Tracing the evolution of novel features of human Toll-like receptor 4

Jeremy A. Anderson,^{1,2} Andrea N. Loes,^{1,2} Grace L. Waddell,^{1,2} and Michael J. Harms^{1,2*}

¹Institute for Molecular Biology, University of Oregon, Eugene, Oregon 97403

²Department of Chemistry and Biochemistry, University of Oregon, Eugene, Oregon 97403

Received 1 April 2019; Accepted 8 May 2019

DOI: 10.1002/pro.3644

Published online 3 June 2019 proteinscience.org

Abstract: Toll-like receptor 4 (TLR4) is a critical innate immune protein that activates inflammation in response to extracellular cues. Much of the work to understand how the protein works in humans has been done using mouse models. Although human and mouse TLR4 have many shared features, they have also diverged significantly since their last common ancestor, acquiring 277 sequence differences. Functional differences include the extent of ligand-independent activation, whether lipid IVa acts as an antagonist or agonist, and the relative species cross-compatibility of their MD-2 cofactor. We set out to understand the evolutionary origins for these functional differences between human and mouse TLR4. Using a combination of phylogenetics, ancestral sequence reconstruction, and functional characterization, we found that evolutionary changes to the human TLR4, rather than changes to the mouse TLR4, were largely responsible for these functional changes. Human TLR4 repressed ancestral ligand-independent activity and gained antagonism to lipid IVa. Additionally, mutations to the human TLR4 cofactor MD-2 led to lineage-specific incompatibility between human and opossum TLR4 complex members. These results were surprising, as mouse TLR4 has acquired many more mutations than human TLR4 since their last common ancestor. Our work has polarized this set of transitions and sets up work to study the mechanistic underpinnings for the evolution of new functions in TLR4.

Keywords: Toll-like receptor 4; ancestral sequence reconstruction; innate immunity; protein evolution; comparative immunology; opossum

Introduction

Toll-like receptor 4 (TLR4) is a critical innate immune receptor that upregulates NF- κ B in response

to extracellular pro-inflammatory signals.^{1,2} It is primarily activated by lipopolysaccharide (LPS), a cell wall component from Gram-negative bacteria.³ The cofactor CD14 loads LPS into the hydrophobic pocket of the MD-2 cofactor of TLR4, promoting TLR4 dimerization and activating both Myd88 and TRIF/TRAM inflammatory pathways [Fig. 1(a)].⁴ This recruits neutrophils to sites of infection and tissue damage,^{5,6} but also contributes to inflammatory disorders and cancer.^{7,8} Most destructively, excessive activation of TLR4 can lead to septic shock—a leading cause of death in hospitals worldwide.^{9,10}

Like many innate immune proteins, TLR4 has maintained its core function even as its amino acid sequence has evolved. Humans and chickens—which diverged 320 million years ago—use TLR4 and the same core cofactors to recognize and respond to LPS.^{11–13} Despite this conservation of function, human and

Additional Supporting Information may be found in the online version of this article.

Jeremy A. Anderson and Andrea N. Loes contributed equally to this work.

Significance statement: We used an evolutionary lens to compare human and mouse innate immune proteins. We found several ways that both the human and mouse proteins have changed significantly since their last common ancestor. This provides context for interpreting studies in mice to understand human biology and health.

Grant sponsor: National Institute of General Medical Sciences T32GM007759; Grant sponsor: Pew Charitable Trusts.

*Correspondence to: Michael J. Harms, Institute for Molecular Biology, University of Oregon, Eugene, OR 97403-1229. E-mail: harms@uoregon.edu

Figure 1. Functional differences between human and mouse TLR4 arose by an evolutionary process. (a) The cartoon shows how LPS activates the TLR4/MD-2 complex. In the absence of LPS, TLR4 and MD-2 form individual complexes (pink and blue). When the cofactor CD14 delivers LPS to the hydrophobic pocket in MD-2, a dimer of TLR4/MD-2 forms and activates NF- κ B signaling. (b) Phylogenetic tree schematic showing the relationship between human (h), mouse (m), opossum (o), and the human/mouse ancestral TLR4 proteins (a). The table to the right shows several functional features that differ between the human and mouse proteins. The information in the gray region was filled in by the current study. (c) Rates of sequence change for the extracellular domain of TLR4 (top), the intracellular domain of TLR4 (middle), and MD-2 (bottom). The numbers give the average number of substitutions per site along each evolutionary branch.

chicken TLR4 only share 46% sequence identity, differing at 459 of 847 sites. This turnover in sequence has allowed TLR4 orthologs in different species to diversify in function,^{11–16} even while maintaining the central LPS response. For example, another small molecule produced in Gram-negative bacteria, lipid IVa, activates mouse TLR4 but inhibits human TLR4.^{17–19} There are quantitative differences in function as well. Mouse TLR4 possesses a degree of ligand-independent activity, inducing NF- κ B even in the absence of agonist, human exhibits very little ligand-independent activity.¹³

These functional differences have important consequences. Much of the work done to understand TLR4 biology has, of necessity, relied on mouse models. While such studies have powerfully illuminated human biology,

they are not always reliable guides for human. One notable example is the anti-sepsis drug Eritoran, which performed well in preclinical mouse trials but then failed in clinical trials,²⁰ possibly because human and mouse TLR4 have different responses to lipid A molecules.¹⁵ Likewise, the anticancer drug taxol has radically different potency against mouse and human TLR4.^{21,22}

Identifying the molecular basis for these functional differences is quite challenging. Even though mice and humans only diverged 90 million years ago,²³ mouse and human TLR4 now differ at 277 sites. Sifting through this number of mutations can be daunting. Further, a direct comparison between two proteins ignores the fact that the sequence differences arose by two parallel evolutionary processes diverging from a common ancestral protein. On the other hand, directly accounting for that evolutionary process can help one make sense of the functional differences between the modern proteins.²⁴

We applied this evolutionary perspective to three functional differences between human and mouse TLR4. For each function, our primary aim was to polarize the evolutionary transition: did functional differences arise by innovation on the human lineage, innovation on the mouse lineage, or innovation on both lineages? We experimentally investigated the features of four proteins: human TLR4 (hTLR4), mouse TLR4 (mTLR4), opossum TLR4 (oTLR4), and a reconstructed version of TLR4 from the last common ancestor of mice and humans (ancTLR4). We found that all three functions we studied—ligand-independent activity, lipid IVa antagonism, and TLR4/MD-2 incompatibility—involved functional changes on the human lineage. We also found that ligand-independent activity increased along the mouse lineage. Identification of human-specific features of TLR4, such as these, is essential to development of rational therapeutic strategies to regulate inflammation and combat sepsis.

Results

We began the study by constructing a phylogeny of TLR4 proteins from across bony vertebrates. We built a sequence alignment containing 199 bony vertebrate TLR4 sequences, and then arranged them into a phylogeny using known species relationships.^{25–28} We then optimized the branch lengths of this tree using a discrete gamma model to capture differential evolutionary rates across sites.²⁹

With the tree in hand, we then estimated the sequence of ancTLR4, the last common ancestor of the human and mouse TLR4. The protein was reconstructed with high confidence. The mean posterior probability across sites was 85.6%, with 65.7% of sites having posterior probabilities $\geq 90\%$. Only 9.8% of sites had a plausible alternate reconstruction; the remainder was completely unambiguous. The full alignment, tree, ancestral reconstruction statistics, and ancestral protein sequence are available in Files S1–S4.

To set our expectations for our functional studies, we then estimated the rates of evolutionary change for the intracellular and the extracellular domains of TLR4 [Fig. 1(a)]. We split the alignment in two at the domain boundary (residue 672 in the human numbering), and then estimated the branch lengths for each domain phylogeny. For comparison, we also constructed an evolutionary tree for these three species for the protein MD-2. The branch lengths for the human, mouse, opossum, and ancestral sequences are shown in Figure 1(c).

We found that the extracellular domain evolved approximately five times faster than the intracellular domain [Fig. 1(c)]. If we calibrate to the time of divergence between the human and opossum—160 million years ago²³—the extracellular domain has accumulated ~3 mutations every million years along each lineage. In addition to evolving faster overall, the extracellular domain exhibits a much wider distribution of rates across sites, with some sites evolving rapidly and others being highly conserved. This is evidenced by the α shape parameter in our phylogenetic model (see Methods).²⁹ We estimated α as 0.64 for the intracellular domain and 1.18 for the extracellular domain. The MD-2 cofactor evolved similarly to the extracellular domain of TLR4 (~3 mutations/million years).

This analysis also revealed that both domains of mouse TLR4, as well as MD-2, evolved approximately twice as fast as human TLR4 over the 90 million years since they diverged from ancTLR4 [Fig. 1(c)]. As a result, hTLR4 differs from ancTLR4 at 130 sites (84.6% identity), while mTLR4 differs from ancTLR4 at 244 sites (71.1% identity). From this, we predicted that the functional differences between the human and mouse TLR4 represent innovations on the mouse, rather than the human lineage.

The TLR4/MD-2 interface has two surfaces with different evolutionary signals

We next asked how the conserved and evolutionarily variable sites in TLR4 and MD-2 mapped to the structure of the protein complex. We found that the TLR4/MD-2 interface has two evolutionarily distinct surfaces. The first is the interaction surface between TLR4 and MD-2 that is present even in the absence of agonist [Fig. 2(a,b)]. This surface is dominated by conserved amino acids on both the TLR4 and MD-2 portion of the interface. This can be seen by the high degree of overlap between the red signal (indicating conservation) and the blue shading (indicating protein–protein contacts). The second surface is between two TLR4/MD-2 complexes,

Figure 2. The TLR4/MD-2 complex uses one conserved and one variable protein–protein interface. (a) Schematic of the TLR4/MD-2 complex indicating the views shown in panel b (view 1) and panel c (view 2). Points of interest are indicated on the structure as numbers from 1 to 3. (b) The primary TLR4/MD-2 interface, taken from PDB 3FXI.³⁰ The top figure shows the orientation of the structure below using the same color scheme as the schematic in panel a. LPS is shown as spheres, with oxygen indicated in brown. The pink subunits are not shown in the lower subpanel. The bottom subpanel shows MD-2 pulled away from the TLR4 interface. The protein surfaces are colored from white (low conservation, score = 0) to red (high conservation, score < −1.5) as calculated using con-surf.³¹ The blue points indicate atoms that are within 5 Å of the partner protein. (c) The second view shows the interface between a dimer of TLR4/MD-2 heterodimers. The “view 2” (shown in panel a) is now oriented directly into the page. The pink subunits are not shown in the lower subpanel. The colors in the lower subpanel (red, white, and blue) are as in panel b.

and thus only forms when the complex is active and dimerized [Fig. 2(c)]. In contrast to the first surface, this second interface has very few conserved positions. Only Pro-88 and Arg-90 of MD-2 appear to be both strongly conserved and within 5 Å of an atom of TLR4 in the structure. The remainder of this second surface is variable across vertebrates.

Inhibition of ligand-independent activity evolved along the human lineage

With this evolutionary context, we then turned our attention to trying to polarize functional transitions in TLR4 function. The first difference between the hTLR4 and mTLR4 we investigated was the degree to which the output depends on the presence of agonist. In transactivation assays, mTLR4 activates even when no agonist is present, while human hTLR4 does not [Fig. 1(b)].¹³ mTLR4 then activates on top of the ligand-independent activity. Tighter regulation of hTLR4 relative to mTLR4 is a potentially confounding factor when using mouse models to study human disease states.

To study this functional difference, we first reproduced the previously published observation of differential ligand-independent activity of hTLR4 and mTLR4 using in a well-established NF-κB transactivation assay in HEK293T cells. As expected, the mouse protein exhibited strong, dose-dependent activity even in the absence of agonist [Fig. 3(a)]. In contrast, the human protein exhibited only weak activity even at the highest doses of the protein [Fig. 3(a)]. To verify that this was robust to the assembly of the entire TLR4/MD-2 complex, we repeated these experiments in the presence of MD-2 and CD14—cofactors known to be critical for the activation of the complex in the presence of agonists [Fig. 1(a)]. We observed the same basic pattern with the MD-2 and CD14 cofactors present: mTLR4 activates to a much greater level than hTLR4 in the absence of agonist [Fig. 3(b)].

Figure 3. Human TLR4 acquired low ligand-independent activation relative to other mammals. Panels show the ligand-independent activation of TLR4 alone (panel a) and co-transfected with MD-2 and CD14 (panel b) as a function of amount of transfected plasmid. The colors denote different species: human (blue), mouse (orange), the human/mouse ancestor (red), and opossum (green).

We then characterized the oTLR4. This protein exhibited slightly higher ligand-independent activity than the human protein, both alone [Fig. 3(a)] with the full TLR4 complex [Fig. 3(b)]. Of this collection of proteins, the human protein exhibited the lowest ligand-independent activity and the mouse the highest.

We then tested the activity of ancTLR4. We found that its ligand-independent activity was significantly greater than the human protein, but less than the mouse protein [Fig. 3(a)]. This ligand-independent activity is suppressed if we included the human MD-2 and CD14 cofactors, but it still remained higher than the human protein [Fig. 3(b)]. These results suggest that changes occurred on both the human and mouse lineages. Mouse TLR4 gained greater ligand-independent activity over the last 90 million years, while human TLR4 lost ligand-independent activity over the same interval.

Human TLR4 gained lipid IVa antagonism

The second functional difference we explored was the interaction between TLR4 and lipid IVa. While lipid IVa is chemically similar to LPS, it has radically different effects on hTLR4 and mTLR4. Lipid IVa activates the mouse protein, but inhibits the human protein [Fig. 1(b)].^{18,32}

As we did for the ligand-independent activity of mTLR4, we asked whether the innovation occurred on the human or mouse lineage. We characterized the response of the human, mouse, and opossum proteins to lipid IVa in the presence of species-matched MD-2 and CD14 [Fig. 4(a)]. We found that lipid IVa did not activate the human complex above a buffer control, but did activate the mouse and opossum complexes [Fig. 4(c)]. We then measured the ability of lipid IVa to inhibit activation of LPS by adding the two molecules together. LPS activation was strongly inhibited by the addition of lipid IVa for the human TLR4/MD-2 complex, but not for the mouse and opossum complexes [Fig. 4(d)]. This suggested that the ancestral state of the receptor was to be activated by lipid IVa, and that antagonism evolved along the human lineage.

We then tested this hypothesis using ancTLR4. Because the ligand-induced activity of TLR4 requires the MD-2 coreceptor, we needed to include an MD-2 protein in the assay. We thus tested the response of ancTLR4 in the presence of the human, mouse, and opossum MD-2 proteins. We found that the ancTLR4/mouse MD-2 complex exhibited strong activation with lipid IVa [Fig. 4(c)]. This response was observed to a lesser extent for the opossum MD-2 complex and was undetectable for the human MD-2. We also observed mild inhibition of LPS activity by lipid IVa in the ancTLR4/hMD-2 complex [Fig. 4(d)]. This is consistent with the antagonism with lipid IVa being encoded in the MD-2 coreceptor rather than in the TLR4 component of the signaling complex.

Overall, our data from both the opossum and the ancTLR4 proteins are consistent with an ancestral protein that activated inflammation in response to lipid

Figure 4. Lipid IVa antagonism evolved along the human lineage. (a) Bar plots show relative activity of TLR4 measured for combinations of TLR4 and MD-2 from different species, challenged with buffer (gray), LPS (red), lipid IVa (blue), or LPS and lipid IVa together (purple). Conditions are indicated schematically below each bar. Bars for each TLR4 protein (h, m, o, or a) are normalized to the species-matched TLR4/MD-2 + LPS value. For the ancestral protein, values are normalized to the ancTLR4/hMD-2 + LPS value. Error bars are standard error on the mean of three biological replicates. (b) Heat map showing the log₂ fold activation induced by LPS above the buffer control for each pairwise combination of TLR4 and MD-2. (c) Heat map showing the log₂ fold activation induced by lipid IVa above the buffer control for each pairwise combination of TLR4 and MD-2. (d) Heat map showing the log₂ fold inhibition of LPS activity (LPS + IVa sample divided by the LPS sample) for each pairwise combination of TLR4 and MD-2. For panels b–d, the color scale is indicated below the plot. **P* < 0.05; ***P* < 0.005 using a one-tailed *t*-test.

IVa. hTLR4 then gained lipid IVa antagonism after the divergence of the human and mouse lineages.

Mutations to a loop in human MD-2 altered the human TLR4 LPS response

In addition to changes in ligand-independent activity and interaction with lipid IVa, the human TLR4/MD-2 and mouse TLR4/MD-2 complexes exhibit differences in their canonical LPS response. For example, we previously found that oTLR4 could activate using mouse MD-2 but not human MD-2.¹³ The inverse also held: hTLR4 could activate using mouse MD-2 but not opossum MD-2 [Fig. 5(a,b)]. Such incompatibility points to lineage-specific coevolution on either the human or opossum lineage (or both). Such lineage-specific changes arise from ratcheting changes in the structure–function relationships in each molecule,²⁴ thus complicating attempts to compare proteins between two species. Investigating such coevolution can also reveal unexpected residues that mediate function.

We suspected that the lineage-specific incompatibility evolved using the variable TLR4/MD-2 surface [Fig. 2(c)] rather than the conserved TLR4/MD-2 surface [Fig. 2(b)], as this is the collection of positions in which sequence changes can be tolerated. A careful look the human, mouse, and opossum MD-2 sequences revealed a sequence difference between human MD-2 and other mammalian MD-2

proteins. At positions 118 and 127 (human numbering), human MD-2 has Ser; other mammals have Pro [Fig. 5(c)]. These Pro to Ser substitutions flank a loop containing a mixture of conserved and variable sites. Crucially, this loop is in direct contact with both LPS and interacts across the TLR4/MD-2 complex dimer [Fig. 5(d)]. Finally, if we overlay the structure of the TLR4/MD-2 complex from human and mouse, we see that the loop is indeed in a different position [Fig. 5(d)].

We therefore introduced the S118P and S127P mutations into the human protein and measured its cross-compatibility with TLR4 taken from mouse and opossum. While wild-type human MD-2 is incompatible with oTLR4, human MD-2 S118P/S127P now allows significant activation of oTLR4. This suggests that these two lineage-specific changes to human MD-2 contributed to the incompatibility between the proteins. Interestingly, these changes do not compromise activity with hTLR4 and mTLR4.

There are two main possibilities that might explain the effects of these mutations. The first possibility is that this is a direct interaction, as residue 127 forms a part of the interface with the opposite TLR4/MD-2 complex. The second is that these mutations are inducing a structural or dynamic rearrangement in the loop, like the mouse TLR4/MD-2 crystal structure [Fig. 5(d)]. To explore the feasibility of latter hypothesis, we ran 300 ns

Figure 5. Mutations to MD-2 loop determine cross-compatibility between MD-2 and TLR4 between species. (a) Bar plots show activation of TLR4/MD-2 complexes with buffer (gray) or LPS (red). Conditions are indicated schematically below each bar. Bars for each TLR4 protein (h, m, o, or a) are normalized to the species-matched TLR4/MD-2 + LPS value. For the ancestral protein, values are normalized to the ancTLR4/hMD-2 + LPS value. Error bars are standard error on the mean of three biological replicates. (b) Heat map showing the log₂ fold activation induced by LPS above the buffer control for each pairwise combination of TLR4 and MD-2. * $P < 0.05$; ** $P < 0.005$ using a one-tailed t -test. (c) Sequence comparison between the 118–127 loops in human, mouse, and opossum MD-2. (d) The MD-2 loop populates a slightly different conformation in crystal structures of human and mouse TLR4/MD-2/LPS complexes (3FXI³⁰ and 3VQ2³³). The colors of each subunit are the same as the cartoon in Figure 1(a). The blue loop is taken from the human structure; the yellow loop is taken from the mouse structure. The locations of positions 118 and 127 are indicated. (e) Snapshots of this loop taken at 5 ns intervals from 300 ns MD simulations of human MD-2/LPS complexes with Ser (left) or Pro (right) at positions 118 and 127. The spheres indicate the mutated positions, the blue or yellow region the intervening loop. The gray structure is from the start of the simulations.

molecular dynamics simulations on human MD-2 bound to LPS. We then introduced Pro at positions 118 and 127 and then re-ran the simulations. We found that the loop populated many more conformations with the wild-type Ser at these positions than it did with the newly introduced Pro residues [Fig. 5(e)]. This suggests that the historical mutations may have altered TLR4/MD-2 cross-compatibility by quenching the dynamics and/or altering the structure of this MD-2 loop.

Discussion

The primary goal of this study was to “polarize” the evolutionary transitions for a collection of functional differences between human and mouse TLR4/MD-2. Did the functional differences arise due to changes on the human or on the mouse lineage? Based solely on the rate of evolution, we would have predicted that the change arose on the mouse lineage, as mouse TLR4 acquired nearly twice as many mutations as human TLR4 since they diverged 90 million years ago (Fig. 1).

We found examples of evolutionary changes on both lineages. Humans and mice differ in their ligand-independent activity because mice have increased this activity dramatically since the human/mouse ancestor (Fig. 3). Mice have diversified, while humans have maintained the ancestral state. In contrast, the antagonism exhibited by lipid IVa arose on the human lineage (Fig. 4): humans have diversified, while mice have maintained the ancestral state.

The molecular basis for the difference in ligand-independent activity between the human and mouse remains unclear. It is known that the intracellular domains of human TLR4 dimerize and activate NF- κ B in the absence of the extracellular domain.³⁴ We hypothesize that the changes important for increasing the ligand-independent activity of mouse TLR4 therefore involved changes in the extracellular domain, making it less effective at inhibiting dimerization. These changes would likely be on the highly variable TLR4/MD-2 dimerization [Fig. 5(c)]. Identifying the sequence changes and their mechanism is future line of research.

The molecular basis for differences in lipid IVA agonism and antagonism is better understood than that of ligand-independent activity. A collection of differences between the residues directly adjacent to lipid IVA in the MD-2 binding pocket likely reposition the small molecule to prevent it from driving dimerization of the TLR4/MD-2 complex.^{16,33} Intriguingly, three of the human/mouse sequence differences that have been observed to alter lipid IVA binding are more similar between the human and opossum than between the mouse and opossum. For example, position 58 is an Asn in mouse, but Lys and Arg in human and opossum. The fact that introducing the human states into the mouse pocket alters lipid IVA specificity contrasts with our observation that the human protein is the one that evolved lipid antagonism through as-yet-unknown mutations. This implies that there are important sequence determinants of lipid IVA antagonism in humans that have yet to be identified.

Finally, the evolution of incompatibility between the human and oTLR4/MD-2 complex shows a different mode of evolution. There is no obvious functional divergence between the LPS activation of human, mouse and oTLR4 (Fig. 4). Despite this, human MD-2 is compatible with mouse TLR4, but not oTLR4. In contrast, mouse MD-2 is compatible with both human and oTLR4. Cryptic changes have led to incompatibility and divergence within the conserved function. We were able to identify two human-specific mutations flanking a loop between TLR4 and MD-2 that, when reverted to their ancestral state, allowed full compatibility with the opossum protein [Fig. 5(b)]. This suggests that relatively few mutations can lead to incompatibility in these rapidly evolving proteins. Further work will be required to understand the nature of the changes on oTLR4 that mediate the incompatibility.

Our analysis focused on a relatively recent evolutionary transition within mammalian proteins. The activation of TLR4 by LPS has, however, been conserved for hundreds of millions of years. Studies of chicken TLR4 have revealed that it both activates with LPS and uses the same cofactors as its human ortholog.^{11,13} The gene itself is even more ancient. TLR4 exists in zebrafish—though, intriguingly, the essential cofactors MD-2 and CD14 have not been described outside of amniotes.^{35,36} It could be quite powerful to determine when the core functional features of the TLR4 complex evolved and how they have diversified in more distantly related mammals, tetrapods, and even bony fishes. With a longer evolutionary lens, looking beyond mice and men,³⁷ the molecular mechanisms for the core features of this complex and additional species-specific differences may be revealed.

Materials and Methods

Phylogenetics and ancestral state reconstruction

We constructed a curated database of TLR4 protein sequences using an NCBI BLAST with the human

protein sequence as a query. Additional amphibian sequences were identified in amphibian transcriptomes available on dryad. We constructed a multiple sequence alignment using MSAPROBS³⁸ followed by manual editing in MEGA.³⁹ Pilot analyses revealed that the JTT substitution model with eight rate categories and a floating gamma distribution parameter yielded the highest likelihood trees.^{29,40} We species-corrected these trees using Archaeopteryx to reflect published vertebrate phylogenies.⁴¹ We used PAML to reconstruct the sequence of ancTLR4.⁴² We assigned gaps manually using parsimony. Our alignment, species-corrected tree, and full posterior probabilities for each reconstructed site are available in the Supporting Information (Files S1–S3). We performed conservation calculations on our full alignment using consurf with the default parameters.³¹

Cell culture and transfection conditions

Human embryonic kidney cells (HEK293T/17, ATCC CRL-11268) were maintained up to 30 passages in Dulbecco's Modified Eagle Media (DMEM) supplemented with 10% fetal bovine serum (FBS) at 37°C with 5% CO₂. For each transfection, a confluent 100 mm plate of HEK293T/17 cells was treated at room temperature with 0.25% Trypsin–EDTA in Hank's Balanced Salt Solution and resuspended with an addition of DMEM +10% FBS. This was diluted fourfold into fresh medium and 135 µL aliquots of resuspended cells were transferred to a 96-well cell culture treated plate. Normal transfection mixes were made with 10 ng of TLR4, 1 ng of CD14, 0.5 ng of MD-2, 0.1 ng of Renilla, 20 ng of ELAM-Luc, and 68.4 ng pcDNA3.1(+) per well for a total of 100 ng of DNA and diluted in OptiMEM to a volume of 10 µL/well. For constitutive TLR4 experiments, these studies were performed in the absence of TLR4 cofactors MD-2 and CD14. Dilution series of TLR4 only transfections were made with 50, 25, 10, 5, and 2.5 ng of TLR4, no MD-2 or CD14, 0.1 ng of Renilla, 20 ng of ELAM-Luc, and a corresponding amount of pcDNA3.1(+) per well to total 100 ng of DNA. Dilution series of TLR4, MD-2, and CD14 were prepared for 50, 25, 10, 5, and 2.5 ng of TLR4 and the corresponding ratio of MD-2 and CD14 used above and 0.1 ng of Renilla, 20 ng of ELAM-Luc, and a corresponding amount of pcDNA3.1(+) per well to total 100 ng of DNA. Large batches of DNA mixes were prepared in advance and aliquots used within a week of mixing. To the DNA mix, 0.5 µL per well of PLUS reagent was added followed by a brief vortex and room temperature (RT) incubation for 10 min. Lipofectamine was diluted 0.5 µL into 9.5 µL OptiMEM per well. This was added to the DNA + PLUS mix, vortexed briefly and incubated at RT for 15 min. The transfection mix was diluted to 65 µL/well in OptiMEM and aliquoted onto a plate. Cells were incubated with transfection mix for 24 h and then treated with the buffer condition (25% phosphate buffered saline and 75% DMEM), LPS at 100 ng/well prepared in the same buffer condition, LPS dilutions indicated in figure legends, lipid IVA at 1 µg/well

concentrations, or mixtures of LPS and lipid IVa at concentrations of 100 ng/well and 1 μ g/well, respectively. *Escherichia coli* K-12 LPS (tlrl-eklps, Invivogen) was dissolved at 5 mg/mL in endotoxin-free water, aliquots were stored at -20°C . To avoid freeze–thaw cycles, working stocks of LPS were prepared at 10- μ g/mL and stored at 4°C . Cells were incubated with treatments for 4 h. The Dual-Glo Luciferase Assay System (Promega) was used to assay Firefly and Renilla luciferase activity of individual wells. Each NF- κ B induction value shown represents the average Firefly luciferase activity/Renilla luciferase activity from three biological replicates that are each the average of two technical replicates per 96-well plate.

Molecular dynamics simulations

We generated models of the MD-2/LPS complex using chain “C” taken from PDB 3FXI.⁴³ Amino acid replacements were made using PyMOL.⁴⁴ Molecular dynamics simulations were performed using the GROMACS 2016.4 package using the CHARMM27⁴⁵ force field with TIP3P water model.⁴⁶ The LPS topology and parameters were generated using the SwissParam server.⁴⁷ Calculations were done at 300 K and 1 bar. The MD-2/LPS model was placed in the center of a cubic water box 20 Å larger than the maximum protein dimension, followed by energy minimization. Velocities were assigned from a Maxwell distribution and the system was equilibrated for 1 ns with heavy protein atoms (except for mutated residues) fixed. This was followed by 300 ns of unrestrained MD. The initial 10 ns were excluded as burn-in. The trajectory time step was 2 fs, with frames recorded every 5 ps. Final analyses were performed on frames taken every 12.5 ps. All bonds were treated as constraints and fixed using LINCS.⁴⁸ Electrostatics were treated with the Particle Mesh Ewald model,⁴⁹ using an Fast Fourier Transform spacing of 12 Å, interpolation order of 4, tolerance of $1\text{e-}5$, and a Coulomb cutoff of 9 Å. van der Waals forces were treated with a simple cutoff at 9 Å. We used velocity rescaled temperature coupling with a τ of 0.1 ps and Berendsen pressure coupling with a τ of 0.5 ps and a compressibility of $4.5\text{e-}5\text{ bar}^{-1}$. Analyses were performed using VMD 1.9.3⁵⁰—with its built-in TCL scripting utility—as well as a set of in-house Python and R scripts.

Acknowledgments

We would like to thank members of the Harms lab for helpful conversations. This work was funded by T32GM007759 (ANL). M.J.H. is a Pew Scholar in the Biomedical Sciences, supported by The Pew Charitable Trusts.

References

1. Ghosh S, May MJ, Kopp EB (1998) NF- κ B and Rel proteins: evolutionarily conserved mediators of immune responses. *Annu Rev Immunol* 16:225–260.
2. Hayashi T, Nakamura T, Takaoka A (2011) Pattern recognition receptors. *Jpn J Clin Immunol* 34:329–345.

3. Lu Y-C, Yeh W-C, Ohashi PS (2008) LPS/TLR4 signal transduction pathway. *Cytokine* 42:145–151.
4. Saitoh S, Akashi S, Yamada T, Tanimura N, Matsumoto F, Fukase K, Kusumoto S, Kosugi A, Miyake K (2004) Ligand-dependent Toll-like receptor 4 (TLR4)-oligomerization is directly linked with TLR4-signaling. *J Endotoxin Res* 10: 257–260.
5. Sabroe I, Prince LR, Jones EC, Horsburgh MJ, Foster SJ, Vogel SN, Dower SK, Whyte MKB (2003) Selective roles for Toll-like receptor (TLR)2 and TLR4 in the regulation of neutrophil activation and life span. *J Immunol* 170: 5268–5275.
6. Clark SR, Ma AC, Tavener SA, McDonald B, Goodarzi Z, Kelly MM, Patel KD, Chakrabarti S, McAvoy E, Sinclair GD, Keys EM, Allen-Vercoe E, DeVinney R, Doig CJ, Green FHY, Kubes P (2007) Platelet TLR4 activates neutrophil extracellular traps to ensnare bacteria in septic blood. *Nat Med* 13:463–469.
7. Szajnik M, Szczepanski MJ, Czysowska M, Elishaev E, Mandapathil M, Nowak-Markwitz E, Spaczynski M, Whiteside TL (2009) TLR4 signaling induced by lipopolysaccharide or paclitaxel regulates tumor survival and chemoresistance in ovarian cancer. *Oncogene* 28:4353–4363.
8. Liu Y, Yin H, Zhao M, Lu Q (2014) TLR2 and TLR4 in autoimmune diseases: a comprehensive review. *Clinic Rev Allergy Immunol* 47:136–147.
9. Gallay P, Heumann D, Le Roy D, Barras C, Glauser MP, Mauri D, Burns K, Riederer BM, Akira S, Calandra T (1993) Lipopolysaccharide-binding protein as a major plasma protein responsible for endotoxemic shock. *Proc Natl Acad Sci U S A* 90:9935–9938.
10. Poltorak A (1998) Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* 282:2085–2088.
11. Keestra AM, van Putten JPM (2008) Unique properties of the chicken TLR4/MD-2 complex: selective lipopolysaccharide activation of the MyD88-dependent pathway. *J Immunol* 181:4354–4362.
12. Wu Z, Rothwell L, Hu T, Kaiser P (2009) Chicken CD14, unlike mammalian CD14, is trans-membrane rather than GPI-anchored. *Dev Comp Immunol* 33:97–104.
13. Loes AN, Bridgham JT, Harms MJ (2018) Coevolution of the toll-like receptor 4 complex with calgranulins and lipopolysaccharide. *Front Immunol* 9:304.
14. Lizundia R, Sauter K-S, Taylor G, Werling D (2008) Host species-specific usage of the TLR4-LPS receptor complex. *Innate Immun* 14:223–231.
15. Oblak A, Jerala R (2015) The molecular mechanism of species-specific recognition of lipopolysaccharides by the MD-2/TLR4 receptor complex. *Mol Immunol* 63:134–142.
16. Vasl J, Oblak A, Gioannini TL, Weiss JP, Jerala R (2009) Novel roles of lysines 122, 125, and 58 in functional differences between human and murine MD-2. *J Immunol* 183: 5138–5145.
17. Walsh C, Gangloff M, Monie T, Smyth T, Wei B, McKinley TJ, Maskell D, Gay N, Bryant C (2008) Elucidation of the MD-2/TLR4 interface required for signaling by lipid IVa. *J Immunol* 181:1245–1254.
18. Muroi M, Ohnishi T, Tanamoto K-I (2002) MD-2, a novel accessory molecule, is involved in species-specific actions of Salmonella lipid A. *Infect Immun* 70:3546–3550.
19. Akashi S, Nagai Y, Ogata H, Oikawa M, Fukase K, Kusumoto S, Kawasaki K, Nishijima M, Hayashi S, Kimoto M, Miyake K (2001) Human MD-2 confers on mouse Toll-like receptor 4 species-specific lipopolysaccharide recognition. *Int Immunol* 13:1595–1599.
20. Opal SM, Laterre P-F, Francois B, LaRosa SP, Angus DC, Mira J-P, Wittebole X, Dugernier T, Perrotin D, Tidswell M, Jauregui L, Krell K, Pacht J, Takahashi T, Peckelsen C,

- Cordasco E, Chang C-S, Oeyen S, Aikawa N, Maruyama T, Schein R, Kalil AC, Van Nuffelen M, Lynn M, Rossignol DP, Gogate J, Roberts MB, Wheeler JL, Vincent J-L (2013) Effect of eritoran, an antagonist of MD2-TLR4, on mortality in patients with severe sepsis: the ACCESS randomized trial. *J Am Med Assoc* 309:1154–1162.
21. Kawasaki K, Gomi K, Nishijima M (2001) Cutting edge: Gln22 of mouse MD-2 is essential for species-specific lipopolysaccharide mimetic action of taxol. *J Immunol* 166:11–14.
 22. Zimmer SM, Liu J, Clayton JL, Stephens DS, Snyder JP (2008) Paclitaxel binding to human and murine MD-2. *J Biol Chem* 283:27916–27926.
 23. Hedges SB, Dudley J, Kumar S (2006) TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* 22:2971–2972.
 24. Harms MJ, Thornton JW (2010) Analyzing protein structure and function using ancestral gene reconstruction. *Curr Opin Struct Biol* 20:360–366.
 25. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2009) GenBank. *Nucleic Acids Res* 37:D26–D31.
 26. Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Edgar R, Federhen S, Feolo M, Geer LY, Helmberg W, Kapustin Y, Landsman D, Lipman DJ, Madden TL, Maglott DR, Miller V, Mizrachi I, Ostell J, Pruitt KD, Schuler GD, Sequeira E, Sherry ST, Shumway M, Sirotkin K, Souvorov A, Starchenko G, Tatusova TA, Wagner L, Yaschenko E, Ye J (2009) Database resources of the National Center for biotechnology information. *Nucleic Acids Res* 37:D5–D15.
 27. Murphy WJ, Eizirik E, O'Brien SJ, Madsen O, Scally M, Douady CJ, Teeling E, Ryder OA, Stanhope MJ, de Jong WW, Springer MS (2001) Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* 294:2348–2351.
 28. Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR (2015) A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 526:569–573.
 29. Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J Mol Evol* 39:306–314.
 30. Park BS, Song DH, Kim HM, Choi B-S, Lee H, Lee J-O (2009) The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex. *Nature* 458:1191–1195.
 31. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N (2016) ConSurf2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 44:W344–W350.
 32. Anwar MA, Panneerselvam S, Shah M, Choi S (2015) Insights into the species-specific TLR4 signaling mechanism in response to *Rhodobacter sphaeroides* lipid A detection. *Sci Rep* 5:7657.
 33. Ohto U, Fukase K, Miyake K, Shimizu T (2012) Structural basis of species-specific endotoxin sensing by innate immune receptor TLR4/MD-2. *Proc Natl Acad Sci U S A* 109:7421–7426.
 34. Godfroy JJ, Roostan M, Moroz YS, Korendovych IV, Yin H (2012) Isolated Toll-like receptor transmembrane domains are capable of oligomerization. *PLoS One* 7:e48875.
 35. Sullivan C, Charette J, Catchen J, Lage CR, Giasson G, Postlethwait JH, Millard PJ, Kim CH (2009) The gene history of zebrafish *tlr4a* and *tlr4b* is predictive of their divergent functions. *J Immunol* 183:5896–5908.
 36. Sepulcre MP, Alcaraz-Pérez F, López-Muñoz A, Roca FJ, Meseguer J, Cayuela ML, Mulero V (2009) Evolution of lipopolysaccharide (LPS) recognition and signaling: fish TLR4 does not recognize LPS and negatively regulates NF-kappaB activation. *J Immunol* 182:1836–1845.
 37. Ernst PB, Carvunis A-R (2018) Of mice, men and immunity: a case for evolutionary systems biology. *Nat Immunol* 19:421–425.
 38. Liu Y, Schmidt B, Maskell DL (2010) MSAProbs: multiple sequence alignment based on pair hidden Markov models and partition function posterior probabilities. *Bioinformatics* 26:1958–1964.
 39. Tamura K, Stecher G, Peterson D, Filipowski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729.
 40. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 8:275–282.
 41. Han MV, Zmasek CM (2009) phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinform* 10:356.
 42. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591.
 43. Spoel DVD, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) GROMACS: fast, flexible, and free. *J Comput Chem* 26:1701–1718.
 44. Schrödinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.8.
 45. Foloppe N, MacKerell AD Jr (2000) All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* 21:86–104.
 46. Neria E, Fischer S, Karplus M (1996) Simulation of activation free energies in molecular systems. *J Chem Phys* 105:1902–1921.
 47. Zoete V, Cuendet MA, Grosdidier A, Michielin O (2011) SwissParam: a fast force field generation tool for small organic molecules. *J Comput Chem* 32:2359–2368.
 48. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: a linear constraint solver for molecular simulations. *J Comput Chem* 18:1463–1472.
 49. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092.
 50. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14:33–38.