptide, TLR2 can still form a stable heterodimer with TLR6 because the lack of the lipid–protein interaction is compensated for by expanding the hydrophobic protein–protein interface.

TLR3 has been shown to recognize double-strand RNA (dsRNA), which is produced by various viruses at some point in their life cycle and is a potent danger signal of virus infection.<sup>64</sup>



**Figure 4** TIR domain structures. (a) The figure was drawn using coordinate files with PDB codes as follows: 1FYV (TLR1), 1FYW (TLR2), 2J67 (TLR10), 2Z5V (MyD88) and 2Y92 (Mal). The backbones of the TIR domains are blue and yellow; positive charges are blue and negative charges are red. (b) Sequence alignment of TIR domains. The BB and DD loops important in TLR signaling are boxed in red.



**Figure 5** Structures and a phylogenetic tree of TLRs. Crystal structures of TLR4–MD-2–LPS, TLR2–TLR1–Pam<sub>3</sub>CSK<sub>4</sub>, TLR2–TLR6–Pam<sub>2</sub>CSK<sub>4</sub>, TLR5–flagellin, TLR3–dsRNA, TLR8–CL097 are shown. The ligands are colored red, and TLRs are blue and green.

TLR3 has the largely uniform and flat LRR structure typical of the subfamily, but lacks the subdomain structure seen in TLR1, TLR2, TLR4 and TLR6.<sup>44,65</sup> It is heavily glycosylated, with 15 predicted N-glycosylation sites, of which 11 are visible in the crystal structure. The dsRNA interacts with both N-terminal and C-terminal sites on the lateral side of the convex surface of TLR3.<sup>44</sup> The positively charged residues near the N- and C-termini of the extracellular domain of TLR3 make the major contribution to the interaction with the sugar-phosphate backbones of the dsRNA. Mutational analyses have established that simultaneous binding to the N- and C-terminal sites is required for stable binding of the dsRNA. TLR3 interacts with the ribose-phosphate backbone of dsRNA, thus accounting for the lack of RNA sequence specificity. The TLR3 structure can also explain the strong pH dependence of RNA binding: several histidines make crucial bonds with the phosphate backbones of the RNA, and their protonation appears to be essential to stabilize the interaction.

TLR5 is one of the few TLRs that recognize a protein pathogen-associated molecular pattern, namely bacterial flagellin, which is the monomeric subunit of the flagellar fiber.<sup>66</sup> The crystal structure of zebrafish TLR5 has been determined in a complex with a truncated fragment of *Salmonella* FliC.<sup>45</sup> The overall shape and curvature of the TLR5 LRR is better categorized as TLR3-like than TLR4- or TLR2-like, which is consistent with the evolutionary proximity of TLR3 and TLR5 (Figure 5). The structure of the TLR5–FliC complex shows that the flagellin D1 domain has the dominant role in the binding and dimerization of TLR5. The D1 domain

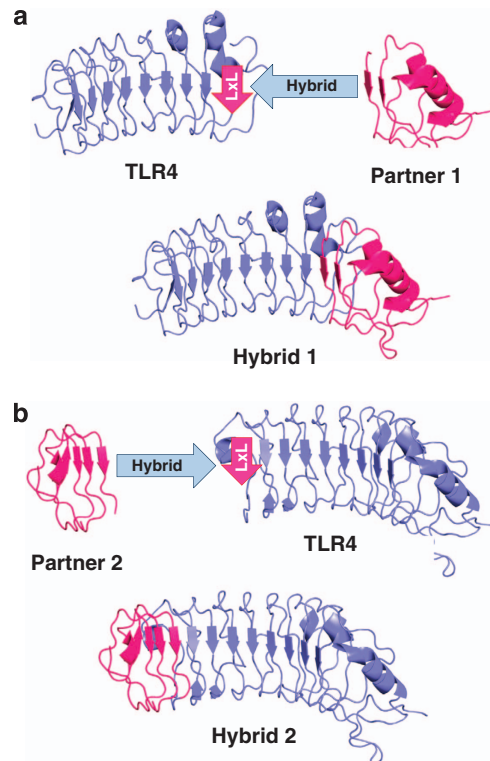
consists of four elongated segments that assemble vertically into a long, largely helical, rod. The four segments consist of two N-terminal  $\alpha$  helices, one C-terminal  $\alpha$  helix and a segment of  $\beta$  hairpin. TLR5 interacts primarily with the three helices of the D1 domain via its side, which is formed by the segment between LRRNT and LRR10. At the same time, the D1 domain interacts with the second TLR5 molecule in the dimer and therefore bridges the two TLRs. The two TLR5 molecules in the dimer also closely resemble the m-like structures of the other TLR dimers.

TLR7 and TLR8 recognize the single-strand RNA of various viruses, as well as small-interfering RNAs. They also mediate the recognition of self RNA released from dead or dying cells and are responsible for certain autoimmune diseases.<sup>67</sup> In phylogenetic analyses, TLR7, 8 and 9 form a subfamily with an amino-acid sequence longer than other TLRs (Figure 5). In 2013, the crystal structures of unliganded and ligand-induced activated human TLR8 dimers were reported.<sup>46</sup> Small molecule agonists, R848, CL097 and CL075, were used in that crystallographic work. The overall TLR8 monomer structure resembles a ring and contains 26 LRR, which is the largest number of LRR modules among the TLRs. It contains a long region of  $\sim 40$  amino acids between LRR14 and LRR15, which is referred to as the Z-loop. This can be cleaved by unidentified proteases, splitting TLR8 into half and generating fragments of 60 and 50 kDa within the cell. This proteolytic step has been proposed to be essential for the maturation and sorting of TLR9.<sup>68–70</sup>

Unlike other TLRs, TLR8 forms homodimers even before binding agonistic ligands. In the absence of bound ligand, the C-terminal ends of the two TLR8 subunits in the dimer are separated by  $\sim 53$  Å; upon ligand binding, the two C-termini are brought into closer proximity, and so should be able to promote close approximation of the intracellular TIR domains. This structural change is induced by several interactions between TLR8 and the bound ligand. The most important residues in this interaction are Phe405 and Asp543, which form a  $\pi$  stack interaction with benzene rings of the imidazoquinoline group of R848 or the thiazoloquinolone groups of CL097 and CL075, and form hydrogen bonds with the nitrogen atoms of the imidazole and thiazole moieties. Thus, acting as a kind of molecular glue, the bound ligands activate TLR8 by rearranging the preformed receptor dimer.

### THE HYBRID LRR TECHNIQUE

Structure determination of TLR family proteins has been a major challenge in the study of innate immunity because crystallization of these proteins is not trivial. This obstacle was overcome by a novel crystallization technique named the 'hybrid LRR technique' (Figure 6).<sup>27,30</sup> In this method, a panel of hybrids of target and partner LRR family proteins is generated. The hybridization partner proteins are chosen from a set of LRR family proteins that are easy to produce and crystallize. The LRR modules of the pairs of proteins are joined while conserving the sequence of the modules. Due to the highly modular nature of LRR family proteins this



**Figure 6** The hybrid LRR technique. Two LRR family proteins are fused together while maintaining the sequence conservation pattern of the LRR modules. (a) The functional domain of TLR4 fused with a partner LRR protein at its C-terminus. (b) Alternatively, the partner protein can be fused to the N-terminal region of TLR4.

simple-minded fusion of the two proteins was surprisingly successful, with a success rate exceeding 50%. In a typical fusion protein procedure in molecular biology, the two proteins are connected by a flexible linker. As a result, the hydrophobic cores of the two proteins remain separate and their tendency to crystallize is reduced due to the flexibility of the linker. However, the crystallization behavior of hybrid LRR proteins is often superior because the hydrophobic cores of the two proteins are connected directly and the resulting hybrid has a rigid structure. Because one can generate a practically unlimited number of hybrid proteins for a single target protein by switching partner proteins, some of them should be able to crystallize even though the original target is not. This technique was successfully applied to the crystallographic study of TLRs 1, 2 and 4–6 described in this study, and to RP105, a close relative of the TLRs.<sup>27,29,30,45,71</sup>

### CONCLUSIONS AND PERSPECTIVES

LPS is a potent stimulator of the innate immune response. Its structure and immunological activity has been studied for over a century. In the last decade, crystal structures of most of the receptors and accessory proteins involved in LPS recognition have been determined. These structures explain how our immune system recognizes the common structural pattern in diverse LPS molecules. They also provide an initial clue as to

how LPS binding initiates the aggregation of intracellular signaling proteins that leads to cytokine production and the initiation of inflammatory signaling. TLR4 is the key receptor involved in LPS recognition and signal initiation. Up to now, the structures of six of the ten human TLRs in complex with their cognate ligands have been determined. Although the locations and structures of the ligand interaction sites of these TLRs differ greatly, they all form m-shaped dimers with the two C-termini in the center, and the N-termini on the outside. All this structural information should contribute to the development of novel therapeutic agents that either attenuate or enhance TLR-mediated signaling.

## ACKNOWLEDGEMENTS

This work was supported in part by a grant from Eulji University in 2012.

- Janeway CA Jr, Medzhitov R. Innate immune recognition. *Annu Rev Immunol* 2002; **20**: 197–216.
- Medzhitov R, Janeway C Jr. Innate immune recognition: mechanisms and pathways. *Immunol Rev* 2000; **173**: 89–97.
- Raetz CR, Whitfield C. Lipopolysaccharide endotoxins. *Annu Rev Biochem* 2002; **71**: 635–700.
- Pugin J, Schurer-Maly CC, Leturcq D, Moriarty A, Ulevitch RJ, Tobias PS. Lipopolysaccharide activation of human endothelial and epithelial cells is mediated by lipopolysaccharide-binding protein and soluble CD14. *Proc Natl Acad Sci USA* 1993; **90**: 2744–2748.
- Chow JC, Young DW, Golenbock DT, Christ WJ, Gusovsky F. Toll-like receptor-4 mediates lipopolysaccharide-induced signal transduction. *J Biol Chem* 1999; **274**: 10689–10692.
- Cohen J. The immunopathogenesis of sepsis. *Nature* 2002; **420**: 885–891.
- Akira S, Takeda K. Toll-like receptor signalling. *Nat Rev Immunol* 2004; **4**: 499–511.
- Gay NJ, Gangloff M. Structure and function of Toll receptors and their ligands. *Annu Rev Biochem* 2007; **76**: 141–165.
- Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. *Cell* 2006; **124**: 783–801.
- Maeshima N, Fernandez RC. Recognition of lipid A variants by the TLR4-MD-2 receptor complex. *Front Cell Infect Microbiol* 2013; **3**: 3.
- Trent MS, Stead CM, Tran AX, Hankins JV. Diversity of endotoxin and its impact on pathogenesis. *J Endotoxin Res* 2006; **12**: 205–223.
- Fujihara M, Muroi M, Tanamoto K, Suzuki T, Azuma H, Ikeda H. Molecular mechanisms of macrophage activation and deactivation by lipopolysaccharide: roles of the receptor complex. *Pharmacol Ther* 2003; **100**: 171–194.
- Miyake K. Innate recognition of lipopolysaccharide by CD14 and toll-like receptor 4-MD-2: unique roles for MD-2. *Int Immunopharmacol* 2003; **3**: 119–128.
- Mulero JJ, Boyle BJ, Bradley S, Bright JM, Nelken ST, Ho TT *et al*. Three new human members of the lipid transfer/lipopolysaccharide binding protein family (LT/LBP). *Immunogenetics* 2002; **54**: 293–300.
- Bingle CD, Craven CJ. Meet the relatives: a family of BPI- and LBP-related proteins. *Trends Immunol* 2004; **25**: 53–55.
- Beamer LJ, Carroll SF, Eisenberg D. Crystal structure of human BPI and two bound phospholipids at 2.4 angstrom resolution. *Science* 1997; **276**: 1861–1864.
- Qiu X, Mistry A, Ammirati MJ, Chrnyk BA, Clark RW, Cong Y *et al*. Crystal structure of cholesteryl ester transfer protein reveals a long tunnel and four bound lipid molecules. *Nat Struct Mol Biol* 2007; **14**: 106–113.
- Huuskonen J, Wohlfahrt G, Jauhiainen M, Ehnholm C, Teleman O, Oikkonen VM. Structure and phospholipid transfer activity of human PLTP: analysis by molecular modeling and site-directed mutagenesis. *J Lipid Res* 1999; **40**: 1123–1130.
- Desrumaux C, Labeur C, Verhee A, Tavernier J, Vandekerckhove J, Rosseneu M *et al*. A hydrophobic cluster at the surface of the human plasma phospholipid transfer protein is critical for activity on high density lipoproteins. *J Biol Chem* 2001; **276**: 5908–5915.
- Kleiger G, Beamer LJ, Grothe R, Mallick P, Eisenberg D. The 1.7A crystal structure of BPI: a study of how two dissimilar amino acid sequences can adopt the same fold. *J Mol Biol* 2000; **299**: 1019–1034.
- Ulevitch RJ, Tobias PS. Receptor-dependent mechanisms of cell stimulation by bacterial endotoxin. *Annu Rev Immunol* 1995; **13**: 437–457.
- Haziot A, Ferrero E, Kontgen F, Hijiya N, Yamamoto S, Silver J *et al*. Resistance to endotoxin shock and reduced dissemination of gram-negative bacteria in CD14-deficient mice. *Immunity* 1996; **4**: 407–414.
- Kim JI, Lee CJ, Jin MS, Lee CH, Paik SG, Lee H *et al*. Crystal structure of CD14 and its implications for lipopolysaccharide signaling. *J Biol Chem* 2005; **280**: 11347–11351.
- Kelley SL, Lukk T, Nair SK, Tapping RI. The crystal structure of human soluble CD14 reveals a bent solenoid with a hydrophobic amino-terminal pocket. *J Immunol* 2013; **190**: 1304–1311.
- Kobe B, Kajava AV. The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol* 2001; **11**: 725–732.
- Dziarski R. Recognition of bacterial peptidoglycan by the innate immune system. *Cell Mol Life Sci* 2003; **60**: 1793–1804.
- Jin MS, Kim SE, Heo JY, Lee ME, Kim HM, Paik SG *et al*. Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* 2007; **130**: 1071–1082.
- Park BS, Song DH, Kim HM, Choi BS, Lee H, Lee JO. The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* 2009; **458**: 1191–1195.
- Kang JY, Nan X, Jin MS, Youn SJ, Ryu YH, Mah S *et al*. Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. *Immunity* 2009; **31**: 873–884.
- Kim HM, Park BS, Kim JI, Kim SE, Lee J, Oh SC *et al*. Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell* 2007; **130**: 906–917.
- Ohto U, Fukase K, Miyake K, Satow Y. Crystal structures of human MD-2 and its complex with antiendotoxic lipid IVa. *Science* 2007; **316**: 1632–1634.
- Inohara N, Nunez G. ML—a conserved domain involved in innate immunity and lipid metabolism. *Trends Biochem Sci* 2002; **27**: 219–221.
- Mata-Haro V, Cekic C, Martin M, Chilton PM, Casella CR, Mitchell TC. The vaccine adjuvant monophosphoryl lipid A as a TRIF-biased agonist of TLR4. *Science* 2007; **316**: 1628–1632.
- Casella CR, Mitchell TC. Putting endotoxin to work for us: monophosphoryl lipid A as a safe and effective vaccine adjuvant. *Cell Mol Life Sci* 2008; **65**: 3231–3240.
- Rietschel ET, Kirikae T, Schade FU, Mamat U, Schmidt G, Loppnow H *et al*. Bacterial endotoxin: molecular relationships of structure to activity and function. *FASEB J* 1994; **8**: 217–225.
- Ohto U, Yamakawa N, Akashi-Takamura S, Miyake K, Shimizu T. Structural analyses of human Toll-like receptor 4 polymorphisms D299G and T399I. *J Biol Chem* 2012; **287**: 40611–40617.
- Rossignol DP, Lynn M. TLR4 antagonists for endotoxemia and beyond. *Curr Opin Investig Drugs* 2005; **6**: 496–502.
- Piazza M, Calabrese V, Damore G, Cighetti R, Gioannini T, Weiss J *et al*. A synthetic lipid A mimetic modulates human TLR4 activity. *ChemMedChem* 2012; **7**: 213–217.
- Barochia A, Solomon S, Cui X, Natanson C, Eichacker PQ. Eritoran tetrasodium (E5564) treatment for sepsis: review of preclinical and clinical studies. *Expert Opin Drug Metab Toxicol* 2011; **7**: 479–494.
- Shirey KA, Lai W, Scott AJ, Lipsky M, Mistry P, Pletneva LM *et al*. The TLR4 antagonist Eritoran protects mice from lethal influenza infection. *Nature* 2013; **497**: 498–502.
- Raetz CR, Reynolds CM, Trent MS, Bishop RE. Lipid A modification systems in gram-negative bacteria. *Annu Rev Biochem* 2007; **76**: 295–329.
- Ohto U, Fukase K, Miyake K, Shimizu T. Structural basis of species-specific endotoxin sensing by innate immune receptor TLR4/MD-2. *Proc Natl Acad Sci USA* 2012; **109**: 7421–7426.
- Kang JY, Lee JO. Structural biology of the Toll-like receptor family. *Annu Rev Biochem* 2011; **80**: 917–941.
- Liu L, Botos I, Wang Y, Leonard JN, Shiloach J, Segal DM *et al*. Structural basis of toll-like receptor 3 signaling with double-stranded RNA. *Science* 2008; **320**: 379–381.



- 45 Yoon SI, Kurnasov O, Natarajan V, Hong M, Gudkov AV, Osterman AL *et al.* Structural basis of TLR5-flagellin recognition and signaling. *Science* 2012; **335**: 859–864.
- 46 Tanji H, Ohto U, Shibata T, Miyake K, Shimizu T. Structural reorganization of the Toll-like receptor 8 dimer induced by agonistic ligands. *Science* 2013; **339**: 1426–1429.
- 47 O'Neill LA, Bowie AG. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nat Rev Immunol* 2007; **7**: 353–364.
- 48 Kenny EF, O'Neill LA. Signalling adaptors used by Toll-like receptors: an update. *Cytokine* 2008; **43**: 342–349.
- 49 Watters TM, Kenny EF, O'Neill LA. Structure, function and regulation of the Toll/IL-1 receptor adaptor proteins. *Immunol Cell Biol* 2007; **85**: 411–419.
- 50 Janssens S, Beyaert R. A universal role for MyD88 in TLR/IL-1R-mediated signaling. *Trends Biochem Sci* 2002; **27**: 474–482.
- 51 Song DH, Lee JO. Sensing of microbial molecular patterns by Toll-like receptors. *Immunol Rev* 2012; **250**: 216–229.
- 52 Xu Y, Tao X, Shen B, Horng T, Medzhitov R, Manley JL *et al.* Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature* 2000; **408**: 111–115.
- 53 Nyman T, Stenmark P, Flodin S, Johansson I, Hammarstrom M, Nordlund P. The crystal structure of the human toll-like receptor 10 cytoplasmic domain reveals a putative signaling dimer. *J Biol Chem* 2008; **283**: 11861–11865.
- 54 Khan JA, Brint EK, O'Neill LA, Tong L. Crystal structure of the Toll/interleukin-1 receptor domain of human IL-1RAPL. *J Biol Chem* 2004; **279**: 31664–31670.
- 55 Valkov E, Stamp A, Dimaio F, Baker D, Verstak B, Roversi P *et al.* Crystal structure of Toll-like receptor adaptor MAL/TIRAP reveals the molecular basis for signal transduction and disease protection. *Proc Natl Acad Sci USA* 2011; **108**: 14879–14884.
- 56 Ohnishi H, Tochio H, Kato Z, Orii KE, Li A, Kimura T *et al.* Structural basis for the multiple interactions of the MyD88 TIR domain in TLR4 signaling. *Proc Natl Acad Sci USA* 2009; **106**: 10260–10265.
- 57 Monie TP, Moncrieffe MC, Gay NJ. Structure and regulation of cytoplasmic adapter proteins involved in innate immune signaling. *Immunol Rev* 2009; **227**: 161–175.
- 58 Gautam JK, Ashish Comeau LD, Krueger JK, Smith MF Jr. Structural and functional evidence for the role of the TLR2 DD loop in TLR1/TLR2 heterodimerization and signaling. *J Biol Chem* 2006; **281**: 30132–30142.
- 59 Bovijn C, Ulrichts P, De Smet AS, Cateeuw D, Beyaert R, Tavernier J *et al.* Identification of interaction sites for dimerization and adapter recruitment in Toll/interleukin-1 receptor (TIR) domain of Toll-like receptor 4. *J Biol Chem* 2012; **287**: 4088–4098.
- 60 Lin SC, Lo YC, Wu H. Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling. *Nature* 2010; **465**: 885–890.
- 61 Takeuchi O, Kawai T, Muhlradt PF, Morr M, Radolf JD, Zychlinsky A *et al.* Discrimination of bacterial lipoproteins by Toll-like receptor 6. *Int Immunol* 2001; **13**: 933–940.
- 62 Takeuchi O, Sato S, Horiuchi T, Hoshino K, Takeda K, Dong Z *et al.* Cutting edge: role of Toll-like receptor 1 in mediating immune response to microbial lipoproteins. *J Immunol* 2002; **169**: 10–14.
- 63 Hutchings MI, Palmer T, Harrington DJ, Sutcliffe IC. Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold 'em, knowing when to fold 'em. *Trends Microbiol* 2009; **17**: 13–21.
- 64 Alexopoulou L, Holt AC, Medzhitov R, Flavell RA. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* 2001; **413**: 732–738.
- 65 Choe J, Kelker MS, Wilson IA. Crystal structure of human toll-like receptor 3 (TLR3) ectodomain. *Science* 2005; **309**: 581–585.
- 66 Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC, Goodlett DR *et al.* The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* 2001; **410**: 1099–1103.
- 67 Heil F, Hemmi H, Hochrein H, Ampenberger F, Kirschning C, Akira S *et al.* Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science* 2004; **303**: 1526–1529.
- 68 Park B, Brinkmann MM, Spooner E, Lee CC, Kim YM, Ploegh HL. Proteolytic cleavage in an endolysosomal compartment is required for activation of Toll-like receptor 9. *Nat Immunol* 2008; **9**: 1407–1414.
- 69 Ewald SE, Lee BL, Lau L, Wickliffe KE, Shi GP, Chapman HA *et al.* The ectodomain of Toll-like receptor 9 is cleaved to generate a functional receptor. *Nature* 2008; **456**: 658–662.
- 70 Kim YM, Brinkmann MM, Paquet ME, Ploegh HL. UNC93B1 delivers nucleotide-sensing toll-like receptors to endolysosomes. *Nature* 2008; **452**: 234–238.
- 71 Yoon SI, Hong M, Wilson IA. An unusual dimeric structure and assembly for TLR4 regulator RP105-MD-1. *Nat Struct Mol Biol* 2011; **18**: 1028–1035.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>