

# A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin

Erica Andersen-Nissen,<sup>1,2</sup> Kelly D. Smith,<sup>1,3</sup> Richard Bonneau,<sup>4,5</sup>  
Roland K. Strong,<sup>2,6</sup> and Alan Aderem<sup>1</sup>

<sup>1</sup>Institute for Systems Biology, Seattle, WA 98103

<sup>2</sup>Department of Immunology and <sup>3</sup>Department of Pathology, University of Washington, Seattle, WA 98195

<sup>4</sup>Biology Department and <sup>5</sup>Courant Computer Science Department, New York University, New York, NY 10003

<sup>6</sup>Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109

**The molecular basis for Toll-like receptor (TLR) recognition of microbial ligands is unknown. We demonstrate that mouse and human TLR5 discriminate between different flagellins, and we use this difference to map the flagellin recognition site on TLR5 to 228 amino acids of the extracellular domain. Through molecular modeling of the TLR5 ectodomain, we identify two conserved surface-exposed regions. Mutagenesis studies demonstrate that naturally occurring amino acid variation in TLR5 residue 268 is responsible for human and mouse discrimination between flagellin molecules. Mutations within one conserved surface identify residues D295 and D367 as important for flagellin recognition. These studies localize flagellin recognition to a conserved surface on the modeled TLR5 structure, providing detailed analysis of the interaction of a TLR with its ligand. These findings suggest that ligand binding at the  $\beta$  sheets results in TLR activation and provide a new framework for understanding TLR-agonist interactions.**

## CORRESPONDENCE

Alan Aderem:  
aderem@systemsbiology.org

Abbreviations used: HA, hemagglutinin; LRR, leucine-rich repeat; TLR, Toll-like receptor.

Toll-like receptors (TLRs) constitute an important family of innate immune receptors that recognize pathogen-associated molecular patterns, evolutionarily conserved structures required for microbial fitness that are absent or underrepresented in the host (1, 2). We previously demonstrated that TLR5 recognizes bacterial flagellin (3), the most abundant protein found in the whip-like structures that propel bacteria (4). We defined flagellin as a prototypical pathogen-associated molecular pattern by identifying the 13 amino acids on flagellin that are involved in TLR5 recognition (5). The TLR5 recognition site on flagellin is located on the convex surface of the flagellin monomer that contacts adjacent monomers when they stack in the flagellar filament. TLR5 is activated by flagellins from a wide variety of bacteria, with the exception of two clades of flagellated bacteria possessing unusual flagellin sequences that permit TLR5 evasion while preserving bacterial motility (6).

In contrast to the well-defined TLR5 recognition site on flagellin, the complementary region on TLR5 that interacts with flagellin has not been established. TLRs are type I inte-

gral membrane glycoproteins with a series of 19–25 tandem leucine-rich repeat (LRR) motifs in the extracellular domain, each of which is 24–29 amino acids in length (7). Multiple LRR proteins have been crystallized and have been shown to form curved solenoid structures, but until recently no TLR structure had been solved. In 2005, the structure of the TLR3 extracellular domain was determined by two groups (8, 9) who demonstrated that the LRR of TLR3 also forms a curved solenoid. A recent study has suggested that the lateral, glycan-free surface of TLR3 recognizes double-stranded RNA (10).

Because TLR5 recognizes a protein agonist, a detailed molecular analysis of its interaction with flagellin is tractable. Two groups previously attempted to define this interaction but produced conflicting results, and neither group took into consideration the three-dimensional structure of TLR5 (11, 12). In the present study, we construct a model of TLR5 based on other LRR proteins. We exploit differences in flagellin recognition between human and mouse TLR5 to define a conserved surface on the concave  $\beta$  sheets of TLR5 that is responsible for flagellin recognition. These data provide detailed molecular analysis of a

The online version of this article contains supplemental material.

TLR–agonist interaction and suggest general rules for TLR/ligand recognition that may guide future studies in this important area.

**RESULTS**

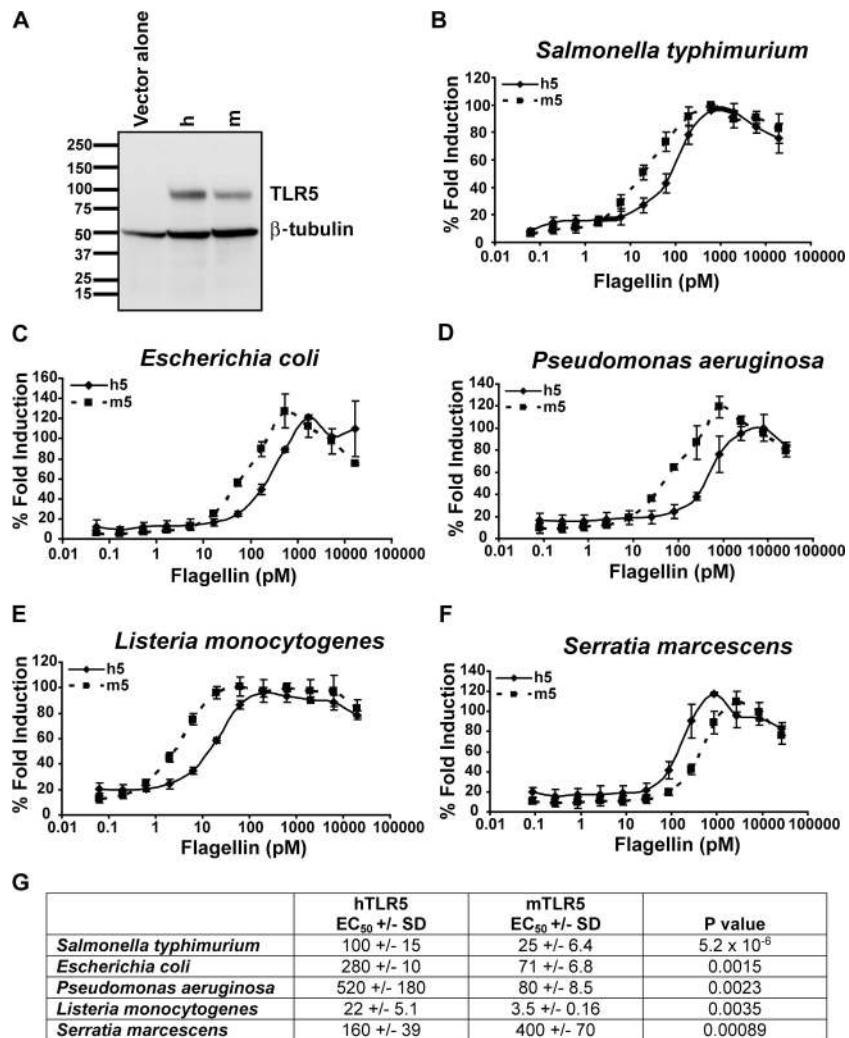
**Mouse TLR5 detects most flagellins better than human TLR5**

We expressed epitope-tagged human or mouse TLR5 and NF-κB luciferase reporter constructs in CHO-K1 cells, and similar TLR5 expression levels were detected in the two stably transfected populations (Fig. 1 A). We purified flagellin from *Salmonella typhimurium*, *Escherichia coli*, *Pseudomonas aeruginosa*, *Listeria monocytogenes*, and *Serratia marcescens* and tested the ability of each flagellin to stimulate either human or

mouse TLR5. Mouse TLR5 detected *S. typhimurium*, *E. coli*, *P. aeruginosa*, and *L. monocytogenes* flagellins at lower doses than human TLR5 (Fig. 1, B–E). These differences could not be explained by variability in receptor expression or signaling, as human TLR5 recognized *S. marcescens* flagellin better than mouse TLR5 (Fig. 1 F). The effective concentrations of flagellin that elicit 50% maximal response (EC<sub>50</sub>) for human and mouse TLR5 are listed in Fig. 1 G.

**The TLR5 extracellular domain is responsible for flagellin recognition**

In our previous study in which we identified the TLR5 recognition site on flagellin, we generated and characterized alanine point mutants of *Salmonella* flagellin (5). One mutant in



**Figure 1. Mouse TLR5 detects most flagellins better than human TLR5.** (A) Immunoblot of CHO cells stably expressing vector alone or human (h) or mouse (m) TLR5. 100 μg of cellular cytoplasmic extracts were loaded per lane, and TLR5 expression was detected by immunoblotting for the HA epitope tag. Equal loading was verified by immunoblotting for β-tubulin. Kilodalton values are shown. (B–F) Shown is the percent fold induction of NF-κB luciferase activity for flagellin purified

from each indicated bacterial species relative to maximal stimulation achieved with *S. typhimurium* flagellin for cells expressing either human (continuous line) or mouse (dashed line) TLR5 at the indicated flagellin doses. Data are representative of at least three independent experiments. Error bars represent the mean ± SD. (G) Table listing the effective flagellin concentrations needed to achieve half maximal activation of TLR5 (EC<sub>50</sub>). p-values were calculated using a two-tailed Student's t test.

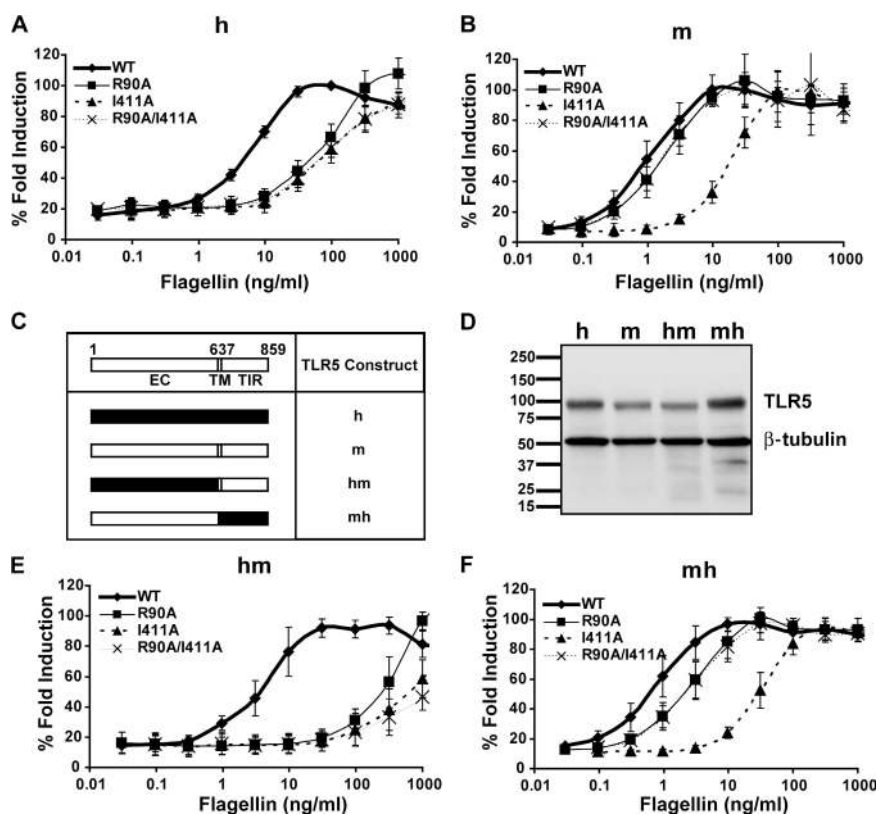
particular, I411A, showed the greatest reduction in recognition by human TLR5 (5). In comparison with wild-type flagellin, I411A showed an  $\sim 20$ -fold increase in  $EC_{50}$  for both human and mouse TLR5 (Fig. 2, A and B; and Table I). In contrast, mutation of an arginine residue centrally located within the TLR5 recognition site, R90A, affected recognition by human TLR5 (13-fold increase in  $EC_{50}$ ) but not by mouse TLR5 (Fig. 2, A and B; and Table I). Interestingly, combining both point mutations in flagellin, R90A/I411A, completely restored recognition by mouse TLR5 to wild-type *Salmonella* flagellin levels but failed to restore recognition by human TLR5 (16-fold reduction in  $EC_{50}$ ; Fig. 2, A and B; and Table I). Thus, mouse TLR5, like human TLR5 (5), discriminates between flagellin molecules that differ by only a single amino acid in the previously defined TLR5 recognition site. These flagellin point mutants provided a tool to investigate species-specific TLR5 recognition of bacterial flagellin.

To determine whether flagellin recognition was localized to the extracellular domain of TLR5, we constructed

epitope-tagged chimeric human and mouse TLR5 molecules by swapping the extracellular domains (Fig. 2 C). When stably transfected, both chimeric constructs were expressed in CHO cells at levels similar to those observed for the parental constructs (Fig. 2 D). We tested the chimeric TLR5 molecules for recognition of the wild-type form and point mutants of *Salmonella* flagellin and found that recognition was dictated by the extracellular domain (Fig. 2, E and F; and Table I).

### The central 228 amino acids of the TLR5 extracellular domain confer species-specific flagellin recognition

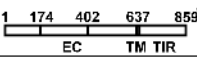
To identify the region of the TLR5 extracellular domain that confers species-specific recognition of flagellin, we aligned available TLR5 sequences (Fig. S1, available at <http://www.jem.org/cgi/content/full/jem.20061400/DC1>). TLR5 residues that vary among different species are scattered throughout the extracellular LRR domain, and we found no single region that was an obvious candidate for species-specific recognition. We therefore made a series of chimeric human



**Figure 2. The extracellular domain of TLR5 is responsible for flagellin recognition.** (A and B) Dose-response curves of CHO cells transfected with human (A) and mouse (B) TLR5 to wild-type flagellin from *S. typhimurium*, point mutants I411A and R90A, and the double mutant R90A/I411A. Fold induction is relative to the maximal induction seen for wild-type flagellin. (C) Table showing a linear schematic of TLR5, with mouse TLR5 amino acid numbers of the domain boundaries shown above the molecule. EC, extracellular domain; TM, transmembrane domain; TIR, Toll/IL-1 receptor domain. (D) Immunoblot of CHO cells stably expressing

human (h) and mouse (m) TLR5 or the extracellular domain TLR5 chimeras (hm and mh). 100  $\mu$ g of cellular cytoplasmic extracts were loaded per lane, and TLR5 expression was detected by immunoblotting for the HA epitope tag. Equal loading was verified by immunoblotting for  $\beta$ -tubulin. Kilodalton values are shown. (E and F) Dose-response curves of the hm (E) and mh (F) chimeric TLR5 receptors to wild-type flagellin from *S. typhimurium*, point mutants R90A and I411A, and the double mutant R90A/I411A.

**Table I.** EC<sub>50</sub> values for human and mouse chimeric TLR5 molecules stimulated with wild-type and mutant flagellins

TLR5	Flagellin								Flagellin Recognition Phenotype <sup>d</sup>
	WT	R90A		I411A		R90A/I411A			
 1 174 402 637 859 <sup>a</sup> EC TM TIR	EC <sub>50</sub> +/- s.d. <sup>b</sup>	EC <sub>50</sub> +/- s.d.	P value vs. WT <sup>c</sup>	EC <sub>50</sub> +/- s.d.	P value vs. WT	EC <sub>50</sub> +/- s.d.	P value vs. WT		
<b>h</b>	5.2 +/- 1.4	70 +/- 16	1.7 x 10 <sup>-5</sup>	90 +/- 18	4.7 x 10 <sup>-5</sup>	84 +/- 22	3.7 x 10 <sup>-5</sup>	<b>h</b>	
<b>m</b>	0.88 +/- 0.74	2.2 +/- 1.2		21 +/- 6.5	6.0 x 10 <sup>-4</sup>	1.8 +/- 0.77		<b>m</b>	
<b>hm</b>	5.2 +/- 2.7	270 +/- 160	0.02	620 +/- 210	0.001	820 +/- 360	0.004	<b>h</b>	
<b>mh</b>	1.05 +/- 0.52	3.3 +/- 1.4		25 +/- 11	0.009	3.3 +/- 1.7		<b>m</b>	
<b>hmh</b>	4.4 +/- 1.7	14 +/- 6.8		140 +/- 48	0.009	17 +/- 10		<b>m</b>	
<b>mhh</b>	19 +/- 7.5	1300 +/- 630	0.007	>10000	9.7 x 10 <sup>-20</sup>	>10000	9.7 x 10 <sup>-20</sup>	<b>h</b>	
<b>mmhm</b>	2.3 +/- 0.67	82 +/- 14	1.3 x 10 <sup>-6</sup>	610 +/- 260	0.001	110 +/- 27	2.5 x 10 <sup>-5</sup>	<b>m</b>	
<b>mhmm</b>	2.4 +/- 1.0	30 +/- 4.9	4.9 x 10 <sup>-5</sup>	230 +/- 160	0.006	200 +/- 110	0.002	<b>h</b>	
<b>m Q407P/M408D</b>	1.1 +/- 0.57	1.3 +/- 0.84		13 +/- 5.9	4.8 x 10 <sup>-4</sup>	1.1 +/- 0.91		<b>m</b>	
<b>m L298R/Q293H</b>	0.70 +/- 0.49	1.1 +/- 0.70		8.9 +/- 3.7	8.6 x 10 <sup>-5</sup>	1.1 +/- 0.87		<b>m</b>	
<b>m P268A</b>	0.91 +/- 0.63	5.8 +/- 2.8	0.002	14.3 +/- 5.0	7.2 x 10 <sup>-5</sup>	10 +/- 3.1	2.8 x 10 <sup>-5</sup>	<b>h</b>	
<b>h A268P</b>	1.8 +/- 0.44	58 +/- 15	1.7 x 10 <sup>-4</sup>	1030 +/- 75	0.03	95 +/- 50	0.007	<b>m</b>	

<sup>a</sup>Schematic of chimeric TLR5 molecules with amino acid numbers corresponding to mouse TLR5.

<sup>b</sup>EC<sub>50</sub> values and standard deviations are listed in ng/ml, and values were calculated from at least three independent experiments, each run in duplicate.

<sup>c</sup>EC<sub>50</sub> values for each flagellin point mutant were compared to wild-type flagellin, and p-values were calculated using a two-tailed Student's t test. p-values >0.05 are not shown.

<sup>d</sup>The pattern of response for each TLR5 chimera to the mutant flagellins was compared to parental human and mouse TLR5, and the overall flagellin recognition phenotype is listed.

and mouse TLR5 constructs and expressed these chimeras in CHO cells (Fig. 3, A and B).

We first tested the TLR5 chimeras containing the largest swap for recognition of the wild-type and mutated forms of *Salmonella* flagellin. The chimera containing central and membrane-proximal portions of the mouse TLR5 extracellular domain (hmh) showed a pattern of recognition similar to full-length mouse TLR5 (compare Fig. 3 C with Fig. 2 B). Conversely, the chimera containing the central and membrane-proximal portions of human TLR5 (mhh) showed a pattern of recognition similar to the parental human TLR5, with recognition of the double mutant resembling that of I411A (compare Fig. 3 D with Fig. 2 A). EC<sub>50</sub> values for each flagellin detected by the chimeras are listed in Table I.

To determine whether the central (residues 174–401) or membrane-proximal (residues 402–637) portion of the extracellular domain was responsible for species-specific flagellin recognition, we made additional chimeras in which these two domains were swapped. A TLR5 chimera containing the central portion of mouse TLR5 (mmhm) showed reduced recognition of all point mutants, but recognition of the double mutant was equivalent to R90A, a pattern resembling mouse TLR5 (compare Fig. 3 F with Fig. 2 B). In contrast, recognition of the flagellin point mutants by the chimera containing only the central portion of the human TLR5 extracellular domain (mhmm) closely resembled that of the parental full-length human TLR5 (compare Fig. 3 E with Fig. 2 A). EC<sub>50</sub> values for each flagellin detected by the chimeras are listed in Table I.

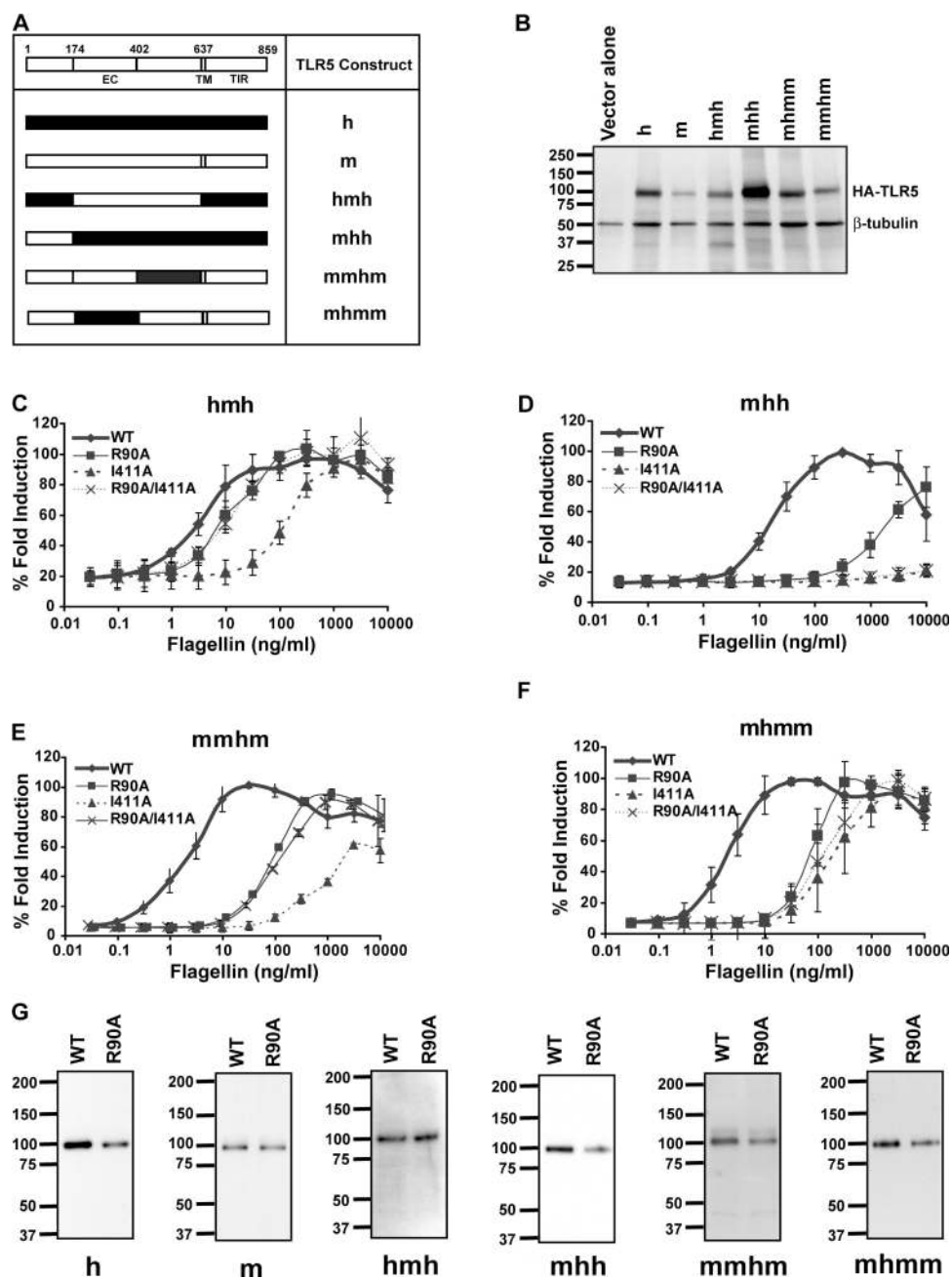
To test whether the decreased recognition of wild-type flagellin and the flagellin mutants reflected decreased associa-

tion of the TLR5 constructs with flagellin, we performed coprecipitations of the chimeric TLR5 receptors with wild-type flagellin or the R90A flagellin mutant. Wild-type flagellin precipitated human TLR5 more efficiently than the R90A flagellin mutant, whereas wild-type and R90A flagellins precipitated mouse TLR5 equivalently (Fig. 3 G). Thus, the pattern of coprecipitation paralleled the dose-response data, suggesting that the dose-response data reflect flagellin's association with TLR5. The pattern of coprecipitation also paralleled the dose-response curves for the chimeric receptors: receptors in which the EC<sub>50</sub> for the wild type and the R90A mutant were similar showed equivalent precipitation by both flagellins, whereas receptors with reduced recognition of R90A showed reduced precipitation by R90A relative to wild-type flagellin (Fig. 3 G). Control precipitation studies demonstrated that the interaction between flagellin and TLR5 was specific, because precipitations with wild-type flagellin failed to pull down TLR2 (5), and biotinylated ovalbumin did not precipitate TLR5 (not depicted).

Our chimera studies demonstrated that the central 228 amino acids of the TLR5 extracellular domain (residues 174–401) were responsible for species-specific flagellin recognition. We next sought to identify specific amino acids responsible for human and mouse TLR5 discrimination of flagellin molecules.

### Modeling of TLR5

Because the crystal structure of TLR5 has not yet been solved, we modeled the structure of the mouse and human TLR5 ectodomains based on known structures of other LRR proteins. Using consensus fold recognition, the best



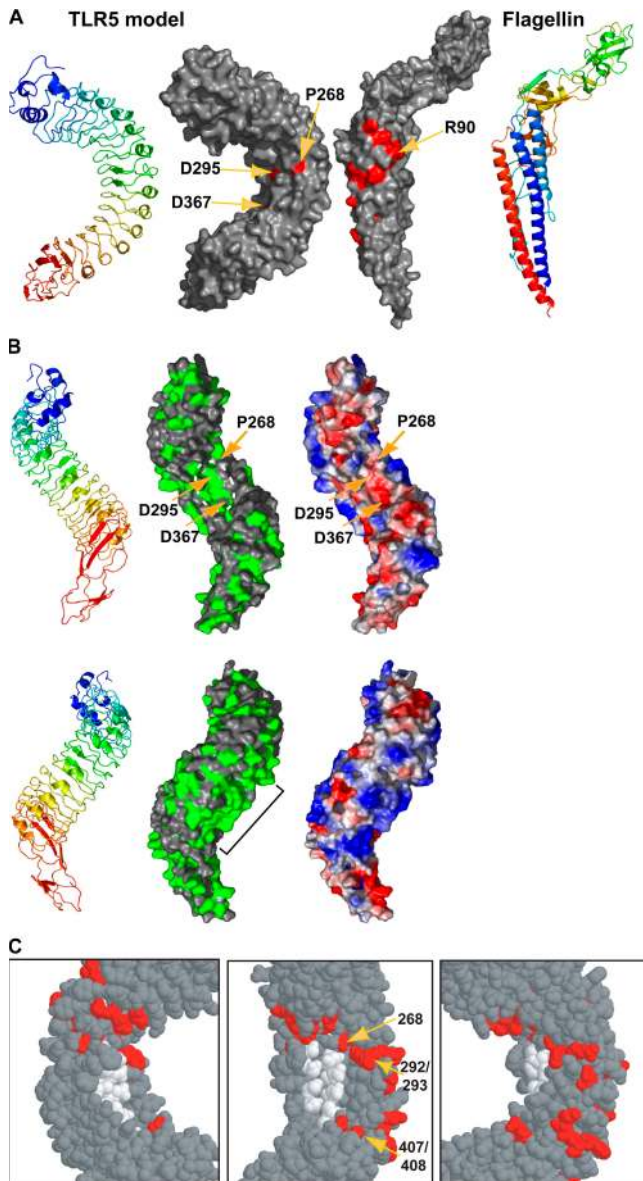
**Figure 3. The central 228 amino acids of the TLR5 extracellular domain confer species-specific flagellin recognition.** (A) Table showing a linear schematic of TLR5, with amino acid numbers of the domain boundaries shown above the molecule. EC, extracellular domain; TM, transmembrane domain; TIR, Toll/IL-1 receptor domain. (B) Immunoblot of CHO cells stably expressing human (h) and mouse (m) TLR5 or the extracellular domain TLR5 chimeras (hmh, mhh, mhmm, and mmhm). 100  $\mu$ g of cellular cytoplasmic extracts were loaded per lane, and TLR5 expression

was detected by immunoblotting for the HA epitope tag. Equivalent loading was verified by immunoblotting for  $\beta$ -tubulin. Kilodalton values are shown. (C–F) Dose-response curves of the hmh (C), mhh (D), mmhm (E), and mhmm (F) chimeric TLR5 receptors to wild-type flagellin from *S. typhimurium*, point mutants R90A and I411A, and the double mutant R90A/I411A. The percent fold induction is relative to the maximal induction seen for wild-type flagellin. (G) Immunoprecipitation of chimeric TLR5 molecules with the wild type or the R90A flagellin mutant.

match was determined to be *L. monocytogenes* internalin A (PDB 1o6s). Modeling of LRR-containing proteins is inherently difficult (13), with the resulting overall global structure varying considerably with even small changes in individual LRR or with relative orientations of adjacent LRR. We

evaluated several models (generated by varying method parameters), which resulted in solenoids of different global structure but largely uniform local structure. The model of the mouse TLR5 ectodomain (amino acids 52–615) that we present (Fig. 4 A, left) conforms well to ideal peptide





**Figure 4. Model of the TLR5 extracellular domain.** (A) Ribbon representations (far left and far right) and molecular surface representations (middle left and middle right) of the best model of the TLR5 ectodomain (left) and the structure of flagellin (PDB 1luc; right). The ribbon is colored sequentially from the N terminus (blue) to the C terminus (red). Amino acids important for flagellin recognition are shown on the TLR5 model in red. Amino acids on flagellin previously determined to be important for TLR5 recognition are shown in red on the flagellin structure (reference 5). (B) Ribbon representations (left) and molecular surface representations (middle and right) of the best model of the TLR5 ectodomain, oriented  $\sim 90^\circ$  to the view in A. Views  $180^\circ$  apart are shown in the top and bottom rows. The ribbon is colored sequentially from the N terminus (blue) to the C terminus. Molecular surface representations are colored by conservation (center: residues  $\geq 90\%$  similar among TLR5 sequences are green) or by electrostatic potential (right: positive, blue; negative, red). The conserved concavity and lateral patch regions are indicated with a dotted white oval and bracket, respectively. Positions of residues important for flagellin recognition are indicated with arrows. (C) Space-filling representation of the TLR5 model. Residues of the conserved concavity (white)

and protein geometry while also generating an overall fold suited to binding large macromolecular ligands such as flagellin (Fig. 4 A, right). The model was further validated by its similarity to the recently deposited structure of the extracellular domain of TLR3 (PDB 1ziw). The real value of the TLR5 model was not in its high resolution atomic details but in guiding selection of residues for further study based on proximity to conserved surface residues in the TLR5 model. We identified amino acids that were similar among at least 90% of vertebrate TLR5 sequences and highlighted their positions on our structural model (Fig. 4 B, green). This analysis revealed two regions that were conserved among all species: an apparent concavity located on the inner surface of the solenoid structure predicted for the TLR5 ectodomain and a patch located on the lateral face of the molecule (Fig. 4 B). Both the concavity and the lateral patch were contained within the central 228–amino acid region that confers species-specific flagellin recognition.

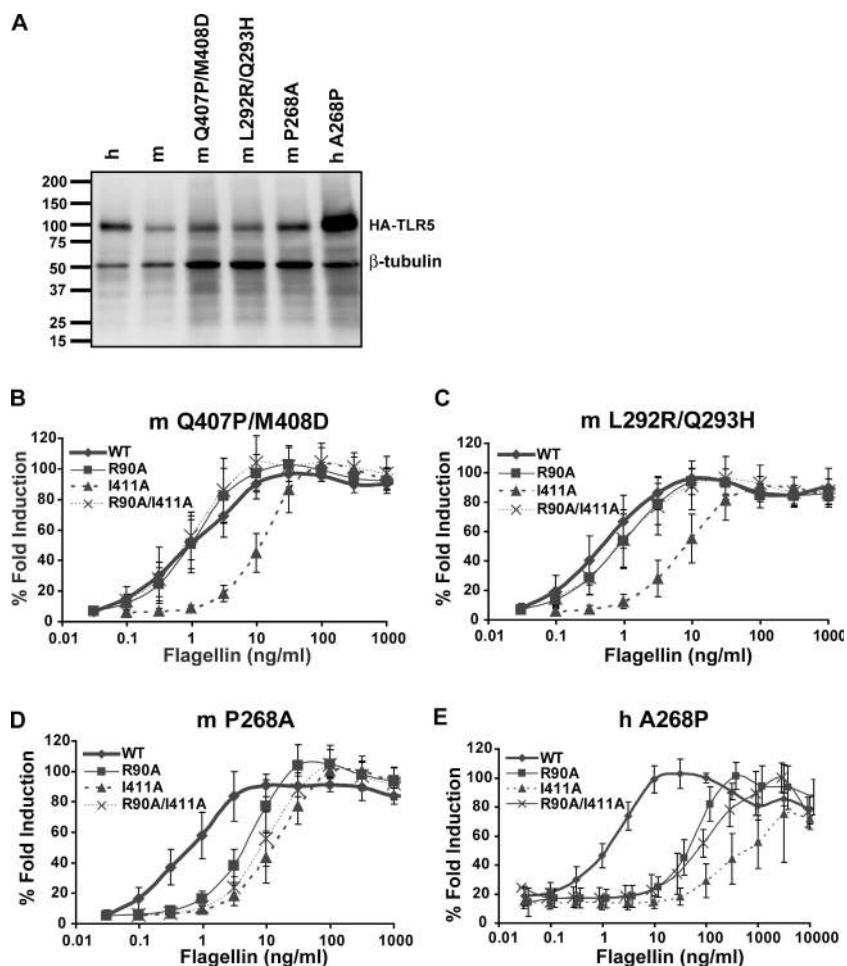
#### TLR5 residue 268 imparts species-specific recognition of flagellin point mutants

We hypothesized that either the conserved concavity or the lateral patch was the flagellin recognition site on TLR5, and that species-specific TLR5 sequence differences adjacent to one of these regions were responsible for differences in flagellin recognition by human and mouse TLR5. Several amino acids that differed between mouse and human TLR5 clustered around the concavity and were good candidates for amino acid changes that mediate species-specific TLR5 recognition (Fig. 4 C). We mutated mouse TLR5 residues in this region to the corresponding human TLR5 amino acids at three sites surrounding the concavity: P268A and L292R/Q293H (both located within the central 228-residue region) and Q407P/M408D (located in the membrane-proximal region).

The mutant mouse TLR5 molecules were expressed in CHO cells at levels similar to parental human and mouse TLR5 (Fig. 5 A). Mutating residues 407/408, which are located in the membrane-proximal region that is not implicated in species-specific flagellin recognition, did not affect the pattern of flagellin recognition (compare Fig. 5 B with Fig. 2 B; Table I). In addition, mutating central region residues 292/293 to the human sequence did not affect recognition of the flagellin point mutants (compare Fig. 5 C with Fig. 2 B; Table I). In contrast, mutation of a single central region amino acid, P268A, transformed the specificity of mouse TLR5 flagellin recognition to the pattern seen for human TLR5 (compare Fig. 5 D with Fig. 2 A; Table I).

To confirm this result, we performed the converse mutation, A268P, in human TLR5. This mutation resulted in a receptor with reduced recognition of all flagellin molecules

and residues that differ between human and mouse TLR5 in the central 228–amino acid region, as well as 407 and 408 (red), are highlighted. Amino acids surrounding the concavity that were mutated are indicated with arrows.



**Figure 5. P268 confers species-specific recognition of flagellin point mutants.** (A) Immunoblot of CHO cells stably expressing human (h) and mouse (m) TLR5 or the extracellular domain TLR5 mutants (mQ407P/M408D, mL292R/Q293H, mP268A, or hA268P). 100  $\mu$ g of cellular cytoplasmic extracts were loaded per lane, and TLR5 expression was detected by immunoblotting for the HA epitope tag. Equal loading was verified by

immunoblotting for  $\beta$ -tubulin. Kilodalton values are shown. (B–E) Dose-response curves of the mQ407P/M408D (B), mL292R/Q293H (C), mP268A (D), and hA268P (E) mutant TLR5 receptors to wild-type flagellin from *S. typhimurium*, point mutants R90A and I411A, and the double mutant R90A/I411A.

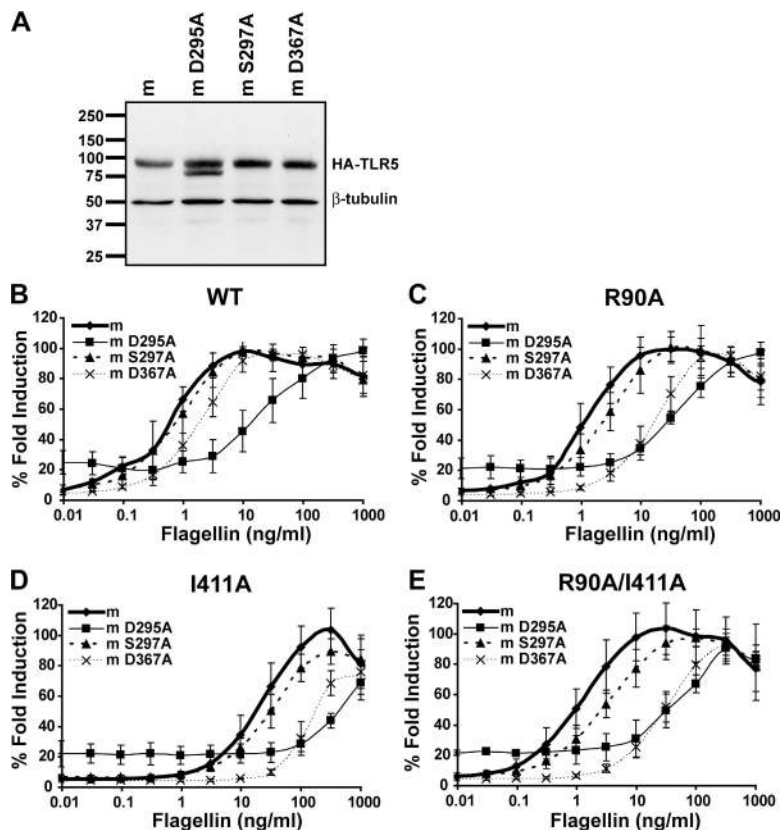
(Fig. 5 E). Notably, the human TLR5 A268P receptor recognized the R90A/I411A double mutant as efficiently as the R90A single mutant and better than the I411A point mutant (Fig. 5 E).

These results suggest that residue 268 determines, at least in part, species-specific flagellin recognition by TLR5. P268 is adjacent to the conserved concavity in our TLR5 model (Fig. 4 C), which suggested that this surface may interact with the TLR5 recognition site on flagellin (5, 6).

#### TLR5 residues within the conserved concavity contribute to flagellin recognition

The concavity consists of 11 conserved and surface-exposed amino acids in our model (Fig. S1). We selected three of these conserved residues and mutated each one to alanine to determine their role in flagellin recognition. The three TLR5 alanine mutants were stably expressed in CHO cells (Fig. 6 A).

D295A showed a second, sharper band of lower molecular mass in addition to a band of equal molecular mass to the other TLR5 proteins, suggesting that the D295A mutation may alter glycosylation of this protein, as has been described for other TLRs (14–17). We tested the TLR5 alanine point mutants for recognition of wild-type *Salmonella* flagellin and the flagellin mutants. Mutating TLR5 residue S297 to alanine had no effect on flagellin recognition, except that it slightly raised the  $EC_{50}$  for the R90A/I411A flagellin double mutant relative to parental mouse TLR5 (Fig. 6, B–E; and Table II). Mutating TLR5 residue D295 increased the  $EC_{50}$  for all flagellins tested relative to parental mouse TLR5, suggesting that this residue is important for flagellin recognition (Fig. 6, B–E; and Table II). In contrast, mutating TLR5 residue D367 to alanine only slightly reduced recognition of wild-type flagellin but greatly decreased recognition of the R90A, I411A, and R90A/I411A flagellin mutants. Thus, mutating TLR5 residue



**Figure 6. The conserved concavity on TLR5 interacts with flagellin.** (A) Immunoblot of CHO cells stably expressing mouse TLR5 (m) or the TLR5 concavity point mutants (mD295A, mS297A, and mD367A). 100  $\mu$ g of cellular cytoplasmic extracts were loaded per lane, and TLR5 expression was detected by immunoblotting for the HA epitope tag. Equal loading

was verified by immunoblotting for  $\beta$ -tubulin. Kilodalton values are shown. (B–E) Dose-response curves for wild-type flagellin (B) from *S. typhimurium* and flagellin point mutants R90A (C), I411A (D), and R90A/I411A (E) for each TLR5 point mutant.

367 specifically weakens the recognition of flagellins mutated in residues R90 and I411 as compared with wild-type flagellin; such nonadditive effects further argue for a direct molecular interaction between TLR5 and flagellin (Fig. 6, B–E; and Table II). Because flagellin residue I411 is largely buried beneath the TLR5 recognition site, the contribution of I411 is most likely indirect and may be transmitted through R90 (6). The locations of TLR5 residues D295 and D367, which are predicted to interact with flagellin, are shown on our model (Fig. 4).

## DISCUSSION

Species-specific differences in TLR agonist recognition have previously been reported for other TLRs. Human but not mouse TLR2 can discriminate between tripalmitoylated and trioleoylated peptides (18). The lipid A analogues, lipid IVa and *Rhodobacter sphaeroides* lipid A, are potent LPS antagonists to human TLR4 yet act as agonists for hamster TLR4 (19, 20). Human TLR4 also distinguishes between *P. aeruginosa* penta- and hexaacylated lipid A structures, whereas mouse TLR4 cannot (21). Finally, mouse and human TLR9 have been shown to recognize distinct synthetic CpG motifs (22). Our findings for TLR5 parallel these previous studies, but the

biological significance of differential flagellin recognition by mouse and human TLR5 remains unknown. Species preferences for certain flagellin molecules by TLRs may have arisen through the long evolutionary history of host–pathogen interactions (23) and may reflect host adaptation to pathogenic and commensal bacteria, as well as selective pressures on bacteria to establish niches within hosts.

In this study, we identify three TLR5 residues, P268, D295, and D367, that are important for flagellin recognition and thus provide further evidence for a direct interaction between flagellin and TLR5. Because proline residues often affect the local structure of the polypeptide chain, the P268A mutation may influence flagellin recognition indirectly. TLR5 residues D295/D367 and flagellin R90, however, are surface exposed and may form salt bridges or other polar interactions. This possibility is supported by the fact that polar–polar and charge–polar interactions account for the majority of molecular contacts between adjacent monomers in the flagellar filament, and a subset of these residues also forms the TLR5 recognition site on flagellin (24).

Two previous reports have addressed the region of TLR5 that mediates flagellin recognition. One study, comprised



**Table II.** EC<sub>50</sub> values for alanine point-mutant mouse chimeric TLR5 molecules stimulated with wild-type and mutant flagellins

TLR5	Flagellin							
	WT		R90A		I411A		R90A/I411A	
	EC <sub>50</sub> +/- s.d. <sup>b</sup>	P value vs. WT <sup>c</sup>	EC <sub>50</sub> +/- s.d.	P value vs. WT	EC <sub>50</sub> +/- s.d.	P value vs. WT	EC <sub>50</sub> +/- s.d.	P value vs. WT
<b>m</b>	0.57 +/- 0.29		1.5 +/- 1.1		23 +/- 10		1.3 +/- 0.60	
<b>m D295A</b>	15 +/- 5.7	8.3 x 10 <sup>-5</sup>	43 +/- 16	6.6 x 10 <sup>-5</sup>	700 +/- 260	9.5 x 10 <sup>-5</sup>	64 +/- 29	0.01
<b>m S297A</b>	0.94 +/- 0.52		2.8 +/- 1.3		31 +/- 8		3.2 +/- 1.1	0.05
<b>m D367A</b>	2.2 +/- 0.89	0.004	17 +/- 7.1	0.0006	180 +/- 73	0.0006	37 +/- 15	0.01

<sup>a</sup>Schematic of alanine point-mutant mouse TLR5 molecules.

<sup>b</sup>EC<sub>50</sub> values and standard deviations are listed in ng/ml, and values were calculated from at least three independent experiments, each run in duplicate.

<sup>c</sup>EC<sub>50</sub> values for each TLR5 construct were compared to parental TLR5, and p-values were calculated using a two-tailed Student's *t* test. p-values >0.05 are not shown.

solely of theoretical modeling, hypothesized that human TLR5 residues 552–560 were responsible for flagellin recognition (11). This region is outside of the central 228 amino acids that contain the conserved concavity and patch, and there is no experimental data to support the 552–560 region in flagellin recognition. In addition, an examination of the TLR5 sequence alignment reveals that five amino acids in this nine amino acid stretch vary among TLR5 sequences, implying that only a subset of these amino acids could have been involved in the interaction with flagellin. A second report used coimmunoprecipitation studies to conclude that human TLR5 residues 386–407 (mouse TLR5 386–406 in our sequence; Fig. S1) were responsible for flagellin binding (12). Residues 386–401 of this block are located within the central 228 stretch; however, none of these amino acids contribute to the conserved concavity and lateral patch and only 9 out of 16 amino acids in this block are >90% conserved among vertebrates. In addition to the lack of sequence conservation in the regions defined by these two studies, the possibility that a linear stretch of amino acids (composing at most a single LRR) could be solely responsible for flagellin recognition is not well supported by our structural model and the recent TLR3 structures (8, 9). Our modeling and mutagenesis studies of TLR5 have led to the hypothesis that the highly conserved concavity on TLR5 formed by  $\beta$  sheets on one face of the LRR structure serves as the flagellin recognition site, which is distinct from the 386–407 and 552–561 regions previously implicated in flagellin recognition.

Our model of TLR5 resembles the recently determined TLR3 structure (8, 9), with some strong similarities of particular note. The conserved lateral patch we identified on TLR5 is located on the same lateral face of the molecule as the nonglycosylated face of TLR3. This surface on TLR3 offers the largest area for intermolecular interactions and contains the site of the dimer interface observed in the TLR3 crystals (8, 9). We therefore hypothesize that the conserved lateral patch on TLR5 is involved in TLR5 dimerization. Several candidate binding sites for double-stranded RNA were identified on the TLR3 structure, one being the concave surface of the solenoid (9) that is analogous to the conserved concavity on TLR5. Indeed, the concave surface of the solenoid is a common site of ligand binding for multiple

LRR proteins, including molecules such as ribonuclease inhibitor (25), internalin A (26), and platelet glycoprotein Ib $\alpha$  (27). A recent mutational analysis of TLR3 reported that double-stranded RNA binds TLR3 on the lateral glycan-free surface of the molecule (10). This study used size-exclusion chromatography to demonstrate that two TLR3 point mutants no longer formed high molecular mass aggregates with polyI:C. These TLR3 amino acid residues may influence ligand-dependent oligomerization either through direct interactions with polyI:C or by influencing receptor–receptor interactions.

The similarities between our data for TLR5 recognition of flagellin and structural models of TLR3 recognition of double-stranded RNA (9) suggest a common mechanism for agonist recognition by TLRs, where the TLR agonist binds to the concave surface of the solenoid and thereby influences receptor dimerization and/or signaling. Flagellin binding to the TLR5 concavity could induce receptor signaling through several mechanisms. TLR5 may be expressed as a dimer, as suggested by the crystal structure of TLR3 (8, 9), and binding of flagellin may induce a conformational change in the receptor, resulting in productive signaling. Alternatively, binding of flagellin to a single TLR5 molecule may induce a conformational change that allows dimerization with other flagellin-bound TLR5 molecules, as occurs with the epidermal growth factor receptor (28). Yet another possibility is that one flagellin monomer may bind two TLR5 molecules, which would require TLR5 to recognize two distinct sites on flagellin, as occurs with binding of human growth hormone to its receptor (29). These questions will only be resolved through more detailed biochemical and structural analyses of TLR/agonist recognition. Our current results define the likely TLR5–flagellin interaction interface and suggest general rules for TLR–ligand interactions, providing an important step toward the rational design of therapeutic interventions.

## MATERIALS AND METHODS

**Cell lines and bacterial strains.** CHO-K1 cells (American Type Culture Collection) were grown in Ham's F-12 medium supplemented with penicillin, streptomycin, L-glutamine, and 10% cosmic calf serum (Hy-Clone). The following bacteria were grown overnight, shaking in Luria-Bertani: *S. typhimurium* strain TH4778 (FljB<sup>-</sup>/FljC<sup>+</sup>; a gift from K. Hughes,

University of Washington, Seattle, WA), *E. coli*, clinical isolate H9049 (University of Washington), *P. aeruginosa* strain PAK (a gift from D. Speert, University of British Columbia, Vancouver, Canada), and *L. monocytogenes* strain 10403 (a gift from D. Portnoy, University of California, San Francisco, San Francisco, CA). *S. marcescens* (clinical isolate; University of Washington, Seattle, WA) was grown on Luria-Bertani agar plates.

**NF- $\kappa$ B luciferase reporter assays.** CHO-K1 cells were electroporated with hemagglutinin (HA)-tagged TLR5 cDNA cloned into the pEFIN vector (the pEF6 V5/His TOPO vector [Invitrogen] with an IRES-Neo fragment cloned into the NotI-XbaI sites) and ELAM-LUC (30) plasmids, and selected with neomycin. Stable populations were stimulated for 4 h and assayed for luciferase activity using an assay system (Luciferase 1000; Promega). All assays were done in duplicate, and each experiment was repeated at least three times. The percent fold induction was calculated by dividing the luciferase values for the test flagellins by the luciferase value for the wild-type *Salmonella* flagellin.

#### CHO-K1 cell cytoplasmic extracts and TLR5 immunoblots.

Approximately  $10^7$  CHO cells expressing either pEFIN alone or pEFIN TLR5 constructs were lifted with PBS containing 1mM EDTA, 1mM dithiothreitol, and protease inhibitors (cocktail set 1; Calbiochem). Cytoplasmic extracts were prepared by washing cells with hypotonic buffer solution (10 mM Hepes, pH 7.9, 1.5 mM MgCl<sub>2</sub>, 10 mM KCl, 1 mM dithiothreitol, and protease inhibitors) and resuspending in hypotonic buffer solution for 10 min on ice. NP-40 was added to a final concentration of 0.2%, and nuclei were removed by centrifuging at 3,300 g for 15 min. Protein concentrations were determined using the DC Protein Assay (Bio-Rad Laboratories), and 100  $\mu$ g of each cytoplasmic extract was run on an SDS-PAGE gel. HA-tagged TLR5 proteins were detected by immunoblotting with mouse anti-HA1.1 ascites (Covance) and horseradish peroxidase-conjugated rabbit anti-mouse (Zymed Laboratories). Loading controls for each immunoblot were detected using mouse anti- $\beta$  tubulin (Sigma-Aldrich) and horseradish peroxidase-conjugated rabbit anti-mouse (Zymed Laboratories).

**Purification of native bacterial flagellin.** *S. typhimurium*, *E. coli*, *P. aeruginosa*, *L. monocytogenes*, and *S. marcescens* were grown overnight and pelleted by centrifugation. Cell pellets were washed once in PBS, resuspended in PBS, and sheared for 2 min at high speed in a Waring blender. The sheared suspension was centrifuged for 10 min at 8,000 g, and the supernatant was collected and centrifuged at 100,000 g for 1 h to pellet flagellar filaments. Filaments were resuspended in PBS at 4°C overnight and centrifuged at 100,000 g for 1 h. This wash step was repeated twice. The resulting pellet of washed flagellar filaments was resuspended in PBS and heated to 70°C for 20 min to depolymerize filaments into flagellin monomers. Protein concentration was determined using the BCA assay (Pierce Chemical Co.), and purity was confirmed by SDS-PAGE and Coomassie blue staining.

**Creation and purification of alanine point mutant flagellins.** The *fliC* gene, which encodes flagellin, was cloned into the *NcoI* and *HindIII* sites of *ptc99a* plasmid. Single amino acid mutations were made using a standard PCR mutagenesis strategy (31). All mutations were verified by DNA sequencing. The mutant plasmids were transformed into the BC696 (*fliB*<sup>-</sup>/*fliC*<sup>-</sup>) strain of *S. typhimurium* SL1344 (5), and flagellin was purified as for the native flagellin described in the previous section, without the final wash steps.

**Coprecipitation studies.** Wild-type or R90A flagellins were biotinylated with EZ-Link Sulfo-NHS-LC-Biotin (Pierce Chemical Co.) and dialyzed against PBS. CHO cells stably expressing HA-tagged TLR5 constructs were lysed in 20 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, and 0.5% NP-40, pH 7.5. Nuclei were cleared by 5 min of centrifugation at 3,000 g. Cleared lysate was incubated with 10  $\mu$ g/ml biotinylated flagellin for 30 min at 4°C, followed by incubation with streptavidin sepharose (GE Healthcare) for 30 min at 4°C. Beads were washed with PBS, and proteins were eluted by boiling in SDS-PAGE loading buffer. Equivalent cell portions of the immuno-

precipitation were separated by SDS-PAGE, and TLR5 was immunoblotted using mouse anti-HA1.1, as described in CHO-K1 cell cytoplasmic extracts and TLR5 immunoblots.

**TLR5 sequence alignments.** TLR5 sequences from the following species were used in the TLR5 sequence alignment and are available from GenBank/EMBL/DBJ under the indicated accession numbers: *Rattus norvegicus* (rat, XP\_223016), *Mus musculus* (mouse, Q9JLF7), *Homo sapiens* (human, NP\_003259), *Canis familiaris* (dog [23]), *Monodelphis domestica* (opossum [23]), *Xenopus laevis* (African clawed frog, AAH84773), *Takifugu rubripes* (pufferfish, AAW69374), *Oncorhynchus mykiss* (trout, BAD38860), *Gallus gallus* (chicken, AF186107), and *Sus scrofa* (pig, BAD91800). TLR5 sequences were aligned using ClustalW (<http://www.ch.embnet.org/software/ClustalW.html>) and displayed with Boxshade ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)).

**Modeling of TLR5.** Models for TLR5 were generated using consensus fold recognition. Alignments were detected to multiple proteins using three-dimensional jury (which found good hits via ORFeus and FFAS) (32, 33). These alignments to multiple proteins were used to make a complete model using Rosetta (34, 35) via the Robetta server (<http://rosetta.bakerlab.org/>). Locally, this model is of high accuracy, and error in this model is likely concentrated near the N and C termini of the protein, where errors in the modeling of long loops can cause register shifts with respect to the repeat structure. Models were analyzed and figures were prepared with PyMOL (DeLano Scientific LLC) and Protein Explorer (<http://www.umass.edu/microbio/chime/pe/protexpl/fmtdoor.htm>).

**Online supplemental material.** Fig. S1 shows a sequence alignment of the central 228 amino acids of TLR5's extracellular domain determined to be important for flagellin recognition. Online supplemental material is available at <http://www.jem.org/cgi/content/full/jem.20061400/DC1>.

We thank Katie Strobe for help in generating plasmid constructs, Mary Brunkow and Colleen Sheridan for critical reading of the manuscript, and Jared Roach for assistance with TLR5 sequences.

This work was supported by a Cancer Research Institute predoctoral training grant (to E. Andersen-Nissen) and by National Institutes of Health grants R01 AI062859 (to K.D. Smith), R01 AI48675 (to R.K. Strong), R01 AI52286 (to A. Aderem), and U54 AI54523 (A. Aderem).

The authors have no conflicting financial interests.

Submitted: 30 June 2006

Accepted: 11 January 2007

#### REFERENCES

- Janeway, C.A., Jr., and R. Medzhitov. 2002. Innate immune recognition. *Annu. Rev. Immunol.* 20:197–216.
- Medzhitov, R., and C.A. Janeway Jr. 1997. Innate immunity: the virtues of a nonclonal system of recognition. *Cell.* 91:295–298.
- Hayashi, F., K.D. Smith, A. Ozinsky, T.R. Hawn, E.C. Yi, D.R. Goodlett, J.K. Eng, S. Akira, D.M. Underhill, and A. Aderem. 2001. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature.* 410:1099–1103.
- Harshey, R.M., and A. Toguchi. 1996. Spinning tails: homologies among bacterial flagellar systems. *Trends Microbiol.* 4:226–231.
- Smith, K.D., E. Andersen-Nissen, F. Hayashi, K. Strobe, M.A. Bergman, S.L. Barrett, B.T. Cookson, and A. Aderem. 2003. Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility. *Nat. Immunol.* 4:1247–1253.
- Andersen-Nissen, E., K.D. Smith, K.L. Strobe, S.L. Barrett, B.T. Cookson, S.M. Logan, and A. Aderem. 2005. Evasion of Toll-like receptor 5 by flagellated bacteria. *Proc. Natl. Acad. Sci. USA.* 102:9247–9252.
- Akira, S., S. Uematsu, and O. Takeuchi. 2006. Pathogen recognition and innate immunity. *Cell.* 124:783–801.
- Choe, J., M.S. Kelker, and I.A. Wilson. 2005. Crystal structure of human toll-like receptor 3 (TLR3) ectodomain. *Science.* 309:581–585.

9. Bell, J.K., I. Botos, P.R. Hall, J. Askins, J. Shiloach, D.M. Segal, and D.R. Davies. 2005. The molecular structure of the Toll-like receptor 3 ligand-binding domain. *Proc. Natl. Acad. Sci. USA*. 102:10976–10980.
10. Bell, J.K., J. Askins, P.R. Hall, D.R. Davies, and D.M. Segal. 2006. The dsRNA binding site of human Toll-like receptor 3. *Proc. Natl. Acad. Sci. USA*. 103:8792–8797.
11. Jacchieri, S.G., R. Torquato, and R.R. Brentani. 2003. Structural study of binding of flagellin by Toll-like receptor 5. *J. Bacteriol.* 185:4243–4247.
12. Mizel, S.B., A.P. West, and R.R. Hantgan. 2003. Identification of a sequence in human toll-like receptor 5 required for the binding of Gram-negative flagellin. *J. Biol. Chem.* 278:23624–23629.
13. Kajava, A.V., and B. Kobe. 2002. Assessment of the ability to model proteins with leucine-rich repeats in light of the latest structural information. *Protein Sci.* 11:1082–1090.
14. da Silva Correia, J., and R.J. Ulevitch. 2002. MD-2 and TLR4 N-linked glycosylations are important for a functional lipopolysaccharide receptor. *J. Biol. Chem.* 277:1845–1854.
15. Weber, A.N., M.A. Morse, and N.J. Gay. 2004. Four N-linked glycosylation sites in human toll-like receptor 2 cooperate to direct efficient biosynthesis and secretion. *J. Biol. Chem.* 279:34589–34594.
16. de Bouteiller, O., E. Merck, U.A. Hasan, S. Hubac, B. Benguigui, G. Trinchieri, E.E. Bates, and C. Caux. 2005. Recognition of double-stranded RNA by human toll-like receptor 3 and downstream receptor signaling requires multimerization and an acidic pH. *J. Biol. Chem.* 280:38133–38145.
17. Sun, J., K.E. Duffy, C.T. Ranjith-Kumar, J. Xiong, R.J. Lamb, J. Santos, H. Masarapu, M. Cunningham, A. Holzenburg, R.T. Sarisky, et al. 2006. Structural and functional analyses of the human Toll-like receptor 3. Role of glycosylation. *J. Biol. Chem.* 281:11144–11151.
18. Grabiec, A., G. Meng, S. Fichte, W. Bessler, H. Wagner, and C.J. Kirschning. 2004. Human but not murine toll-like receptor 2 discriminates between tri-palmitoylated and tri-lauroylated peptides. *J. Biol. Chem.* 279:48004–48012.
19. Lien, E., T.K. Means, H. Heine, A. Yoshimura, S. Kusumoto, K. Fukase, M.J. Fenton, M. Oikawa, N. Qureshi, B. Monks, et al. 2000. Toll-like receptor 4 imparts ligand-specific recognition of bacterial lipopolysaccharide. *J. Clin. Invest.* 105:497–504.
20. Poltorak, A., P. Ricciardi-Castagnoli, S. Citterio, and B. Beutler. 2000. Physical contact between lipopolysaccharide and toll-like receptor 4 revealed by genetic complementation. *Proc. Natl. Acad. Sci. USA*. 97:2163–2167.
21. Hajjar, A.M., R.K. Ernst, J.H. Tsai, C.B. Wilson, and S.I. Miller. 2002. Human Toll-like receptor 4 recognizes host-specific LPS modifications. *Nat. Immunol.* 3:354–359.
22. Hartmann, G., and A.M. Krieg. 2000. Mechanism and function of a newly identified CpG DNA motif in human primary B cells. *J. Immunol.* 164:944–953.
23. Roach, J.C., G. Glusman, L. Rowen, A. Kaur, M.K. Purcell, K.D. Smith, L.E. Hood, and A. Aderem. 2005. The evolution of vertebrate Toll-like receptors. *Proc. Natl. Acad. Sci. USA*. 102:9577–9582.
24. Samatey, F.A., K. Imada, S. Nagashima, F. Vonderviszt, T. Kumasaka, M. Yamamoto, and K. Namba. 2001. Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. *Nature*. 410:331–337.
25. Kobe, B., and J. Deisenhofer. 1996. Mechanism of ribonuclease inhibition by ribonuclease inhibitor protein based on the crystal structure of its complex with ribonuclease A. *J. Mol. Biol.* 264:1028–1043.
26. Schubert, W.D., C. Urbanke, T. Ziehm, V. Beier, M.P. Machner, E. Domann, J. Wehland, T. Chakraborty, and D.W. Heinz. 2002. Structure of internalin, a major invasion protein of *Listeria monocytogenes*, in complex with its human receptor E-cadherin. *Cell*. 111:825–836.
27. Huizinga, E.G., S. Tsuji, R.A. Romijn, M.E. Schiphorst, P.G. de Groot, J.J. Sixma, and P. Gros. 2002. Structures of glycoprotein Ibalph and its complex with von Willebrand factor A1 domain. *Science*. 297:1176–1179.
28. Burgess, A.W., H.S. Cho, C. Eigenbrot, K.M. Ferguson, T.P. Garrett, D.J. Leahy, M.A. Lemmon, M.X. Sliwkowski, C.W. Ward, and S. Yokoyama. 2003. An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. *Mol. Cell*. 12:541–552.
29. Wells, J.A. 1996. Binding in the growth hormone receptor complex. *Proc. Natl. Acad. Sci. USA*. 93:1–6.
30. Underhill, D.M., A. Ozinsky, A.M. Hajjar, A. Stevens, C.B. Wilson, M. Bassetti, and A. Aderem. 1999. The Toll-like receptor 2 is recruited to macrophage phagosomes and discriminates between pathogens. *Nature*. 401:811–815.
31. Smith, K.D., A. Valenzuela, J.L. Vigna, K. Aalbers, and C.T. Lutz. 1993. Unwanted mutations in PCR mutagenesis: avoiding the predictable. *PCR Methods Appl.* 2:253–257.
32. Ginalski, K., A. Elofsson, D. Fischer, and L. Rychlewski. 2003. 3D-Jury: a simple approach to improve protein structure predictions. *Bioinformatics*. 19:1015–1018.
33. Ginalski, K., and L. Rychlewski. 2003. Detection of reliable and unexpected protein fold predictions using 3D-Jury. *Nucleic Acids Res.* 31:3291–3292.
34. Rohl, C.A., C.E. Strauss, D. Chivian, and D. Baker. 2004. Modeling structurally variable regions in homologous proteins with rosetta. *Proteins*. 55:656–677.
35. Chivian, D., D.E. Kim, L. Malmstrom, P. Bradley, T. Robertson, P. Murphy, C.E. Strauss, R. Bonneau, C.A. Rohl, and D. Baker. 2003. Automated prediction of CASP-5 structures using the Robetta server. *Proteins*. 53(Suppl. 6):524–533.