Structure and dynamics of pathway activation by the Toll immunoreceptor from the viral mosquito vector Aedes aegypti

Monique Gangloff (✉ mg308@cam.ac.uk)  
University of Cambridge

Yoann Saucereau  
University of Cambridge  https://orcid.org/0000-0002-7175-2416

Thomas Wilson  
University of Cambridge

Matthew Tang  
University of Cambridge

Martin Moncrieffe  
University of Cambridge

Steven Hardwick  
Department of Biochemistry, University of Cambridge  https://orcid.org/0000-0001-9246-1864

Dimitri Chirgadze  
University of Cambridge  https://orcid.org/0000-0001-9942-0993

Sandro Soares  
University of Cambridge

Maria Marcaida  
École Polytechnique Fédérale de Lausanne  https://orcid.org/0000-0003-0663-5309

Nicholas J. Gay  
University of Cambridge

Article

Keywords: Toll receptor, Spaetzle cytokine, Aedes aegypti, immunity, Dengue, vectorial capacity, cryoEM, RNAseq

Posted Date: December 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1082903/v1

License: ☑  This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Yoann Saucereau1, Tom H. Wilson1, Matthew C. K. Tang1, Martin C. Moncrieffe1, Steven W. Hardwick1, Dimitri Y. Chirgadze1, Sandro G. Soares1, Maria Jose Marcaida2, Nick Gay1, Monique Gangloff1*

1Department of Biochemistry; University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA; United Kingdom. 2Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne; Lausanne, Switzerland.

*Corresponding Author: Dr Monique Gangloff. Department of Biochemistry, University of Cambridge. 80 Tennis Court Road, Cambridge CB2 1GA. United Kingdom. Correspondence: mg308@cam.ac.uk.

ABSTRACT

Aedes aegypti has evolved to become an efficient vector for arboviruses such as Dengue but the mechanisms of host-pathogen immune tolerance are unknown. Toll receptors and Spaetzle (Spz) ligands have undergone duplication raising the possibility of neofunctionalization and mutualism to develop between arboviruses and mosquitoes. Here we present cryo-EM structures and biophysical characterisation of low affinity Toll5A-Spz1C complexes that display transient but specific interactions. Binding of the first ligand alters receptor-receptor interactions and promotes asymmetric contacts in the vicinity of the Z-loop in Toll5A. This conformation then restricts binding of a second ligand, while temporarily bridging the C-termini that promote signalling. Increased receptor concentrations promote inactivating head-to-head receptor assemblies. Furthermore, the transcriptional signature of Spz1C differs from other Spz cytokines in the control of genes involved in innate immunity, lipid metabolism and tissue regeneration. Given the remarkable DENV-induced expression patterns of these proteins, our data rationalises how Spz1C upregulation might promote antimicrobial defence in the midgut, and Toll5A upregulation, viral tolerance in the salivary glands.

KEYWORDS

Toll receptor; Spaetzle cytokine; Aedes aegypti; immunity; Dengue; vectorial capacity; cryo-EM; RNAseq
INTRODUCTION

Insect Toll receptors perform critical functions in both embryogenesis and innate immunity. They are part of an ancient defence system that also includes the immunodeficiency (IMD) pathway and are conserved in vertebrates as the Toll-like receptors. Tolls and TLRs have a modular structure with an ectodomain made up of leucine rich repeats (LRRs) and associated capping structures, a single transmembrane helix and a cytosolic Toll/Interleukin 1 receptor (TIR) domains\(^1\). Most studies of innate immune function in insects have focussed on the fruit fly *Drosophila melanogaster*\(^2\). The *Drosophila* genome encodes 9 Toll receptors that have diverse roles in development and other areas of cellular regulation as well as in immunity. Immune function is mediated mainly by the Toll receptor 1 (Toll1), while Toll6 and Toll7 are expressed in the nervous system and have equivalent functions to vertebrate neurotrophin receptors\(^3\).

Toll receptors are activated by a complementary family of six cytokine-like molecules called Spaetzle (Spz)\(^4\). Spz proteins are secreted as larger dimeric precursors consisting of a natively unstructured pro-domain, which is proteolytically cleaved upon activation and a C-terminal cystine-knot fold similar to that found in human neurotrophins, such as nerve growth factor\(^5\). In the case of Spz1, Lys-type peptidoglycan from Gram-positive bacteria is detected by recognition proteins PGRP-SA and GNBP1 and these complexes activate a cascade of serine proteases\(^6\). The terminal proteinase, Spaetzle processing enzyme (SPE), has trypsin-like specificity and cleaves the Spz precursor specifically at the junction between the pro-domain and cystine-knot which remain associated by non-covalent interactions\(^7\). Spz2 (also called neurotrophin-1, NT1) and Spz5 (NT2) are cleaved during secretion by furin-like proteases\(^8\) and, in the case of Spz2, a truncated pro-domain and the cystine knot remain associated non-covalently. By contrast, Spz5 is secreted as a mature form without the pro-domain. Growing evidence suggest promiscuity in ligand binding, with Toll1 and Toll7 recognising Spz1, Spz2 and Spz5 as well as vesicular stomatitis virus (VSV) virions\(^9\)–\(^12\), while Toll6 senses Spz2 and Spz5\(^8\).

Mosquitoes and fruit flies are both dipterans but they diverged in evolution about 260 million years ago\(^13,14\). Comparative genomic analysis of *Aedes aegypti*, *Anopheles gambiae* and *Drosophila melanogaster* reveals that most genes involved in innate immunity are conserved in the mosquito including the Toll and Spaetzle families\(^15\). Nevertheless, the Toll family has undergone significant diversification with the loss of orthologs in mosquitoes (no Toll2 or Toll3 in *A. aegypti*, for example) and species-specific expansion of two more (Toll10 and Toll11). The underlying driving force of gene duplication is likely interconnected with the mosquito’s change to a hematophagous diet and the evolutionary arms-race between pathogens and insects. In *Drosophila* Toll1 mediates most immune functions while in *Aedes* two gene reduplications have occurred to produce an orthologous group of four closely related receptors Toll1A, Toll1B, Toll5A and Toll5B. The phylogeny does not provide unequivocal evidence as to which of these receptors function in immunity or if they have acquired new functions upon diversification.

Although there is a clear evolutionary relationship between insect Tolls and vertebrate TLRs the latter appear to have arisen from the atypical insect Toll9 and adapted to directly recognise a diverse range of pathogen associated molecular patterns (PAMPS) such a bacterial lipopolysaccharide\(^16\). Thus, vertebrate TLRs are bona fide pattern recognition receptors but insect Tolls are activated by endogenous cytokines or growth factors, with the exception of VSV recognition. Another distinction between Toll and TLRs is the kinetics of receptor activation. TLRs display positive cooperativity being activated over a narrow range of ligand
concentrations. By contrast, the function of Toll1 in development requires a diffused gradient of active Spaetzle with different threshold concentrations specifying cell fates along the embryonic dorso-ventral axis.

A third function of insect Tolls that is not shared with TLRs is cell adhesion. However, a number of other leucine rich repeat proteins including human platelet glycoprotein 1B and Drosophila Chaoptin are cell adhesion molecules. Adhesion mediated by Toll1 may be either homo- or heterotypic although a partner for heterotypic binding has not been described and adhesion does not activate signal transduction. In the embryo Toll1 is transiently expressed in a specific subset of muscle fibres and may mediate formation of neuromuscular junctions together with another non-receptor LRR protein Connectin.

Structural studies have provided insight into the molecular mechanism by which PAMPS activate the TLRs. A common feature is stimulus-induced dimerization of the receptor ectodomains or alternatively conformational rearrangement of a preformed inactive dimer. For example, double stranded RNA crosslinks two TLR3 ectodomains causing the juxtamembrane regions to move into close proximity and promoting dimerization of the TIR domains in the cytosol. By contrast, TLR8 is expressed as pre-formed dimer and binding of small drug molecules or single stranded RNA induces a conformational rearrangement of the dimer interface that causes the two ectodomains to tilt together and, like TLR3, moves the juxtamembrane C-termini into close proximity. It is likely that activation of insect Toll receptors by Spaetzle ligands involves similar conformational changes. In that regard biochemical evidence suggests that the active complex is a 2:2 heterotetramer with two molecules of receptor and two molecules of Spaetzle-C106. However, to date the only crystal structures of the Toll–C106 complex reveal a 1:1 complex with a binding mode that is reminiscent of mammalian neurotrophins. The covalent cystine-knot dimer forms asymmetric contacts at the concave side of the N-terminal cap and within the first ten LRRs. In contrast to TLRs, Spaetzle does not induce dimerization of the receptor in the crystal structure possibly because negative cooperativity requires that the active signalling complex is unstable.

Drosophila has served as a model system for insects but adaptation to blood-borne pathogens by mosquitoes seems to have driven the evolution of their innate immune system. The Toll pathway has been involved in resistance against the entomopathogenic fungus Beauveria bassiana and Dengue, with fungal infections limiting viral replication. However, differential expression of duplicated Toll5A and Spz1C suggests that this pathway is regulated in a tissue and microbe-specific manner. Here, we use purified proteins to further characterise Toll5A-Spz1C interactions and the impact of their relative concentrations on stoichiometry. Single particle cryo-EM reveals that the receptor is crosslinked in an asymmetric complex, in which C-terminal regions are brought into proximity upon Spz1C binding, which is a structural requirement for signal transduction. Along with the functional characterisation of Spz1C-induced A. aegypti Aag2 cell signalling, our study sheds light on the effect of duplication and shows the importance of revisiting Toll signalling in hematophagous insects that impact human health.
RESULTS

CryoEM of Toll5A-Spz1C single particles.

To gain mechanistic insight into the activation of Toll5A by ligand Spz1C, we carried out single particle cryo-EM to determine the structure of the stable complexes formed by the receptor ectodomain and the activated ligand. Recombinant Toll5A ectodomain was mixed with a 3-fold molar excess of Spz1C and further purified by size exclusion chromatography immediately prior to grid preparation as described in Methods. We then visualised 3 oligomeric states of the receptor at near atomic resolutions, suggesting that Toll5A is highly dynamic in the presence of Spz1C. The resolution of the original maps was extended by density modification with ResolveCryoEM. Homodimers of Toll5A at 3.41 Å maximum resolution had 85,810 particles, 2:1 heterodimers of Toll5A bound to Spz1C at 4.23 Å had 40,153 particles, and a 3:1 heterotrimer at 3.74 Å had 42,866 particles (Fig. 1 and S1). While all three types of particles spread randomly across the grid, according to their large Euler angular distribution, they display significant variations in local resolutions, likely a consequence of their stoichiometric and conformational heterogeneity. Overall, the N-terminal LRR domains of each receptor chain achieve better resolutions than their C-terminal moieties. The latter were partially truncated upon density modification. The three Toll5A molecules observed are referred to as chain A (light blue), chain B (dark blue) and chain C (cyan). The Spz1C homodimer is composed of chain E (yellow) and chain F (orange), attached to Toll5A chain C at the concave side and to chain B at its “back” or convex side in superimposable conformations in both the heterodimer and trimer.

Structure of Toll5A ectodomain and how it compares to Drosophila Toll1.

The structure of Toll5A homodimers (chain A and chain B) was solved at the highest overall resolutions between 3.4 to 4.4 Å, with a head-to-head arrangement maintaining the C-terminal LRR domains far apart (Fig. 1A). The Ca atoms of Cys-783 at the C-termini of each receptor chain are separated by about 200 Å. Such a structural arrangement –if sterically possible when the receptor is expressed on the same cell– would ensure that Toll5A is locked in an inactive state preventing TIR domain association and signalling. If receptors are situated on neighbouring cells, such an orientation might be relevant to cell adhesion.

The overall structure of Toll5A is comparable to DmToll1 with a conserved number of leucine-rich repeats in both the N- and C-terminal domains, and conserved cysteine-rich capping structures surrounding these domains. The sequences of Toll5A and DmToll1 are 30 % identical and 50 % similar. In particular, the extended N-terminal cap formed by two hairpins and a parallel beta-sheet, which is involved in Spz binding, is conserved in A. aegypti Toll5A but is not found in Drosophila Toll-5, also known as Tehao, which remains an orphan receptor. Superimposition Toll5A and DmToll reveals that the diameter of the N-terminal LRR solenoid of Toll5A is 5-10 Å greater compared to the Drosophila Toll1 receptor, with an inner diameter of about 50 Å and an outer diameter of 90 Å. Toll5A has six glycosylation sites that restrict access to its surface. In contrast, DmToll has thirteen sites. Hence, most glycosylation sites are not conserved, except for Asn 481 on the ascending flank of LRR15 and Asn 521 on the descending flank of LRR17. Glycans restrict protein-protein contacts on the concave surface and the flanks of the receptor. In the homodimer, LRRNT1 cap residues between Thr-34 and Tyr-79 bind the concave surface in the vicinity of residues Tyr-202 (LRR4) to Asn-419 (LRR13). In particular, Tyr 56 is hydrogen bonded to Tyr 226 for each chain (Fig. 2). The
buried surface area is about 1,000 Å². This surface is also used for Spz ligand binding suggesting a direct competition between receptor-receptor and receptor-ligand interactions. The presence of unliganded receptor in the Toll5A-Spz1C sample suggests that formation of the complex is reversible.

Spz1C binding breaks receptor symmetry and brings the receptor juxtamembrane regions into proximity.

We observed a ligated heterodimer of Toll5A with a stoichiometry of two receptor ectodomains to one Spz1C ligand (2:1 complex) at a lower resolution between 4.2 to 8.4 Å. Spz1C is a covalent dimer stabilised by two intermolecular disulphide bonds instead of one in Drosophila Spz. The central disulphide bond between Cys-94 residues is conserved in DmSpz, while Spz1C has an additional intermolecular bond between the Cys-59 residues located basally in an area directly involved in binding to the N-terminal concave surface of the receptor. Cys-59 in chain F of Spz1C is located within hydrogen bonding distance (3.4 Å) of the hydroxyl group of Toll5A Tyr-226 in chain C LRR5 (Fig. 2B). The concave interface formed by the N-terminus up to LRR7 buries only ~ 890 Å² of accessible surface area, which represents less than half that observed in the Drosophila complex. Nevertheless, Spz1C adopts an asymmetric binding mode, in which chain F contributes about 610 Å² and chain E, 280 Å² to the concave binding site (Fig. 2B). It overlaps but does not completely match the site of Drosophila Toll1 and Spz. The interactions of Spz1C chain F at the concave site involves mostly residues from the first and third β-strand of Spz1C (Gln-14 and Leu-16, and Tyr-71, respectively) with the LRRNT1 of Toll5A (chain C: Tyr-79, Tyr-65 and His-54, respectively). Hydrogen bonds occur between His-18 in the Spz1C Trp-loop and the Phe-60 main-chain carboxy group in the receptor. Spz1C Tyr-38 and Toll5A Tyr-56 interact with side chain-mediated contacts (Fig. 2B). There is a salt bridge between Arg-156 in Toll5A LRR2 and Glu-11 in Spz1C chain F. The same residue in chain E is solvent-exposed and not well resolved in density. Glycans linked to Asn-521 on LRR17 restrict Spz1C spatially in the vicinity of Arg-50 and Gln-51 in chain E (Fig. 2C).

The overall structure of Spz1C is very similar to DmSpz, if the flexible Trp-loops between residues His-18 and Gln-40, and the β-wings between residues Tyr-71 and Val-89, are omitted. Indeed, the R.M.S.D of superimposed Cys-knot domains is about 2.2 Å over 53 Ca atoms in the absence of these flexible regions. Interestingly, in the mosquito Toll5A-Spz1C complex, these loops define asymmetric contacts at the dimerization interface with the “back” of chain B (Fig. 2C). Chain E of Spz1C binds extensively to the dimerization interface with over 1,300 Å² of its accessible surface area, while chain F contributes only 180 Å², confined to a loop that protrudes from the convex side of LRR14. Remarkably, the integrity of the LRR14 loop seems to determine the dynamic interactions between Toll5A and Spz1C, while its cleavage stabilises the complex (Fig. S3). This property is reminiscent of the Z-loop of nucleic-sensing TLRs. We will therefore refer to it as the Z-loop with conserved Asn and Asp residues suitable for Asparagine endopeptidase (AEP) processing, while noting the absence of a cathepsin site in Toll5A.

In a crystal structure of DmSpz, the β-wings are displaced from the Cys-knot framework by 90° (Hoffmann et al., 2008). By contrast in our cryo-EM structure, the β-wings adopt a different conformation contacting the Z-loop in a pincer like grasp (Fig. S3A). As most side chains are poorly resolved in these particles we generated a complete structure by side chain modelling, which is suitable for surface analysis (Fig. S2). This analysis suggests that the electrostatic charge distribution and shape complementarity are suitable for receptor-ligand coupling upon conformational selection. The dimerization interface of 1,480 Å² is one and a half times the area of the concave binding surface, suggesting a higher affinity binding compared to the
concave interface, while the absence of particles with a stoichiometry of 1:1 may indicate a sequential mechanism of assembly.

Interestingly, the N-terminal region of Toll5A chain B deviates by an angle of ~ 60°, compared to the axis traversing chain C from N- to C-terminus (Fig. 1). There is evidence of extra density in the original 2:1 heterodimer map that matches the dimensions of a Cys-knot domain without its loops. However, there is no possible contact at the dimer interface for the second ligand in such a configuration. Hence, the asymmetric disposition of the receptor chains is the molecular basis of negative cooperativity in the system, where binding of the first ligand makes binding of a second molecule energetically unfavourable.

**Toll5A ectodomain competes with ligand binding.**

A third particle was observed with a stoichiometry of 3 receptors and 1 ligand (Fig. 1). It is a hybrid of the homodimer (Toll5A chains A and B) and the 2:1 Toll5A-Spz1C heterodimer (Toll5A chains B and C) leading to a trimer of A, B and C chains, with Spz1C bound to the back of chain B and the concave side of chain C without any further contacts with chain A. Some minor differences are observed in main-chain and side-chains positions, which might be due to flexibility in both receptor and ligand molecules and the intermediate resolution of the heterotrimer map at 3.7 - 4.7 Å, compared to the homo- and heterodimer. While binding of a first Spz1C ligand shifts the receptor to adopt an asymmetric dimer configuration, the second binding site remains predominantly associated to another receptor chain.

**Receptor specificity is achieved with low ligand binding affinity**

The interactions between Spz1C, and the ectodomain of Toll5A were characterised using a range of biophysical techniques. We used surface-plasmon resonance to measure the kinetics of association and dissociation between mosquito and *Drosophila* Toll and Spz proteins (Fig. 3A and S4) and found that Toll5A only binds Spz1C. Surprisingly, *A. aegypti* Spz1A did not bind Toll1A, despite being the closest structural homologues of DmToll1 and DmSpz1 in the mosquito. Spz1A did not bind Toll5A, and neither did DmSpz1. Furthermore, Spz1C did not bind Toll1A or DmToll1, suggesting receptor-ligand specificity. By contrast, Toll5A and Spz1C interact with a *K*_D value of ~ 2μM whereas *Drosophila* Toll binds Spz with a much higher affinity of 30-80 nM, consistent with previous studies. Hence, the Toll5A-Spz1C complex is species- and paralogue specific despite being of low affinity.

**Ligand binding triggers a conformational change within Toll5A dimers in solution**

We then used SEC-MALS experiments in the presence and absence of Spz1C to characterise the stoichiometry of the receptor and its complex. This technique revealed a concentration-dependent shift in the stoichiometry of the receptor, from a monomer to a homodimer in the absence of ligand when the concentration was increased from 20 to 50 µM prior to size-exclusion chromatography (Fig. 3). Complexes were prepared in the presence of a three molar excess ligand and gel-filtered before analysis. In the presence of 20 µM Toll5A saturated with Spz1C, a complex formed that was polydisperse but at 50 µM Toll5A saturated with Spz1C, the mixture appeared monodisperse with a mass corresponding to 206 kDa. SEC-MALS indicates a mass of 175 kDa for the unbound receptor dimer so the 206 kDa form is consistent with a 2:1 complex of Toll5A and a disulphide linked Spz1C dimer (26 kDa).

We then explored this 2:1 complex further using SEC-SAXS at 50 µM prepared as above and compared it to the receptor homodimer at the same concentration and in the same buffer. SAXS allows rapid assessment of structural changes in response to ligand binding and can quantitatively characterize flexible molecules. The Guinier plot demonstrates the aggregation-
free state of the receptor and its complex (Fig. S5). As one of the few structural techniques amenable to dynamical systems, SAXS analysis suggests that the 2:1 complex is less flexible than the homodimer, despite its slightly larger dimensions (Table S2). Hence, the receptor likely undergoes conformational changes upon ligand binding.

**Toll5A-Spz1C complexes undergo dynamic exchange.**

The linearity of the Guinier plot (Fig. S5) does not ensure the ideality of the sample, hence further direct methods were used in solution and under native conditions. We used Analytical Ultracentrifugation (AUC) to determine relative concentrations, sedimentation coefficients, molecular weight and shape (frictional ratio) of Toll5A and Toll5A-Spz1C complexes (Fig. S6). These experiments reveal that Toll5A ectodomain alone is in equilibrium between monomers (≤ 3 μM) and dimers that prevail at concentrations ≥ 30 μM (Fig. 4A). The concentration-dependent stoichiometry differs from SEC-MALS, for which monomeric receptor was observed at 20 μM, most likely as a consequence of sample dilution or the effect of the matrix on protein-protein interactions during the size-exclusion chromatography step that precedes MALDS.

The shape of the AUC curves indicates that receptor-receptor interactions undergo slow exchanges, defined by discrete peaks at 5.6S, and 7.3S, respectively. In contrast, the presence of Spz1C at equimolar concentration causes the formation of Toll5A-Spz1C complexes with sedimentation coefficients ranging between 5.5 and 8.6S (Fig. 4B). Toll5A forms heterogenous complexes with its ligand, which may include a 1:1 complex at 6.7 S, as well 7.3 S (homodimers) and 7.8 S species (2:1 heterodimers) in different conformational states. However, in the presence of excess Spz1C, better resolved molecular species sediment as two distinct populations of 6.7 S and 8.5 S. These complexes are reminiscent of the 1:1 and 2:2 complexes formed by DmToll and Spz in similar experimental conditions27,29.

To determine whether these conformers are present for mosquito proteins at more physiological concentrations of ligand, we measured the composition of Toll5A-Spz1C mixtures at nanomolar concentrations using mass photometry (Fig. 4C-D). Our experimental setting was able to measure accurately molecules and complexes above 60 kDa, while free Spz1C was below the threshold, which resulted in erroneous mass determination. At 25 nM Toll5A in the presence of 25 nM Spz1C, Toll5A is predominantly monomeric but a 2:1 complex is detected along a species at 48 kDa, corresponding most likely to Spz1C. By contrast, at 50 nM Toll5 in the presence of 50 nM Spz1C, 1:1 ligated monomer and 2:2 heterodimer complexes appear in addition to Toll5A monomer and 2:1 complex. Hence, mass photometry detects stoichiometric conversions in non-equilibrium conditions. Toll5A homodimer is not detected consistent with the AUC data that suggests a KD value for this interaction in the μM range. None of these techniques detected the heterotrimer, which is therefore most likely a side effect of cryo-EM’s capacity of visualising transition intermediates.

**Spz1C activates the production of antimicrobial peptides in Aag2 cells.**

In order to assay the activity of Spz1C, we stimulated the *Aedes aegypti* cell line Aag2, which constitutively expresses Toll5A (Fig. S7), with either full-length Spz1C proprotein or processed forms and Gram-negative bacteria for IMD pathway stimulation, as a positive control for a potent innate immune response. RT-qPCR was used to measure the induction of a range of antimicrobial peptides.

Aag2 cells do not up-regulate antimicrobial peptides (AMP) upon stimulation with pro-Spz1C. In contrast, processed Spz1C potently stimulated the production of several antimicrobial
peptides including Defensin 1, Cecropin A, Glycine-rich repeat protein (GRRP) holotricin and Attacin B, (Fig. 5A). According to RT-PCR, we found all members of the Spz family to be constitutionally expressed in Aag2 cells, except Spz2 (Fig. S7). Given that this cell line was shown to be unresponsive to microbial stimulation of the Toll pathway\textsuperscript{45,46}, our result endorses full Toll signalling capacity when provided the active form of the ligand. Aag2 cells may therefore be defective either in microbial sensing or within the protease cascade activating Spz.

Upon activation, we found that Spz1C signalling overlaps but is distinct to that induced by a soluble extract of heat inactivated Gram-negative bacteria (GNB), which in addition activates the production of Gambicin, a general-purpose AMP. Neither GNB or Spz1C triggers the production of Vago, which is regulated by the RNA interference pathway, or Diptericin, an AMP potently induced by GNB in \textit{Drosophila} via the IMD pathway\textsuperscript{47}.

Next, we assayed production of GRRP holotricin, an AMP that is strongly induced by Spz1C, in a dose-response experiment (Fig. 5B). Activation of holotricin expression occurs over a wide range of concentrations with an EC\textsubscript{50} at sub-nanomolar concentrations. Furthermore, Spz1C signalling displays the same hallmarks of negative cooperativity observed in the \textit{Drosophila} pathway, with 10 % to 90 % maximal signalling requiring an increase in ligand concentration of about 600-fold.

Interestingly, we found that different Spz paralogues have different signalling activities, as suggested by the capacity of \textit{Aedes aegypti} Spz5 to potently activate gambicin (GAM) (Fig. 5C). While Drosophila Spz1 is able to stimulate GAM moderately in Aag2 cells, this property is not shared by Spz1C. In \textit{Drosophila}, Spz5 is recognised by Toll1, Toll6 and Toll\textsuperscript{78–10}. Given that Aag2 cells express other members relating to the Toll family of receptors (Fig. S7), it is conceivable that one of them triggers GAM activation reflecting its ligand specificity as opposed to the documented promiscuity of the Drosophila system.

**Spz1C regulates expression of genes involved in immunity and homeostasis**

We have used RNASeq to define the transcriptomic signature linked to Spz1C and compare it to activation of the immune system by Gram-negative bacteria and purified DAP-PGN, which in \textit{Drosophila} activates the IMD signalling pathway. As shown in Fig. 6 there are 85 genes regulated by Spz1C, of which 83 are up-regulated and only 2 are down-regulated. These include GRRP holotricin, confirming its role as a marker of Spz1C, as well as other molecules linked to complement and immunity. Many genes encoding Clip Serine proteases, Serpins and Gram-negative Binding Protein (GNBP) are upregulated. In \textit{Drosophila} these proteins function in peptidoglycan recognition and proteolytic cascades that activate Spz.

The Patched 1 (Ptc1) 7-TM receptor is strongly induced with potential implications for tissue regeneration\textsuperscript{48}. The ligand for Ptc1 is Hedgehog, a morphogen involved in embryonic segmentation. This signalling pathway also plays a crucial role in maintaining adult tissue homeostasis. It is not known whether Ptc1 functions in insect immunity. Another upregulated gene encodes a putative ecdysone inducible protein ortholog L2 and is homologous to \textit{Drosophila} Imaginal morphogenesis protein-Late 2 (IMPL2)\textsuperscript{49}, which is involved in the regulation of metabolism, growth, reproduction and lifespan. Two genes are down-regulated: Fatty Acyl-CoA Reductase (FAR) and the prostaglandin EP4 receptor. Interestingly, eicosanoid biosynthesis is controlled by the Toll pathway in some insects and also human PGE2 triggers a negative feedback loop, in which TLR4 signalling is restricted (see discussion).

**Bioinformatic analysis supports neofunctionalization.**
In mosquitoes, immune-related gene families involved in pattern recognition and effector activity have increased compared to *Drosophila* as a result of gene duplication and family expansion. In contrast, cytosolic signal transducers such as MyD88 tend to have a stricter orthology. At the interface between both are transmembrane receptors and their ligands. Spz1 is the only ligand that has diversified in *A. aegypti* with genetic evidence pointing towards a specific role for Spz1C in mosquito immunity. In particular, Spz1C together with Toll5A mediate anti-fungal immunity. More importantly, Dengue virus upregulates Spz1C in the midgut and Toll5A in the salivary glands, hinting at a role in vector-virus interactions.

At a protein level, mosquito Toll-1 and 5 receptors are between 32 – 41 % identical to *Drosophila* Toll (DmToll-1) and between 36 to 53 % identical within the *Aedes* Toll-1/5 paralogue group. However, the extracellular domains alone are less well conserved (27 – 35 %). This contrasts with the non-duplicated receptors, which have sequence identities typically over 60 %. Upon closer inspection, *Aedes aegypti* Toll-1A, 1B and 5A have a precisely conserved number of leucine-rich repeats (LRRs) and Cys-rich capping structures, comparable to the prototypical *Drosophila* Toll-1 receptor (Fig. S8A). On the other hand, Toll5B differs, with the gene being located on chromosome III and a shorter ectodomain with 15 LRRs instead of 17 at the N-terminus. The duplicated mosquito Toll-5 receptors also differ markedly from the *Drosophila* orthologue DmToll-5 called Tehao, which has 8 N-terminal LRRs and fewer cysteine residues. No ligand for DmToll-5 has been identified so far. We note that its ectodomain lacks features involved in DmToll-1 ligand binding. Nevertheless DmToll-5 can induce the production of antimicrobial peptides when overexpressed in cell culture and forms heterodimers with DmToll-1.

As for the Toll ligand, the duplicated Spz1 genes in *Aedes* have also diverged considerably from *Drosophila* with only 21-31% sequence identity among orthologues. Non-duplicated Spz vary from 35 % for Spz2 to 89 % for Spz6. The C-terminal active fragment containing the Cys-knot domain follows the same trend and displays sequence identities around 64-88 % for 1-to-1 orthologues, and 25-33 % for duplicated Spz1. The direct orthologue of DmSpz1 is predicted to be SpzX in *A. aegypti*. SpzX also conserves the patterns of alternative splicing seen for DmSpz1 including transcripts with a truncated Cys-knot (SpzX-G) (Fig. S8B). However, SpzX represents the amalgamation of two different Cys-knot sequences previously annotated Spz1A and Spz1B, now gathered in one locus that undergoes alternative splicing. Remarkably, four splice forms SpzX-D/E/I/J, have Cys-knot domain sequences identical to Spz1A but gained an additional C-terminal furin cleavage site. SpzX-D and E acquired an additional furin site in the pro-domain (not shown). SpzX-F, Spz1C and DmSpz1 do not have furin recognition sites, which are the hallmark of neurotrophins, such as Spz2 and Spz5.

Structurally, SpzX-F/C/H have lost the conserved Cys residue involved in the formation of the intermolecular disulphide bonding (Fig. S8B). Instead, they have a cysteine residue in an alternative location that is predicted to form an inter-molecular disulphide bond located at the ligand-receptor interface. This residue is shared by Spz3/4/5 and Spz1C. The structure of Spz1C confirms the presence of two intermolecular disulphide bridges that constrain the covalent cystine-knot dimer doubly.
Interestingly, OrthoMCL-DB congregated *Aedes* SpzX and *Drosophila* Spz1 genes into an ortholog group based on their sequence similarity (OG6_117926). Spz1C however is excluded from this group and forms a group on its own (OG6_426965). Next, we checked if Spz1C was restricted to the *A. aegypti* lineage. When this study was initiated, it seemed to have no counterpart in any of the sequenced organisms. It therefore appeared to be a lineage-restricted gene. *A. albopictus*, which diverged only ~71 Mya from *A. aegypti*, has recently been re-sequenced. The updated genome of *A. albopictus* contains 3 copies of Spz1C-like genes at loci LOC109412940, LOC109424625 and LOC115260685. Hence, Spz1C has undergone *Aedes*-specific duplication and a rapid expansion in *A. albopictus*, whereas Toll1/5 sequences all appear to cluster in the same ortholog group (OG6_106857).

**Discussion**

In this study we show that the haematophagous mosquito *A. aegypti* has a duplicated Toll ligand, Spz1C, apparently unique to the *Aedes* genus, that specifically activates the Toll5A paralogue with low affinity. We provide direct evidence of their interaction using purified proteins and show that ligand binding may be restricted by receptor self-association in a dose-dependent manner, resulting in conformational intermediates that we visualise by single particle cryo-EM. Functionally, Spz1C and Toll5A are known sensors of fungal infections in the fat body of adult mosquitoes. Interestingly, anti-dengue activity has been assigned to entomopathogenic fungus *Beauveria bassiana*, and to the Toll pathway, without elucidating the potential roles of individual Toll receptors or their respective ligands.

Here we provide evidence of a robust Spz1C response in Aag2 cells that differs from its *Drosophila* orthologue and mosquito paralogue. We also note that duplicated Toll and Spz are more divergent in sequence than 1-to-1 orthologues, with Spz1 the only duplicated cytokine within its family, while duplication occurred for members of the Toll family including Toll1, Toll5 and Toll9 and expansion of the Toll6/7/8 family with two additional members, Toll10 and Toll11 with potential anti-plasmodial properties. *Aedes aegypti* diverged from *Anopheles gambiae*, which is the main vector for malaria, about 217 Myr ago and adapted to become an efficient vector for flaviviruses such as Dengue, as evident from the large number of non-retroviral integrated RNA viruses present in the *A. aegypti* but not the *A. gambiae* genome. Interestingly, the Asian tiger mosquito *A. albopictus*, which diverged from *A. aegypti* more recently, further expanded Spz1C genes. In contrast, more divergent species from the *Anopheles* and *Culex* genera, which are not (or less) competent to vector these viruses, do not possess this paralog. Assuming the available mosquito genomes have been reliably sequenced, the presence of Spz1C genes might therefore be an indicator of vectorial competence enabling flaviviruses, such as DENV to exist symbiotically at high titre in mosquitoes.

Our biophysical characterisation shows that Spz1C binds specifically to Toll5A with micromolar affinity compared to the nanomolar binding of DmSpz1 by DmToll1. DmSpz1 is also able to bind promiscuously to DmToll6 and DmToll7 and so it may be that low affinity binding confers signalling specificity for Toll5A. Consistent with this idea the *Aedes* Spz5 paralog but not Spz1C activates the AMP gambiaicin in Aag2 cells as strongly as bacterial extracts while DmSpz1 induces a partial activation (Fig. 5C). Thus *A. aegypti* has evolved a tiered Toll mediated immune system compared to *Drosophila* where Toll1 alone fulfils most immune functions.

Low affinity ligand binding and its ramifications in terms of signalling has been extensively characterised for mammalian cytokine receptors for interferons and interleukin ligands. Such cytokines display pleiotropic effects while inducing a spectrum of redundant and yet
distinct cellular functions. Receptor-ligand association and dissociation rates (kon and koff) have been found to be key in determining signalling outcomes. On-rates determine the amount of STAT transcription factor activation upon controlling the number of ligand-receptor complexes formed at the plasma membrane. In contrast, off-rates correlate to the kinetics of STAT activation depending on the half-life of ligand-receptor complexes. Alternatively, cell surface abundance of cytokine receptors plays a major role in triggering different transcriptional programs and cell fates, through obeying the mass action law and titrating cytokine concentrations. Our hypothesis in the case of Aedes Toll5A and Spz1C is that cell fate will be regimented in two ways depending on receptor and ligand concentrations: (i) ligand concentration will change the intensity of the signal (increased Spz1C for increased antimicrobial peptide production as illustrated in Fig.5); and (ii) increased receptor density will promote inhibitory self-association leading to immune quiescence.

The structures we present provide a plausible molecular explanation for negatively cooperative signalling. Binding of the first Spz1C homodimer to form the asymmetric complex can occur transiently at relatively low ligand concentrations but the 2:2 form would be even lower affinity, and hence only form at higher ligand concentrations. There are two classical theoretical models of sequential binding that lead to negatively cooperative receptor signalling as initially proposed by Koshland in the 1960’s. Thus, binding of the first monomeric ligand to a dimeric receptor partially activates signalling but full activation requires the less favourable binding of the second ligand monomer. In the second model the binding of two monomeric ligands to form a 2:2 complex is required for signal transduction. The two models lead to somewhat different theoretical stimulus-response curves and our experiments appear to fit better with Model 1 for the production of antimicrobial peptides (Fig. 5). If the ligand is limiting there is an ultrasensitive response with a pronounced threshold, more reminiscent of positive than negative cooperativity. In our cell culture assays, Spz1C is likely not limiting, however this phenomenon may be relevant for development and homeostasis controlled by Spz in the whole organism.

The activation profile of Spz1C differs from Drosophila Spz1 and Aedes Spz5, despite their shared cystine-knot fold, having a transcriptomic profile that indicates functions that go beyond immunity. Spz1C specifically activates the production of Hedgehog receptor, insect insulin-binding protein L2 and genes involved in reprogramming fatty acid metabolism. If Spz1C and Toll5A were both upregulated in the salivary glands, and achieved the double ligated state as suggested by Bonizzoni et al., it is also conceivable that Spz1C and its receptor could be involved in viral tolerance by promoting lipid metabolism and homeostasis. The mechanism of action may differ in the midgut where Spz1C expression is induced to limit viral invasion, an immunological role. By contrast in the salivary gland, where Toll5A up-regulation may increase avidity for the ligand, a double ligated and hence fully saturated receptor complex might lead to a signal that differs from low density single-ligated receptors. Hence, future studies need to tackle the signalling capacities of 2:1 and 2:2 complexes via detecting the recruitment of MyD88 first, and if so, by checking if different sets of genes are controlled by both type of complexes. Fluorescence microscopy experiments will be best suited to detect such complexes along the presence of unliganded receptor homodimers to distinguish between these models.

Remarkably and despite a significantly lower affinity, Spz1C is as potent in stimulating mosquito Aag2 cells, as DmSpz is at stimulating Drosophila S2 cells. An alternative explanation is that cleavage of the Z-loop promotes stable ligand binding and thus processing might account for the discrepancy between protein affinity and cellular potency of Spz1C. On
the other hand, Spz1C binding prevents cleavage of Toll5A of the Z-loop, a finding that is not consistent with the Z-loop impeding ligand binding. Additionally, we expressed Toll5A in insect cells without detecting such processing, which likely rules out its spontaneous occurrence in our view. However, Drosophila Toll1 can be cleaved at an equivalent residue, Asp 458 and forms a stable dimer with the ectodomain remaining intact27. Taken together, we conclude that this region is critical for promoting ligand-induced dimerization while the physiological relevance of its processing requires further examination. Of note, the nucleic-acid binding TLRs that are also activated by endoproteolytic cleavage display positive cooperativity61.

In light of the above it is likely that there will be a degree of synergy between Toll5A signalling and the vectorial capacity of Aedes aegypti. Our study sheds new light on Toll signalling, while raising fundamental questions. Do Toll receptors undergo endoproteolytic processing to regulate their activity? Which oligomeric forms occur in vivo and what are their respective signalling outputs? More importantly, can Spz1C signalling be exploited to fight mosquito-borne diseases? If so, the structural data in hand can guide future transmission-blocking strategies.

**Figures.**

![Figure 1: Cryo-EM reveals three particles with variable stoichiometries that reflect the dynamic nature of the system. 2:0 ligand-free Toll5A homodimer (A), 2:1 single-ligated](image)
heterodimer (B), and (C) 3:1 single-ligated heterotrimer. Density is colour-coded according to protein content. A different orientation rotated by 90° along a vertical axis defined by the dimerization axis is shown for the 3:1 (D) and the 2:1 complexes (E). The latter is also shown with a top view to emphasize the deviation of chain B by 60° from the dimerization axis (F). This configuration prevents a poorly resolved second ligand (grey density) bound at the concave side of chain B to contact the dimerization interface.

**Figure 2:** Receptor-receptor interactions compete with receptor-ligand interactions between Toll5A and Spz1C. N-terminal receptor-receptor interactions at LRRNT1-LRR14 with accessible LRR14 Z-loops (A) directly compete with ligand binding at the concave side of LRRNT1-LRR7 (B). Spz1C mediates extensive contacts between its β-wings and Toll5A Z-loop between LRR14-LRR15, with the Trp-loop of the distal chain E nesting between LRR domains (C). The Trp-loop of the proximal chain F caps the β-wings and participates in Z-loop interactions.
Figure 3: Spz1C low affinity binding to Toll5A homodimer decreases protein complex flexibility. SPR binding analysis (A). Sensorgrams of Toll5A (acronym, T5A) run over a range of concentrations (0.1 – 7.5 μM) on an amine coupled Spz1C-chip (acronym, S1C). SEC-SAXS dimensionless Kratky plot analysis at 50 μM (B). SEC-MALS analysis at 20 μM (C) and 50 μM (D) of Toll5A on its own (blue) and in complex with Spz1C (orange). Both MALS and SAXS were carried out upon loading ~50 μl samples at the given concentrations onto a Superose 6 size-exclusion chromatography column (GE Healthcare) in 50 mM Tris pH 7.5, 50 mM NaCl.
Figure 4: Slow exchanges in receptor-receptor interactions in the absence of ligand contrast with fast dynamics in the presence of ligand. AUC sedimentation velocity profiles and the shifts in sedimentation coefficients are indicative of dynamical behaviour or the receptor in the presence and the absence of its ligand (A-B). Mass photometry (C-D) reveals the concentration-dependent stoichiometries of Toll5A and Spz1C. The single-ligated Toll5A dimer is detected at the lowest measurable concentration of 25 nM (C), whereas saturated dimer and ligated monomer appear at 50 nM and above (D).
Figure 5: Spz1C activates a set of antimicrobial peptides (A), over a large range of concentrations (B), and in an isoform-specific manner (C). Expression of antimicrobial peptides in Aag2 cells at 12 hours after stimulation with Spz1C (proS1C and S1C: before and after cleavage of the pro-domain, respectively) upon addition of purified protein in the media and RT-qPCR. Expression values were determined using the ΔΔCT method with normalisation to mRNA levels of the eEFG1a housekeeping gene. The grey line indicates a value of Fold Change=2 as we determined to the limit to have change in expression profile. Statistical analyses were performed using Student T-test or two-sided Kruskal–Wallis test (according to application condition of each) to compare result versus a FC >2. Data are mean ± s.e.m. (n=3). (B) Kinetics curves of GRRP expression in Aag2 cells after stimulation with different concentration of Spz1C (S1C). The fitted curve (in orange) represents smoothed conditional means calculated by ggPlot2 on R. The grey line indication a FC=2 and Data are mean ± s.e.m. (n=3). Heat-inactivated E.coli extract (GNB) was used as a positive control and conditioning buffer was used as a negative control.
**Figure 6:** Spz1C transcriptional signature in Aag2 cells includes innate immune and metabolic genes that confer homeostasis and possibly, viral tolerance. RNAseq analysis. (A) Genes exhibiting different response to stimulation with S1C heat map with indication of gene names normalized on MOCK stimulated Aag2 cells expression. These genes expression where also showed for Gram- Bacteria and PGN stimulation. (B) Enriched GO categories associated with DEGs shown in (C). BP-biological processes, CC-cellular components, MF-molecular functions. (D) Significantly up-regulated (in red) and downregulated mosquito genes (in blue) upon stimulation with S1C compared with mock. Labelled genes where unique to S1C condition compared to other condition. All statistical tests were performed on R with a p-value > 0.05. Stars indicate significance: * = p-value<0.05, ** = p-value<0.01, and *** = p-value<0.001.

**ACKNOWLEDGMENTS**

We thank Drs Katherine Stott, Biophysics Facility manager, and Paul Brear for training in biophysical techniques and continuous support; Drs Marko Hyvönen and Joseph Maman, for helpful discussions. We also thank Joe for training in SEC-MALS and Dr Nathan Cowieson, for supporting the SEC-SAXS at B21 Diamond Light Source. We thank Shilo Dickens, DNA sequencing facility manager, and her team, in particular Dr Markiyan Samborskyy, for RNAseq analysis. We thank Tom Dendooven for sharing the protocol for cryo-EM sample preparation used in this study, and Lee Cooper for Vitrobot training and cryo-EM grid preparation. This work is supported by a New Investigator grant from the Medical Research Council to M.G. (MR/P02260X/1).

**AUTHOR CONTRIBUTIONS**

M.G. and Y.S. designed biochemical experiments. M.G, S.G.S and Y.S. planned cell assays. M.G., Y.S., and T.H.W. purified and assembled complexes. M.C.K.T. expressed Spz5. M.C.M performed AUC experiments. S.H. and D.Y.C. collected and processed cryo-EM data. M.G. performed model building and structure refinement with the help of M.C.M. M.J.M. carried out MP experiments and analysed SEC-SAXS data. M.G. analysed structural and functional data and wrote the manuscript with N.J.G.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.

**REFERENCES**

1. Gay, N. J. J. N. J., Symmons, M. F. F. M. F., Gangloff, M. & Bryant, C. E. E. C. E. Assembly and localization of Toll-like receptor signalling complexes. *Nat. Rev. Immunol.* **14**, 546–558 (2014).
2. Hoffmann, J. A. The immune response of Drosophila. *Nature* **426**, (2003).
3. McIlroy, G. *et al.* SI - Toll-6 and Toll-7 function as neurotrophin receptors in the Drosophila melanogaster CNS. *Nat. Neurosci.* **16**, 1248–56 (2013).
4. Parker, J. S., Mizuguchi, K. & Gay, N. J. A family of proteins related to Spatzle, the toll receptor ligand, are encoded in the Drosophila genome. *Proteins* **45**, 71–80 (2001).
5. Arnot, C. J., Gay, N. J. & Gangloff, M. Molecular mechanism that induces activation of Spatzle, the ligand for the Drosophila toll receptor. *J. Biol. Chem.* **285**, 19502–19509 (2010).
6. Wang, L. et al. Peptidoglycan recognition protein-SD provides versatility of receptor formation in Drosophila immunity. *Proc. Natl. Acad. Sci. U. S. A.* **105**, (2008).
7. Jang, I. H. et al. A Spätzle-processing enzyme required for toll signaling activation in drosophila innate immunity. *Dev. Cell* **10**, (2006).
8. Foldi, I. et al. Three-tier regulation of cell number plasticity by neurotrophins and Tolls in Drosophila. *J. Cell Biol.* **216**, (2017).
9. Chowdhury, M. et al. Toll family members bind multiple Spätzle proteins and activate antimicrobial peptide gene expression in Drosophila. *J. Biol. Chem.* (2019). doi:10.1074/jbc.RA118.006804
10. Nonaka, S. et al. Characterization of Spz5 as a novel ligand for Drosophila Toll-1 receptor. *Biochem. Biophys. Res. Commun.* (2018). doi:10.1016/j.bbrc.2018.10.096
11. Shelly, S., Lukinova, N., Bambina, S., Berman, A. & Cherry, S. Autophagy Is an Essential Component of Drosophila Immunity against Vesicular Stomatitis Virus. *Immunity* **30**, 588–598 (2009).
12. Nakamoto, M. et al. Virus Recognition by Toll-7 Activates Antiviral Autophagy in Drosophila. *Immunity* **36**, 658–667 (2012).
13. Nene, V. et al. Genome sequence of Aedes aegypti, a major arbovirus vector. *Science* **316**, 1718–23 (2007).
14. Chen, X. G. et al. Genome sequence of the Asian tiger mosquito, aedes albopictus, reveals insights into its biology, genetics, and evolution. *Proc. Natl. Acad. Sci. U. S. A.* (2015). doi:10.1073/pnas.1516410112
15. Christophides, G. K. et al. Immunity-related genes and gene families in Anopheles gambiae. *Science* **298**, 159–65 (2002).
16. Gay, N. J. & Gangloff, M. Structure and function of Toll receptors and their ligands. *Annu Rev Biochem* **76**, 141–165 (2007).
17. Gay, N. J., Gangloff, M. & Weber, A. N. Toll-like receptors as molecular switches. *Nat Rev Immunol* **6**, 693–698 (2006).
18. Keith, F. J. & Gay, N. J. The Drosophila membrane receptor Toll can function to promote cellular adhesion. *EMBO J.* **9**, 4299–4306 (1990).
19. Uff, S., Clemetson, J. M., Harrison, T., Clemetson, K. J. & Emsley, J. Crystal structure of the platelet glycoprotein Ibα N-terminal domain reveals an unmasking mechanism for receptor activation. *J. Biol. Chem.* **277**, (2002).
20. Krantz, D. E. & Zipursky, S. L. Drosophila chaoptin, a member of the leucine-rich repeat family, is a photoreceptor cell-specific adhesion molecule. *EMBO J.* **9**, (1990).
21. Rose, D. et al. Toll, a muscle cell surface molecule, locally inhibits synaptic initiation of the RP3 motoneuron growth cone in Drosophila. *Development* (1997).
22. Inaki, M., Shinza-Kameda, M., Ismat, A., Frasch, M. & Nose, A. Drosophila Tey represses transcription of the repulsive cue Toll and generates neuromuscular target specificity. *Development* (2010). doi:10.1242/dev.046672
23. Nose, A., Mahajan, V. B. & Goodman, C. S. Connectin: A homophilic cell adhesion molecule expressed on a subset of muscles and the motoneurons that innervate them in Drosophila. *Cell* (1992). doi:10.1016/0092-8674(92)90426-D
24. Kang, J. Y. & Lee, J. O. Structural biology of the toll-like receptor family. *Annu. Rev. Biochem.* **80**, (2011).
25. Liu, L. et al. Structural basis of toll-like receptor 3 signaling with double-stranded RNA. *Science* **320**, 379–81 (2008).
26. Tanji, H., Ohto, U., Shibata, T., Miyake, K. & Shimizu, T. Structural reorganization of the Toll-like receptor 8 dimer induced by agonistic ligands. *Sci. (Washington, DC, United States)* **339**, 1426–1429 (2013).
27. Weber, A. N. R., Moncrieffe, M. C., Gangloff, M., Imler, J. L. & Gay, N. J. Ligand-
receptor and receptor-receptor interactions act in concert to activate signaling in the Drosophila toll pathway. *J. Biol. Chem.* **280**, 22793–22799 (2005).

28. Lewis, M. *et al.* Cytokine Spatzle binds to the Drosophila immunoreceptor Toll with a neurotrophin-like specificity and couples receptor activation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20461–6 (2013).

29. Parthier, C. *et al.* Structure of the Toll-Spatzle complex, a molecular hub in Drosophila development and innate immunity. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 6281–6 (2014).

30. Shin, S. W., Bian, G. & Raikhel, A. S. A toll receptor and a cytokine, Toll5A and Spz1C, are involved in toll antifungal immune signaling in the mosquito Aedes aegypti. *J. Biol. Chem.* **281**, 39388–39395 (2006).

31. Ramirez, J. L. & Dimopoulos, G. The Toll immune signaling pathway control conserved anti-dengue defenses across diverse Ae. aegypti strains and against multiple dengue virus serotypes. *Dev. Comp. Immunol.***34**, 625–629 (2010).

32. Xi, Z. *et al.* The Aedes aegypti toll pathway controls dengue virus infection. *PLoS Pathog.* **4**, e1000098 (2008).

33. Dong, Y., Morton, J. C., Ramirez, J. L., Souza-Neto, J. A. & Dimopoulos, G. The entomopathogenic fungus Beauveria bassiana activate toll and JAK-STAT pathway-controlled effector genes and anti-dengue activity in Aedes aegypti. *Insect Biochem. Mol. Biol.* (2012). doi:10.1016/j.ibmb.2011.11.005

34. Wang, J. M. *et al.* Aedes aegypti HPX8C modulates immune responses against viral infection. *PLoS Negl. Trop. Dis.* (2019). doi:10.1371/journal.pntd.0007287

35. Luplertlop, N. *et al.* Induction of a Peptide with Activity against a Broad Spectrum of Pathogens in the Aedes aegypti Salivary Gland, following Infection with Dengue Virus. *PLoS Pathog.* **7**, e1001252 (2011).

36. Terwilliger, T. C., Ludtke, S. J., Read, R. J., Adams, P. D. & Afonine, P. V. Improvement of cryo-EM maps by density modification. *Nat. Methods* (2020). doi:10.1038/s41592-020-0914-9

37. Luo, C., Shen, B., Manley, J. L. & Zheng, L. Tehao functions in the Toll pathway in Drosophila melanogaster: Possible roles in development and innate immunity. *Insect Mol. Biol.* **10**, 457–464 (2001).

38. Hoffmann, A. *et al.* Biophysical characterization of refolded Drosophila spätzle, a cystine knot protein, reveals distinct properties of three isoforms. *J. Biol. Chem.* (2008). doi:10.1074/jbc.M801815200

39. Ewald, S. E. *et al.* The ectodomain of Toll-like receptor 9 is cleaved to generate a functional receptor. *Nature* **456**, 658–62 (2008).

40. Park, B. *et al.* Proteolytic cleavage in an endolysosomal compartment is required for activation of Toll-like receptor 9. *Nat. Immunol.* **9**, (2008).

41. Ewald, S. E. *et al.* Nucleic acid recognition by Toll-like receptors is coupled to stepwise processing by cathepsins and asparagine endopeptidase. *J. Exp. Med.* **208**, (2011).

42. Maschalidi, S. *et al.* Asparagine Endopeptidase Controls Anti-Influenza Virus Immune Responses through TLR7 Activation. *PLoS Pathog.* (2012). doi:10.1371/journal.ppat.1002841

43. Sepulveda, F. E. *et al.* Critical Role for Asparagine Endopeptidase in Endocytic Toll-like Receptor Signaling in Dendritic Cells. *Immunity* (2009). doi:10.1016/j.immu.2009.09.013

44. Weber, A. N. R. R. *et al.* Binding of the Drosophila cytokine Spätzle to Toll is direct and establishes signaling. *Nat. Immunol.* **4**, 794–800 (2003).

45. Zhang, R. *et al.* Regulation of antimicrobial peptides in Aedes aegypti Aag2 Cells. *Front. Cell. Infect. Microbiol.* (2017). doi:10.3389/fcimb.2017.00022
46. Russell, T. A., Ayaz, A., Davidson, A. D., Fernandez-Sesma, A. & Maringer, K. Imd pathway-specific immune assays reveal NF-κB stimulation by viral RNA PAMPs in Aedes aegypti Aag2 cells. bioRxiv (2020). doi:10.1101/2020.06.30.179879

47. Tanji, T., Hu, X., Weber, A. N. R. & Ip, Y. T. Toll and IMD Pathways Synergistically Activate an Innate Immune Response in Drosophila melanogaster. Mol. Cell. Biol. 27, (2007).

48. Lum, L. & Beachy, P. A. The hedgehog response network: Sensors, switches, and routers. Science (2004). doi:10.1126/science.1098020

49. Roed, N. K. et al. Structures of insect Imp-L2 suggest an alternative strategy for regulating the bioavailability of insulin-like hormones. Nat. Commun. (2018). doi:10.1038/s41467-018-06192-3

50. Waterhouse, R. M. et al. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. Science (80-. ). 316, 1738–43 (2007).

51. Dudchenko, O. et al. De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science (80-. ). (2017). doi:10.1126/science.aal3327

52. Palatini, U. et al. Improved reference genome of the arboviral vector Aedes albopictus. Genome Biol. (2020). doi:10.1186/s13059-020-02141-w

53. Redmond, S. N. et al. Association mapping by pooled sequencing identifies TOLL 11 as a protective factor against Plasmodium falciparum in Anopheles gambiae. BMC Genomics 16, 779 (2015).

54. Moraga, I. et al. Instructive roles for cytokine-receptor binding parameters in determining signaling and functional potency. Sci. Signal. 8, (2015).

55. Moraga, I., Harari, D., Schreiber, G., Uzé, G. & Pellegrini, S. Receptor Density Is Key to the Alpha2/Beta Interferon Differential Activities. Mol. Cell. Biol. 29, (2009).

56. Levin, D., Harari, D. & Schreiber, G. Stochastic Receptor Expression Determines Cell Fate upon Interferon Treatment. Mol. Cell. Biol. 31, (2011).

57. Levitzki, A. & Koshland, D. E. Negative cooperativity in regulatory enzymes. Proc. Natl. Acad. Sci. U. S. A. 62, (1969).

58. Koshland, D. E. The structural basis of negative cooperativity: Receptors and enzymes. Curr. Opin. Struct. Biol. 6, (1996).

59. Ha, S. H. & Ferrell, J. E. Thresholds and ultrasensitivity from negative cooperativity. Science (80-. ). 352, (2016).

60. Bonizzoni, M. et al. Complex Modulation of the Aedes aegypti Transcriptome in Response to Dengue Virus Infection. PLoS One 7, (2012).

61. Leonard, J. N. et al. The TLR3 signaling complex forms by cooperative receptor dimerization. Proc. Natl. Acad. Sci. U. S. A. 105, 258–63 (2008).

62. Bateman, A. et al. UniProt: The universal protein knowledgebase in 2021. Nucleic Acids Res. 49, (2021).

63. Giraldo-Calderón, G. I. et al. VectorBase: An updated Bioinformatics Resource for invertebrate vectors and other organisms related with human diseases. Nucleic Acids Res. 43, (2015).

64. Edgar, R. C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics (2004). doi:10.1186/1471-2105-5-113

65. Webb, B. & Sali, A. Comparative Protein Structure Modeling Using MODELLER. Curr. Protoc. Bioinformatics 47, 5.6.1-5.6.32 (2014).

66. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics (2009). doi:10.1093/bioinformatics/btp033

67. Pettersen, E. F. et al. UCSF Chimera--a visualization system for exploratory research
68. Gangloff, M. et al. Structural insight into the mechanism of activation of the toll receptor by the dimeric ligand Spätzle. *J. Biol. Chem.* **283**, (2008).

69. Schuck, P.Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* (2000). doi:10.1016/S0006-3495(00)76713-0

70. Laue TM, Shah BD, Ridgeway TM, P. S. Computer-aided interpretation of analytical sedimentation data for proteins. Analytical Ultracentrifugation in Biochemistry and Polymer Science. (Royal Society of Chemistry, Cambridge, UK, 1992).

71. Cowieson, N. P. et al. Beamline B21: High-throughput small-angle X-ray scattering at Diamond Light Source. *J. Synchrotron Radiat.* **27**, (2020).

72. Franke, D. et al. ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* (2017). doi:10.1107/S1600576717007786

73. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* (1992). doi:10.1100/0021889892001663

74. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* (2019). doi:10.1038/s41592-019-0580-y

75. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* (2017). doi:10.1038/nmeth.4169

76. Terwilliger, T. C., Ludtke, S. J., Read, R. J., Adams, P. D. & Afonine, P. V. Improvement of cryo-EM maps by density modification. *bioRxiv* (2019). doi:10.1101/845032

77. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2010). doi:10.1107/S0907444910007493

78. Kidmose, R. T. et al. Namdinator - Automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ* (2019). doi:10.1100/S2052252519007619

79. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D Struct. Biol.* (2018). doi:10.1100/S2059798318006551

80. Afonine, P. V. et al. New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. Sect. D Struct. Biol.* (2018). doi:10.1100/S2059798318009324

81. Xiao, X. et al. Complement-Related Proteins Control the Flavivirus Infection of Aedes aegypti by Inducing Antimicrobial Peptides. *PLoS Pathog.* (2014). doi:10.1371/journal.ppat.1004027

82. Dzaki, N., Ramli, K. N., Azlan, A., Ishak, I. H. & Azzam, G. Evaluation of reference genes at different developmental stages for quantitative real-time PCR in Aedes aegypti. *Sci. Rep.* (2017). doi:10.1038/srep43618

83. Luna, C. et al. Characterization of three Toll-like genes from mosquito Aedes aegypti. *Insect Mol. Biol.* **12**, 67–74 (2003).

84. Rosenthal, P. B. & Rubinstein, J. L. Validating maps from single particle electron cryomicroscopy. *Current Opinion in Structural Biology* **34**, (2015).

85. Rambo, R. P. & Tainer, J. A. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **95**, (2011).
MATERIALS AND METHODS

Bioinformatic analysis

Mosquito Toll and Spz sequences were retrieved via BLAST searches in Uniprot \(^{62}\) and Vectorbase \(^{63}\). Sequence alignments were performed using Muscle \(^{64}\). Homology modelling was carried out using Modeller version 9 \(^{65}\). Alignments were visualised in Jalview \(^{66}\) and 3D-models in PyMol (Molecular Graphics System) and Chimera \(^{67}\).

DNA Constructs

Constructs were either derived from genomic or synthetic DNA upon codon optimisation to improve the protein production yields for Toll1A (Vectorbase identifier AAEL026297), Toll5A (AAEL007619), Spz1A (UniProt accession code Q17P53), Spz1C (UniProt accession code Q16J57; Vectorbase identifier AAEL013433) and SpzX (AAEL013434). Constructs were cloned into baculovirus transfer vector pFast-Bac1 within BamHI and NotI, and into pMT-V5-His\(^{A}\) within KpnI and NotI (ThermoFisher).

Protein production and purification

**Spz Proprotein and Cys-knot domain preparation**

*Aedes aegypti* Spz paralogues were produced in a baculovirus expression system (Bac-to-Bac, ThermoFisher) with a C-terminal Strep-tag\(^{®}\) II with or without an engineered TEV-cleavage site between the pro-domain (Spz1C pro-domain residues 43-218) and the Cystine-knot domain (Spz1C residues 219-320) after establishing the domain boundaries using limited trypsin proteolysis.

Typically, 4 litres Sf9 at 2 million cells per ml were infected at a MOI = 2.0 and cultured under agitation at 19 °C instead of 27 °C for optimal expression over 5 days. The supernatant was harvested after removing cells by centrifugation at 2,000 g for 10 min and filtered on a Sartobran P sterile capsule of 0.45 μm (Sartorius). The buffer was then exchanged to buffer A (150 mM NaCl, 100 mM Tris-HCl, pH 7.5, 1 mM EDTA) and concentrated to 500 ml using a Centramate tangential flow filtration system (Pall). It was loaded on a Strep-Tactin\(^{®}\)XT Superflow\(^{®}\) resin (IBA) following manufacturer recommendations. The resin was equilibrated in 10 column volumes of buffer A prior to use and subsequently washed with 10 column volumes of the same buffer to remove non-specifically bound proteins. Strep-tagged protein was eluted in buffer A containing 50 mM biotin. Peak fractions were pooled and purified by anion-exchange on a 5 ml Hitrap Q (Cytiva, formerly GE Healthcare Life science) in a NaCl gradient from 50 mM to 1 M, 50 mM Tris-HCl, pH 7.5, followed by size exclusion chromatography on a Superdex 75 10-300 GL column (Cytiva) at 0.5 ml/min in buffer C: 50 mM NaCl, 50 mM Tris-HCl, pH 7.5. Fractions were analysed by Coomassie-stained SDS-PAGE. Protein concentrations were quantified by absorption at 280 nm. A typical yield was 3 mg of purified protein per litre of cell culture. *Drosophila* Spz production has been described elsewhere \(^{68}\).

**Toll ectodomain preparation**

A stable Schneider 2 cell line containing pMT-V5/His-A *Aedes aegypti* Toll1A (AAEL026297) ectodomain (residues 1-835) and Toll5A (AAEL007619) ectodomain (residues 1- 789), a C-terminal TEV-cleavable Protein A fusion and a Flag-tag was induced with copper sulfate 0.5
mM at 3 million cells/ml. Cells were cultured at 27 °C under agitation and harvested after 3-4 days. Culture supernatant were filtered, buffer-exchanged and concentrated to 0.5 L in 150 mM NaCl, 50 mM Tris-HCl pH 7.5, 0.05% Tween 20. Protein A Flag tagged proteins were isolated with IgG Sepharose® 6 Fast Flow (Cytiva), incubated with TEV protease at a 1:10 (w/w) ratio at 4°C overnight in 150 mM NaCl, 50 mM Tris-HCl, pH 7.5 (buffer D). TEV-cleaved proteins lacking Protein A-Flag tag were collected in the flow through upon washing in buffer D. Protein A eluted in 0.1 M Na Acetate pH 3.4. Fractions of interest were pooled and further purified by anion-exchange (as above) and size exclusion chromatography on a Superdex 200 10-300 GL column (Cytiva) at 0.5 ml/min in buffer C.

**Surface plasmon resonance**

SPR experiments were performed on a Biacore T200 instrument with dextran-based Sensor Chip CM5 (Cytiva) in 100 mM NaCl, 20 mM HEPES pH 7 running buffer and a flowrate of 30 μl/minute. The chips were activated by 1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride and N-hydroxysuccinimide, and Spz1A, Spz1C and DmSpz proproteins and cyst-knot ligands were immobilized by amine coupling at pH 4.5. Sensorgrams were recorded and corrected by subtraction of control signal from an empty flow cell. Purified ectodomain analytes of Toll1A, Toll5A and DmToll were injected at concentrations between 0.1 to 7.5 μM in tenfold dilution series. Kinetic analysis was performed by fitting sensorgrams to a two-state reaction model.

**Analytical ultracentrifugation**

Analytical ultracentrifugation experiments were performed on an Optima XL-A/I (Beckman Coulter) centrifuge equipped with a four-hole titanium rotor, double-sector centrepieces, and an interference optical system for data acquisition. Sedimentation velocity runs were performed at 50,000 rpm with 3-min intervals between scans for a total of 190 scans at 20 °C. The sample volume was 400 μL. Data were analysed using Sedfit software 69. The partial specific volumes, buffer density, and viscosity were estimated using SEDNTERP software 70.

**SEC-MALS**

Size exclusion chromatography-coupled multi-angle light scattering (SEC-MALS) was used to analyse protein monodispersities and molecular weights. SEC was performed using an Äkta Purifier (GE Healthcare) and a Superose 6 10/300 GL column (GE Healthcare) in 50 mM Tris-HCl pH 7.5, 50 mM NaCl. For each measurement, 50 μL of protein at a given concentration was injected and gel filtrated at a flow rate of 0.5 ml/minute. Light scattering was recorded using a miniDAWN TREOS instrument (Wyatt Technology). Protein concentration in each elution peak was determined using differential refractive index (dRI). The data were analysed using the ASTRA 6.2 software (Wyatt Technology).

**SEC-SAXS**

SAXS measurements were performed at Diamond Light Source (Oxfordshire, UK), beamline 21 (B21) 71 at a wavelength 0.89 – 1.3 Å with a sample to detector distance of 3.7 m and a Eiger 4M (Dectris) detector, covering a momentum transfer of 0.0026 < q > 0.34 Å⁻¹ [q = 4πsinθ/λ, 2θ is the scattering angle]. The proteins were analysed by size-exclusion chromatography in line with small-angle X-Ray scattering (SEC-SAXS) to avoid the signal from possible aggregates. The samples were applied to a Superose 6 Increase 3.2/300 column (Cytiva) at a concentration
of 5 mg/ml and run at a flow rate of 0.075 ml/min in 50 mM Tris-HCl, pH 7.5, 50 mM NaCl. SAXS measurements were performed at 20°C using an exposure time of 3 s frame\(^{-1}\). SAXS data were processed and analysed using the ATSAS program package version 2.8.3\(^{72}\) and ScAtter (www.bioisis.net). The radius of gyration \(R_g\) and forward scattering \(I(0)\) were calculated by Guinier approximation. The maximum particle dimension \(D_{\text{max}}\) and \(P(r)\) function were evaluated using the program GNOM\(^{73}\).

**Mass Photometry**

All mass photometry measurements were executed on a Refeyn OneMP instrument. The calibration was done with a native marker protein standard mix (NativeMark Unstained Protein Standard, Thermo Scientific), which contains proteins ranging from 20 to 1,200 kDa. Coverslips (24 × 50mm, No. 1.5H, Marienfeld) were cleaned by sequential sonication in Milli-Q water, isopropanol and Milli-Q-water, followed by drying with nitrogen. For each acquisition 2 \(\mu\)L of protein solution was applied to 18 \(\mu\)L PBS buffer, pH 7.4 in a gasket (CultureWellTM Reusable Gasket, Grace Bio-Labs) on a coverslip. Increasing working concentrations tested included 25, 50, 75 to 100 nM. Movies were recorded at 999 Hz with an exposure time of 0.95 ms by using the AcquireMP software. All mass photometry movies were processed and analysed in the DiscoverMP software. Samples were measured in duplicates.

**Cryo-EM**

Holey carbon grids (Quantifoil Cu R1.2/1.3, 300 mesh) were glow discharged for 60 sec at current of 25 mA in PELCO Easiglow (Ted Pella, Inc). Aliquots of 3 \(\mu\)l of between 3-6 mg/ml of Toll5A-Spz1C gel-filtered complex mixed with 8 mM CHAPSO (final concentration, Sigma) were applied to the grids, which were immediately blotted with filter paper once to remove any excess sample, and plunge-frozen in liquid ethane using a FEI Vitrobot Mark IV (Thermo Fisher Scientific Ltd) at 4 °C and 95 % humidity. All cryo-EM data presented were collected eBIC (Harwell, UK) and all data collection parameters are given in Table S1.

Cryo-EM images were processed using Warp\(^{74}\) and CryoSPARC\(^{75}\). In short, CTF correction, motion correction, and particle picking were performed using Warp. These particles were subjected to two-dimensional (2D) classification in CryoSPARC followed by \textit{ab initio} reconstruction to generate initial 3D models. Particles corresponding to different classes were selected and optimised through iterative rounds of heterogeneous refinement as implemented in CryoSPARC. The best models were then further refined using homogenous refinement and finally non-uniform refinement in CryoSPARC. Finally, all maps were further improved using ResolveCryoEM in PHENIX\(^{76}\).

The final cryo-EM maps following density modification were used for model building. The crystal structures of \textit{Drosophila} Toll and Spz were used to generate the homology model of Aedes Toll5A and Spz1C using Modeller\(^{65}\). The initial model was then rigid-body fitted into the cryo-EM density for the highest resolution map of the apo-dimer of Toll5A in UCSF chimera\(^{67}\) and manually adjusted and rebuilt in Coot\(^{77}\). NAMDingator\(^{78}\) was used to adjust the structure and several rounds of real space refinement were then performed in PHENIX\(^{79}\) before the final model was validated using Molprobity\(^{80}\). For Toll5A in the holo-dimer and the 3:1 trimer cryo-EM maps, the initial model for the receptor was taken from the apo-dimer refined structure, while Spz1C was the \textit{Drosophila} Spz homology model and the same strategy was applied. All structures were refined and validated before being deposited into the PDB.

**RNA sequencing and RT-qPCR**
Total cellular RNA was extracted using the RNeasy Mini Kit (Qiagen, Germany) following manufacturer recommendations. Contaminating DNA was removed using TURBO DNA-free Kit (ThermoFisher Scientific, USA). The quantity and quality of RNA was checked using a NanoDrop spectrophotometer (ThermoFisher) and Bioanalyzer for RNA sequencing samples. Library preparation was performed with TruSeq Stranded mRNA Library Prep Kit (Illumina). Sample sequencing was performed on an Illumina NextSeq 500 using mRNA derived from different conditions (Mock Aag2 cells and stimulated with 100nM of Spz1C, GNB and PGN). The reads obtained by RNA-seq were analysed using the Cufflinks RNA-Seq workflow against Aedes aegypti genome on VectorBase (www.vectorbase.org). All graphic representations were made using CummeRbund package (v3.10) on R. RT-qPCR was conducted in a Rotor-Gene Q system (Qiagen) for over 40 cycles with an annealing temperature of 60 °C, with each well containing 2 µl RNA (10 ng/µl), 0.8 µl 10 µM specific primers (final concentration, 300 nM), 5.4 µl H₂O, 1 µl 20X Luna WarmStart RT Enzyme Mix (NEB) and 10 µl 2X Luna Universal One-step Reaction mix (NEB). Samples were measured in triplicates. Assessment of the expression of each target gene was based on relative quantification (RQ) using the comparative critical threshold (CT) value method. The RQ of a specific gene was evaluated in each reaction by normalization to the CT obtained for endogenous control gene elongation factor 1 alpha (eEF1α). Three independent infection experiments were conducted. The primers for quantitative RT-PCR used in this study are presented (Table S3). Data were analysed by t-test, Mann-Whitney, Krustal Wallis or ANOVA test depending on the application conditions and P value of <0.05 was considered significant on R software.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification and statistical analyses pertain to the analysis of cryo-EM data are integral parts of algorithms and software used.

DATA AND CODE AVAILABILITY

SAXS data were deposited at the Small Angle Scattering database with accession numbers SASDKX8 for Toll5A alone and, SASDKY8 for Toll5A with Spz1C, respectively. The cryo-EM 3D maps corresponding to the apo-dimer, the holo-dimer and the holo-trimer were deposited in EMDB database with accession codes EMD-11984; EMD-11982 and EMD-11983, respectively. The corresponding atomic models were deposited in PDB with accession codes 7B1D, 7B1B and 7B1C, respectively.
### Supplemental tables

**Table S1:** Cryo-EM Data Collection, Refinement, and Validation Statistics.

| T5A+S1C mixture                  | 2:0 Homodimer | 2:1 Heterodimer | 3:1 Heterotrimer |
|----------------------------------|---------------|-----------------|-----------------|
| PDB entry ID                     | 7B1D          | 7B1B            | 7B1C            |
| EMDB entry ID                    | EMD-11984     | EMD-11982       | EMD-11983       |
| **DATA COLLECTION AND PROCESSING** |               |                 |                 |
| Microscope                       | TFS KRIOS     |                 |                 |
| Voltage (kV)                     | 300           |                 |                 |
| Detector                         | GATAN K3 (6K X 4K) |           |                 |
| Nominal magnification            | 105,000 x     |                 |                 |
| Electron dose, total (e⁻/Å²)     | 51.10         |                 |                 |
| Defocus range (μm)               | -2.7 to -1.0  |                 |                 |
| Exposure (s)                     | 2             |                 |                 |
| Pixel size (Å)                   | 0.83          |                 |                 |
| Frames                           | 52            |                 |                 |
| **Software**                     | WARP (CTF correction, motion correction, and particle picking) | CryoSPARC (2D-classification and ab initio reconstruction) | Phenix (density modification and real space refinement) | Coot (model building and fitting) |
| **RECONSTRUCTION**               |               |                 |                 |
| Number of particles, total picked| 507,268       |                 |                 |
| Number of particles used         | 85,810        | 40,153          | 42,866          |
| Final resolution, FSC₀.₁₄₃ (Å)   | 3.41          | 4.23            | 3.74            |
| Map-sharpening B factor (Å²)     | -82.8         | -93.5           | -54.0           |
| **MODEL COMPOSITION**            |               |                 |                 |
| Non-hydrogen atoms               | 10,967        | 12877           | 17197           |
| Protein residues                 | 1505          | 1697            | 2440            |
| Glycans                          | 12            | 22              | 35              |
### REFINEMENT

|                                |     |     |     |
|--------------------------------|-----|-----|-----|
| Correlation coefficient, masked| 0.82| 0.70| 0.78|
| Correlation coefficient, box   | 0.88| 0.85| 0.88|

### VALIDATION (PROTEINS)

|                                |     |     |     |
|--------------------------------|-----|-----|-----|
| MolProbity score               | 1.79| 2.41| 2.47|
| Clash score, all atoms         | 6.82| 7.84| 14.83|
| Poor rotamers (%)              | 0.29| 10.24| 2.17|

### RAMACHANDRAN PLOT STATISTICS

|                                |     |     |     |
|--------------------------------|-----|-----|-----|
| Favoured, overall (%)          | 93.80| 96.63| 90.49|
| Allowed, overall (%)           | 6.20| 3.37| 9.51|
| Disallowed, overall (%)        | 0.00| 0.00| 0.00|

### R.M.S. DEVIATIONS

|                                |     |     |     |
|--------------------------------|-----|-----|-----|
| Bond lengths (Å)               | 0.004| 0.013| 0.009|
| Bond angles (°)                | 0.693| 1.811| 1.201|

| T5A+S1C mixture                | 2:0 Homodimer | 2:1 Heterodimer | 3:1 Heterotrimer |

**Table S2: Solution scattering of Toll5A alone and in complex with Spz1C.**

Synchrotron SAXS data from solutions of Toll5A ectodomain alone and in complex with Spz1C Cys-knot domain in 50 mM NaCl, 50 mM Tris-HCl, pH 7.5 were collected on the B21 beamline at the Diamond Light Source (Didcot, UK). In-line size-exclusion chromatography (SEC) SAS was employed. The SEC parameters were as follows: the samples were injected at a 0.075 ml/min flow rate onto a Superose 6 (3.2/300) column at 20°C. 31 successive 3 second frames were collected through the SEC elution peak of the sample. The data were normalized to the intensity of the transmitted beam and radially averaged; the scattering of the solvent-blank was subtracted.

### INSTRUMENT

|                                | Diamond -B21 |
|--------------------------------|---------------|
| Beam size at sample (µm)       | 1102 x 240    |
| Wavelength (Å)                 | 0.89 – 1.3 (equivalent to 9.5 – 14 keV) |
| Q range (Å⁻¹)                  | 0.0026 to 0.34 |
| Detector                       | Eiger 4M (Dectris) |
| Detector distance (m)          | 3.7           |
| Exposure (s per image)         | 3             |
| Column                         | Superose 6 Increase 3.2/300 |
Flow rate (ml/min) 0.075
Sample volume (µl) 55
Sample concentration (mg/ml) 5
Temperature (K) 293

**STRUCTURAL PARAMETERS**

| Parameter       | Guinier        | P(r)          |
|-----------------|----------------|---------------|
| R\(g\) (Å)      | 53.56 ± 0.05   | 55.37         |
| D\(_{\text{max}}\) (Å) | 205           | 215           |

Porod volume (Å\(^3\)) 370,000 567,000

**MOLECULAR MASS DETERMINATION**

**Theoretical MW (kDa)**

| MW (kDa) | Description                        |
|----------|------------------------------------|
| 104 kDa  | 88 kDa amino acids; 16 kDa glycans |
| 130 kDa  | (T5A\(_1\)/S1C\(_1\)), S1C 26 kDa |
| 234 kDa  | (T5A\(_2\)/S1C\(_1\))              |
| 260 kDa  | (T5A\(_2\)/S1C\(_2\))              |

**MALLS MW (kDa)**

| MW (kDa) | 175.5 | 206.5 |

**DATPOROD MW (kDa)**

| MW (kDa) | 231   | 354   |

**DATVC MW (kDa)**

| MW (kDa) | 196   | 271   |

**DATMOW MW (kDa)**

| MW (kDa) | 202   | 267   |

**DATA ANALYSIS SOFTWARE**

- Data reduction: PRIMUS & ScAtter
- Ab initio modelling: DAMMIF
- Homology modelling: Modeller
- Computation of model fitting to data: Pepsi-SAXS

Table S3: RT-PCR oligonucleotides used.

| RT-PCR primers       | Reference | Sequence                          |
|----------------------|-----------|------------------------------------|
| Fw Aa_Defensin_A     | B1        | CTATCAGGCCGCGGTGGAG                |
| Rev Aa_Defensin_A    | B1        | CAATGAGCAGCACAAGCAGCTATC           |
| Fw Aa_Dipterin_1     | B1        | GCACATGTCGGACCGATTCA               |
| Rev Aa_Dipterin_1    | B1        | GTCCTTCGCTGCTGTTGATGG             |
| Fw Aa_Cecropin_A     | This paper | CAAAGTTATTTCTCTTGATCGCG         |
| Rev Aa_Cecropin_A    | This paper | CTGCACCTTCACAATTTCTCTTCC         |
| Fw Aa_Gambicin_1     | This paper | GTTCCTCTTGCAAGGCATATG            |
| Rev Aa_Gambicin_1    | This paper | GACAGTCACGCGACTTTTCATTG           |
| Fw Aa_GRRP           | This paper | GCCGTCTGGCAGTTTCTCTTC           |
| Rev Aa_GRRP          | This paper | TTTTACAAACGCTACCTTGACC           |
| Fw Aa_Attacin_B      | This paper | TGGTTCAGCGGCCAAAAGGAT            |
| Rev Aa_Attacin_B     | This paper | GGTTCGACGGGGTTTGAAC              |
| Fw Aa_Vago           | This paper | CCTGGAAGTGTCATCGATCC            |
| Rev Aa_Vago          | This paper | TGGAAACGATGTCGGCTCGGTG           |
| Fw eEF1a             | B2        | AGGAATTGCGTGGATGAC                |
| Rv eEF1a             | B2        | GGTTCCTGCGTGACTGTTG               |
| Fw Aa_Toll1A         | This paper | GACGTAGGTGTTCTACCTTGAGG             |
| Rev Aa_Toll1A        | This paper | CGTACTCGCACAAGGACGA                |
|       | This paper |                                     |
|-------|------------|-------------------------------------|
| Fw Aa_Toll1B |           | GGCAACCTCAACGCAAACC                 |
| Rev Aa-Toll1B |           | CTTGGGACACCTTTGGTTGA                |
| Fw Aa_Toll5A  | 83         | TGGAATATGATCCGATGAACCTTC           |
| Rev Aa_Toll5A  | 83         | GTCCAGGCTGTCGATGTCCT               |
| Fw Aa_Toll5B |           | GACGTGTGCAGATGTAATGG               |
| Rev Aa_Toll5B |           | GTTACCGAGCTACCGATTC                |
| Fw Aa_Toll6   |           | GGGTCCTTTGTTGGCGACAT              |
| Rev Aa_Toll6   |           | ACGCATCACAAGACGTC                |
| Fw Aa_Toll7   |           | GCATAGTGGAGCGCAATGTG             |
| Rev Aa_Toll7   |           | GCGAAGCTCCGTTTTTCA             |
| Fw Aa_Toll8   |           | TTTCCGGACGTTGATGAGCC           |
| Rev Aa_Toll8   |           | CTCAAATTCCTTCAGCCCGT            |
| Fw Aa_Toll9A  |           | TATAAGGCTCCCGTTTGC               |
| Rev Aa_Toll9A  |           | TAGATCAGCAGGCCAAGTCC            |
| Fw Aa_Toll9B  |           | GACGACTACGTCGATCC                |
| Rev Aa_Toll9B  |           | TCAGATTGTCAGCAGCGTT          |
| Fw Aa_Toll10  |           | ACATATCCGGGCGGACATTTC          |
| Rev Aa_Toll10  |           | GCAGCGTAAATGACGTAAAC         |
| Fw Aa_Toll11  |           | GTAACCGATTAACGCGAG              |
| Rev Aa_Toll11  |           | TCAGTTTGACCGCAAGTCGAGG         |
Supplemental figures

A

6054 movies collected → WARP
Motion correction
CTF estimation
507,268 particles picked
Import to Cryosparc

Selected 261,344 particles ab initio reconstruction generated "good" volumes
Excluded 245,924 particles ab initio reconstruction generated "junk" volumes

Heterogenous refinement
All 749,185 particles
Multiple copies of "good" and "junk" volumes

Multiple iterations

167,530 initial particles
81,002 initial particles
55,185 initial particles

85,810 final particles
42,866 final particles
40,153 final particles

Non-uniform refinement
3.41 Å
Non-uniform refinement
3.74 Å
Non-uniform refinement
4.23 Å

PDB: 7B1D
EMD: 11984

PDB: 7B1C
EMD: 11983

PDB: 7B1B
EMD: 11982

Multiple iterations
**Figure S1: Cryo-EM workflow.** (A) Data were collected on a Titan Krios. Representative cryo-EM micrographs are shown. Scale bars, 50 nm. Cryo-EM data analysis was performed in Warp\(^74\) and CryoSparc\(^75\). FSC threshold criteria at 0.143\(^84\) is used to determine the resolution of the reconstructions and indicates the size of the smallest reliable detail. The atomic model refinement process was halted when iterations stopped improving. In the end, the maps were masked, sharpened, and a final resolution was determined. (B) Toll5A homodimer, (C) Spz1C-bound Toll5A trimer. (D) Spz1C-bound Toll5A heterodimer. For each particle the following are shown: First column: the angular distributions for particles projections estimated by cryoSPARC\(^75\). Heat maps show number of particles for each viewing angle from low, in dark blue, to high, in red. Second column: the gold standard FSC, calculated by comparing the two independently determined half maps from cryoSPARC. The purple line represents the 0.143 FSC cut-off, which indicates a nominal resolution after correction. Third column: density coloured according to a local resolution gradient from high at 3 Å, in dark blue; to low resolution, with 7 Å in green, and 11 Å in red.

**Figure S2: Shape and electrostatic complementarity of Toll5A and Spz1C.** All-atom models were generated in Modeller\(^65\). Molecular surface charge distributions were generated in PyMol (https://pymol.org/) with electronegative patches in red, neutral in white and electropositive in blue. Focusing on the concave binding site in (A) looking up towards Toll5A and down onto Spz1C interacting surface representations. Focusing on the convex lateral interface in (B) with complementary views that reveal Spz1C and Toll5A interfaces. (C) Close up cartoon representation with density map (blue wire) of the concave interface.
Figure S3: Spz1C binding mode to Toll5A LRR14 Z-loop. Electron density of unprocessed Toll5A-Spz1C 2:1 heterodimer with Spz1C crosslinking the Z-loop at the heterodimerization interface. Spz1C β-loops surround the Z-loop, while the W-loops connect the N-terminal region of chain C and the hinge region of chain B, respectively. The location of Toll5A Asp 464 is indicated by a magenta sphere (A). Spz1C protects Toll5A ECD from endoproteolytical degradation. In contrast, spontaneous Z-loop degradation stabilises Spz1C binding with a clear band shift as evidenced by native gel electrophoresis (B). Sequence alignment of duplicated mosquito Tolls and Drosophila Toll1 and 5 at LRR14 and its Z-loop (C). DmToll5 (DMEL-T5) does not have a Z-loop. A. aegypti Toll5A (AAEL-T5A) Asp 464 (highlighted in magenta) is partially conserved in most mosquito receptors and potential target site for Asparaginyl endopeptidase (AEP) procession. An alternative location is described for Drosophila Toll at Asp 458. Some receptors have an Asn substitution instead of Asp, as for A. gambiae Toll1B and 5B, while the site’s position does not appear conserved for A. aegypti Toll5B, with its Asn at +1 a position and Asp at +3. Drosophila Toll Asp 458 cleavage is involved in constitutive receptor dimerization. Sequence alignment generated by Muscle software and illustrated with ClustalX colour scheme above 30% conservation in Jalview. A. albopictus AALF-T1A (AALF013620); AALF-T1B (AALF013620); AALF-T5A (AALF027669); A. aegypti AAEL-T5A (AAEL026297); AAEL-T1B (AAEL027951); AAEL-T5A (AAEL007619, Uniprot A0A618TEX2); AAEL-T5B (AAEL000057); A. gambiae AGAP-T1A (AGAP001004; UniProt Q8WRE5); AGAP-T1B (AGAP010636; Q7Q545); AGAP-T5A (Q5TW82; AGAP000999); AGAP-T5B (Q7Q545); AGAP010669); D. melanogaster Toll1 (DMEL-T1: UniProt P08953) and Toll5 (Q9NBK8).
| Toll5A-Spz1C | $k_{on}$ (1/Ms) | $k_{off}$ (1/s) | $K_D$ (μM) | Chτ² | $R_{max}$ (RU) |
|-------------|----------------|---------------|-------------|-------|---------------|
| Global      | $k_{on} = 16.930$ | $k_{off} = 0.03239$ | $K_D = 1.913$ | 1.65  | $R_{max} = 93.45$ |
| heterogeneous | $k_{on} = 2.548$ | $k_{off} = 0.005325$ | $K_D = 2.090$ |       | $R_{max} = 66.40$ |

**Figure S4:** Toll and Spz binding assessed by surface plasmon resonance suggests paralog and orthologue specificity. Sensorgrams and global fitting analysis of heterogeneous ligand with two binding sites of Toll5A and Spz1C. Both binding sites have similar affinities despite different on and off rate constants.

**Figure S5:** Solution scattering data of Toll5A alone and in complex with Spz1C. (A) 1-dimensional intensity plot with (B) its Guinier fit. (C) Normalized Pair-distance distribution function, used to determine the maximum dimension of the particles ($D_{max}$). (D) Dimensionless Kratky Plot, which quantifies the conformational flexibility of proteins. The dotted lines are drawn at $qR_g = 1.73$ and $I(q)/I(0)\cdot(qR_g)^2 = 1.1$. Folded proteins have a local maximum where the two lines intersect. 

---

36
Figure S6: Analytical ultracentrifugation profiles. AUC sedimentation data are shown with their fit and residuals after fitting to a c(S) model in SEDFIT and the distribution of sedimentation coefficients. From left to right: 3 μM Toll5A; 30 μM Toll5A; a mixture of 30 μM Toll5A and 30 μM Spz1C (equimolar); and a mixture of 30 μM Toll5A and 74 μM Spz1C (excess). There is a contaminant at ~ 2S in the Toll5A preparation, which is only visible in the 3 μM Toll5A dataset given its lower c(S) scale.
Figure S7: Transcriptional profile in unstimulated Aag2 cells of members of the Toll (top) and Spz family (bottom panel). RT-PCR reactions were analysed on 1.5% agarose gel using SYBR™ Safe DNA gel stain. RT-PCR primers are listed in Table S3.
Figure S8: Similarities and differences in duplicated Toll and Spz in *Aedes aegypti*. Gene localisation and exon/intron organisation in the Toll1/5 family (A) with Toll1A (AAEL026297), Toll1B (AAEL027951), and Toll5A (AAEL007619) on chromosome I, Toll5B (AAEL000057) on chromosome III. Organisation of duplicated Spz1 genes (B) with all family members located on chromosome II. Spz1A and Spz1B have been reannotated as mutually exclusive splice forms of a single gene labelled SpzX (AAEL003434), while Spz1C sequence and location (AAEL013433) was confirmed by high throughput sequencing techniques. (C) Toll ligand-binding ectodomain, and (D) Spz C-terminal active domain homology models generated by Modeller and visualised by PyMol.