Supporting Information Appendix

Valkov et al.

Supporting Information: Detailed Materials and Methods

Recombinant expression, purification and crystallization of the TIR domain of MAL. The construct comprising residues 79-221 of human MAL, including the predicted TIR domain and an N-terminal hexahistidine tag, was expressed in large-scale under the control of a T7 promoter in *E. coli* BL21 (DE3) cells for 18 h at 28 °C. Following cell lysis and lysate clarification, the protein was purified by nickel affinity chromatography, and the tag removed by digestion with human rhinovirus 14 (HRV14) 3C protease. Subsequent size exclusion chromatography on a HiLoad 26/60 Superdex-75 column (GE Healthcare) in 500 mM NaCl, 10 mM Tris-HCl, pH 7.5 and 2 mM DTT, yielded pure protein. For storage, the protein was concentrated to 7 mg ml^{-1}, supplemented with 10% v/v glycerol, frozen in liquid nitrogen, and kept at −80 °C.

Following high-throughput evaluation of sparse matrix crystallization conditions, a few very small crystals were found in several conditions after 3-4 weeks incubation at room temperature. After extensive screening and optimization, the best crystals (with dimensions of $10\times10\times50$ μm^3) were obtained in 10-12% PEG 10,000, 5% PEG 3,350, 0.2 M NaCl, 0.1 M Tris-HCl pH 7.3 and 20 mM DTT. For cryoprotection prior to flash-cooling in liquid nitrogen, the crystals were briefly transferred to a drop containing the crystallization solution supplemented with 20% v/v glycerol.

Protein-protein docking analysis. The coordinates of the three-dimensional structures of the MAL-TIR and the MyD88-TIR (PDB code 2z5v) were submitted as receptor and ligand, respectively, to GRAMM-X protein-protein docking server (http://vakser.bioinformatics.ku.edu/resources/gramm/grammx) (1). The computation of the global search of the best rigid-body conformations was conducted using default parameters and the FireDock server (http://bioinfo3d.cs.tau.ac.il/FireDock) (2) was used for further refinement and scoring for global energy values. The top ten models with the lowest energy were selected for further visual inspection.

Immunoprecipitation and immunoblot analysis. HEK293T cells (2×10^6 cells/10-cm dish) were transfected using Fugene 6 (Roche Applied Science) with the total amount of DNA (2.5 μg/dish) kept constant. Twenty-four hours later, cells were lysed in KalB buffer (50 mM Tris (pH 7.4), 1.0% Triton X-100, 150 mM NaCl, 1 mM EDTA, 2 mM Na_3VO_4, 10 mM NaF, 1 mM PMSF and protease cocktail inhibitor mixture; Roche). Antibodies (2 μg) or anti-FLAG-Sepharose beads (20 μl, 50% slurry) were incubated with the cell lysates for 2 h, followed by the addition of 40 μl of 50% protein G slurry for 1 h. The immune complexes were precipitated, washed, eluted by the addition of sample buffer followed by SDS-PAGE and boiled for 5 min. Protein samples were separated by SDS-PAGE, transferred to nitrocellulose, and immunoblotted. Immunocomplexes were visualized by using SuperSignal West Pico chemiluminescent substrate solution (Pierce) followed by exposure to x-ray film (Hyperfilm ECL; Amersham Biosciences).

Gene reporter assays. HEK293 cells ($2x10^4$) were seeded in 96-well plates 24 h before transfection. Transfections were performed with Fugene 6 (Roche) transfection reagent

according to the manufacturer's instructions, in conjunction with TK-renilla construct to normalize for transfection efficiency and an appropriate empty vector plasmid to maintain a constant amount of DNA. We determined NF-κB–dependent gene expression using the 5x κB-luciferase reporter construct (Stratagene). Cells were grown for a further 24 h after transfection, lysed (passive lysis buffer; Promega) and assayed for luciferase and TK-renilla activity using FLUOstar Optima (BMG Technologies) after incubation in luciferase assay reagent (Promega) or coleonterazine assay reagent. We corrected the luminescence readings for renilla activity and expressed them as factor increases over nonstimulated control values.

Confocal imaging analysis. HEK293 cells were cotransfected with MAL-GFP fluorophore-conjugated proteins for 48 h at 32 °C. The cells were washed in phosphate-buffered saline, fixed in 4% paraformaldehyde, and permeabilized in 0.1% Triton X-100. The cells were immunohistostained with FITC-conjugated cholera toxin B subunit. Mounted cells on coverslips were examined with 40× oil lens objective on Leica TCS NT upright confocal microscope (Leica Microsystems GmbH, Welzlar, Germany) using Leica TCS NT LASAF software.

Supporting Information: Tables

Table S1. Crystallographic data collection and refinement statistics.

Data collection statistics	
Wavelength (Å)	1.0722
Data collection temperature (K)	100
Space group	$P4_32_12$
Unit cell parameters: a, b, c (Å); α, β, γ (°)	88.1, 88.1, 78.8; 90, 90, 90
Resolution range (outer shell in brackets; Å)	48.9 – 3.01 (3.37 – 3.01)
Unique reflections	6,390
Total observations	95,039
<I / σ(I)>: all (outer shell)	15.5 (2.0)
R_{merge}: all (outer shell)*	11.6 (169.2)
Completeness: all (outer shell) (%)	99.6 (97.9)
Multiplicity	13.7
Wilson B-factor (Å^2)	112.8
Refinement statistics	
Amino acid residues	79-111, 124-167, 172-221
Non-hydrogen protein atoms	989
Non-protein ligands	1 × DTT
Solvent content (%)	77.6
Bond length deviation from ideal values (Å)	0.007
Bond angle deviation from ideal values (°)	0.89
Peptide omega torsion angles (°)	1.56
Other torsion angles (°)	19.86
Average B-factor (Å^2)	95.0
Ramachandran favoured/outliers (%)	97.5/0
Random reflections assigned for cross-validation	306
R_{work}/R_{free} (%)†	23.1/25.1
MolProbity score‡ (100th percentile)	1.80
All-atom clashscore‡ (100th percentile)	6.7
Luzzati plot estimate of coordinate error (Å)	0.75

*$R_{merge} = \sum_{hkl}(\sum_i(|I_{hkl,i}-\langle I_{hkl}\rangle|))/\sum_{hkl,i}\langle I_{hkl}\rangle$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices h, k and l, and $\langle I_{hkl}\rangle$ is the mean intensity of that reflection. Calculated for $I > -3\sigma(I)$.

†$R_{work} = \sum_{hkl}(||Fobs_{hkl}|-|Fcalc_{hkl}||)/|Fobs_{hkl}|$, where $|Fobs_{hkl}|$ and $|Fcalc_{hkl}|$ are the observed and calculated structure factor amplitudes. R_{free} is equivalent to R_{work} but calculated with reflections (5 %) omitted from the refinement process.

‡MolProbity score is defined as follows: 0.42574*log(1+clashscore)+0.32996*log(1+max(0,pctRotOut-1))+0.24979*log(1+max(0,100-pctRamaFavored-2))+0.5; *100th percentile is the best among structures of comparable resolution; 0th percentile is the worst (3).

Table S2. List of observed dimeric interfaces in the crystal structures of TIR domain-containing proteins.

Molecule	PDB ID	Area buried in dimeric interface (Å^2)*	Secondary structure elements involved in the interface	Symmetry	Supporting experimental evidence	Special features	Reference
TLR1	1fyv	1610	αB, αC, αD	Yes	-	Disufide linkage between C707 and its symmetry mate across the putative dimeric interface	(4)
TLR2	1fyw	2190	αB, αC	Yes	-	Thiol groups of cysteine residues modified by cacodylate	(4)
TLR2-C713S	1o77	1310	A: αB, αC, αC', αD B: αB	No	Site-directed mutagenesis combined with functional signaling assays	Disulfide linkage between C640 and C750 of its symmetry mate across the dimeric interface	(5)
TLR10	2j67	1430	αB, αC	Yes	-	-	(6)
PdTIR	3h16	1530	αD, αE	Yes	Hydrogen-deuterium exchange and proteolytic mapping studies	-	(7)
L6	3ozi	1100	αD, αE, βE	Yes	Site-directed mutagenesis combined with yeast-two-hybrid assays; biophysical solution studies	-	(8)
MAL	2y92	1510	αC′	Yes	-	Covalent DTT adduct between the thiol groups of C91 and C157	This study
IL1-RAPL	1t3g	2100	αB, αC, αC′, αD	Yes	-	-	(9)

*Calculated using "assemblies buried area" function in *PISA* (10).

Table S3. Predicted effect of amino acid substitutions in MAL on protein stability.

Residue substitution	Secondary structure environment of the residue	Change in solvent accessibility	Change in hydrogen bonding of side-chain	*Pseudo* $\Delta\Delta G$ (kcal mol^{-1})*	Predicted effect on protein stability**
W156A	α-helix	–20.3%	H-bonds removed	+2.36	Highly stabilizing
P155A	α-helix	+16.5%	No change	+2.21	Highly stabilizing
Y159A	α-helix	–27.4%	No change	+0.13	Neutral
L162A	α-helix	–44.4%	No change	+0.29	Neutral
L165A	α-helix	+13%	No change	–0.71	Slightly destabilizing
T166A	Loop	–8.7%	No change	–0.81	Slightly destabilizing
P189A	3_{10}-helix	+6.6%	No change	+2.21	Highly stabilizing
E190A	3_{10}-helix	–0.7%	No change	+0.40	Neutral
F193A	Loop	+5.0%	No change	+0.04	Neutral
M194A	β-turn	+9.1%	No change	–1.15	Destabilizing
D96N	Loop	+2.5%	No change	–0.33	Neutral
S180L	Loop	+4.9%	No change	+2.13	Highly stabilizing

*Predicted stability score, analogous to the free energy difference between the wild-type and mutant protein. Positive value indicates an increase in the thermodynamic protein stability, while negative values indicate a decrease in stability. The stability score was evaluated with the SDM server (http://www-cryst.bioc.cam.ac.uk/~sdm/sdm.php) using MAL-TIR crystal structure coordinates as input. SDM (11) utilizes a statistical approach by fitting variables such as substitution frequencies, distance potentials and residue environments to a potential energy function and this method has been shown to have an accuracy of 74% in predicting the sign of stability change and a linear correlation coefficient of 0.60 between predicted and observed $\Delta\Delta G$ values (11).

**A cut-off of 2.0 kcal mol^{-1} in *pseudo* $\Delta\Delta G$ value indicates a significant effect on protein stability. None of the substitutions are likely to result in a significantly destabilized, misfolded protein. We consider it is unlikely that the mutants lead to unstable proteins for a number of other reasons: (i) the mutants were all designed using the crystal structure as a guide and care was taken to choose only those that are surface-exposed and not likely to influence protein folding or stability, and those predicted to lead to subtle effects that can demonstrate specificity of individual mutations; (ii) most of the mutations showed no phenotype (Fig. 6A), demonstrating they did not affect structural integrity; (iii) the L162A, L165A and T166A mutants showed a partial and not complete loss of binding to MyD88 (Fig. 6A), consistent with an effect restricted to the side-chain substitution that does not cause gross structural changes; (iv) like wild-type MAL, the D96N and S180L variants are able to form homodimers (Fig. 6B), consistent with them adopting a native-like fold; (vi) the mutant D96N that shows the most prominent phenotype corresponds to a rather small isosteric change.

Supporting Information: Figures

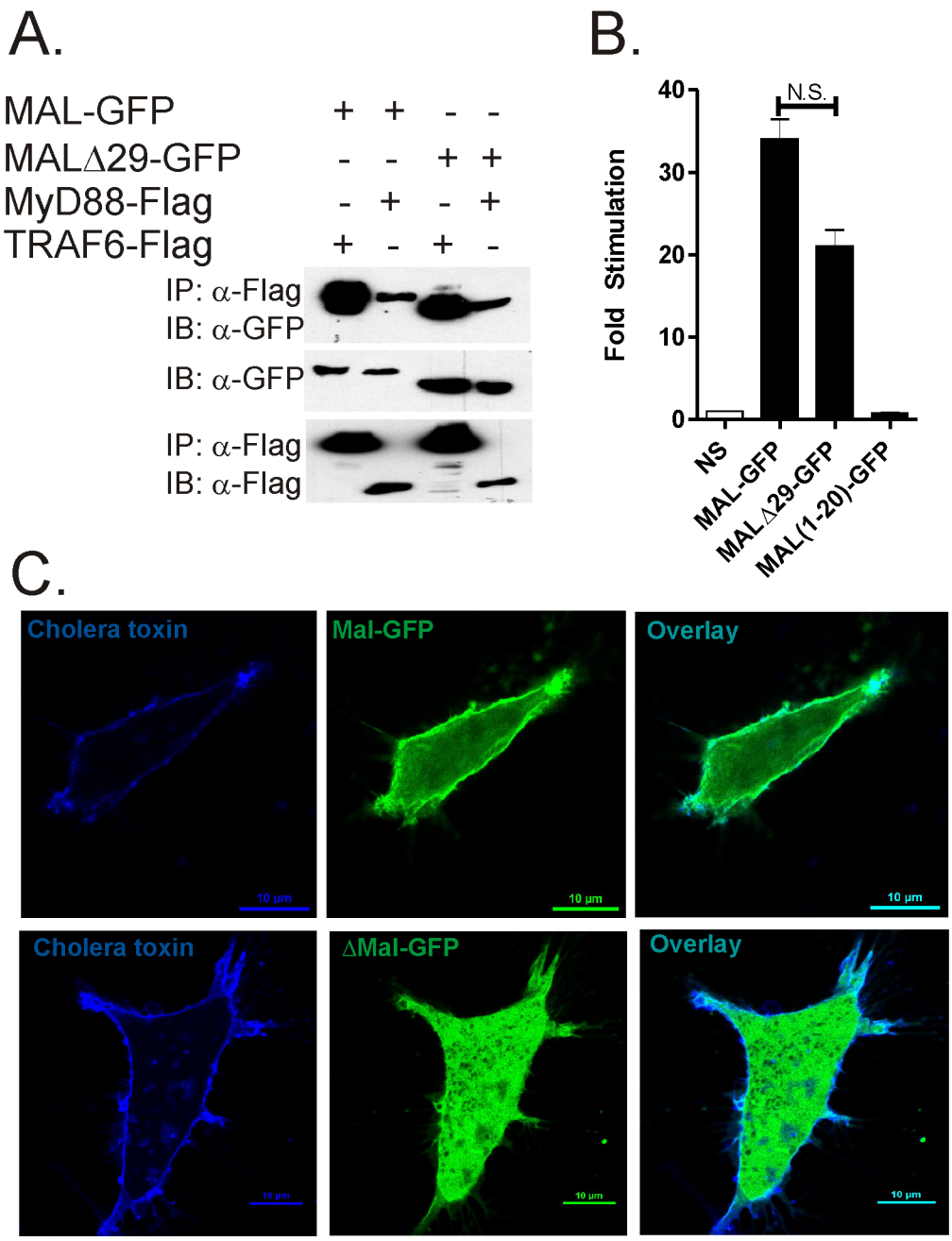


Figure S1. Functional properties of MAL lacking N-terminal 29 residues. Both wild-type MAL and MALΔ29 mutant were *(A)* able to interact with both coexpressed MyD88 and TRAF6 by ectopic immunoprecipitation, and *(B)* induce comparable activation of the κB-luciferase reporter construct, suggesting the deletion of the first 29 amino acids did not affect functional properties of MAL. (*C*) As expected however, the truncation altered the ability of MAL to localize to the plasma membrane due to the lack of PIP2-interaction motif found at the N-terminus of MAL (12, 13).

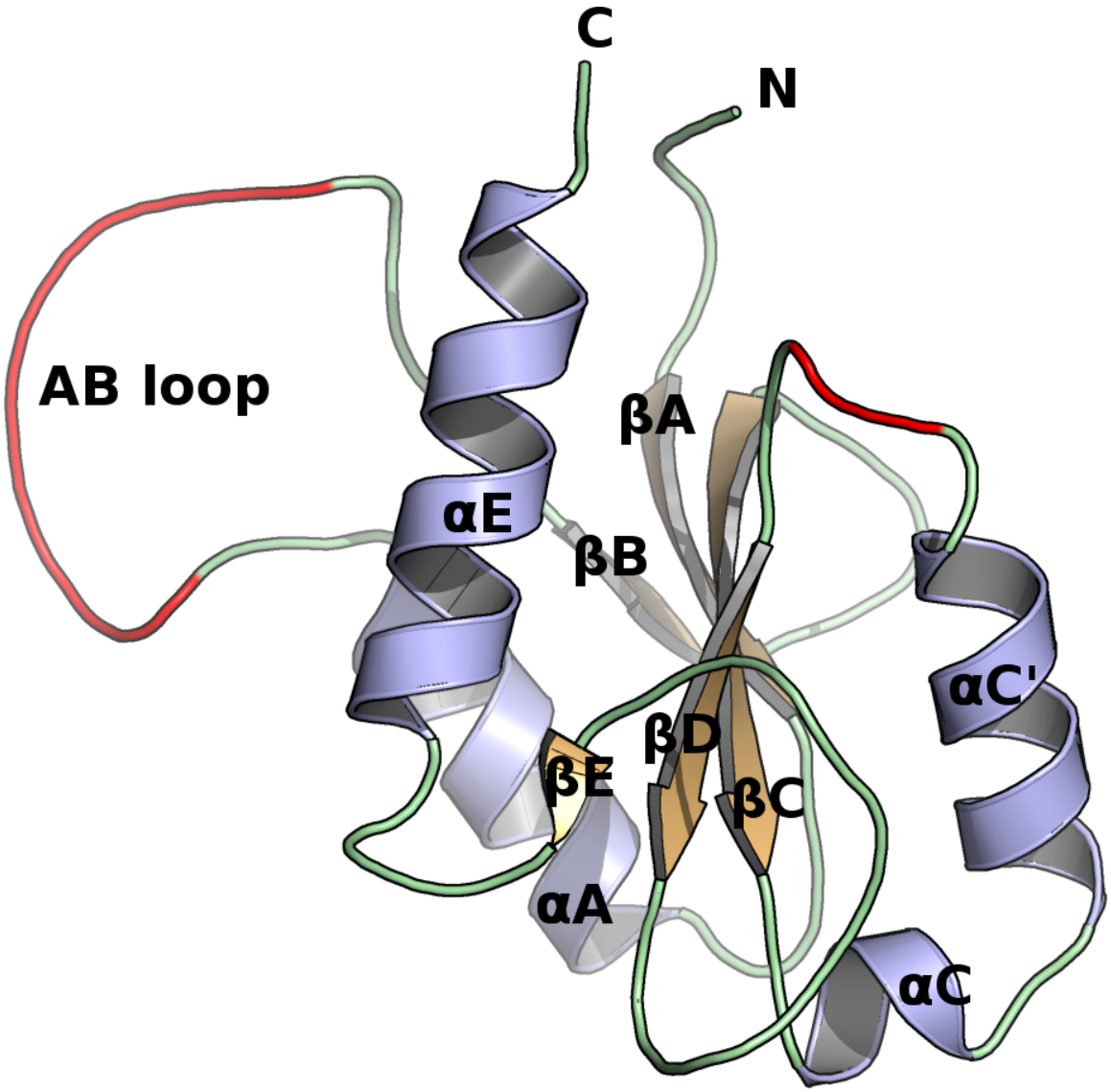


Figure S2. Crystal structure of MAL with modelled disordered regions. The regions comprising residue 112-123 and 168-171, which were disordered and not visible in the electron density, were modelled using stereochemical constraints and energy minimized in *COOT*. They are shown in red to distinguish them from the loop segments that were supported by electron density in the structure (shown in green).

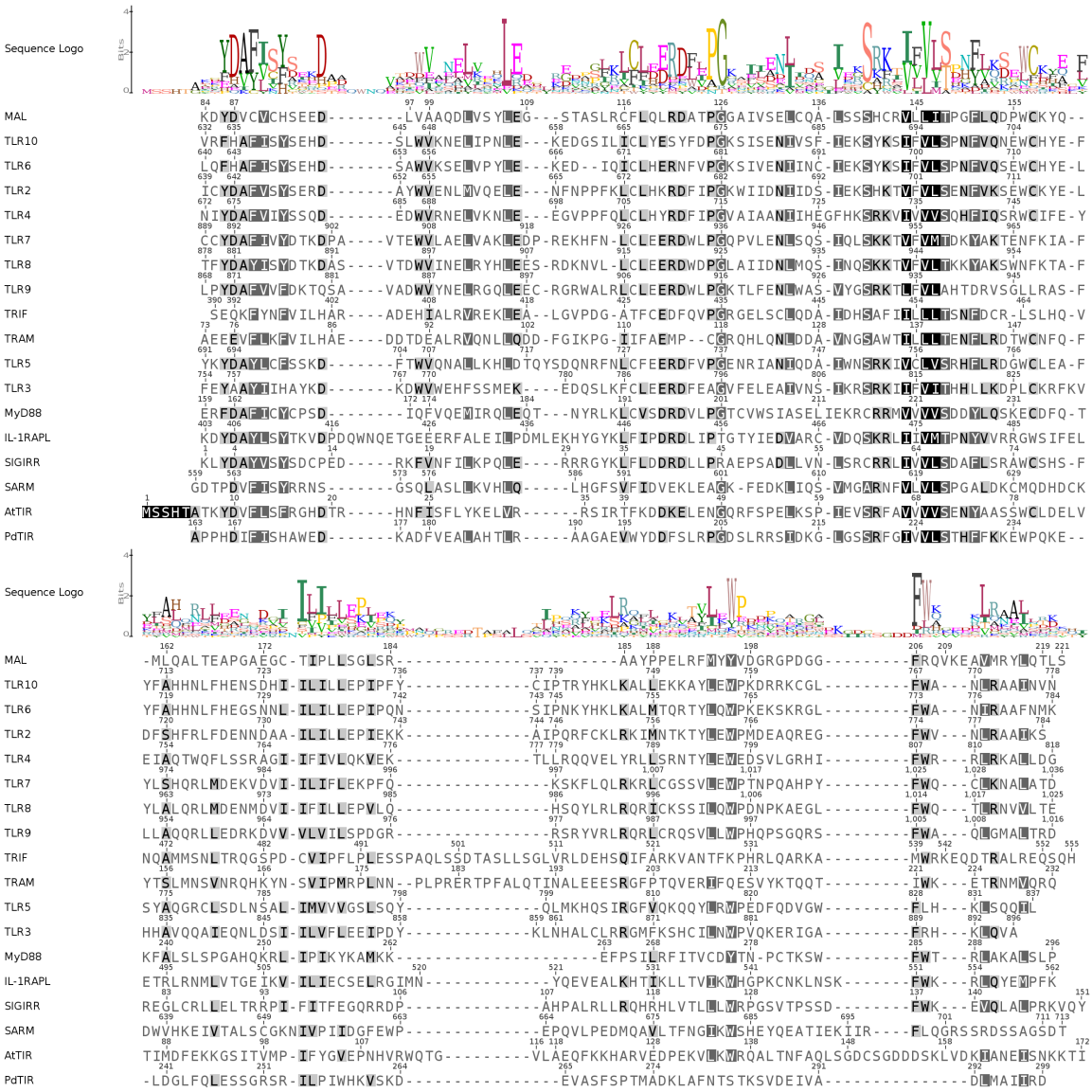


Figure S3. Sequence-based alignment of TIR domains. The alignment was generated using default parameters with the ClustalW program (14). The residue numbering corresponds to the full-length proteins. Conserved residues are shaded. Figure was prepared using the *Geneious* software suite (http://www.geneious.com).

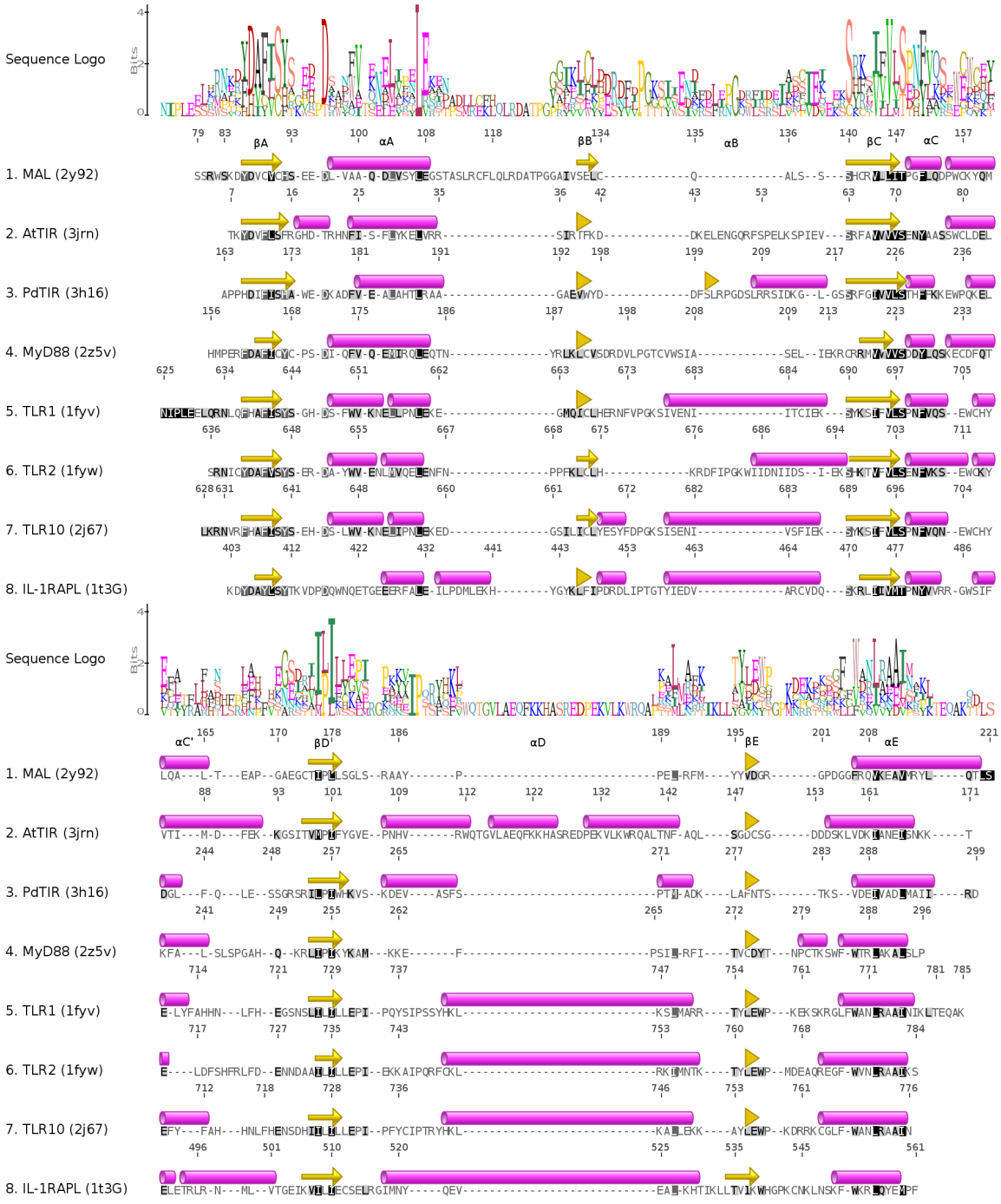


Figure S4. Structure-based sequence alignment of TIR domains with known structures. The alignment was prepared using the DALI server (15, 16). The residue numbering corresponds to the full-length proteins. Secondary structure elements are shown schematically above each sequence. The naming convention of the secondary structure elements follows the established convention first described for the TLR1 and TLR2 TIR domain structures. Figure was prepared using the *Geneious* software suite (http://www.geneious.com).

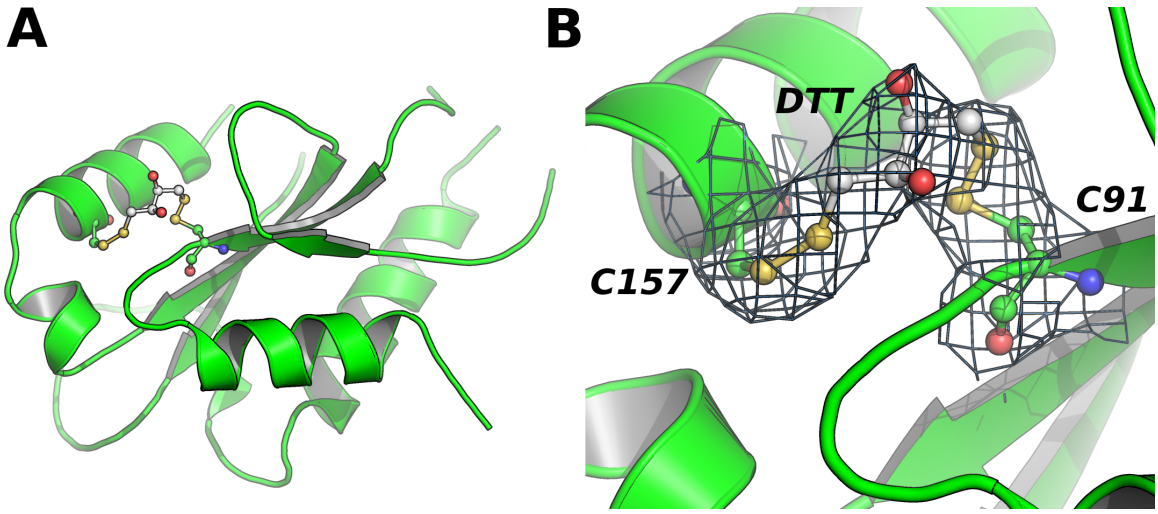


Figure S5. C91 and C157 form an unusual adduct with DTT in the crystal structure of MAL-TIR. *(A)* Overall view of the structure with the protein in cartoon representation and the C157-DTT-C91 adduct shown in ball-and-stick representation. Continuous electron density was observed between the thiol groups of residues C91 and C157; as the distance between the thiols was more than twice the typical distance for a disulfide bond (2.05 Å), we modelled it as a single molecule of DTT, which was included at all stages of protein purification as well as crystallization. The formation of unusual adducts with DTT by two surface cysteines has previously been observed in protein crystal structures and correlates with the reactivity of these residues towards covalent modification by sulfhydryl-containing small molecules in post-translational modification phenomena such as glutathionylation (17, 18). *(B)* Close-up view of the adduct. The electron density is the final refined $2F_o$ - F_c map contoured at 1σ level.

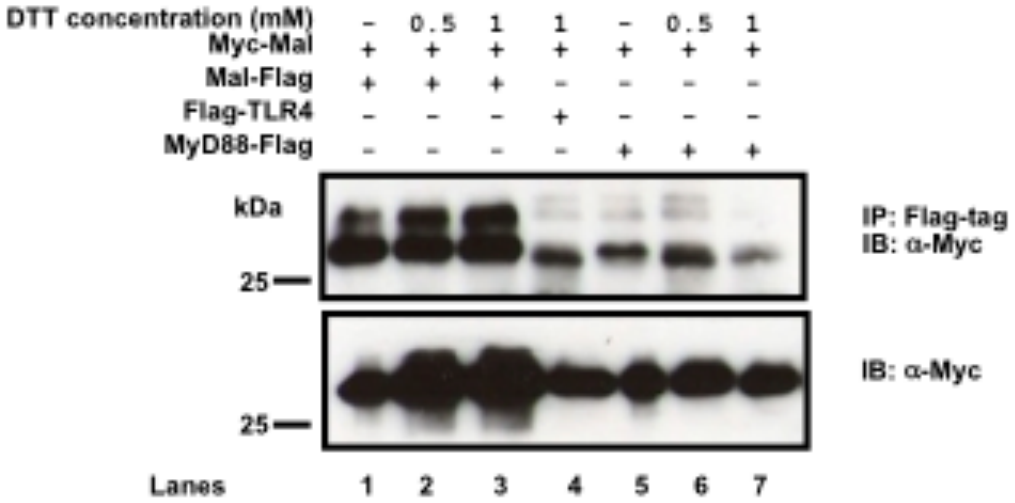


Figure S6. DTT reduces the interaction of MAL with MyD88 but not with TLR4 or itself. Immunoblot analysis of transfected full-length MAL interacting with either itself, TLR4 or MyD88 in the presence or absence of DTT. HEK293T cells were plated at a cell density of 2×10^6 per well in a 6-well plate 24 hours prior to co-transfection with 3 μg of total DNA containing either: 1.5 μg of Myc-tagged MAL and 1.5 μg MAL-Flag (lanes 1-3); 1.5 μg of Myc-tagged MAL and 1.5 μg Flag-TLR4 (lane 4); or 1.5 μg of Myc-tagged MAL and 1.5 μg MyD88-Flag (lanes 5-7). Whole-cell lysates were harvested, pre-absorbed with Protein A-Sepharose and immunoprecipiated with α-M2 Flag agarose beads (Sigma) (top panel). Expression levels of each protein along with detection of immunoprecipitated complexes were visualised by immunoblotting with anti-Myc antibody (middle and top panels). Results are representative of three independent experiments.

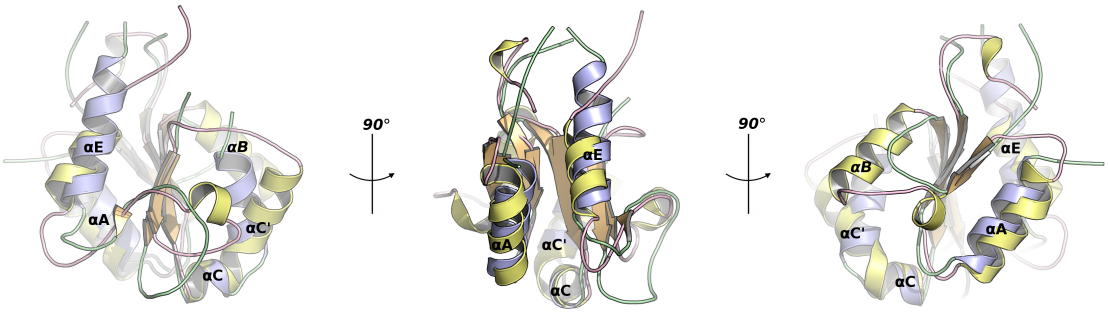


Figure S7. Comparison of the overall fold of the theoretical model of the TIR domain of MAL and the crystal structure. The helical secondary structure elements of the homology model based on the structures of TIR domains from TLR1 and TLR2 (19) are in yellow cartoon representation and those of the crystal structure are in blue. The loop regions of the model are in pink and the loops of the crystal structure are in green. The β-sheet core is in good structural agreement and the individual strands are not labelled for clarity. Major differences are seen for the conformations of the loops, as well as different orientation of the αC' helix. The αB helix (label italicized) is not present in the crystal structure.

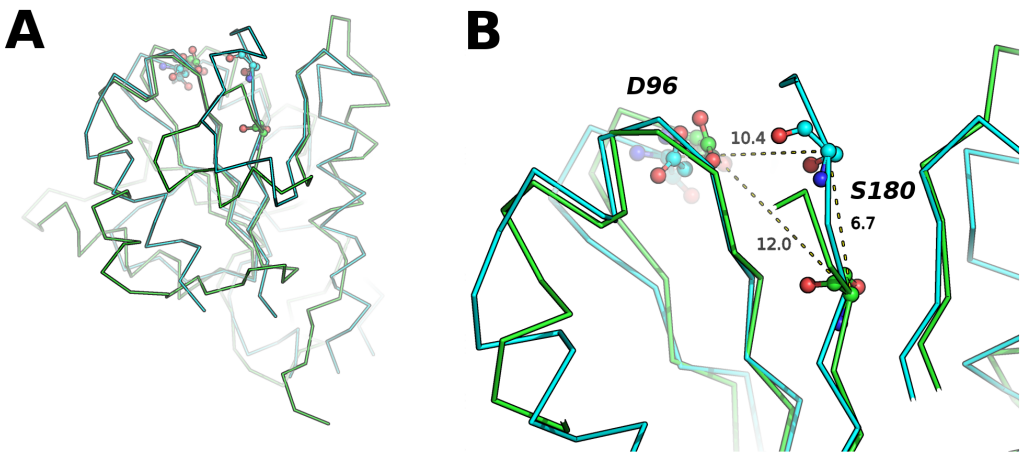


Figure S8. Comparison of the structural environments of the D96 and S180 in the homology model and the crystal structure of the TIR domain of MAL. (*A*) The Cα traces of the homology model based on the structures of TIR domains from TLR1 and TLR2 (19) (green) and the crystal structure (cyan) are shown. The D96 and S180 residues are shown in ball-and-stick representations. (*B*) Close-up view of the D96 and S180 residues. Distances shown are in Å between the Cα atoms of the residues. The structural position of D96 is accurately predicted in the model. However, S180 was predicted to be nearly 7 Å away from its actual position in the structure, where it is exposed on the surface of the protein. The distance between D96 and S180 is approximately 10 Å in the crystal structure.

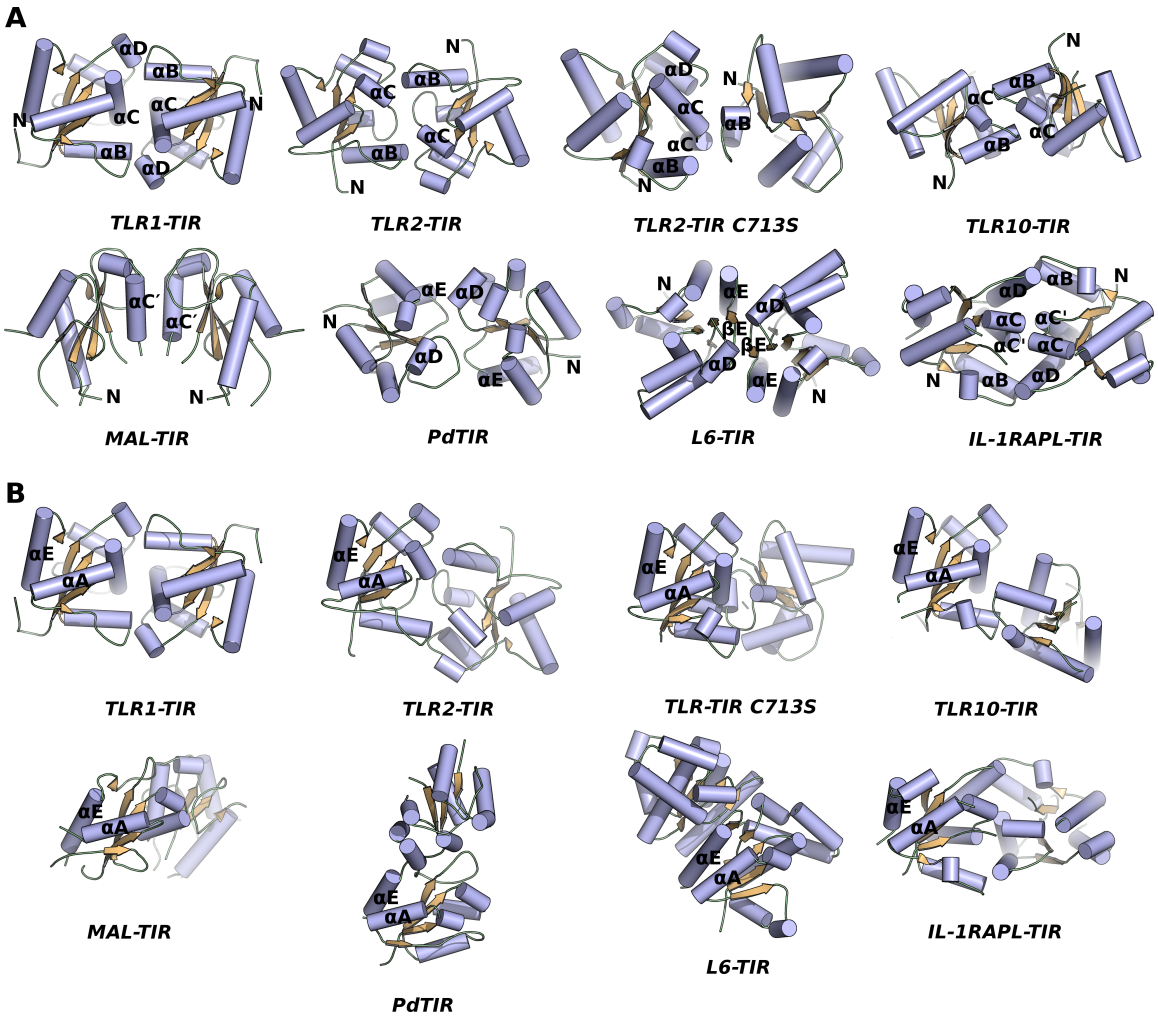
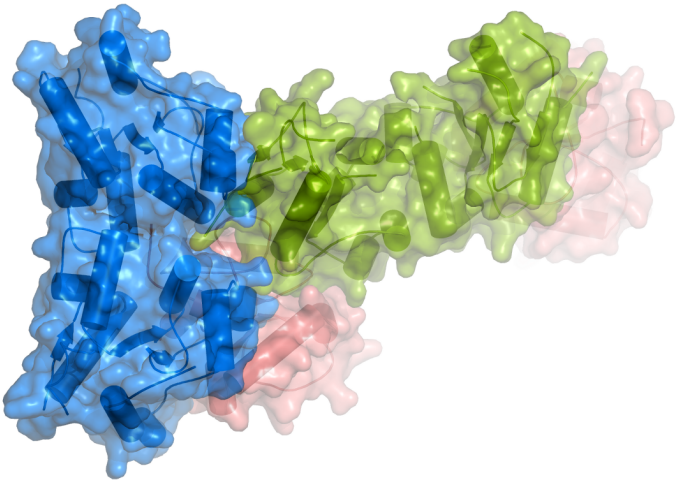


Figure S9. Comparison of the putative dimeric interfaces observed in the structures of the TIR domain-containing proteins. (*A*) The helical segments are shown as cylinders and the key secondary structure elements mediating the interface are labelled. (*B*) To illustrate that the interfaces all occur at distinct spatial locations, one monomer of the assembly was structurally superimposed to the reference monomer of the TLR1-TIR dimer. The first helix (αA) and the last (αE) of the reference protomer are labelled. While it is clear that the TIR domain-containing proteins from plant and bacterial sources (L6-TIR and PdTIR, respectively) show the most drastic differences in the dimeric arrangement, all proposed dimerization interfaces differ.

A

B



C

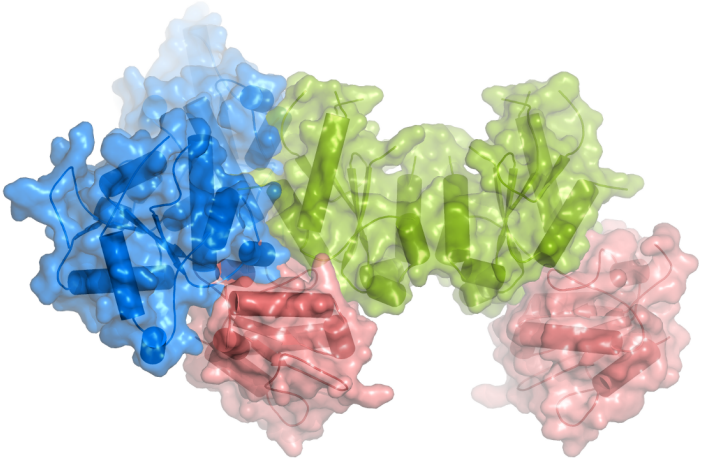


Figure S10. A structural model of the TLR4-TIR:MAL-TIR:MyD88-TIR signalosome complex predicted through docking calculations. The docking calculations were performed using the same procedure as for the MAL-TIR:MyD88-TIR complex. The TLR4-TIR dimer homology model was from a previous study (19) and was used as the 'receptor' in docking. The symmetric dimer of MAL-TIR formed by interaction 1 in the crystal interface was used as the 'ligand'. *(A)* The ten top-scoring solutions in ribbon representation; MAL-TIR in light green, TLR4-TIR in light blue. Eight out of ten best scoring solutions place the MAL-TIR dimer within the structural vicinity of the interface containing the BB loops and C747 of TLR4-TIR. *(B)* The model of the TLR4-TIR:MAL-TIR:MyD88-TIR complex assembled from independent docking calculations of the TLR4-TIR:MAL-TIR and MAL-TIR:MyD88-TIR complexes. The TLR4-TIR is light blue, MAL-TIR in light green and MyD88-TIR is in light red. The view down the crystallographic 2-fold symmetry axis of the TLR4-TIR dimer. The interface containing the AB loop motif occupies the cleft in the TLR4-TIR dimer that contains the BB loops and C747 residues. *(C)* The view across the crystallographic 2-fold symmetry axis of the MAL-TIR dimer. It can be seen that the N-termini of the MAL-TIR are unobstructed in the complex, supporting the notion that the N-terminal region interacts with the cell membrane. The MyD88-TIR is also seen to form the second heterotypic TIR-TIR interaction within the structural context of the complex, suggesting that the initiation of TLR signalling event involves cooperative interactions amongst the receptor and adaptor TIR domains.

Supporting Information: References

1. Tovchigrechko A, Vakser IA (2006) GRAMM-X public web server for protein-protein docking. *Nucleic Acids Res* 34:W310-314.
2. Mashiach E, Schneidman-Duhovny D, Andrusier N, Nussinov R, Wolfson HJ (2008) FireDock: a web server for fast interaction refinement in molecular docking. *Nucleic Acids Res* 36:W229-232.
3. Chen VB, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66:12-21.
4. Xu Y, et al. (2000) Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature* 408:111-115.
5. Tao X, Xu Y, Zheng Y, Beg AA, Tong L (2002) An extensively associated dimer in the structure of the C713S mutant of the TIR domain of human TLR2. *Biochem Biophys Res Commun* 299:216-221.
6. Nyman T, et al. (2008) The crystal structure of the human Toll-like receptor 10 cytoplasmic domain reveals a putative signaling dimer. *J Biol Chem* 283:11861-11865.
7. Chan SL, et al. (2009) Molecular mimicry in innate immunity: crystal structure of a bacterial TIR domain. *J Biol Chem* 284:21386-21392.
8. Bernoux M, et al. (2011) Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling and autoregulation. *Cell Host Microbe*.
9. Khan JA, Brint EK, O'Neill LA, Tong L (2004) Crystal structure of the Toll/interleukin-1 receptor domain of human IL-1RAPL. *J Biol Chem* 279:31664-31670.
10. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774-797.
11. Worth CL, Preissner R, Blundell TL (2011) SDM--a server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res*.
12. Kagan JC, Medzhitov R (2006) Phosphoinositide-mediated adaptor recruitment controls Toll-like receptor signaling. *Cell* 125:943-955.

13. Verstak B, et al. (2009) MyD88 adapter-like (Mal)/TIRAP interaction with TRAF6 is critical for TLR2- and TLR4-mediated NF-kappaB proinflammatory responses. *J Biol Chem* 284:24192-24203.
14. Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics* Chapter 2:Unit 2 3.
15. Holm L, Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* 38 Suppl:W545-549.
16. Holm L, Kaariainen S, Wilton C, Plewczynski D (2006) Using Dali for structural comparison of proteins. *Curr Protoc Bioinformatics* Chapter 5:Unit 5 5.
17. Lowther WT, Brot N, Weissbach H, Matthews BW (2000) Structure and mechanism of peptide methionine sulfoxide reductase, an "anti-oxidation" enzyme. *Biochemistry* 39:13307-13312.
18. Haikarainen T, Chen WQ, Lubec G, Kursula P (2009) Structure, modifications and ligand-binding properties of rat profilin 2a. *Acta Crystallogr D Biol Crystallogr* 65:303-311.
19. Nunez Miguel R, et al. (2007) A dimer of the Toll-like receptor 4 cytoplasmic domain provides a specific scaffold for the recruitment of signalling adaptor proteins. *PLoS One* 2:e788.