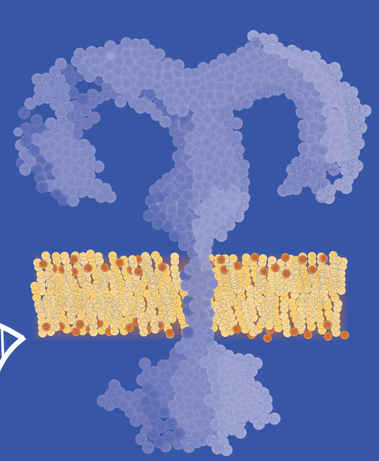
TOLL-LIKE RECEPTOR BIOLOGY FROM EVOLUTION TO FUNCTION

Carlos G.P. Voogdt

2019

Toll-like receptor biology
From evolution to function



Carlos G.P. Voogdt

UITNODIGING

Voor het bijwonen van
de openbare verdediging
van het proefschrift:

Toll-like receptor biology
From evolution to function

Donderdag 6 juni 2019
om 14:30u in de senaatszaal
van het academiegebouw
van Universiteit Utrecht.
Domplein 29, 3512 JE Utrecht

Na afloop bent u van harte welkom
op de receptie ter plaatse

Carlos Voogdt
Pallasburg 14
3437GS Nieuwegein
c.voogdt@uu.nl

Paranimfen:

Guus van Muijlwijk
g.h.vanmuijlwijk@uu.nl
0614889131

Xinyue Li
x.li5@uu.nl
0643773205

Toll-like receptor biology
From evolution to function

Carlos G.P. Voogdt

Toll-like receptor biology – From evolution to function

ISBN: 978-94-6380-375-5

Lay-out and printing: Carlos Voogdt and Proefschriftmaken.nl

Printing of this thesis was financially supported by: Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University and Infection & Immunity Utrecht

Toll-like receptor biology

From evolution to function

Toll-like receptor biologie
Van evolutie tot functie
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

donderdag 6 juni 2019
des middags te 2.30 uur

door

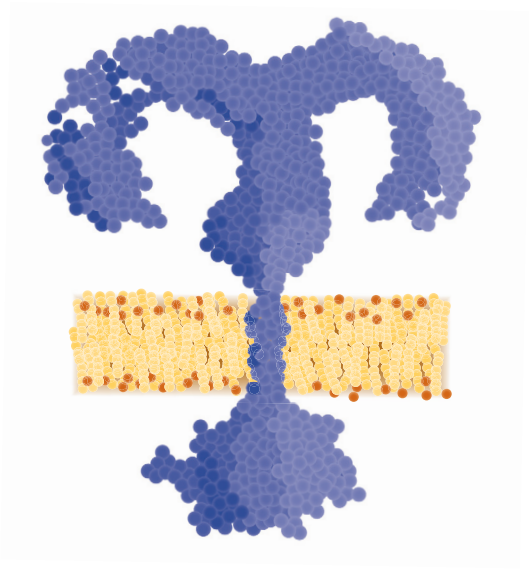
Carlos Geert Pieter Voogdt

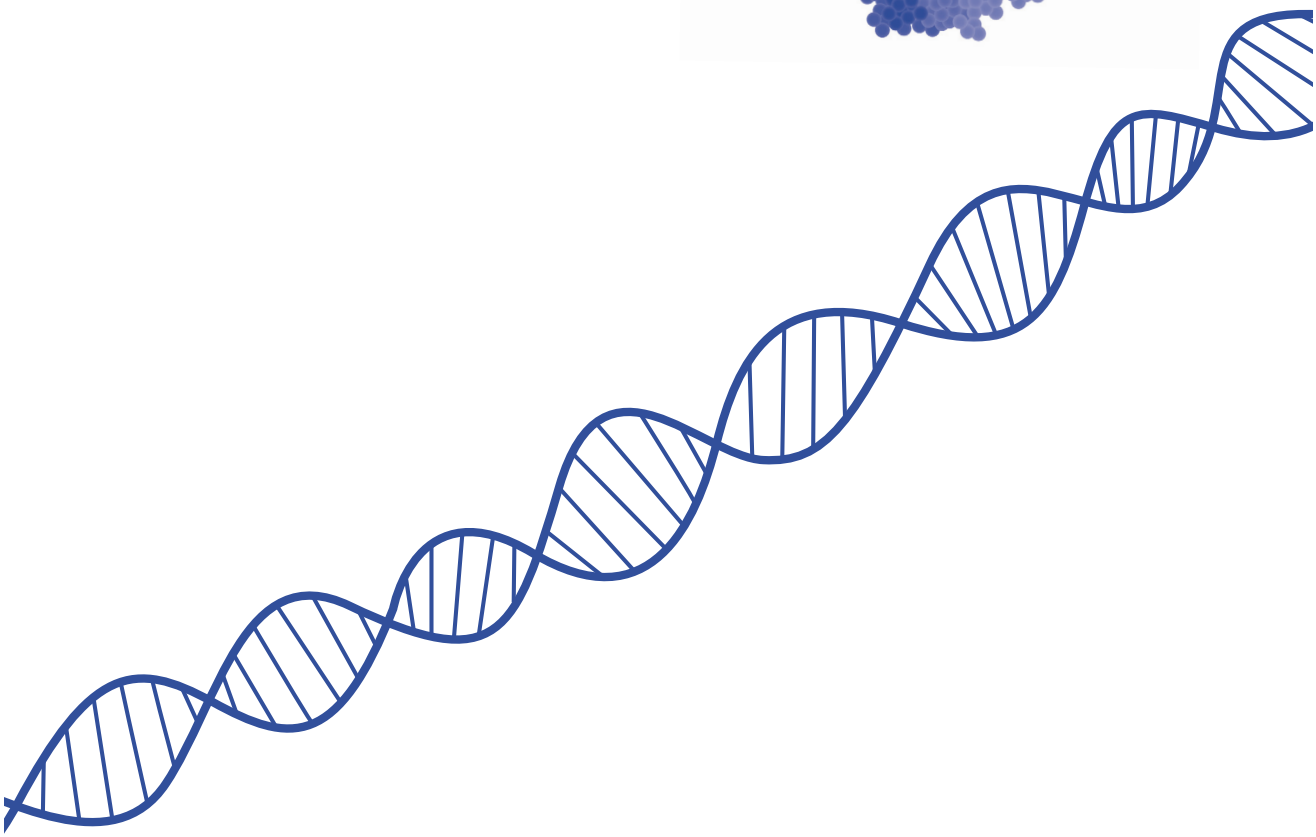
geboren op 18 januari 1987
te Zeist

Promotoren:

Prof.dr. J.P.M. van Putten
Prof.dr. J.A. Wagenaar

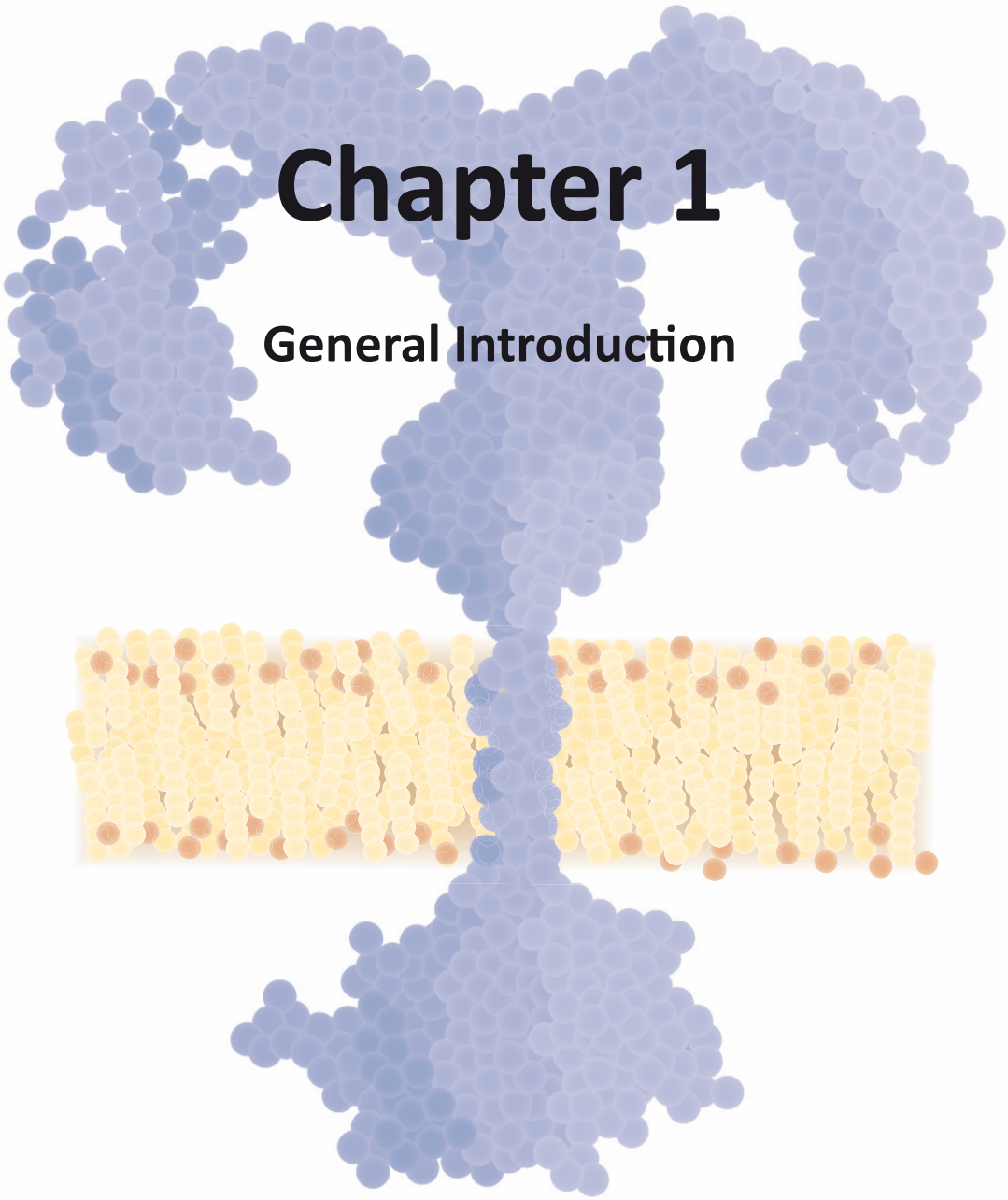
Contents





Chapter 1

General Introduction



Introduction

Throughout evolution animals have developed an elaborate collection of specialized molecules and cells that together form the immune system. The immune system defends against microorganisms and plays a key role in maintaining homeostasis. A response by the immune system to an environmental threat such as a pathogenic microorganism, typically consists of two steps which involve immune receptors and effectors. The receptors detect a disturbance and induce the expression of effectors. The effectors subsequently neutralize the microorganism and resolve the infection. During the co-evolution of hosts with microorganisms, numerous immune receptors have emerged. One group of receptors that has been proven to be crucial in the detection of infectious agents is the family of Toll-like receptors (TLRs), which is reviewed extensively in **Chapter 2** of this thesis. As the name implies, TLRs are similar to the Toll receptor which was discovered during studies on embryogenesis in the fruit fly *Drosophila melanogaster*[1]. Besides playing a role in embryogenesis, the fruit fly Toll receptor also proved essential in anti-fungal immunity[2]. Soon after the discovery of Toll, the first functional description of a vertebrate ortholog of the Toll receptor was published in 1997[3]. In 1998 TLR4 was identified as the receptor for bacterial endotoxin (lipopolysaccharide)[4]. Since then, hundreds of publications have demonstrated the importance of TLRs in the detection of microbes and regulation of the immune system. To emphasize the significance of these receptors, the discovery of the Toll and Toll-like Receptor was awarded with the 2011 Nobel prize in Physiology or Medicine[5].

TLRs form a family of structurally highly conserved, membrane bound receptors of which each family member recognizes a distinct ligand. TLRs are best known for their detection of ligands derived from microorganisms and, collectively, the TLR family members within an animal's immune system ensure detection of a broad variety of microbial structures. Dynamics in ligand detection are added through distinct spatial receptor expression as some TLRs are expressed at the surface of cells while others are expressed in endolysosomes. The importance of detecting microbial ligands by TLRs is underlined by strong evolutionary conservation of these receptors in almost all animals and is most compellingly illustrated by the often drastic increase in susceptibility to infection upon loss of TLR function[6–10]. Apart from microbial ligands, TLRs can also detect endogenous ligands that may be released during tissue damage. Hyperresponsiveness to endogenous ligands by TLRs, function-affecting mutations in *tlr* genes or aberrant expression of TLRs can drive the development of

debilitating sterile auto-immune diseases such as arthritis and systemic lupus erythematosus (SLE)[11–13], cardiovascular diseases including atherosclerosis[14,15] and some types of cancer[16–18]. Because of their major role in infectious and non-infectious diseases, TLRs have been marked as targets for therapeutics that aim to modulate TLR function during or prior to disease. These therapeutics include for example, antagonistic or agonistic ligands such as bacterial cell wall mimetics, modified nucleic acid motifs and derivates of bacterial flagellin or TLR targeting antibodies[19–21]. In addition, to prophylactically protect against infections, vaccines can be supplemented with TLR ligands as adjuvants which may induce a stronger, TLR-mediated, antibody response[22–24]. Finally, as infectious diseases also form a major burden for food production systems, suggestions have been made to improve disease resistance in farm animals through directed breeding using TLRs as high potential selection markers[25,26].

TLR challenges

A little over twenty years after their discovery, a tremendous body of knowledge about TLR molecules and their regulation of the immune system has been generated and applied research aimed at therapeutically modulating TLRs has taken off. Yet, despite intense efforts it is still poorly understood how exactly multiple highly similar TLRs are involved in very diverse infectious and non-infectious disease phenotypes. In addition, although progress is being made, medically modulating TLR function is still in its infancy as TLR-targeting therapeutics show limited clinical success[27–29]. These challenges request a better understanding of fundamental TLR biology and indeed many unanswered questions concerning multiple aspects of TLR biology remain. Such aspects include (but are not limited to): modes of regulating receptor expression, the precise mechanism of TLR activation, and receptor signals involved in intracellular transport and degradation. From a broader perspective, other fundamental TLR aspects that are incompletely understood include differences in ligand detection by TLRs from different animals and the consequences of this for the use of animals as models. The ultimate question in unraveling the essential features that determine TLR function, may be what drives and limits the continuous adaptation of *tlr* genes throughout evolution.

Aim of this study

Comprehending the key role of TLRs as immune sensors of the environment in health and disease and exploiting their potential as therapeutic targets, demands a deep understanding of the receptor family at the molecular level. The aim of this thesis is to uncover missing elementary principles of TLR biology using an evolutionary approach. New knowledge gained from this should not only increase the appreciation of the role of TLRs throughout evolution, but also indicate new essential features of TLRs that provide a more solid scientific basis for the choice of suitable animal models and future drug design. To achieve this, the function of distinct TLRs from a wide variety of vertebrate animals will be compared. This approach is based on the philosophy that speciation (the formation of different species) is the result of an evolution driven necessity to adapt functions to meet specific environmental challenges. Comparison of differences or similarities among TLRs from different species thus holds power to uncover specific features that are important for receptor function.

The work described in this thesis focuses on two different TLRs that display distinct characteristics with respect to evolutionary history and the mechanism of receptor activation. TLR5 is a highly conserved receptor that in mammalian and avian species functions as a homodimeric receptor to detect the conserved bacterial protein flagellin. Whether this detection of flagellin by TLR5 is conserved in species other than mammals and birds, and whether these TLR5 orthologs display specificity in the recognition of different flagellin proteins, is unknown. TLR5 of the zebrafish, for example, has never been reported to be activated directly by flagellin. Yet, a crystal structure of zebrafish TLR5 in complex with flagellin is used by the scientific community as a model for TLR5-flagellin interactions of other species and as a basis for drug development, but whether this extrapolation is valid has remained unresolved. In contrast to TLR5, TLR15 is present in birds and reptiles but absent in other vertebrates. TLR15 has thus been suggested to have evolved *de novo* in the reptilian lineage, suggesting unique evolutionary properties of this receptor. Unlike any other TLR, TLR15 is not activated by conserved microbial structures but by microbial proteases that activate the receptor by proteolytic cleavage. The receptor sites sensitive for cleavage, the microbial proteases that activate TLR15 and their possible species-specific mode of TLR15 activation are all still unknown. In addition, the factors that determine TLR15 expression and its evolutionary history are incompletely understood. The unique properties of TLR15 compared to other TLRs may shed

light on novel mechanistic elements that are important in the biology of other members of the TLR family.

Outline of this thesis

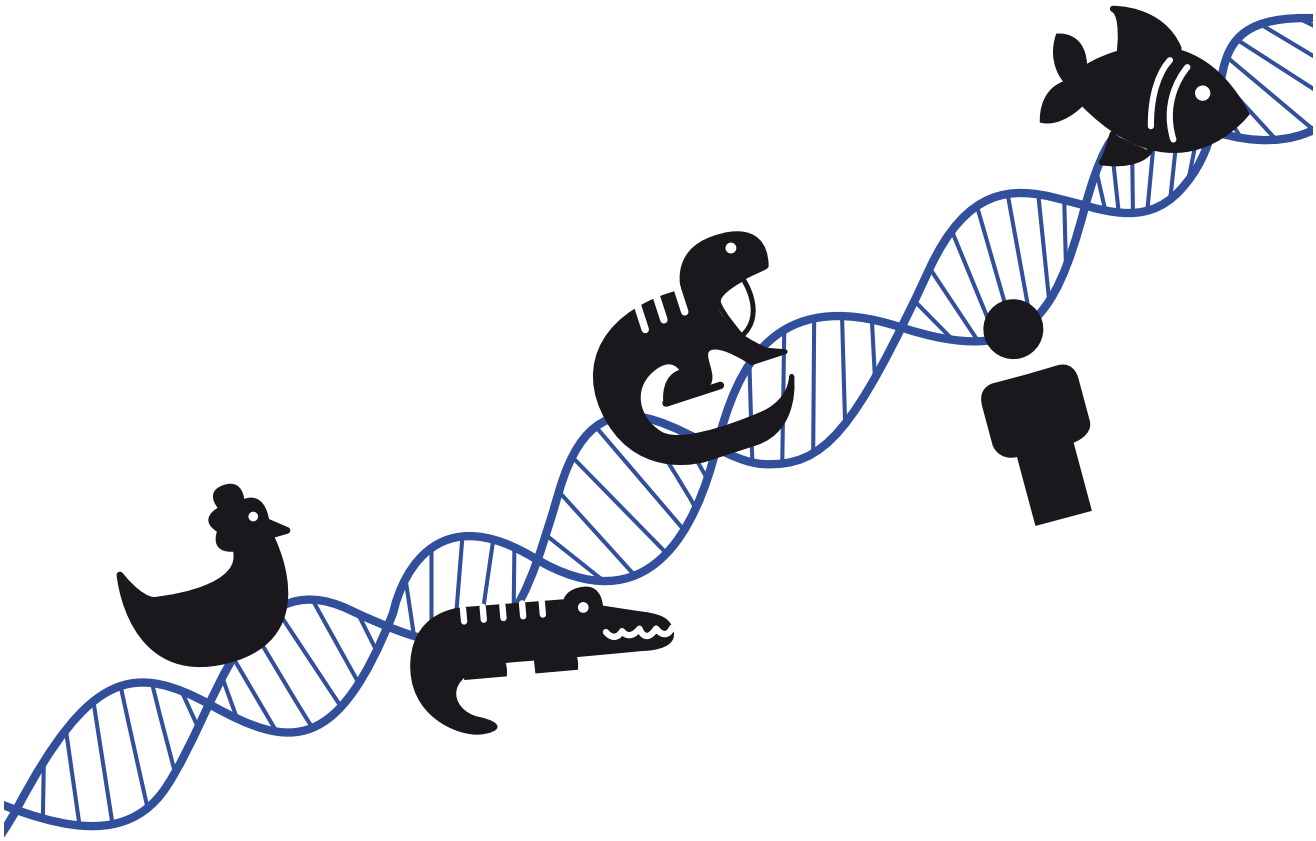
The main body of this thesis starts in **Chapter 2** with an extensive review on TLR structure, function and evolution throughout the animal kingdom. **Chapter 3** follows with the first functional characterization of a reptilian TLR, namely TLR5 of the green anole lizard. This receptor is then employed to identify differences in the detection of different bacterial flagellins between reptile and human TLR5. **Chapter 4** describes the unexpected finding that TLR5 of the zebrafish functions as a heterodimer. The broader implications of this for the concept of TLR activation as well as for the use of zebrafish TLR5 as a model for other TLR5 species are discussed. **Chapter 5** reports on the previously unknown function of the evolutionarily conserved C-terminal tail of TLR5 and its importance for receptor function and intracellular transport. In **Chapter 6** evidence is provided that TLR15 is actually evolutionarily much older than expected. The work also addresses for the first time the importance of codon usage bias in *tlr15* genes for receptor expression and the species specificity of TLR15 codon usage bias. **Chapter 7** is a brief follow-up on Chapter 6 and describes the first identification of a specific bacterial protease as activator of TLR15 as well as that the activation of TLR15 is highly species-specific. Finally, in **Chapter 8**, the findings reported in this thesis are placed in the context of the current knowledge about TLR biology and future perspectives on TLR research are discussed.

References

1. Anderson, K. V., Jürgens, G. & Nüsslein-Volhard, C. Establishment of dorsal-ventral polarity in the *Drosophila* embryo: Genetic studies on the role of the Toll gene product. *Cell* **42**, 779–789 (1985).
2. Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.-M. & Hoffmann, J. A. The Dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell* **86**, 973–983 (1996).
3. Medzhitov, R., Preston-Hurlburt, P. & Janeway Jr, C. A. A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. *Nature* **388**, 394–397 (1997).
4. Poltorak, A. *et al.* Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* **282**, 2085–2088 (1998).
5. O'Neill, L. A. J., Golenbock, D. & Bowie, A. G. The history of Toll-like receptors — redefining innate immunity. *Nature Reviews Immunology* **13**, 453–460 (2013).
6. Böhland, M. *et al.* Lack of Toll-like receptor 2 results in higher mortality of bacterial meningitis by impaired host resistance. *Journal of Neuroimmunology* **299**, 90–97 (2016).

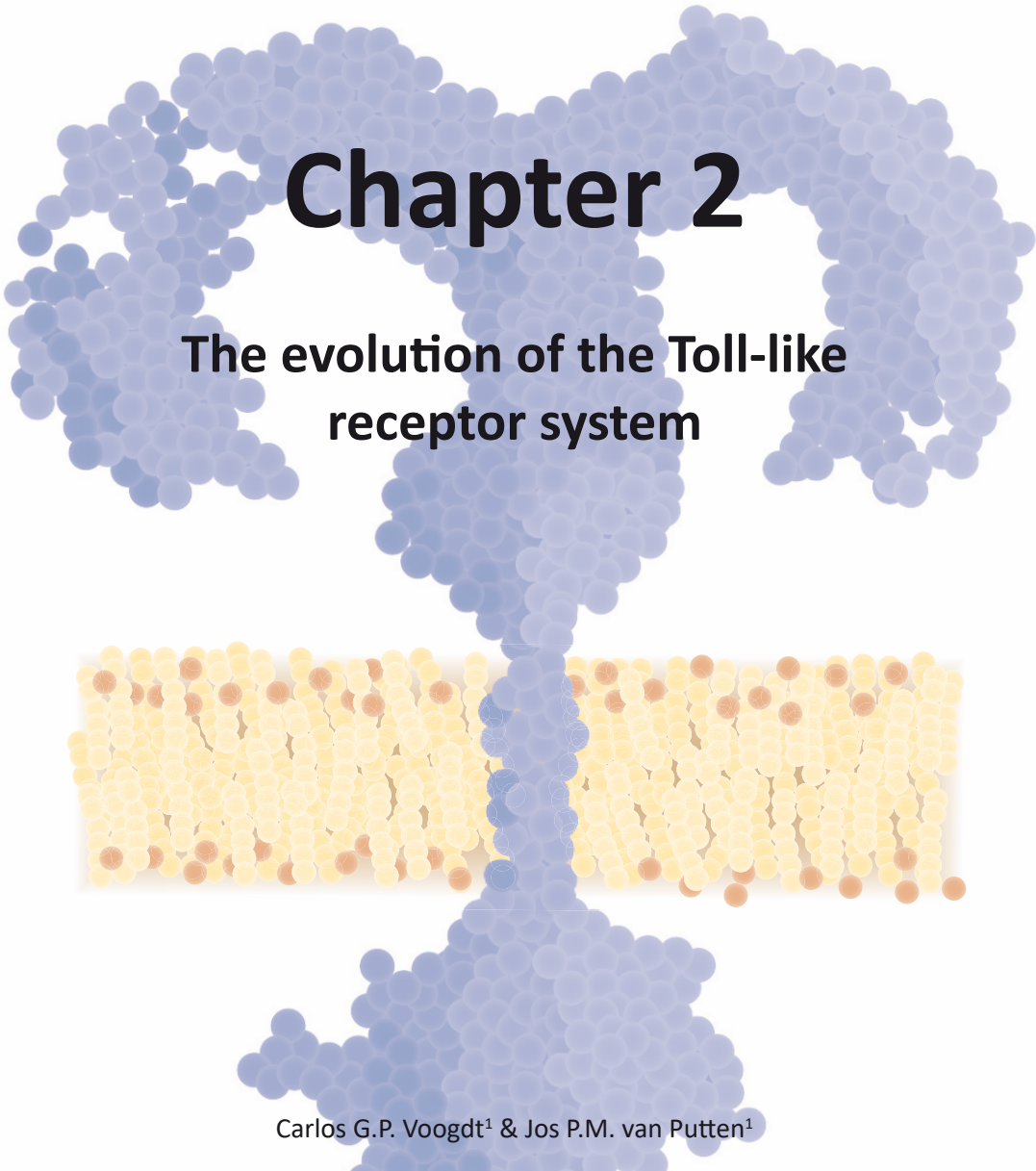
7. Koblansky, A. A. *et al.* Recognition of profilin by Toll-like receptor 12 is critical for host resistance to *Toxoplasma gondii*. *Immunity* **38**, 119–130 (2013).
8. Sivick, K. E. *et al.* Toll-like receptor-deficient mice reveal how innate immune signaling influences *Salmonella* virulence strategies. *Cell Host & Microbe* **15**, 203–213 (2014).
9. Takeuchi, O. *et al.* Differential roles of TLR2 and TLR4 in recognition of gram-negative and gram-positive bacterial cell wall components. *Immunity* **11**, 443–451 (1999).
10. Andersen-Nissen, E. *et al.* Cutting edge: Tlr5−/− mice are more susceptible to *Escherichia coli* urinary tract infection. *The Journal of Immunology* **178**, 4717–4720 (2007).
11. Joosten, L. A. B., Abdollahi-Roodsaz, S., Dinarello, C. A., O'Neill, L. & Netea, M. G. Toll-like receptors and chronic inflammation in rheumatic diseases: new developments. *Nature Reviews Rheumatology* **12**, 344–357 (2016).
12. Tsokos, G. C., Lo, M. S., Reis, P. C. & Sullivan, K. E. New insights into the immunopathogenesis of systemic lupus erythematosus. *Nature Reviews Rheumatology* **12**, 716–730 (2016).
13. Jiménez-Dalmaroni, M. J., Gerswhin, M. E. & Adamopoulos, I. E. The critical role of toll-like receptors — From microbial recognition to autoimmunity: A comprehensive review. *Autoimmunity Reviews* **15**, 1–8 (2016).
14. Seneviratne, A. N., Sivagurunathan, B. & Monaco, C. Toll-like receptors and macrophage activation in atherosclerosis. *Clinica Chimica Acta* **413**, 3–14 (2012).
15. Ward, J. R., Wilson, H. L., Francis, S. E., Crossman, D. C. & Sabroe, I. Translational mini-review series on immunology of vascular disease: inflammation, infections and Toll-like receptors in cardiovascular disease. *Clinical & Experimental Immunology* **156**, 386–394 (2009).
16. Dajon, M., Iribarren, K. & Cremer, I. Toll-like receptor stimulation in cancer: A pro- and anti-tumor double-edged sword. *Immunobiology* **222**, 89–100 (2017).
17. Echizen, K., Hirose, O., Maeda, Y. & Oshima, M. Inflammation in gastric cancer: interplay of the COX-2/prostaglandin E2 and Toll-like receptor/MyD88 pathways. *Cancer Science* **107**, 391–397 (2016).
18. Ou, T., Lilly, M. & Jiang, W. The pathologic role of Toll-like receptor 4 in prostate cancer. *Frontiers in Immunology* **9**, (2018).
19. Achek, A., Yesudhas, D. & Choi, S. Toll-like receptors: promising therapeutic targets for inflammatory diseases. *Archives of Pharmacal Research* **39**, 1032–1049 (2016).
20. Connolly, D. J. & O'Neill, L. A. New developments in Toll-like receptor targeted therapeutics. *Current Opinion in Pharmacology* **12**, 510–518 (2012).
21. Shi, M., Chen, X., Ye, K., Yao, Y. & Li, Y. Application potential of toll-like receptors in cancer immunotherapy: systematic review. *Medicine* **95**, (2016).
22. Li, Q. & Guo, Z. Recent advances in Toll like receptor-targeting glycoconjugate vaccines. *Molecules* **23**, 1583 (2018).
23. Maisonneuve, C., Bertholet, S., Philpott, D. J. & Gregorio, E. D. Unleashing the potential of NOD- and Toll-like agonists as vaccine adjuvants. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 12294–12299 (2014).
24. Vasou, A., Sultanoglu, N., Goodbourn, S., Randall, R. E. & Kostrikis, L. G. Targeting pattern recognition receptors (PRR) for vaccine adjuvantation: from synthetic PRR agonists to the potential of defective interfering particles of viruses. *Viruses* **9**, 186 (2017).
25. Novák, K. Functional polymorphisms in Toll-like receptor genes for innate immunity in farm animals. *Veterinary Immunology and Immunopathology* **157**, 1–11 (2014).
26. Uenishi, H. & Shinkai, H. Porcine Toll-like receptors: the front line of pathogen monitoring and possible implications for disease resistance. *Developmental & Comparative Immunology* **33**, 353–361 (2009).
27. Patra, M. C. & Choi, S. Recent progress in the development of Toll-like receptor (TLR) antagonists. *Expert Opinion on Therapeutic Patents* **26**, 719–730 (2016).
28. Tartey, S. & Takeuchi, O. Pathogen recognition and Toll-like receptor targeted therapeutics in innate immune cells. *International Reviews of Immunology* **36**, 57–73 (2017).

29. Gao, W., Xiong, Y., Li, Q. & Yang, H. Inhibition of Toll-like receptor signaling as a promising therapy for inflammatory diseases: a journey from molecular to nano therapeutics. *Frontiers in Physiology* **8**, (2017).



Chapter 2

The evolution of the Toll-like receptor system



Carlos G.P. Voogdt[1] & Jos P.M. van Putten[1]

[1]Department of Infectious Diseases & Immunology, Utrecht University, Utrecht, the Netherlands

The Evolution of the Immune System – Conservation and Diversification
Elsevier 1st Edition 2016

Summary

Toll-like receptors are considered key elements of the immune system. The receptors scan the environment for mainly microbial danger signals and initiate a signaling cascade that mobilizes the appropriate host defense. Since the discovery of the protein Toll of the *Drosophila melanogaster* fruit fly, numerous (mostly putative) TLR genes have been identified throughout the animal kingdom. TLRs are structurally highly related proteins that belong to a very ancient receptor family. The first prototypical TLR gene appeared approximately 600 million years ago in the eumetazoan ancestor. During evolution TLRs underwent both purifying and diversifying selection, probably in order to adapt to co-evolving microbial danger signals. This review describes the current knowledge of the structure, function, phylogeny and distribution of TLRs across the animal kingdom and the selective microbial pressures that drive TLR adaptation and species-specific TLR function.

Introduction

Ever since their emergence, multicellular hosts evolved strategies to survive in optimal symbiosis with parasitizing micro-organisms. In turn, microbes continuously developed to evade the defensive barricades put up by their hosts. This ongoing evolutionary arms race has led to the development by the host of a sophisticated, germline encoded immune system, commonly referred to as the innate immune system. The innate immune system distinguishes beneficial and harmful microbes and responds to environmental threats through an extensive arsenal of so-called pattern recognition receptors (PRR). Throughout host-microbe co-evolution these PRRs have evolved to recognize highly conserved microbe associated molecular patterns (MAMP). These cell wall or nucleic acid structures are essential for microbial survival and hence difficult for the microbe to modify. Detection of these MAMPs enables the recognition of diverse microbes with a minimum set of receptors.

The best studied family of PRRs are the Toll-like receptors (TLR). TLRs are type I membrane spanning glycoproteins that are typically composed of an extracellular domain, a transmembrane domain and an intracellular signaling domain. Although the presence of TLR genes is conserved across the animal kingdom[1], TLR structure and function have diversified in response to the changing habitat and environmental challenges. The discovery of TLRs started with the identification of a protein called Toll in the fruit fly *Drosophila melanogaster*. Toll was identified as a regulator during embryonic development[2]. Later it was found that during an infection with fungi, the Toll protein was activated by its endogenous ligand Spätzle. This activation initiated the production of antimicrobial peptides thereby conferring immunity to fungi in *D. melanogaster*[3]. A search for proteins resembling Toll in other species resulted in the discovery of a murine Toll-like receptor (TLR4). TLR4 proved to be essential for the innate recognition of bacterial lipopolysaccharide (LPS)[4]. Since the discovery of TLR4 as the LPS receptor many more TLRs with their respective microbial ligands have been identified and characterized in many different animals. The revolutionary advances in whole genome sequencing now allows studies on TLR evolution across diverse phyla. Here we will review current knowledge of the evolution of the TLR structure, diversity and distribution across the animal kingdom, and the functions of TLRs in invertebrate and vertebrate animals, also in the context of ongoing host-microbe co-evolution.

TLR structure

Extracellular domain

The extracellular domain (ECD) of TLRs is composed of multiple consecutive leucine rich repeats (LRR) that each consist of 22 to 29 amino acids with specifically placed hydrophobic residues. Upon folding, the LRRs form an horse-shoe shaped structure in which closely packed β-sheets form the concave surface of the arch. The consecutive LRR motifs are typically flanked by a N- and C-terminal LRR (LRRNT and LRRCT, respectively) that often contain characteristically spaced cysteine residues. The presence of LRR motifs in numerous proteins of animals, plants, fungi, bacteria and viruses indicates that these motifs are of ancient origin and have remained important for protein-protein and receptor-ligand interactions throughout evolution[5]. In 2005 the crystal structure of the first TLR ECD (human TLR3) was resolved and since then additional ECD structures of different vertebrate TLRs and of *D. melanogaster* Toll have been determined[6–11]. Results indicate that in all TLRs the extracellular LRR motifs form an arch shaped structure that directly interacts with a particular microbial ligand in contrast to *D. melanogaster* Toll that interacts with the endogenous cytokine Spätzle rather than directly with microbial ligands.

Throughout evolution, continuous diversification in the number and sequence of LRR motifs has resulted in an extensive family of structurally distinct TLRs[12]. Due to the structural diversity the TLR family is able to recognize a large array of microbial ligands, ranging from lipids to proteins and nucleic acid motifs. Upon interaction with ligand the LRRCT of two monomeric TLR molecules come into close proximity with the LRRNTs spaced far apart. This results in the formation of a homo- or hetero- dimeric TLR complex that obtains a somewhat 'M' shaped configuration (Fig. 1). One striking example of structure based ligand specificity is the interaction of TLR2/1 and TLR2/6 heterodimers with bacterial lipoproteins. After binding triacylated lipoproteins, TLR2 forms a heterodimer with TLR1, while binding of diacylated lipoproteins leads to heterodimerization with TLR6. The structural difference between TLR1 and TLR6 that allows for the discrimination of lipoproteins is the blocking of a lipid channel in TLR6 by only two phenylalanine residues. Substitution of these residues opens the lipid channel in TLR6 and makes the receptor receptive for binding triacylated lipoproteins, just like TLR1[13]. This shows that minute structural differences in the TLR ECD allows TLR family members to recognize ligands with great specificity.

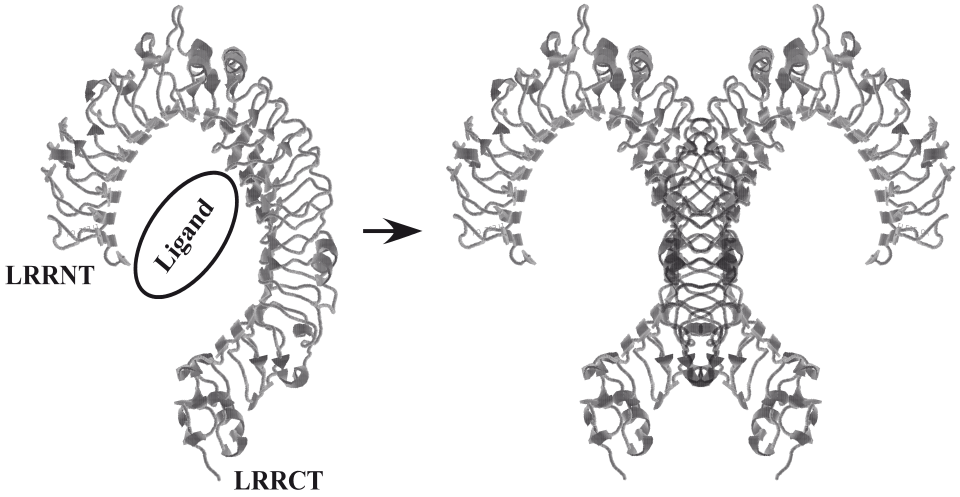


Figure 1. Upon binding of ligand to the arch shaped ECD of a TLR, two TLRs will form a dimeric complex bringing their LRRCT close together while the NTLRR are space far apart.

Transmembrane domain

TLRs are embedded in the membrane via a single membrane spanning region of approximately 20 amino acids. The TLR family members that recognize lipid or protein ligands are generally positioned at the cell surface, whereas TLRs that bind nucleic acid motifs are located in endosomes. Some TLRs appear in a soluble form. The soluble TLRs originate from enzymatic cleavage of the full length receptor (TLR2)[14], alternative splicing of the TLR gene (TLR4)[15], or from a separate gene (TLR5)[16]. Both soluble TLR2 and TLR4 reduce the response of their membrane bound form and thus may act as decoy receptors that prevent an excessive response to their TLR ligands. Soluble TLR5 however, enhances the reactivity of membrane bound TLR5 and might therefore aid in microbial detection. Soluble TLRs thus provide an additional form of structural TLR diversity.

Intracellular domain

On the cytoplasmic side TLRs contain a TIR domain, named after its structural and functional homology with *Drosophila* Toll and the Interleukin-1 receptor. The evolutionary importance of the TIR domain is evidenced by its presence in multiple proteins of animals, plants and even bacteria[17]. The TIR domain of TLRs is structurally composed of five alternating β-sheets and α-helices connected by short loops that fold into a core of β-sheets surrounded by the α-helices[18]. Its function is to initiate downstream signaling upon ligand induced receptor dimerization. Dimerization of TLRs brings their TIR domains in close proximity creating a

docking site for recruited adaptor proteins. The adapter proteins also contain TIR domains and associate with TLRs through TIR-TIR domain interactions. In mammals, TLRs interact with five adapter proteins: MyD88 (myeloid differentiation primary response protein 88), MAL or TIRAP (myeloid differentiation factor-88 adaptor-like protein), TRIF (TIR domain-containing adaptor protein inducing interferon β), TRAM (TIR domain containing adapter protein inducing interferon β related adapter molecule) and SARM (sterile α- and armadillo motif-containing protein). Based on the interaction with two major adapter proteins, TLR signaling can be divided into two signaling routes. The MyD88 dependent route, used by all TLRs except for TLR3, results in early activation of the NF-κB (nuclear factor κB) transcription factors that drive the transcription of pro-inflammatory cytokines. The TRIF pathway used by TLR3 and (often) TLR4, activates IRF (interferon regulatory factor) transcription factors that stimulate transcription of type I interferon cytokines. The MAL and TRAM adapters facilitate the interaction between MyD88 and TRIF and TLRs, while SARM functions as a negative regulator of the TRIF pathway[19,20]. Although considerably different in sequence and receptor make-up, both animals and plants use TIR domains and also LRR motifs in receptors involved in microbial recognition, implying that these structures may have originated before the divergence of the plant and animal kingdom[1,21].

Evolution and distribution of TLR genes

Origin of TLRs

Bioinformatics analysis on whole genome data indicates that prokaryotes and fungi lack TLR orthologs. Within the kingdom of plants, receptors composed of LRR motifs attached to various signaling domains (so called Receptor-like kinases or Nucleotide-binding site LRRs) are present but show only low sequence similarity to TLRs. Functional studies indicate that these plant receptors respond to different microbial motifs and exploit fundamentally different signaling networks compared to animal TLRs. This indicates that the LRR containing plant receptors are not ancient TLR orthologs but rather form a separate type of plant-specific receptors that have adapted a similar function as TLRs through the process of convergent evolution[21,22]. The origin of TLRs therefore lies in the animal kingdom (Metazoa). At the root of the metazoan evolutionary tree is the phylum of sponges (Porifera) (Fig. 2).

The sponge species *Suberites domuncula* and *Amphimedon queenslandica* do not contain typical TLRs but do carry TLR related genes. The predicted proteins contain a TIR and transmembrane domain but have a very short extracellular domain

without canonical LRRs (*S. domuncula*[23]) or instead an extracellular immunoglobulin domain (*A. queenslandica*[24]). TLR related proteins are also present in species of Cnidaria (jellyfish, sea anemones, corals, Hydra), a sister phylum of Porifera. *Hydra magnipapillata* recognizes microbial ligands through interaction of two membrane proteins, one carrying a cytoplasmic TIR domain and the other carrying extracellular LRR motifs[25]. However, in the genome of a different cnidarian, *Nematostella vectensis* (sea anemone), a typical TLR gene (i.e. with TIR and LRR domain), is present[26]. Typical TLR genes are also present in many species within the superphylum of Bilateria. The origin of TLRs therefore dates back to the eumetazoan (all animals except Porifera) ancestor, before the separation of Cnidaria and Bilateria, approximately 600 million years ago[1] (Fig. 2).

Structural difference between protostomian and deuterostomian TLRs

TLRs have been extensively conserved in bilaterian animals. The superphylum of Bilateria is divided in the deuterostomes (animals of which the embryonic blastophore forms the anus) and the protostomes (animals of which the embryonic blastophore forms the mouth). Protostomes (including Ecdysozoa and Lophotrochozoa) carry genes encoding Toll proteins which are structurally distinct from the deuterostomian TLRs. The protostomian Toll proteins generally contain two or more cysteine rich clusters in the midst of their LRR motifs, whereas the deuterostome TLRs generally contain only one or two cysteine rich clusters capping the LRR motifs (the LRRNT and LRRCT)[27]. Exceptions are the *Drosophila melanogaster* and *Anopheles gambiae* (Hexapoda) Toll-9 which LRR motifs are more similar to deuterostomian TLRs than to other Toll proteins[27]. Conversely, some TLRs of the invertebrate deuterostome *Strongylocentrotus purpuratus* (sea urchin) are more similar to protostomian Tolls than to other TLRs[28]. Since the gene in the sea anemone *N. vectensis* (Cnidaria) shows a higher structural similarity to protostomian Toll than to deuterostomian TLR, it is likely that the structure of Toll represents the ancestral form, while in deuterostomes this ancestral form independently evolved to the TLR form[1,27].

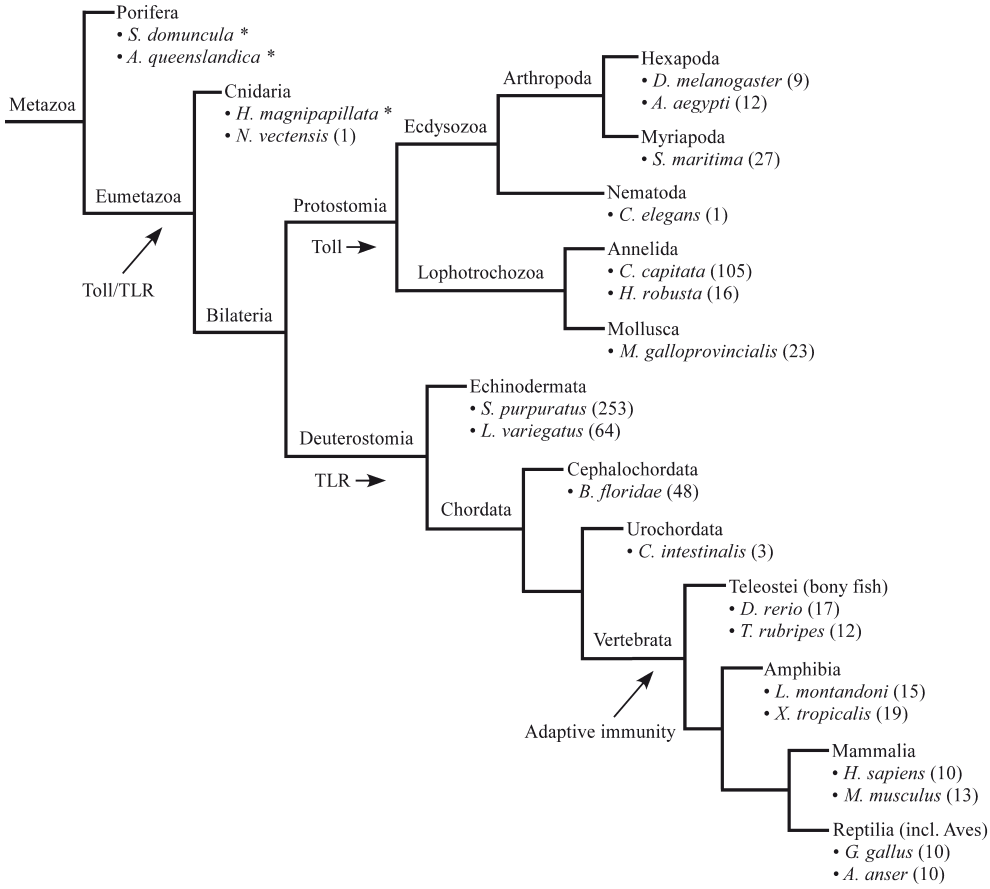


Figure 2. Simplified phylogeny indicating the relative relationship among some metazoans and the number of identified TLR genes in these species given in brackets. * indicates TLR-related proteins in these species. The prototypical Toll/TLR has originated in the eumetazoan ancestor approximately 600 million years ago (Mya). Basal components of the adaptive immune system arose approximately 500 Mya in early vertebrates. Echinoderms and non-vertebrate chordates show in general a large expansion of their TLR gene repertoire, a feature which may have evolved as an alternative to adaptive immunity. This phylogenic representation is not intended to include all the species in which TLR genes have thus far been identified.

Toll and TLR repertoire in protostomes and deuterostomes

Genomic data indicate that protostomes generally contain low to moderate numbers of Toll genes. Most arthropods (insects, crustaceans, myriapods and chelicerates) and some species of molluscs and annelids (superphylum Lophotrochozoa) have between 2 and 27 Toll genes[1,29–33]. No Toll genes have been detected yet in the lophotrochozoan Platyhelminthes which may indicate that these animals have secondarily lost their ancestral Toll genes[1]. Among protostomes, extremes in the abundance of Toll genes exist. The nematode *Caenorhabditis elegans* has only one

Toll gene, whereas the annelid worm *Capitella capitata* has 105 predicted Toll genes, an extensive repertoire likely created via many gene duplication events[33].

Interestingly, vertebrates (deuterostomes) contain roughly equal numbers of TLR genes, ranging from 10 TLRs in humans to 21 TLRs in amphibians. However, the invertebrate deuterostome *S. purpuratus* (Echinodermata) has massively expanded its repertoire to 253 TLR genes. Most of the sequence diversity in these genes exists in the LRR motifs and the TLRs more resemble each other than TLR genes from other animals. This suggests that the vast TLR repertoire in the sea urchin results from gene duplication, conversion and/or recombination events[28,34]. Independent expansion of the TLR repertoire seems to have occurred in the invertebrate deuterostome Amphioxus (*Branchiostoma floridae,* subphylum Cephalochordata), resulting in 48 TLR genes (mixed protostome and deuterostome like TLRs)[35]. Perhaps these animals have compensated their lack of an adaptive immune system by expanding their innate receptor diversity to cope with the encountered diverse array of encountered microbial structures, a feature which may be of particular importance in the aquatic environment these species inhabit[36]. In contrast, the aquatic living ascidian *Ciona intestinalis*, which is an urochordate closely related to vertebrates, has only three TLR genes[37]. Unlike vertebrate TLRs, two of the *C. intestinalis* TLRs recognize more than one type of ligand[38]. This multi-ligand recognition of these TLRs together with an expansion of complement factors[37], may have reduced the need for an expanded TLR repertoire in *C. intestinalis.*

Vertebrate TLR phylogeny

Most knowledge on TLRs is based on studies of vertebrates (Fig. 3). To date, 16 TLR genes have been identified in the lamprey (jawless vertebrate) 13 have been identified in mammals, 10 in birds, 21 in amphibians and 20 in teleost fish[31,39–41]. Reptiles are predicted to have at least 9 TLR genes. Based on sequence homology most vertebrate TLRs can be grouped into six major families[42] that in general have retained the ability to recognize distinct ligands. The large TLR1 family, consisting of TLR1, 2, 6, 10, 14, 15, 16, 18 and 25 recognizes lipoproteins (e.g. di- and tri-acylated lipopeptides). Atypical members of this family are TLR15 that is activated by microbial proteolytic cleavage[43] and TLR10 that functions as a negative regulator of TLR2[44]. The TLR3, 4 and 5 families recognize double stranded RNA, LPS and bacterial flagellin, respectively. The family of TLR7 including TLR7, 8 and 9, recognize nucleic acid motifs. The sixth major family contains TLR11, 12, 13, 19, 20, 21, 22, 23 and 26[42,45]. The receptors in this family that are functionally characterized, sense either protein (TLR11 and TLR12 respond to profilin of the protozoan *Toxoplasma gondii*[46,47]) or nucleic acid motifs, like the TLR7 family[48].

Especially from the large TLR1 and TLR11 families, some TLR genes appear to have been lost in various lineages, perhaps due to functional redundancy. Yet, almost all vertebrate species carry at least one gene from each of the major TLR families, underlining the importance of innate recognition of a diverse array of microbial ligands.

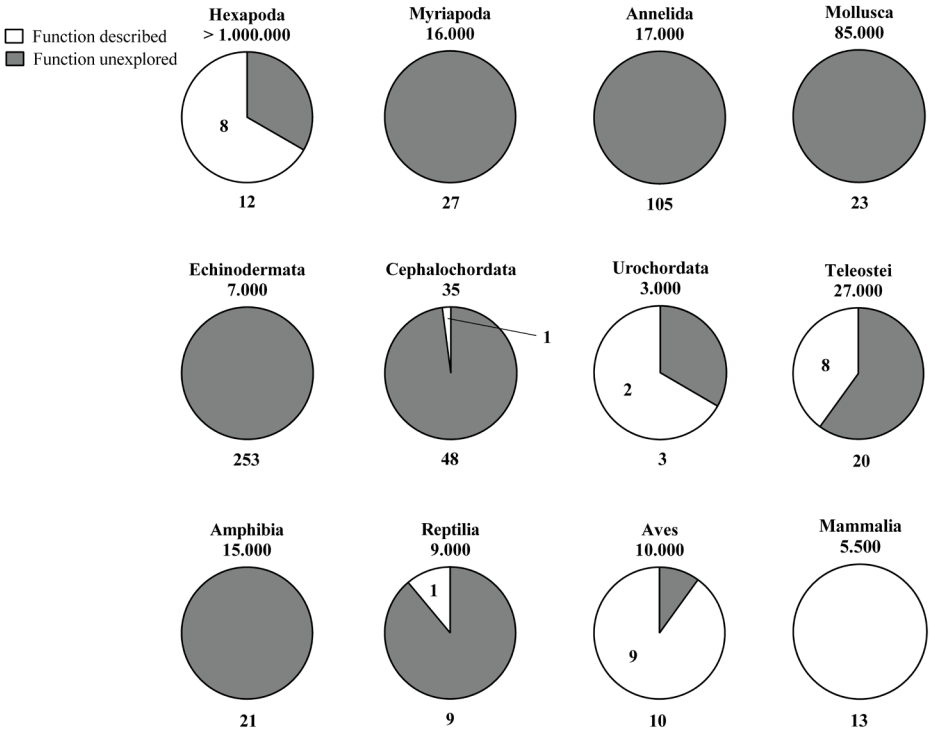


Figure 3. Knowledge on the function of the numerous TLR genes is still marginal and dominated by studies on vertebrate TLRs. The pie charts indicate the number of TLRs per animal group with a known function (white) or without any functional data (gray). The currently highest number of TLR genes identified per animal group is given under each chart. The estimate number of species per group is given under the group name to put our knowledge on TLR function in perspective with species diversity.

An interesting exception in the conservation of the vertebrate TLR repertoire is the lack of TLR4 in some teleost fish like *Takifugu rubripes*. TLR4, combined with its co-receptors MD-2 and CD14, recognizes LPS and this is of critical importance in the mammalian immune response to bacterial infections[4]. Some fish including zebrafish (*Danio rerio*) and common carp (*Cyprinus carpio*) do have multiple TLR4 copies but lack the TLR4 co-receptor genes[49]. As a result, LPS sensing in fish is not mediated by TLR4. Instead, fish TLR4 appears to be a negative regulator of the pro-

inflammatory NF-κB transcription factor[50]. The factors driving this divergent evolution are unknown and may await analysis of TLR4 in intermediate amphibian and reptile species.

Another example of dynamic TLR evolution in vertebrates is TLR15 in the clade of diapsids (reptiles and birds). TLR15 is only present in avian and reptilian genomes and its TIR domain is related to members of the TLR1 family[51]. However, unlike TLR1 family members that recognize lipopeptides, extensive sequence diversification of the TLR15 LRR motifs has led to the unique ability of this receptor to become activated by microbial proteases[43]. Why only diapsid animals have developed this trait and whether its provides these animals a significant immunological benefit is unknown†.

TLR function

Current knowledge on the evolution of Tolls and TLRs mainly results from predictions based on genome analysis. As evolution is primarily function-driven, the challenge is to corroborate the predictions with functional evidence. At this time, TLR related proteins in Porifera have not been functionally characterized and the only functional studies on cnidarian TLRs have been performed in Hydra. Hydra recognizes bacterial flagellin via an intermolecular interaction between the LRR motif and TIR domain containing proteins[25]. In addition, Hydra deficient in the primary TLR adaptor protein MyD88 are more susceptible to infection with *Pseudomonas aeruginosa*[52]. Although limited, these functional studies suggest that ever since its origin in the eumetazoan ancestor the TLR system of Hydra functions in the host response to micro-organisms. After the divergence of bilaterian animals, TLRs evolved independently in the protostomes and deuterostomes which has led to functional diversification of the Toll and TLR system.

Protostome Toll function

Most knowledge of the function of protostomian Tolls comes from studies on the arthropod *D. melanogaster* (Hexapoda; Insecta) which has nine Toll genes. *D. melanogaster* Toll-1 (or simply; Toll) regulates the formation of the dorsoventral axis in the fruit fly embryo[2], but is also involved in other developmental processes including the regulation of organogenesis, alignment and migration of cardioblasts in the embryonic heart and neural network development[1]. In addition, Toll-1 initiates

† In **Chapter 6** we describe the novel finding that TLR15 actually did not originate in diapsid animals but instead was reciprocally lost from most major classes of vertebrate animals.

the synthesis of antimicrobial peptides from the fat body (equivalent of vertebrate liver) in adult flies after fungal infection[3]. Although less well characterized, other *D. melanogaster* Tolls also play a role in either development or immunity; Toll-2, 6 and 8 are involved in regulating the anterior-posterior axis formation in the fly embryo[53], whereas Toll-3, 8 and 9 are involved in the elimination of unfit cells from the developing embryo[54]. Toll-5 is phylogenetically most similar to Toll-1 and shares the ability to activate the promoter of an antimicrobial peptide gene[55]. Interestingly, Toll-9 is structurally more similar to deuterostomian TLRs, suggesting a function in fly immunity but functional studies on this Toll produced inconsistent results[56,57].

Functional studies on Tolls have also been performed in other insects. In the mosquito *Aedes aegypti* Toll-5A, a gene duplicate homologous to *D. melanogaster* Toll-5, is also involved in immunity as deduced from increased susceptibility of the Toll-5A knock down mutant to the fungus *Beauveria bassiana*[58]. Other arthropods including crustaceans, myriapods and chelicerates contain roughly equal numbers of Toll genes compared to insects, but their function has not been determined[30].

Studies on Toll in the protostomian phylum Nematoda have been limited to the model nematode *Caenorhabditis elegans*. *C. elegans* contains only one Toll gene (TOL-1) and lacks the MyD88 adapter protein and NF-κB transcription factor. Loss of function of the TOL-1 gene causes severe developmental defects as well as increased susceptibility to bacterial infection[59–61], suggesting a dual (MyD88 and NF-κB independent) function of *C. elegans* Toll in development and in interaction with microbes, although the latter function appears less pronounced than in other protostomes like *D. melanogaster*.

For species of the protostomian superphylum Lophotrochozoa functional studies on Tolls and TLRs have not been reported, although genes encoding Tolls, TLRs and signaling molecules including MyD88 and NF-κB transcription factor have been identified[31,62] and are subject to regulation upon bacterial and fungal infection[32,63]. As functional studies on arthropod and nematode Tolls have revealed diverse and specialized functions for a structurally highly related family of proteins, the study on Toll function in lophotrochozoans may provide key insights in the original function of Toll proteins in the protostomian ancestor and the events that have driven functional diversification of protostomian Tolls.

Invertebrate Deuterostome TLR function

The function of TLRs and TLR signaling molecules in diverse deuterostomian animals is still in its infancy. TLR gene expression profiling in various tissues during different life stadia or after infection indicate that the identified genes are expressed. The presence of TLR transcript in immune cells but not embryos of sea urchin

(Echinodermata) suggests that TLRs play a role in the sea urchin immune system[64]. Cephalochordates like amphioxus also have many putative TLR genes most of which are not yet functionally explored (Fig. 2). The TLR1 receptor of Chinese amphioxus (*Branchiostoma belcheri tsingtauense*) functionally interacts with an amphioxus MyD88 adapter and activates NF-κB in a human cell background[65]. As the amphioxus genome is predicted to encode NF-κB orthologs[35], this finding suggests that the TLR signaling pathway as known in arthropods may be conserved in early invertebrate deuterostomes and play a role in the amphioxus immune system. The amphioxus TLR is expressed in microbe-interacting tissues including gills and the gut, and also during the gastrula stage of embryogenesis, which may indicate an additional role for this TLR in amphioxus development.

Two of the three TLRs of the urochordate *Ciona intestinalis* (closely related to vertebrates) recognize diverse microbial ligands (incl. flagellin and nucleic acids) and activate NF-κB in human cells. These ligands also induce expression of the pro-inflammatory cytokine Tumor Necrosis Factor alpha (an ancient and highly conserved inflammatory mediator) in the *C. intestinalis* gastrointestinal tract, suggesting that the TLRs have an immunomodulatory function[38]. The role of the TLRs in *C. intestinalis* development has not been investigated. Overall, the apparent role of invertebrate deuterostomian TLRs in immunity together with the immune function of the studied Hydra TLR-like molecules and some protostomian Tolls, suggests that a role in immunity is one of the ancestral functions of these receptors.

Vertebrate TLR function

In vertebrates the different TLRs are variably expressed in virtually all cell types. The TLRs are considered to scan the environment for microbial ligands and to orchestrate an adequate immune defense. These events are relatively rapid and of low specificity and form the basis of the innate immune response. Activation of TLRs ultimately results in enhanced production of antimicrobial peptides and pro-inflammatory cytokines and chemokines that attract and activate professional immune cells like neutrophils and macrophages[66]. T and B cells also express TLRs. TLR activation in diverse T-cell subsets promotes proliferation, migration and pro-inflammatory cytokine production[67]. In B-cells, activation of TLRs also induces the production of cytokines as well as the expression of co-stimulatory molecules and differentiation of B-cells into immunoglobulin producing plasma cells[68]. The adaptive immune system is unique to vertebrates and provides a much slower but more specific immune response than the innate response.

Besides microbial ligands, TLRs also respond to endogenous danger signals, the so-called damage associated molecular patterns (DAMPs). Endogenous activators include, among others, heat shock protein 60 (HSP60)[69], extracellular matrix

components, antimicrobial peptides, and self-nucleic acid motifs. DAMP activation of TLRs can promote a potent inflammatory response that causes excessive and detrimental immune stimulation. Indeed, the presence of high levels of endogenous TLR ligands are associated with several auto immune diseases[70]. The protostomian Toll receptors may also respond to endogenous danger signals. DAMPs released in *D. melanogaster* hemolymph result in the proteolytic cleavage of endogenous pro-Spätzle to form mature Spätzle which binds to Toll-1 and initiates Toll signaling[71].

Overall, TLRs and Toll share a function in the innate immune system. Whereas Toll has expanded its function to regulate development, TLRs in vertebrates seem to have expanded their function to initiate and regulate the vertebrate specific adaptive immune response. However, evidence is growing that vertebrate TLRs may also have additional functions. In cardiomyocytes, TLR9 activation reduces ATP synthesis via a MyD88 independent pathway, increasing stress tolerance in these cells. This could be beneficial during myocardial ischemia[72]. TLR9 may also be involved in correct neuron and muscle development as mice lacking TLR9 show abnormalities in sensitivity, activity and coordination[73]. Other TLRs including TLR2, 3 and 8 negatively regulate the proliferation of neuronal progenitor cells as well as the outgrowth of axons from neurons in the developing mouse brain[74]. Stimulation of these TLRs in neurons does not activate the transcription factors that drive the immune responses, suggesting that TLRs in neurons function via a (unknown) different signaling network[74]. These vertebrate TLR functions may resemble to some extent the functions of Tolls in *D. melanogaster* neuronal development. Besides its role in embryonic axis formation, Toll-1 of *D. melanogaster* is also involved in proper development of motor-neurons and musculature[75] and Toll-6 and Toll-7 are receptors for neurotrophic factors[76].

Microbe driven TLR evolution

Microbial methods to avoid detection by TLRs

Throughout evolution, microbes and their hosts are in a Red Queens race to prevent their own extinction. In order to survive and reproduce microbes have to invent strategies to resist or evade host defenses, while hosts have to retaliate these strategies to prevent becoming over-exploited. This also holds true for TLRs. One of the primary functions of TLRs is to detect microbes and to limit their numbers via activation of the immune system. Microbes on the other hand, have evolved a great variety of tools to evade the TLR system.

One of the microbial evasion strategies is the degradation of TLR ligand. Some bacterial species including the opportunistic pathogen *P. aeruginosa,* are motile by

using a flagellum which is composed of monomeric flagellin subunits. These flagellin subunits are potent activators of TLR5. However to avoid activation of TLR5, *P. aeruginosa* secretes an alkaline protease (AprA) which degrades released monomeric flagellins (but not intact flagella), thus preventing activation of TLR5 and the development of an innate immune response. *P. aeruginosa* AprA cleaves flagellin in a domain that is conserved across bacterial species and therefore also degrades flagellins of other bacterial species[77]. Homologs of AprA have been identified in other flagellated bacteria, suggesting that flagellin degradation may be an evolutionary successful bacterial strategy for TLR5 evasion.

Microbes may also display virulence factors that physically block TLR recognition. Superantigen-like proteins (SSL3 & 4) of *Staphylococcus aureus* directly interact with the ECD of human and murine TLR2. This interaction presumably blocks the TLR2 ligand binding pocket and hence prevents TLR2 from recognizing *S. aureus* cell-wall components. As a result, innate immune cells incubated with the SSLs and sub-sequentially stimulated with TLR2 ligands show greatly impaired production of pro-inflammatory cytokines[78,79].

When microbes fail to escape recognition by TLRs, activation of the immune system may still be prevented by interfering with the TLR signaling cascade. Bacteria may engage host inhibitory receptors which overrule activating receptors like TLRs[80]. Alternatively, *S. aureus* secretes a TIR domain containing protein (TirS) which interferes with TLR signaling and impairs NF-κB activation and cytokine production[81]. Viruses also use molecular mimicry of signaling domains or degradation of molecules to interrupt TLR signaling[82].

The most obvious microbial strategy to avoid detection by TLRs is to alter the structure of main TLR ligand such as flagellins or LPS. Flagellin subunits of β- and γ-proteobacteria (e.g. the genus *Salmonella*) are recognized by TLR5 and evoke an immune response[83]. However, flagellin of α- and ε-proteobacteria (e.g. the genus *Helicobacter*) is structurally different and lacks the TLR5 binding site, thus preventing detection by TLR5[84]. Similarly, bacteria may alter their lipid A which is the part of LPS that is recognized by TLR4, and thus impair detection by the TLR4/MD-2/CD14 receptor complex[85]. At temperatures of its flea host (21 to 27°C) *Yersinia pestis* produces a lipid A structure that contains six acyl chains that potently activates TLR4. After transfer of *Y. pestis* to its 37°C mammalian host via a flea bite, *Y. pestis* produces tetra-acylated lipid A which no longer activates TLR4 and enables *Y. pestis* to evade host immune activation[86]. The abilities of microbes to alter TLR ligand structures are however limited as the ligands typically consist of conserved molecular patterns that are critical for microbial survival. This was nicely illustrated by substituting the TLR5 binding site in *Salmonella* flagellin with amino acids from

Helicobacter flagellin. The *Salmonella* flagellin with *Helicobacter* amino acids could no longer be detected by TLR5 but at the cost of a loss of bacterial motility[84].

Purifying selection on TLRs

The diversification of microbes and their TLR evasion strategies exert a selective pressure on the evolution of the TLR system. The 'direction' of this selective pressure can be determined using phylogeny based analyses of site specific codon substitutions. By comparing TLR sequences between species, a site is identified to undergo positive selection when the ratio of non-synonymous over synonymous codon substitutions is >1. This indicates that a site has remained polymorphic and may provide a fitness advantage through adaptive evolution. When the ratio of non-synonymous over synonymous codon substitutions is <1 the codon shows no or little variation across species and underwent purifying selection. This indicates that polymorphisms in such a site would generally be detrimental and hence the site evolves under functional constraint[87]. Evolution under functional constraint is seen in TLR adapter proteins, especially MyD88 and TRIF, due to their non-redundant role in signal transduction[88,89]. The TLR adapters interact with multiple proteins and therefore polymorphisms would almost certainly impair their interaction with some of these proteins[19]. Maintaining function also dictates the evolution of the TIR domain[88,90] since it shows a high level of identity across diverse species and substituting even a single important site in the TIR domain can render the TLR inactive[4,91]. In addition, the ECD of nucleic acid sensing TLRs (e.g. TLR3, 7, 8 and 9) harbors polymorphisms but these are rarely found in the ligand binding region, indicating that ligand binding by these TLRs also evolves under functional constraint[92–95]. The reason for this constraint is likely the very similar structure of microbial and host nucleic acids which poses the risk of inducing auto-immune responses. Through purifying selection, detrimental polymorphisms that may have increased the affinity to self-nucleic acids have likely been expelled from the population, thereby minimizing the risk of recognizing self-nucleic acids while maintaining adequate sensing of microbial nucleic acids.

Positive selection on TLRs

In contrast to nucleic acid sensing intracellular TLRs, the ECD of surface-exposed TLRs (e.g. TLR2, 4 and 5) display a strong diversifying evolution driven by positive selection of advantageous mutations. Genomic data from diverse species has allowed the identification of positively selected sites in TLR genes among primates[92], cattle[96], rodents[94], pigs[97], birds[39,93], and fish[98]. Most of these sites are located near to or directly in the ligand binding region. The highly polymorphic make-up of TLR

ligand binding regions in different hosts may have been driven by antagonistic co-evolution with host specific pathogens and/or the necessity to discriminate between host specific commensals and pathogens[99]. Support for this diversifying evolution includes crystallography studies on both the human[100] and mouse[101] TLR4/MD2 complex that indicate that binding of LPS involves species-specific residues and functional studies that demonstrate that mouse TLR4 is activated by both hexa- and penta-acylated LPS[102,103] while human TLR4 only responds to hexa-acylated LPS.

Diversifying evolution may also explain the differential TLR5 response to flagellin in different species. Most of the zebrafish TLR5 residues physically interacting with *Salmonella* flagellin[10] are poorly conserved across vertebrates. Furthermore, functional studies indicate that chicken and mouse TLR5 are more sensitive than human TLR5 to flagellin of *S. enterica* serovar Typhimurium[91]. In addition, the TLR5 of the reptile *Anolis carolinensis* (Carolina anole) proved to be more sensitive than human TLR5 to flagellins of *P. aeruginosa*[104] (**Chapter 3**). The species-specific recognition of TLR ligands, which has likely developed as a result of host specific adaptations to co-evolving microbial communities, may form a basis for the differential susceptibility or resistance to infection as seen between host species. Understanding the evolution of species-specific TLR functioning may thus aid in unveiling fundamental concepts behind zoonotic diseases.

Concluding remarks

TLRs are among the most extensively studied innate immune receptors. Yet, there still remains much to be discovered about the evolution of this receptor family. One topic that merits more in-depth investigation is the greatly expanded number of TLR genes present in early invertebrate deuterostomes. Functional characterizations of (combinations of) TLR genes from these animals may provide evidence of independent evolution of a highly specific, sophisticated microbial recognition system based on innate immune receptors that is similar to the immunoglobulin based adaptive immune system of vertebrates.

Another intriguing issue is the role and evolution of deuterostome TLRs beyond the immune system. Deciphering the roles for Tolls and TLRs in different organismal systems may question our view of the original function of the TLR system. Among the protostomes the Lophotrochozoa seem to deserve more attention to elucidate the functions of the TLR genes in this superphylum[1]. Potential roles for TLRs outside the immune system may also be studied in more species of Cnidaria and early deuterostomes so that the positional origin of TLR functions can be more accurately estimated.

Finally, understanding of the evolution of TLRs would be greatly aided by much broader functional studies involving ligands from a large variety of microbes. Phylogeny based analyses of molecular TLR evolution are able to predict residues with potential relevance for TLR function. Functional studies may provide experimental evidence and unveil the selective pressures that are at the basis of the purifying or diversifying selection of TLRs. Combined, these analyses could be instrumental in deciphering the molecular basis for antagonistic host-specific coevolution with microbes and the resistance to disease that naturally follows. Such knowledge may be of particular interest in animal breeding to select for disease resistant genotypes[93] as well as in predicting an individual's proneness to disease.

References

1. Leulier, F. & Lemaitre, B. Toll-like receptors--taking an evolutionary approach. *Nature Reviews Genetics* **9**, 165–78 (2008).
2. Anderson, K. V., Jürgens, G. & Nüsslein-Volhard, C. Establishment of dorsal-ventral polarity in the *Drosophila* embryo: Genetic studies on the role of the Toll gene product. *Cell* **42**, 779–789 (1985).
3. Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.-M. & Hoffmann, J. A. The Dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell* **86**, 973–983 (1996).
4. Poltorak, A. *et al.* Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* **282**, 2085–2088 (1998).
5. Buchanan, S. G. S. C. & Gay, N. J. Structural and functional diversity in the leucine-rich repeat family of proteins. *Progress in Biophysics and Molecular Biology* **65**, 1–44 (1996).
6. Bell, J. K. *et al.* The molecular structure of the Toll-like receptor 3 ligand-binding domain. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10976–10980 (2005).
7. Choe, J., Kelker, M. S. & Wilson, I. A. Crystal structure of human Toll-like receptor 3 (TLR3) ectodomain. *Science* **309**, 581–585 (2005).
8. Ohto, U., Tanji, H. & Shimizu, T. Structure and function of toll-like receptor 8. *Microbes and Infection* **16**, 273–282 (2014).
9. Ohto, U. *et al.* Structural basis of CpG and inhibitory DNA recognition by Toll-like receptor 9. *Nature* **520**, 702–705 (2015).
10. Yoon, S. I. *et al.* Structural basis of TLR5-flagellin recognition and signaling. *Science* **335**, 859–64 (2012).
11. Parthier, C. *et al.* Structure of the Toll-Spätzle complex, a molecular hub in *Drosophila* development and innate immunity. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 6281–6286 (2014).
12. Matsushima, N. *et al.* Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors. *BMC Genomics* **8**, 124 (2007).
13. Kang, J. Y. *et al.* Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. *Immunity* **31**, 873–884 (2009).
14. LeBouder, E. *et al.* Soluble forms of Toll-like receptor (TLR)2 capable of modulating TLR2 signaling are present in human plasma and breast milk. *Journal of Immunology* **171**, 6680–6689 (2003).

15. Iwami, K. *et al.* Cutting edge: naturally occurring soluble form of mouse Toll-like receptor 4 inhibits lipopolysaccharide signaling. *Journal of Immunology* **165**, 6682–6686 (2000).
16. Tsujita, T. *et al.* Sensing bacterial flagellin by membrane and soluble orthologs of Toll-like receptor 5 in rainbow trout (*Onchorhynchus mikiss*). *Journal of Biological Chemistry* **279**, 48588–48597 (2004).
17. Ve, T., Williams, S. J. & Kobe, B. Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains. *Apoptosis* **20**, 250–261 (2014).
18. Xu, Y. *et al.* Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature* **408**, 111–115 (2000).
19. O'Neill, L. A. J. & Bowie, A. G. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nature Reviews Immunology* **7**, 353–364 (2007).
20. Gay, N. J., Symmons, M. F., Gangloff, M. & Bryant, C. E. Assembly and localization of Toll-like receptor signalling complexes. *Nature Reviews Immunology* **14**, 546–558 (2014).
21. Ausubel, F. M. Are innate immune signaling pathways in plants and animals conserved? *Nature Immunology* **6**, 973–979 (2005).
22. Boller, T. & Felix, G. A Renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology* **60**, 379–406 (2009).
23. Wiens, M. *et al.* Toll-like receptors are part of the innate immune defense system of sponges (Demospongiae: Porifera). *Molecular Biology and Evolution* **24**, 792–804 (2007).
24. Gauthier, M. E. A., Pasquier, L. Du & Degnan, B. M. The genome of the sponge *Amphimedon queenslandica* provides new perspectives into the origin of Toll-like and interleukin 1 receptor pathways. *Evolution & Development* **12**, 519–533 (2010).
25. Bosch, T. C. G. *et al.* Uncovering the evolutionary history of innate immunity: The simple metazoan Hydra uses epithelial cells for host defence. *Developmental & Comparative Immunology* **33**, 559–569 (2009).
26. Miller, D. J. *et al.* The innate immune repertoire in Cnidaria - ancestral complexity and stochastic gene loss. *Genome Biology* **8**, R59 (2007).
27. Imler, J.-L. & Zheng, L. Biology of Toll receptors: lessons from insects and mammals. *Journal of Leukocyte Biology* **75**, 18–26 (2004).
28. Rast, J. P., Smith, L. C., Loza-Coll, M., Hibino, T. & Litman, G. W. Genomic insights into the immune system of the sea urchin. *Science* **314**, 952–956 (2006).
29. Waterhouse, R. M. *et al.* Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science* **316**, 1738–1743 (2007).
30. Palmer, W. J. & Jiggins, F. M. Comparative genomics reveals the origins and diversity of arthropod immune systems. *Molecular Biology and Evolution* **32**, 2111–2129 (2015).
31. Rauta, P. R., Samanta, M., Dash, H. R., Nayak, B. & Das, S. Toll-like receptors (TLRs) in aquatic animals: signaling pathways, expressions and immune responses. *Immunology Letters* **158**, 14–24 (2014).
32. Toubiana, M. *et al.* Toll-like receptors and MyD88 adaptors in Mytilus: complete cds and gene expression levels. *Developmental & Comparative Immunology* **40**, 158–166 (2013).
33. Davidson, C. R., Best, N. M., Francis, J. W., Cooper, E. L. & Wood, T. C. Toll-like receptor genes (TLRs) from *Capitella capitata* and *Helobdella robusta* (Annelida). *Developmental & Comparative Immunology* **32**, 608–612 (2008).
34. Buckley, K. M. & Rast, J. P. Dynamic evolution of toll-like receptor multigene families in echinoderms. *Frontiers in Immunology* **3**, 136 (2012).
35. Huang, S. *et al.* Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. *Genome Research* **18**, 1112–1126 (2008).
36. Ward, A. E. & Rosenthal, B. M. Evolutionary responses of innate immunity to adaptive immunity. *Infection, Genetics and Evolution* **21**, 492–496 (2014).

37. Azumi, K. *et al.* Genomic analysis of immunity in a Urochordate and the emergence of the vertebrate immune system: "waiting for Godot". *Immunogenetics* **55**, 570–581 (2003).
38. Sasaki, N., Ogasawara, M., Sekiguchi, T., Kusumoto, S. & Satake, H. Toll-like receptors of the Ascidian *Ciona intestinalis* prototypes with hybrid functionalities of vertebrate Toll-like receptors. *Journal of Biological Chemistry* **284**, 27336–27343 (2009).
39. Alcaide, M. & Edwards, S. V. Molecular evolution of the Toll-like receptor multigene family in birds. *Molecular Biology Evolution* **28**, 1703–1715 (2011).
40. Babik, W. *et al.* Constraint and adaptation in newt Toll-like receptor genes. *Genome Biology and Evolution* **7**, 81–95 (2015).
41. Kasamatsu, J., Oshiumi, H., Matsumoto, M., Kasahara, M. & Seya, T. Phylogenetic and expression analysis of lamprey toll-like receptors. *Developmental & Comparative Immunology* **34**, 855–865 (2010).
42. Roach, J. C. *et al.* The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 9577–9582 (2005).
43. Zoete, M. R. de, Bouwman, L. I., Keestra, A. M. & van Putten, J. P. M. Cleavage and activation of a Toll-like receptor by microbial proteases. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 4968–4973 (2011).
44. Oosting, M. *et al.* Human TLR10 is an anti-inflammatory pattern-recognition receptor. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E4478–E4484 (2014).
45. Quiniou, S. M. A., Boudinot, P. & Bengtén, E. Comprehensive survey and genomic characterization of Toll-like receptors (TLRs) in channel catfish, *Ictalurus punctatus*: identification of novel fish TLRs. *Immunogenetics* **65**, 511–530 (2013).
46. Kucera, K. *et al.* Structure-based analysis of *Toxoplasma gondii* profilin: a parasite-specific motif is required for recognition by Toll-like receptor 11. *Journal of Molecular Biology* **403**, 616–629 (2010).
47. Raetz, M. *et al.* Cooperation of TLR12 and TLR11 in the IRF8-dependent IL-12 response to *Toxoplasma gondii* profilin. *Journal of Immunology* **191**, 4818–4827 (2013).
48. Keestra, A. M., Zoete, M. R. de, Bouwman, L. I. & van Putten, J. P. M. Chicken TLR21 is an innate CpG DNA receptor distinct from mammalian TLR9. *Journal of Immunology* **185**, 460–467 (2010).
49. Kanwal, Z., Wiegertjes, G. F., Veneman, W. J., Meijer, A. H. & Spaink, H. P. Comparative studies of Toll-like receptor signalling using zebrafish. *Developmental & Comparative Immunology* **46**, 35–52 (2014).
50. Sepulcre, M. P. *et al.* Evolution of lipopolysaccharide (LPS) recognition and signaling: fish TLR4 does not recognize LPS and negatively regulates NF-κB activation. *Journal of Immunology* **182**, 1836–1845 (2009).
51. Boyd, A. C. *et al.* TLR15 is unique to avian and reptilian lineages and recognizes a yeast-derived agonist. *Journal of Immunology* **189**, 4930–4938 (2012).
52. Franzenburg, S. *et al.* MyD88-deficient Hydra reveal an ancient function of TLR signaling in sensing bacterial colonizers. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 19374–19379 (2012).
53. Paré, A. C. *et al.* A positional Toll receptor code directs convergent extension in *Drosophila*. *Nature* **515**, 523–527 (2014).
54. Meyer, S. N. *et al.* An ancient defense system eliminates unfit cells from developing tissues during cell competition. *Science* **346**, 1258236 (2014).
55. Tauszig, S., Jouanguy, E., Hoffmann, J. A. & Imler, J.-L. Toll-related receptors and the control of antimicrobial peptide expression in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 10520–10525 (2000).
56. Ooi, J. Y., Yagi, Y., Hu, X. & Ip, Y. T. The *Drosophila* Toll-9 activates a constitutive antimicrobial defense. *EMBO reports* **3**, 82–87 (2002).

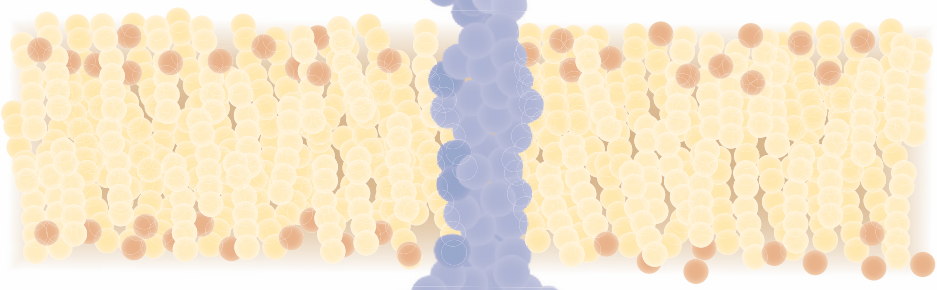
57. Narbonne-Reveau, K., Charroux, B. & Royet, J. Lack of an antibacterial response defect in *Drosophila* Toll-9 mutant. *PLoS ONE* **6**, e17470 (2011).
58. Shin, S. W., Bian, G. & Raikhel, A. S. A Toll receptor and a cytokine, Toll5A and Spz1C, are involved in Toll antifungal immune signaling in the mosquito *Aedes aegypti*. *Journal of Biological Chemistry* **281**, 39388–39395 (2006).
59. Pujol, N. *et al.* A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. *Current Biology* **11**, 809–821 (2001).
60. Tenor, J. L. & Aballay, A. A conserved Toll-like receptor is required for *Caenorhabditis elegans* innate immunity. *EMBO reports* **9**, 103–109 (2008).
61. Pradel, E. *et al.* Detection and avoidance of a natural product from the pathogenic bacterium *Serratia marcescens* by *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 2295–2300 (2007).
62. Halanych, K. M. & Kocot, K. M. Repurposed transcriptomic data facilitate discovery of innate immunity Toll-like receptor (TLR) genes across Lophotrochozoa. *The Biological Bulletin* **227**, 201–209 (2014).
63. Zhang, S.-M. & Coultas, K. A. Identification and characterization of five transcription factors that are associated with evolutionarily conserved immune signaling pathways in the schistosome-transmitting snail *Biomphalaria glabrata*. *Molecular Immunology* **48**, 1868–1881 (2011).
64. Buckley, K. M. & Rast, J. P. Diversity of animal immune receptors and the origins of recognition complexity in the deuterostomes. *Developmental & Comparative Immunology* **49**, 179–189 (2015).
65. Yuan, S. *et al.* An amphioxus TLR with dynamic embryonic expression pattern responses to pathogens and activates NF-κB pathway via MyD88. *Molecular Immunology* **46**, 2348–2356 (2009).
66. Takeuchi, O. & Akira, S. Pattern recognition receptors and inflammation. *Cell* **140**, 805–20 (2010).
67. Reynolds, J. M. & Dong, C. Toll-like receptor regulation of effector T lymphocyte function. *Trends in Immunology* **34**, 511–519 (2013).
68. Buchta, C. M. & Bishop, G. A. Toll-like receptors and B cells: functions and mechanisms. *Immunologic Research* **59**, 12–22 (2014).
69. Ohashi, K., Burkart, V., Flohé, S. & Kolb, H. Cutting edge: heat shock protein 60 is a putative endogenous ligand of the Toll-like receptor-4 complex. *Journal of Immunology* **164**, 558–561 (2000).
70. Piccinini, A. M. & Midwood, K. S. DAMPening inflammation by modulating TLR signalling. *Mediators of Inflammation* **2010**, e672395 (2010).
71. Ming, M., Obata, F., Kuranaga, E. & Miura, M. Persephone/Spätzle pathogen sensors mediate the activation of Toll receptor signaling in response to endogenous danger signals in apoptosis-deficient *Drosophila*. *Journal of Biological Chemistry* **289**, 7558–7568 (2014).
72. Shintani, Y. *et al.* Toll-like receptor 9 protects non-immune cells from stress by modulating mitochondrial ATP synthesis through the inhibition of SERCA2. *EMBO reports* **15**, 438–445 (2014).
73. Khariv, V. *et al.* Toll-like receptor 9 deficiency impacts sensory and motor behaviors. *Brain, Behavior, and Immunity* **32**, 164–172 (2013).
74. Okun, E., Griffioen, K. J. & Mattson, M. P. Toll-like receptor signaling in neural plasticity and disease. *Trends in Neurosciences* **34**, 269–281 (2011).
75. Halfon, M. S., Hashimoto, C. & Keshishian, H. The *Drosophila* Toll gene functions zygotically and is necessary for proper motoneuron and muscle development. *Developmental Biology* **169**, 151–167 (1995).
76. McIlroy, G. *et al.* Toll-6 and Toll-7 function as neurotrophin receptors in the *Drosophila melanogaster* CNS. *Nature Neuroscience* **16**, 1248–1256 (2013).

77. Bardoel, B. W. *et al. Pseudomonas* evades immune recognition of flagellin in both mammals and plants. *PLoS Pathogens* **7**, e1002206 (2011).
78. Bardoel, B. W. *et al.* Evasion of Toll-like receptor 2 activation by staphylococcal superantigen-like protein 3. *Journal of Molecular Medicine* **90**, 1109–1120 (2012).
79. Yokoyama, R. *et al.* Staphylococcal superantigen-like protein 3 binds to the Toll-like receptor 2 extracellular domain and inhibits cytokine production induced by *Staphylococcus aureus*, cell wall component, or lipopeptides in murine macrophages. *Infection and Immunity* **80**, 2816–2825 (2012).
80. Van Avondt, K., van Sorge, N. M. & Meyaard, L. Bacterial immune evasion through manipulation of host inhibitory immune signaling. *PLoS Pathogens* **11**, e1004644 (2015).
81. Askarian, F. *et al.* A *Staphylococcus aureus* TIR domain protein virulence factor blocks TLR2-mediated NF-κB signaling. *Journal of Innate Immunity* **6**, 485-498 (2014).
82. Bowie, A. G. & Unterholzner, L. Viral evasion and subversion of pattern-recognition receptor signalling. *Nature Reviews Immunology* **8**, 911–922 (2008).
83. Hayashi, F. *et al.* The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* **410**, 1099–103 (2001).
84. Andersen-Nissen, E. *et al.* Evasion of Toll-like receptor 5 by flagellated bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 9247–9252 (2005).
85. Maeshima, N. & Fernandez, R. C. Recognition of lipid A variants by the TLR4-MD-2 receptor complex. *Frontiers in Cellular and Infection Microbiology* **3**, (2013).
86. Montminy, S. W. *et al.* Virulence factors of *Yersinia pestis* are overcome by a strong lipopolysaccharide response. *Nature Immunology* **7**, 1066–1073 (2006).
87. Yang, Z. & Nielsen, R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution* **19**, 908–917 (2002).
88. Nakajima, T. *et al.* Natural selection in the TLR-related genes in the course of primate evolution. *Immunogenetics* **60**, 727–735 (2008).
89. Fornarino, S. *et al.* Evolution of the TIR domain-containing adaptors in humans: swinging between constraint and adaptation. *Molecular Biology and Evolution* **28**, 3087–3097 (2011).
90. Mikami, T., Miyashita, H., Takatsuka, S., Kuroki, Y. & Matsushima, N. Molecular evolution of vertebrate Toll-like receptors: evolutionary rate difference between their leucine-rich repeats and their TIR domains. *Gene* **503**, 235–243 (2012).
91. Keestra, A. M., de Zoete, M. R., van Aubel, R. A. M. H. & van Putten, J. P. M. Functional characterization of chicken TLR5 reveals species-specific recognition of flagellin. *Molecular Immunology* **45**, 1298–1307 (2008).
92. Wlasiuk, G. & Nachman, M. W. Adaptation and constraint at Toll-like receptors in primates. *Molecular Biology and Evolution* **27**, 2172–2186 (2010).
93. Vinkler, M., Bainová, H. & Bryja, J. Protein evolution of Toll-like receptors 4, 5 and 7 within Galloanserae birds. *Genetics Selection Evolution* **46**, 72 (2014).
94. Fornůsková, A. *et al.* Contrasted evolutionary histories of two Toll-like receptors (Tlr4 and Tlr7) in wild rodents (MURINAE). *BMC Evolutionary Biology* **13**, 194 (2013).
95. Webb, A. E. *et al.* Adaptive evolution as a predictor of species-specific innate immune response. *Molecular Biology and Evolution* **32**, 1717–1729 (2015).
96. Jann, O. C., Werling, D., Chang, J.-S., Haig, D. & Glass, E. J. Molecular evolution of bovine Toll-like receptor 2 suggests substitutions of functional relevance. *BMC Evolutionary Biology* **8**, 288 (2008).
97. Darfour-Oduro, K. A., Megens, H.-J., Roca, A. L., Groenen, M. A. M. & Schook, L. B. Adaptive evolution of Toll-like receptors (TLRs) in the family Suidae. *PLoS ONE* **10**, e0124069 (2015).
98. Sundaram, A. Y. M., Consuegra, S., Kiron, V. & Fernandes, J. M. O. Positive selection pressure within teleost toll-like receptors tlr21 and tlr22 subfamilies and their response to temperature stress and microbial components in zebrafish. *Molecular Biology Reports* **39**, 8965–8975 (2012).

99. Werling, D., Jann, O. C., Offord, V., Glass, E. J. & Coffey, T. J. Variation matters: TLR structure and species-specific pathogen recognition. *Trends in Immunology* **30**, 124–130 (2009).
100. Park, B. S. *et al.* The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex. *Nature* **458**, 1191–1195 (2009).
101. Ohto, U., Fukase, K., Miyake, K. & Shimizu, T. Structural basis of species-specific endotoxin sensing by innate immune receptor TLR4/MD-2. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 7421–7426 (2012).
102. Hajjar, A. M., Ernst, R. K., Tsai, J. H., Wilson, C. B. & Miller, S. I. Human Toll-like receptor 4 recognizes host-specific LPS modifications. *Nature Immunology* **3**, 354–359 (2002).
103. Steeghs, L. *et al.* Differential activation of human and mouse Toll-like receptor 4 by the adjuvant candidate LpxL1 of *Neisseria meningitidis*. *Infection and Immunity* **76**, 3801–3807 (2008).
104. Voogdt, C. G. P., Bouwman, L. I., Kik, M. J. L., Wagenaar, J. A. & van Putten, J. P. M. Reptile Toll-like receptor 5 unveils adaptive evolution of bacterial flagellin recognition. *Scientific Reports* **6**, 19046 (2016).

Chapter 3

Reptile Toll-like receptor 5 unveils adaptive evolution of bacterial flagellin recognition



Carlos G.P. Voogdt[1], Lieneke I. Bouwman[1], Marja J.L. Kik[2], Jaap A. Wagenaar[1,3], Jos P.M. van Putten[1]

[1]Department of Infectious Diseases & Immunology, Utrecht University, Utrecht, the Netherlands
[2]Department of Pathobiology, Utrecht University, Utrecht, The Netherlands
[3]Wageningen Bioveterinary Research, Lelystad, the Netherlands

Scientific Reports 6:19046 2016

Abstract

Toll-like receptors (TLR) are ancient innate immune receptors crucial for immune homeostasis and protection against infection. TLRs are present in mammals, birds, amphibians and fish but have not been functionally characterized in reptiles despite the central position of this animal class in vertebrate evolution. Here we report the cloning, characterization, and function of TLR5 of the reptile *Anolis carolinensis* (Green Anole lizard). The receptor (acTLR5) displays the typical TLR protein architecture with 22 extracellular leucine rich repeats flanked by a N- and C-terminal leucine rich repeat domain, a membrane-spanning region, and an intracellular TIR domain. The receptor is phylogenetically most similar to TLR5 of birds and most distant to fish TLR5. Transcript analysis revealed acTLR5 expression in multiple lizard tissues. Stimulation of acTLR5 with TLR ligands demonstrated unique responsiveness towards bacterial flagellin in both reptile and human cells. Comparison of acTLR5 and human TLR5 using purified flagellins revealed differential sensitivity to *Pseudomonas* but not *Salmonella* flagellin, indicating development of species-specific flagellin recognition during the divergent evolution of mammals and reptiles. Our discovery of reptile TLR5 fills the evolutionary gap regarding TLR conservation across vertebrates and provides novel insights in functional evolution of host-microbe interactions.

Introduction

Toll-like receptors (TLRs) form a family of evolutionarily highly conserved innate immune receptors that play a crucial role in immune homeostasis and the response to infection[1,2]. TLRs are glycoproteins that typically consist of an extracellular sensor domain (ECD) composed of multiple leucine rich repeats (LRR), a transmembrane domain (TM) and an intracellular Toll/Interleukin-1 receptor (TIR) signaling domain[3]. The ECD senses the presence of conserved microbial structures in the environment and transduces this signal to the TIR domain which acts as a docking station for intracellular adapter proteins like Myeloid differentiation primary response gene 88 (MyD88). The formed complex then initiates a cascade of events that ultimately results in nuclear translocation of transcription factors like Nuclear factor kappa light chain enhancer of activated B cells (NF-κB) that direct the innate and adaptive immune response[4].

Throughout evolution, selective pressures exerted by microbes have driven diversification of the TLR ECD, resulting in a family of distinct receptors that recognize a variety of mainly microbial ligands[5]. For example, TLR4 binds bacterial lipopolysaccharide[6]; TLR9 or 21 recognizes bacterial nucleic acid motifs[7,8] and avian TLR15 is uniquely activated by microbial proteases via cleavage of the receptor ectodomain[9]. TLR5 senses flagellin subunits[10] that make up the flagellum of certain bacterial species including *Salmonella enterica* and *Pseudomonas aeruginosa*. Besides structural diversity between TLR family members, coevolution with microbes has also led to adaptive evolution of individual TLRs[11–13], leading to differential recognition of TLR ligands between animal species[14–16].

Within the animal kingdom, the TLR repertoire varies among species. Regarding vertebrates, genome wide studies have identified 16 TLR types in lampreys compared to 20 in bony fish, 21 in amphibians and 10 in both humans and birds[4,17–20]. The dynamic evolution between and within TLR family members and the conservation of TLRs across highly diverse animals underlines the importance of TLRs throughout vertebrate evolution. However, one major gap in our knowledge on vertebrate TLR evolution is the complete lack of information about the structure, function, and ligand specificity of TLRs in any species of reptile. Reptiles have a unique physiology, being the only poikilothermic amniotes, and take a central position in vertebrate evolution[21]. The first reptiles evolved around 330 to 310 million years ago (Mya) from an amphibian-like ancestor[22]. Development of the amniotic egg and a water impermeable skin allowed these early reptiles to be the first vertebrates that could permanently colonize terrestrial

habitats. This pioneering step must have brought the first reptiles into contact with the prehistorical terrestrial flora, fauna and microbiota that undoubtedly shaped the immune system of reptiles and descendant animals. Yet compared to other vertebrates our knowledge on the reptile immune system, especially concerning molecular insights in reptile microbe interactions, is marginal[21].

In present study we report the molecular cloning, characterization and function of the first reptile TLR namely TLR5 of the 'New world' lizard *Anolis carolinensis* (acTLR5). Evidence is provided that acTLR5 is closely related to other TLR5 orthologs and responds to bacterial flagellin, even when expressed in human cells. Differential sensitivity of acTLR5 compared to human TLR5 to *Pseudomonas aeruginosa* but not *Salmonella* Enteritidis flagellins indicate host specific adaptation of flagellin recognition.

Results

Reptile cells respond to bacterial flagellin

To assess whether reptile cells respond to TLR ligands we first stimulated IgH-2 *Iguana iguana* cells carrying a NF-κB luciferase reporter plasmid with the canonical mammalian TLR ligands; LTA (TLR2), Pam_3CSK_4 (TLR2/1), FSL-1 (TLR2/6), LPS (TLR4), FliC (flagellin of *Salmonella enterica* serovar Enteritidis) (TLR5), CL097 (TLR7), ODN2006 (TLR9) and the avian TLR15 activator Proteinase K. None of these TLR agonists elicited significant NF-κB activity except for bacterial flagellin (Fig. 1). In search for the putative TLR receptor conferring this response, and by absence of the *I. iguana* whole genome sequence, we interrogated the whole genome sequence of the related model organism *Anolis carolinensis*[23,24], using BLAST with mammalian and chicken TLR protein sequences as queries. This search yielded nine putative TLR orthologs including a putative TLR5 ortholog (Genbank accession number: XP003216083.1), which was designated as acTLR5.

Expression and characterization of the ac*tlr5* gene

To verify that the putative acTLR5 ortholog is expressed *in vivo* in the Anolis lizard, we tested total mRNA isolated from different organs of an adult male for the presence of the ac*tlr5* transcript using RT-PCR with glyceraldehyde 3-phosphate dehydrogenase (ac*gapdh*) as a control. Transcripts of ac*tlr5* were detected in all the tissues tested including lung, heart, stomach, liver, spleen, kidneys, intestine and testis (Fig. 2), indicating that the gene product is expressed and may be functional in various tissues.

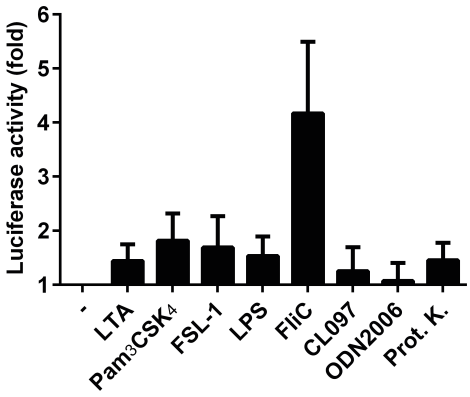


Figure 1. Flagellin stimulation activates NF-κB in IgH-2 cells.
Iguana iguana IgH-2 cells were transfected with a NF-κB luciferase reporter plasmid and stimulated (5 h) with the following TLR ligands: LTA (1 µg/ml), Pam_3CSK_4 (0.1 µg/ml), FSL-1 (0.1 µg/ml), LPS (0.1 µg/ml), FliC (flagellin) (1 µg/ml), CL097 (2 µg/ml), ODN2006 (500 nM) and Proteinase K (2 ng/ml). Data represent the fold increase of luciferase activity compared to the unstimulated control (-). Values are the mean ± SEM of three independent experiments performed in duplicate.

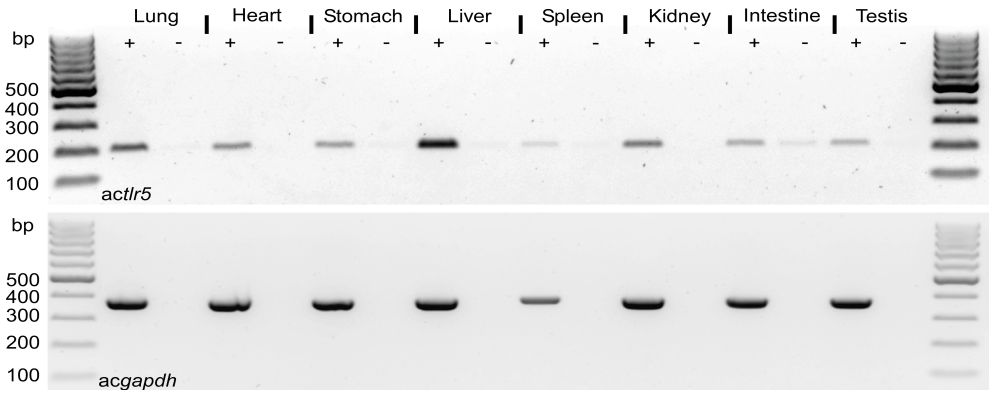


Figure 2. Expression of acTLR5 transcript in multiple tissues of *A. carolinensis*.
RT-PCR analysis on total RNA extracted from the indicated tissues of an *A. carolinensis* lizard after reverse transcription into cDNA (+) or without the reverse transcription step (-). PCR amplified a 216 bp (base pair) fragment of ac*tlr5* or (as control) a 374 bp fragment of *A. carolinensis* glyceraldehyde 3-phosphate dehydrogenase (ac*gapdh*).

In order to examine the function of the acTLR5 we cloned the *tlr5* gene from genomic DNA of an adult male *A. carolinensis*. The gene consisted of a single exon encoding a protein of 871 amino acids that contained typical TLR domains. These included an ECD (residues 28 to 634) containing 24 LRRs (including N- and C-terminal LRR) as found in other TLR5 orthologs[25], a TM domain (residues 647 to 665) and an intracellular TIR signaling domain (residues 697 to 840). The amino

acid sequence differed from the *A. carolinensis* reference sequence at positions: 471 (H471L), 550 (V550A), 642 (S642P) and 658 (F658Y), suggesting the existence of polymorphisms in TLR5 of *A. carolinensis*.

Phylogenetic analysis using full-length protein sequences of different TLR types from several vertebrates including fish, amphibians, birds and mammals clustered acTLR5 with other TLR5 orthologs and in particular with chicken TLR5 (Fig. S1). BLAST analysis with the ECD, TM and TIR domains as separate queries indicated that all three domains of acTLR5 were most similar to (predicted) TLR5 sequences of other reptiles and birds and least similar to fish TLR5 (Table 1), fully in line with the evolutionary relationships among these vertebrates.

Table 1. Similarity (%) of acTLR5 domains with several vertebrate TLR5 orthologs

TLR5 Species	Accession number	ECD[b]	TM[c]	TIR[d]
Python bivittatus (r)[a]	XP_007434471.1	76	68	95
Chelonia mydas (r)	EMP25733.1	74	73	94
Alligator mississippiensis (r)	XP_006270945.1	72	77	93
Gallus gallus (b)	ABW07794.1	70	64	90
Columba livia (b)	AIK67343.1	70	77	92
Anser anser (b)	AFP65787.1	68	72	91
Xenopus leavis (a)	NP_001088449.1	65	47	90
Bos taurus (m)	ABC68311.1	65	69	85
Homo sapiens (m)	NP_003259.2	64	69	85
Mus musculus (m)	AAI25262.1	63	69	84
Oncorhynchus mykiss (f)	NP_001118216.1	58	53	80
Takifugu rubripes (f)	AAW69374.1	57	54	77
Danio rerio (f)	NP_001124067.1	56	57	78

[a] r: reptile; b: bird; a: amphibian; m: mammal; f: fish

[b] ECD: extracellular domain, residues 28 to 634

[c] TM: transmembrane domain, residues 647 to 665

[d] TIR: Toll/interleukin-1 receptor domain, residues 697 to 840

acTLR5 is functional in reptile but also in human cells

Evidence for the function of acTLR5 was sought by introducing an expression vector carrying ac*tlr5* (or a control plasmid without insert) together with a NF-κB luciferase reporter plasmid into reptile IgH-2 cells. Stimulation of the mock-transfected cells with *S.* Enteritidis flagellin (FliC) increased NF-κB activity in these cells, confirming the results depicted in Fig. 1. However, stimulation with *S.* Enteritidis flagellin significantly increased NF-κB activity in acTLR5 transfected cells ($p < 0.05$) (Fig. 3A), indicating that recombinant acTLR5 is functional in the transfected reptile cells and responds to flagellin. To ensure the specificity of this response, cells were stimulated with FSL-1, a synthetic lipoprotein known to be

recognized by TLR2 and TLR6 heterodimers. A high dose of FSL-1 yielded similar responses in empty vector and acTLR5 transfected cells (Fig. 3A) confirming the specificity of the flagellin-induced acTLR5 response.

Reptiles and mammals have evolved independently over more than 300 million years[22]. Yet, a sequence alignment of acTLR5 with human and other vertebrate TLR5 orthologs indicated strong conservation across vertebrates of a critical proline[15] and tyrosine[26] residue as well as a phosphorylation motif[27] in the TLR5 signaling domain (Fig. S2). To determine whether TLR5 signaling has evolved under strong functional constraint, the functioning of acTLR5 was determined in human HeLa-57A cells which do not endogenously express TLRs and stably express the NF-κB luciferase reporter[28]. Stimulation with FliC, and not with other TLR ligands, yielded a strong increase in NF-κB activity in acTLR5 transfected human cells compared to control cells carrying empty vector (Fig. 3B). This functionality of reptile TLR5 in human cells strongly suggests that the expression and trafficking of the receptor and its signaling properties as well as its ligand specificity have been functionally conserved across the reptile and mammalian lineage.

Finally, to verify that acTLR5 was also able to recognize native (non-recombinant) flagellin we incubated acTLR5 transfected cells with live wild-type *S.* Enteritidis (WT) or its isogenic flagellin deficient derivative (Δ*fliC*). Only incubation with wild-type *S.* Enteritidis resulted in NF-κB activation in an acTLR5 dependent manner, confirming that TLR5 is a bonafide reptile receptor for bacterial flagellin (Fig. 3C).

Reptile and human TLR5 recognize the D1 domain in flagellin

Now that we had identified acTLR5 as a specific receptor for bacterial flagellin, we examined the conservation of residues involved in flagellin binding by aligning acTLR5 (and also chicken, African clawed frog and human TLR5) with zebrafish TLR5b of which the crystal structure in complex with flagellin has been determined[29]. The alignment showed that only 40% (18/45) of the zebrafish TLR5b-flagellin interacting residues resemble the residues at the same positions in acTLR5 and the other vertebrate TLR5 sequences (Fig. S2), suggesting a differential basis for the structural recognition of flagellin among these vertebrates.

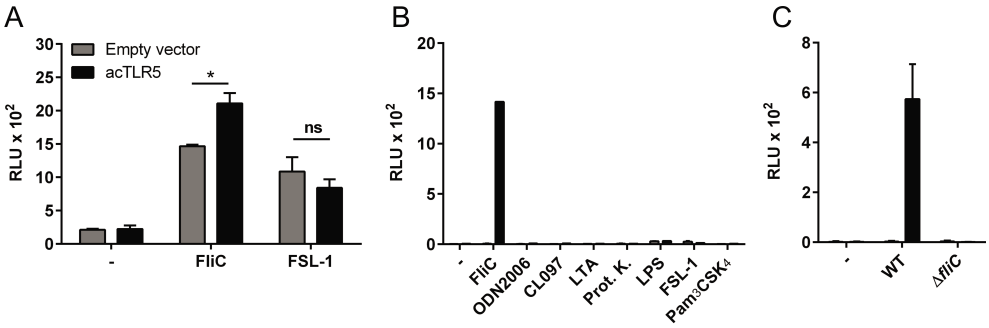


Figure 3. Response of acTLR5 expressed in reptile and human cells.
(A) IgH-2 cells transfected with a NF-κB luciferase reporter plasmid and either empty vector or acTLR5 were stimulated with FliC (5 ng/ml) or FSL-1 (250 ng/ml) (10 h). * $p < 0.05$ by unpaired Students *t*-test. ns; not-significant. **(B)** NF-κB activation in human HeLa-57A cells transfected with empty vector or acTLR5 plasmid after 5 h of stimulation with the following TLR ligands: (-): unstimulated; FliC: 1 μg/ml; ODN2006: 500 nM; CL097: 2 μg/ml; LTA: 1 μg/ml; Proteinase K: 2 ng/ml; LPS: 0.1 μg/ml; FSL-1: 0.1 μg/ml; Pam_3CSK_4: 0.1 μg/ml. **(C)** HeLa-57A cells transfected with empty vector or acTLR5 were incubated (5 h) with sterile LB medium (-), $2.5 \cdot 10^4$ live wild-type *Salmonella* Enteritidis (WT) or the isogenic flagellin deficient strain (Δ*fliC*). Data represent the mean ± SEM luciferase activity in relative light units (RLU) of three independent experiments performed in duplicate **(A and C)** or the mean RLU of a representative of three independent experiments performed in duplicate **(B)**.

To determine whether the structural differences in TLR5 have influenced flagellin recognition throughout the divergent evolution of reptiles and mammals we mapped the domain of flagellin that is recognized by acTLR5. For this, we took advantage of the fact that the D1 domain of flagellins of *γ*- and *β*- Proteobacteria (incl. *Salmonella*, and *Pseudomonas* species) activate TLR5, whereas, due to compositional changes, the D1 domain of flagellins of *α*- and *δ*- Proteobacteria (incl. *Campylobacter* species) escapes recognition by TLR5[30,31]. acTLR5 or human TLR5 (hTLR5) were transfected in HeLa-57A cells and stimulated with purified recombinant *S.* Enteritidis flagellin (FliC) or *Campylobacter jejuni* flagellin (FlaA). This showed that *S.* Enteritidis FliC but not *Campylobacter* FlaA activated NF-κB in both acTLR5 and hTLR5 transfected cells (Fig. 4). To ascertain that the unresponsiveness of acTLR5 and hTLR5 to *Campylobacter* FlaA involved the FlaA D1 domain, we stimulated both TLRs with NHC flagellin. NHC is a chimeric flagellin based on *Campylobacter* FlaA in which the D1 domain was exchanged for the *S.* Enteritidis FliC D1 domain[31]. Indeed, this swapping of the D1 domain restored the activation of hTLR5 and acTLR5 (Fig. 4), indicating that both receptors recognize the D1 region of *Salmonella* but not *Campylobacter* flagellin (Fig. 4) and thus that this ability is conserved between reptiles and humans. The inability of acTLR5 and hTLR5 to recognize *Campylobacter* flagellin may further

indicate that evasion of TLR5 detection by *Campylobacter* developed before the divergence of reptiles and mammals.

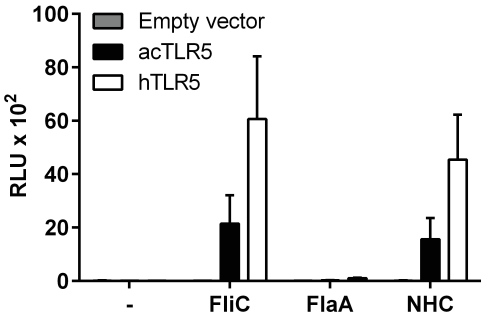


Figure 4. Activation of acTLR5 by the D1 domain of *Salmonella* FliC.
HeLa-57A cells transfected with empty vector, acTLR5 or human TLR5 (hTLR5) were stimulated (5 h) with *S.* Enteritidis FliC, *C. jejuni* FlaA or chimeric NHC flagellin (1 µg/ml). Results represent the mean ± SEM luciferase activity as relative light units (RLU) of three independent experiments performed in duplicate.

Lysate of *Pseudomonas* activates reptile but not human TLR5

Despite recognition of the same flagellin D1 domain, chicken, mouse and human TLR5 respond differently to flagellins of various bacterial species[14,15], suggesting host specific adaptations in bacterial flagellin recognition. To determine whether specific adaptations in flagellin recognition have also occurred in reptiles, we compared the response of acTLR5 transfected HeLa-57A cells with hTLR5 transfected cells to stimulation with bacterial lysate of *S.* Enteritidis, *C. jejuni* and three motile reptile isolates i.e., a *Campylobacter fetus* subsp. *testudinum*[32], *Aeromonas hydrophila* and *Pseudomonas aeruginosa.* For *S.* Enteritidis and *C. jejuni* the use of lysate closely resembled the response to purified flagellins (compare Fig. 5A to Fig. 4). Stimulation of the cells with the reptile *C. fetus* subsp. *testudinum* lysate did not activate acTLR5 or hTLR5 (Fig. 5A) suggesting a similar evasion of TLR5 recognition by this reptile strain as noted for mammalian and chicken derived *Campylobacter* strains[31]. The lysate of *A. hydrophila* activated acTLR5 and hTLR5 equally well (Fig. 5A). However in clear contrast, reptile derived *P. aeruginosa* (isolate 1) potently activated acTLR5 but failed to activate hTLR5 (Fig. 5A). Additional analysis using three extra reptile (isolates 2-4) and also four human *P. aeruginosa* isolates (isolates 1-4) indicated stronger activation of acTLR5 than hTLR5 by *P. aeruginosa* isolates, regardless of their reptile or human origin (Fig. 5B). The opposite response of these TLRs to the *Pseudomonas*

and *Salmonella* lysates indicates that differential recognition of the lysates was not due to variable receptor expression.

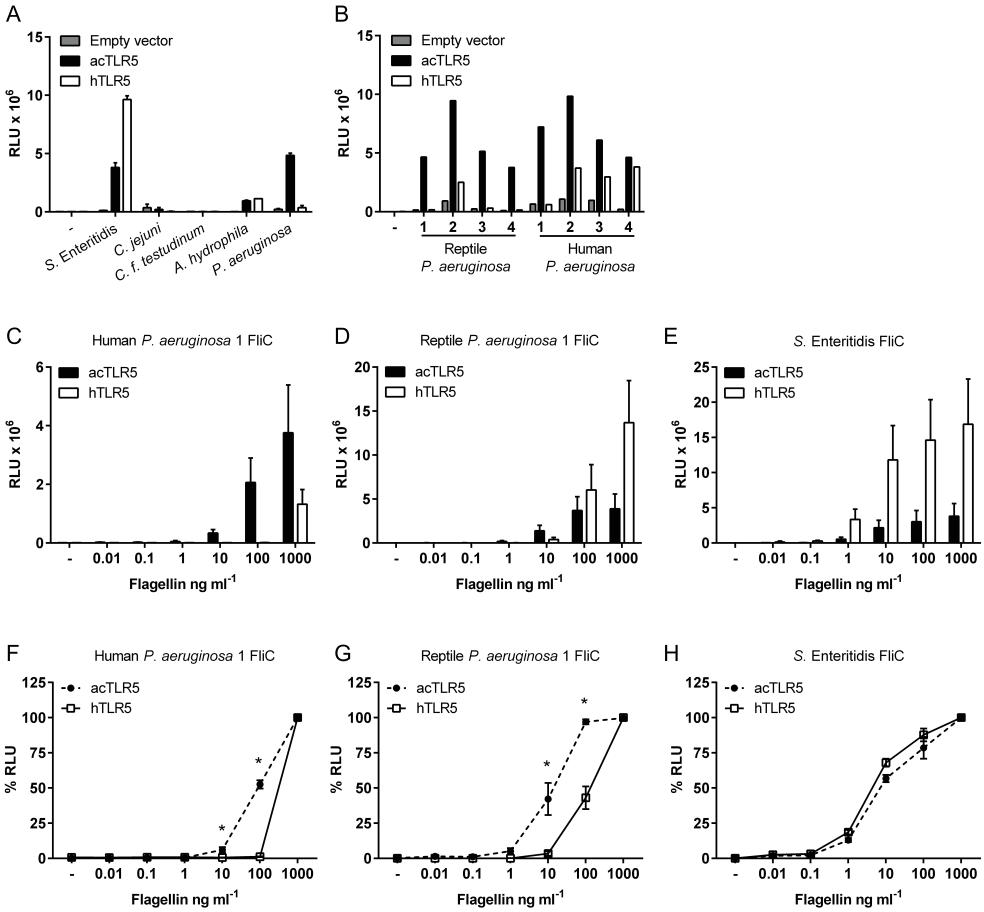


Figure 5. Differential recognition of *Pseudomonas* flagellins by acTLR5 and hTLR5.
(A) HeLa-57A cells transfected with acTLR5, hTLR5 or empty vector were stimulated (5 h) with lysate (1 µg/ml total protein) of the indicated bacterial species, or **(B)** four different reptile and four different human *P. aeruginosa* isolates. **(C, D and E)** acTLR5 or hTLR5 transfected cells were stimulated (5 h) with the indicated concentrations of purified his-tagged flagellin of human *P. aeruginosa* isolate 1 **(C)** or reptile *P. aeruginosa* isolate 1 **(D)** or *S.* Enteritidis **(E)**. **(F, G and H)** Relative sensitivity plots calculated from figures C, D and E showing the % RLU of acTLR5 and hTLR5 for each of the indicated concentrations of flagellin of human *P. aeruginosa* isolate 1 **(F)**, reptile *P. aeruginosa* isolate 1 **(G)** or *S.* Enteritidis **(H)**. The response to 1000 ng/ml flagellin was set at 100%. Values show the mean ± SEM luciferase activity as relative light units (RLU) of two independent experiments **(A)**, a representative of two independent experiments **(B)** or mean ± SEM of three independent experiments **(C to H)** all performed in duplicate. * $p < 0.05$ by unpaired Students *t*-test.

acTLR5 is more sensitive than hTLR5 to *Pseudomonas* flagellin

To verify that flagellin was the key determinant in the differential recognition of *P. aeruginosa* lysates by acTLR5 and hTLR5 and to exclude the destructive effect of flagellin degrading proteases potentially present in the lysates[33], we cloned and purified recombinant flagellin of reptile and human *P. aeruginosa* isolate 1. Stimulation of transfected HeLa-57A cells with low concentrations (0.1-10 ng/ml) of these purified *P. aeruginosa* flagellins revealed again stronger activation of acTLR5 compared to hTLR5 (Fig. 5C and 5D) and an opposite effect for *S.* Enteritidis flagellin (Fig. 5E). At high concentrations (100-1,000 ng/ml), reptile but not human *P. aeruginosa* flagellin did yield a potent hTLR5 response.

The differential dose-dependent responses by acTLR5 and hTLR5 suggested that the receptors recognize the purified flagellins with a different sensitivity. To substantiate the apparent different sensitivity of acTLR5 and hTLR5 to the purified flagellins we set the response to 1,000 ng/ml flagellin at 100%. This revealed that the receptors had a similar relative sensitivity to *S.* Enteritidis flagellin (Fig. 5H). However, compared to hTLR5, acTLR5 showed a higher relative sensitivity to the human *P. aeruginosa* flagellin ($p < 0.05$) (Fig. 5F). Higher relative sensitivity of acTLR5 was also noted for the reptile *P. aeruginosa* flagellin ($p < 0.05$) (Fig. 5G), despite the fact that high doses of this flagellin induced stronger activation of hTLR5. Overall, these results show that acTLR5 is more sensitive than hTLR5 to *P. aeruginosa* but not *S.* Enteritidis flagellin.

Discussion

Reptiles form a large group of vertebrates with a central position in vertebrate evolution and a unique physiology, being the only ectothermic amniotes. Despite this, relatively few studies have investigated the reptile immune system and detailed molecular characterizations of reptile immune molecules are scarce. Here we report a detailed functional characterization of the first TLR in reptiles. Our characterization of TLR5 of the lizard *A. carolinensis* fills the evolutionary gap of functional TLRs across vertebrates and provides a novel view on the reptile immune system at a molecular level. Evidence is provided that acTLR5 is expressed and functional in reptile as well as human cells and responds to bacterial flagellin. Our results indicate that TLR5 structure, function and signaling are highly conserved throughout evolution, although differences in relative sensitivity of reptile and human TLR5 to *Pseudomonas* but not *Salmonella* flagellin point to bacterial species dependent adaptations in flagellin recognition by reptile and human TLR5.

The reptile *tlr5* gene was cloned from an *Anolis carolinensis* lizard. Support for its identification as *tlr5* ortholog included a strong phylogenetic relationship of the full-length protein with the well-characterized chicken TLR5[15,31]. The ECD and TIR domain of the cloned acTLR5 were highly similar to a putative TLR of the Burmese python (snake), suggesting that the gene is present in other reptiles as well. Lizards, snakes and tuatara form the group of Lepidosauria that diverged approximately 270 Mya from their bird and crocodile sister group; the Archosauria[34]. Lizards and snakes thereafter diverged approximately 180 Mya[35,36]. The phylogeny of these species is reflected by the high similarity of acTLR5 with the putative snake and chicken TLR5, suggesting that TLR5 underwent a constrained evolution according to species divergence.

Functional evidence for identifying the cloned Anolis gene as a TLR5 ortholog was provided by the responsiveness of acTLR5 transfected cells to bacterial flagellin, thus far the only known TLR5 ligand. Activation of NF-κB in acTLR5 expressing cells was observed upon stimulation with wild type but not flagellin-deficient *Salmonella* as well as with purified recombinant *Salmonella* and *Pseudomonas* flagellins, thereby excluding non-specific activation of NF-κB. The results indicate that acTLR5 senses flagellins of different bacterial species and is capable of initiating a signaling cascade required to evoke an immune response. In mammals, flagellin recognition by TLR5 is indispensable for an adequate immune response to infection with flagellated bacteria[37–40]. As *A. carolinensis* tissues express the ac*tlr5* gene *in vivo* (Fig. 2) acTLR5 may have a similar function in reptiles.

A striking finding that underpins the evolutionary conservation of the TLR system is the functional expression of reptile TLR5 in a human cell background. The first step in TLR5 mediated NF-κB activation is the recruitment of the intracellular MyD88 adapter protein to the TLR5 TIR domain[41]. Comparison of the TIR domains of reptile and human TLR5 revealed a high overall sequence similarity (85%) and conservation of specific amino acid residues that are critical for TLR5 signaling[15,26,27]. In addition, both the TIR domain and MyD88 have been shown to evolve under strong functional constraint[42–44]. Together, this may explain the successful activation of NF-κB by acTLR5 in human cells. The compatibility of reptile TLR5 with human intracellular proteins suggests that the TLR5 signaling system was already functional in the common ancestor of reptiles and mammals and provides support for the functionally constrained evolution of TLR5 signaling at least throughout the divergent evolution of reptiles and mammals. Here it may be noteworthy that efforts to functionally express intact TLR5 from fish or amphibians in human cells have thus far not been reported.

Bioinformatics analysis indicated that the ECD of acTLR5 contained a N- and C-terminal LRR separated by 22 consecutive LRRs which is a typical feature of chicken[15] and other vertebrate TLR5 orthologs[25]. In line with the apparent conserved structure of the ECD, both reptile and human TLR5 recognized and responded to the D1 domain of *Salmonella* but not *Campylobacter* flagellin. This finding demonstrates that throughout 300 million years of divergent evolution, reptile and human TLR5 have conserved the ability to recognize flagellin at its D1 domain and hence the flagellin D1 domain of certain bacterial species has remained a critical activator of TLR5.

Interestingly, despite different amino acid compositions of the ECD, reptile and human TLR5 showed equal sensitivity to flagellin of *Salmonella enterica* serovar Enteritidis. Pet reptiles are frequently reported as carriers of zoonotic *Salmonella* serovars that can cause salmonellosis in humans but are generally considered non-pathogenic in healthy reptiles[45–49]. The principles underlying resistance or tolerance of reptiles to *Salmonella* are unknown but may relate to the poikilothermic nature of reptiles since *Salmonella* virulence is influenced by environmental temperature[50,51]. Yet, the fact that reptile and human TLR5 show a similar relative sensitivity to *S.* Enteritidis flagellin may suggest that flagellin recognition does not play a significant role in the differential susceptibility to *Salmonella* infection observed between reptiles and humans.

In contrast to *Salmonella* flagellin, reptile and human TLR5 showed a differential sensitivity to flagellin of *P. aeruginosa* clinical isolates. *P. aeruginosa* is a common bacterium that resides in diverse environments including water and soil and is an opportunistic pathogen of both reptiles and humans[52,53]. Why reptile TLR5 is more sensitive to *P. aeruginosa* flagellin than human TLR5 remains to be elucidated but it may suggest that throughout host-microbe coevolution, *P. aeruginosa* has exerted a stronger selective pressure on the evolution of acTLR5 than on hTLR5. Indeed, *in silico* studies indicate that among primates[54] and galloanserae birds[13] TLR5 undergoes diversifying, adaptive evolution through positive selection, a process most likely driven by host specific coevolution with flagellated bacteria. A similar process in reptiles may explain the observed differences in *P. aeruginosa* flagellin recognition between the Anolis and human TLR5.

Methods and Materials

Isolation of *Anolis carolinensis* DNA and RNA

Anolis tissue samples were obtained from a healthy male *Anolis carolinensis* lizard that had been euthanized by intra-coelomic injection of pentobarbital (200 mg kg^{-1} BW, Euthanimal®, Alfasan International, The Netherlands). Organs were directly frozen in liquid nitrogen. Genomic DNA was isolated using the high pure template kit (Roche) according to the manufacturer's instructions. RNA was extracted from tissue lysed with RLT buffer (1% β-mercaptoethanol) (Qiagen) in 1.4 mm Fastprep lysing matrix tubes (MPbio) in a Magna Lyser centrifuge (6,500 x *g*, 40 s, RT) (Roche). Total RNA was isolated using the RNeasy mini kit (Qiagen) following the manufacturer's instructions, treated with DNase I (1 U mg^{-1} RNA, Thermo Scientific) and stored at -80ºC until use.

Ethics statement

Euthanasia of the Anolis lizard was performed by a veterinarian specialized in reptiles (M.K, Diplomat European College Zoological Medicine, herpetology) and was in accordance with the guidelines in the Directive 2010/63/EU of the European Parliament and of the September 2010 Council on the protection of animals used for scientific purposes (http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32010L0063). The procedure was approved by the Animal Ethics Committee of Utrecht University (study number 2014.II.04.031).

Cloning of *A. carolinensis tlr5*

The *A. carolinensis tlr5* gene (ac*tlr5*) was amplified from genomic DNA (500 ng) by PCR in 50 µl volume containing 1X Phusion polymerase buffer, dNTP's (0.2 mM each), $MgCl_2$ (50 mM), Phusion hot start II high fidelity polymerase (1 Unit, Thermo Scientific) and 20 µM of forward (5'-CCGGATCCATGAAAAAGATGCTTCATTATCTCTTC-3') and reverse (5'-CCGCGGCC**GC**AAGAGATTGTGACTACTTT-3') primer (Life Technologies). Underlined sequences in the forward and reverse primer indicate BamHI and NotI restriction sites, respectively. The bold GC in the reverse primer substituted an AG in the *tlr5* gene, thereby replacing the terminal stopcodon for a cysteine. PCR conditions were: one cycle for 1 min at 98°C followed by 35 cycles of 30 s at 98°C, 30 s at 54°C, 90 s at 72°C and one final extension step of 10 min at 72°C. The PCR product was purified from gel using the GeneJet gel extraction kit (Thermo Scientific) and ligated into a pTracer-CMV2ΔGFP/3xFLAG[8] using the BamHI and NotI restriction sites, yielding pTracer 3xFLAG-ac*tlr5* carrying ac*tlr5*

with a C-terminal 3x FLAG tag. The plasmid was propagated in DH5-α. The cloned ac*tlr5* gene sequence was verified by DNA sequencing (Macrogen). The sequence was deposited in Genbank (accession number: KT347095).

Reverse transcriptase PCR on ac*tlr5* mRNA from various tissues

First strand cDNA was created of 1 µg total RNA with oligo $(dT)_{18}$ primers using the RevertAid First Strand cDNA Synthesis kit (Thermo Scientific). Non-reverse transcribed RNA served as control. PCR on cDNA and control samples was performed in 50 µl volume containing 1X DreamTaq polymerase buffer, dNTP's (0.2 mM each), DreamTaq Green DNA polymerase (1.25 Units, Thermo Scientific), template cDNA (10 ng) and 20 µM of ac*tlr5* forward (5'-GCATGAATTCCTTGGGCACTCTG-3') and reverse (5'-GGGCCACATCCCAACCATTAC-3') primer or *A. carolinensis* GAPDH (ac*gapdh*) forward (5'-GAGAGGAGCTTCTCAGAACATC-3') and reverse (5'-GACAATGCGGTTGCTGTATC-3') primer. PCR conditions were: one cycle of 3 min at 95°C followed by 35 cycles of 30 s at 95°C, 30 s at 55°C, 1 min at 72°C and followed by one final extension step of 10 min at 72°C. PCR products were analysed using 2% agarose gels.

acTLR5 bioinformatics analysis

Amino acid sequence comparison was performed using NCBI-BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple sequence alignment was conducted using Clustal W[55] (http://www.ebi.ac.uk/Tools/msa/clustalw2/) with default settings. MEGA6 software[56] was used to construct the phylogenetic tree. Leucine rich repeats in the predicted acTLR5 protein sequence were identified by manual sequence analysis as described by Matsushima et al.[25] and by use of the Leucine Rich Repeat Finder database (http://www.lrrfinder.com/). Transmembrane domains were predicted with the TMpred server (http://embnet.vital-it.ch/software/TMPREDform.html). A TIR domain was predicted by identifying five alternating β-sheets and α-helices[57] using the Proteus Protein Structure Prediction server (http://wks80920.ccis.ualberta.ca/proteus/).

Bacterial strains

The following bacterial strains were grown (37°C, 18 h) on Luria-Bertani (LB) agar plates or in 5 ml of LB broth (Biotrading) at 160 rpm: *Escherichia coli* DH5-α, *E. coli* BL21 star (DE3), *Salmonella enterica* serovar Enteritidis (referred to as *S.* Enteritidis) strain 90-13-706 (CVI, Lelystad, The Netherlands), *S.* Enteritidis 90-13-706 isogenic *fliC* mutant[58], *Aeromonas hydrophila* (turtle isolate, Utrecht

University), four *Pseudomonas aeruginosa* from reptiles (one lizard, three snake isolates, Utrecht University), and four human *P. aeruginosa* isolates (University Medical Center Utrecht). *Campylobacter jejuni* strain 81116 (NCTC: 11828), and *Campylobacter fetus* subsp. *testudinum* (reptile isolate[32]) were grown (37°C, 18 h) under micro-aerobic conditions (80% N_2, 7.5% H_2, 7.5% CO_2, 5% O_2) on saponin agar plates or in LB broth at 160 rpm.

Preparation of bacterial cell lysates

Single colonies of the bacterial species described above were grown (37°C, 20 h) in 5 ml LB broth at 160 rpm and placed on ice. After microscopic confirmation of motility all cultures were normalized to an OD_{550} of 2, pelleted by centrifugation (5,000 x *g*, 30 min, 4°C), washed with 1 ml Dulbecco's phosphate buffered saline (DPBS, Sigma), briefly vortexed, and collected by centrifugation (5,000 x *g*, 30 min, 4°C). Pellets were re-dissolved in 2 ml DPBS and placed for 1 h at 70°C. Heat killed bacteria were sonicated (6 x 15 s, Vibra-cell, Sonics, USA) and centrifuged (14,000 x *g*, 40 min, 4°C). Lysate supernatants were stored at -20°C until use. Protein concentration of lysates was determined by BCA assay (Thermo Scientific).

Construction, expression and purification of recombinant His-tagged flagellins

Construction of recombinant His-tagged flagellin of *S.* Enteritidis (FliC), *C. jejuni* (FlaA) and chimeric NHC has been described previously[15,31]. The flagellin gene of both reptile and human *P. aeruginosa* isolate 1 was amplified from genomic DNA by PCR in 50 µl volume containing 1X Dreamtaq polymerase buffer, dNTP's (0.2 mM each), Dreamtaq polymerase (1 Unit) and 20 µM of forward (5'-AAACCATGGCCTTGACCGTCAACAC-3') and reverse (5'-AAAGAGCTCGCGCAGCAGGCTCAGAAC-3') primer. Underlined sequences in the forward and reverse primer indicate NcoI and SacI restriction sites, respectively. PCR conditions were: one cycle for 3 min at 95°C followed by 35 cycles of 30 s at 95°C, 30 s at 64°C, 2 min at 72°C and a final extension step for 10 min at 72°C. PCR products were ligated into the pET101/D-TOPO (Promega) expression vector using NcoI and SacI restriction enzymes. Ligation into the pET101/D-TOPO vector added a C-terminal His-tag to the flagellin gene and the plasmids were transformed into *E. coli* BL21 star (DE3).

Protein expression was induced by growing log phase cultures in the presence of 1 mM IPTG (Thermo Scientific) for 4 h at 37°C. For flagellin purification bacteria were pelleted (4,400 x *g*, 15 min, 4°C), resuspended in 10 ml cold DPBS with protease inhibitor cocktail (Roche), spun down (4,400 x *g*, 15 min, 4°C) and

incubated (RT) under end-over-end rotation for 16 h in 8 M urea buffer (8 M urea, 100 mM NaH_2PO_4, 100 mM Tris-HCl, pH 8). After removal of cell debris (5,300 x *g*, 30 min, RT) supernatant was incubated with Ni^{2+}-NTA agarose beads (Qiagen). After 2 h the beads were washed with 4 x 4 ml of 8 M urea buffer pH 6.3. Flagellins were eluted with 4 x 0.5 ml of 8 M urea buffer pH 5.9 followed by 4 x 0.5 ml of 8 M urea buffer pH 4.5. Collected fractions were checked for purity on SDS-PAGE and pure fractions were pooled and concentrated using Amicon YM-30 filters (Millipore). Protein concentration was measured by BCA assay. Concentrated flagellins were diluted to the desired concentration and stored (-20°C) as aliquots in 4 M urea, 100 mM NaH_2PO_4, 10 mM Tris-HCl, pH 9.

Cell culture

HeLa-57A cells that are stably transfected with a NF-κB luciferase reporter construct[59] were routinely propagated in Dulbecco's modified eagle medium (DMEM) plus 5% fetal calf serum (FCS, Bodinco) at 37°C and 10% CO_2. Green iguana (*Iguana iguana*) heart cells (IgH-2, ATCC: CCL-108) were grown in minimal essential medium with Hank's salts (MEM) and 10% FCS at 30°C in air. Cells were passaged twice a week.

Transient transfection of cells

HeLa-57A cells and IgH-2 cells were grown to ± 80% confluence in a 6-well plate and transfected using Fugene HD transfection reagent (Promega) according to the manufacturer's instructions. HeLa-57A cells and IgH-2 cells were transfected with 1 µg of pTracer 3xFLAG-ac*tlr5* or human (h)*tlr5*[15] at a lipid to DNA ratio of 3:1 (HeLa-57A) or 4:1 (IgH-2). Empty pTracer 3xFLAG was used for mock transfections. IgH-2 were additionally transfected with 1 µg of NF-κB luciferase reporter plasmid.

Luciferase assay

Twenty-four hours after transfection cells were re-distributed in a 48-well plate. After 24 h cells were washed twice with medium without FCS and stimulated with the indicated TLR ligands or live bacteria in 500 µl medium without FCS (for stimulation with LPS, medium did contain FCS). After 5 h at 37°C (HeLa-57A) or 10 h at 30°C (IgH-2), cells were washed with DPBS and lysed with reporter lysis buffer (100 µl, Promega) at -80°C for at least 1 h. After thawing, cell lysate (20 µl) was mixed with luciferase reagent (50 µl, Promega) and luciferase activity was measured in a luminometer (TD20/20, Turner designs). Experiments with bacterial lysates and purified *P. aeruginosa* flagellins were performed in 96-well plates in

250 µl volumes. Cells were lysed in 50 µl reporter lysis buffer. Luciferase activity in these experiments was measured with a TriStar^2 luminometer (Berthold) by mixing 15 µl cell lysate with 37 µl luciferase reagent. Values obtained from the TriStar^2 were 1000 times higher compared to the values obtained from the TD20/20 but relative sensitivity and accuracy between the two luminometers was equal. Results were expressed in relative light units (RLU) or % RLU in experiments with purified *P. aeruginosa* flagellins. Percent RLU was calculated by dividing the RLU obtained from each concentration of flagellin over the value obtained from stimulation with 1 µg /ml flagellin which was set at 100%.

Statistics

Statistical analysis were performed using Graphpad 6 (Prism) software. Differences between two groups were tested with unpaired Student *t*-tests. A probability (*p*) value of < 0.05 was considered significant.

Acknowledgements

We thank Dr. R.T. Hay (Institute of Biomolecular Sciences, School of Biomedical Sciences, University of St. Andrews, Scotland, UK) for providing the HeLa-57A cell line.

Author contributions

C.V, J.W and J.vP designed research, C.V performed experiments and analysed all data, L.B contributed to cloning of acTLR5, M.K contributed in obtaining *A. carolinensis* tissues, C.V wrote the paper with support from all authors.

References

1. Medzhitov, R. Toll-like receptors and innate immunity. *Nature Reviews Immunology* **1,** 135–45 (2001).
2. Leulier, F. & Lemaitre, B. Toll-like receptors--taking an evolutionary approach. *Nature Reviews Genetics* **9**, 165–78 (2008).
3. Gay, N. J. & Gangloff, M. Structure and function of Toll receptors and their ligands. *Annual Review of Biochemistry* **76,** 141–165 (2007).
4. Takeuchi, O. & Akira, S. Pattern recognition receptors and inflammation. *Cell* **140,** 805–20 (2010).
5. Roach, J. C. *et al.* The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 9577–9582 (2005).

6. Poltorak, A. *et al.* Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* **282,** 2085–2088 (1998).
7. Hemmi, H. *et al.* A Toll-like receptor recognizes bacterial DNA. *Nature* **408,** 740–745 (2000).
8. Keestra, A. M., Zoete, M. R. de, Bouwman, L. I. & van Putten, J. P. M. Chicken TLR21 is an innate CpG DNA receptor distinct from mammalian TLR9. *Journal of Immunology* **185**, 460–467 (2010).
9. Zoete, M. R. de, Bouwman, L. I., Keestra, A. M. & Putten, J. P. M. van. Cleavage and activation of a Toll-like receptor by microbial proteases. *Proceedings of the National Academy of Sciences of the United States of America* **108,** 4968–4973 (2011).
10. Hayashi, F. *et al.* The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* **410,** 1099–1103 (2001).
11. Werling, D., Jann, O. C., Offord, V., Glass, E. J. & Coffey, T. J. Variation matters: TLR structure and species-specific pathogen recognition. *Trends in Immunology* **30**, 124–130 (2009).
12. Wlasiuk, G. & Nachman, M. W. Adaptation and constraint at Toll-like receptors in primates. *Molecular Biology and Evolution* **27,** 2172–2186 (2010).
13. Vinkler, M., Bainová, H. & Bryja, J. Protein evolution of Toll-like receptors 4, 5 and 7 within Galloanserae birds. *Genetics Selection Evolution* **46,** 72 (2014).
14. Andersen-Nissen, E., Smith, K. D., Bonneau, R., Strong, R. K. & Aderem, A. A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. *Journal of Experimental Medicine* **204,** 393–403 (2007).
15. Keestra, A. M., de Zoete, M. R., van Aubel, R. A. M. H. & van Putten, J. P. M. Functional characterization of chicken TLR5 reveals species-specific recognition of flagellin. *Molecular Immunology* **45,** 1298–1307 (2008).
16. Keestra, A. M. & van Putten, J. P. M. Unique Properties of the Chicken TLR4/MD-2 Complex: Selective lipopolysaccharide activation of the MyD88-dependent pathway. *Journal of Immunology* **181,** 4354–4362 (2008).
17. Kasamatsu, J., Oshiumi, H., Matsumoto, M., Kasahara, M. & Seya, T. Phylogenetic and expression analysis of lamprey toll-like receptors. *Developmental and Comparative Immunology* **34,** 855–865 (2010).
18. Alcaide, M. & Edwards, S. V. Molecular evolution of the Toll-like receptor multigene family in birds. *Molecular Biology and Evolution* **28,** 1703–1715 (2011).
19. Rauta, P. R., Samanta, M., Dash, H. R., Nayak, B. & Das, S. Toll-like receptors (TLRs) in aquatic animals: signaling pathways, expressions and immune responses. *Immunology Letters* **158,** 14–24 (2014).
20. Babik, W. *et al.* Constraint and adaptation in newt Toll-like receptor genes. *Genome Biology and Evolution* **7,** 81–95 (2015).
21. Zimmerman, L. M., Vogel, L. A. & Bowden, R. M. Understanding the vertebrate immune system: insights from the reptilian perspective. *Journal of Experimental Biology* **213,** 661–671 (2010).
22. Benton, M. J. & Donoghue, P. C. J. Paleontological evidence to date the tree of life. *Molecular Biology and Evolution* **24,** 26–53 (2007).
23. Merchant, M., Fleury, L., Rutherford, R. & Paulissen, M. Effects of bacterial lipopolysaccharide on thermoregulation in green anole lizards (*Anolis carolinensis*). *Veterinary Immunology and Immunopathology* **125,** 176–181 (2008).
24. Alföldi, J. *et al.* The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature* **477,** 587–591 (2011).
25. Matsushima, N. *et al.* Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors. *BMC Genomics* **8,** 124 (2007).
26. Ivison, S. M. *et al.* A phosphorylation site in the Toll-like receptor 5 TIR domain is required for inflammatory signaling in response to flagellin. *Biochemical and Biophysical Research Communications* **352,** 936–941 (2007).
27. Ivison, S. M. *et al.* Protein kinase D interaction with TLR5 is required for inflammatory signaling in response to bacterial flagellin. *Journal of Immunology.* **178,** 5735–5743 (2007).

28. van Aubel, R. A. M. H., Keestra, A. M., Krooshoop, D. J. E. B., van Eden, W. & van Putten, J. P. M. Ligand-induced differential cross-regulation of Toll-like receptors 2, 4 and 5 in intestinal epithelial cells. *Molecular Immunology* **44,** 3702–3714 (2007).
29. Yoon, S. I. *et al.* Structural basis of TLR5-flagellin recognition and signaling. *Science* **335,** 859–64 (2012).
30. Andersen-Nissen, E. *et al.* Evasion of Toll-like receptor 5 by flagellated bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **102,** 9247–9252 (2005).
31. Zoete, M. R. de, Keestra, A. M., Wagenaar, J. A. & van Putten, J. P. M. Reconstitution of a Functional Toll-like Receptor 5 Binding Site in *Campylobacter jejuni* Flagellin. *Journal of Biological Chemistry* **285,** 12149–12158 (2010).
32. Fitzgerald, C. *et al. Campylobacter fetus* subsp. *testudinum* subsp. nov., isolated from humans and reptiles. *International Journal of Systematic and Evolutionary Microbiology* **64,** 2944–2948 (2014).
33. Bardoel, B. W. *et al. Pseudomonas* evades immune recognition of flagellin in both mammals and plants. *PLoS Pathogens* **7,** e1002206 (2011).
34. Jones, M. E. *et al.* Integration of molecules and new fossils supports a Triassic origin for Lepidosauria (lizards, snakes, and tuatara). *BMC Evolutionary Biology* **13,** 208 (2013).
35. Vidal, N. & Hedges, S. B. The phylogeny of squamate reptiles (lizards, snakes, and amphisbaenians) inferred from nine nuclear protein-coding genes. *Comptes Rendus Biologies* **328,** 1000–1008 (2005).
36. Castoe, T. A. *et al.* Dynamic nucleotide mutation gradients and control region usage in Squamate reptile mitochondrial genomes. *Cytogenetic and Genome Research* **127,** 112–127 (2010).
37. Hawn, T. R. *et al.* A common dominant TLR5 stop codon polymorphism abolishes flagellin signaling and is associated with susceptibility to legionnaires' disease. *Journal of Experimental Medicine* **198,** 1563–72 (2003).
38. Andersen-Nissen, E. *et al.* Cutting edge: Tlr5-/- mice are more susceptible to *Escherichia coli* urinary tract infection. *Journal of Immunology* **178,** 4717–20 (2007).
39. Vijay-Kumar, M. *et al.* Deletion of TLR5 results in spontaneous colitis in mice. *Journal of Clinical Investigations* **117,** 3909–3921 (2007).
40. Carvalho, F. A. *et al.* Transient inability to manage proteobacteria promotes chronic gut inflammation in TLR5-deficient mice. *Cell Host & Microbe* **12,** 139–152 (2012).
41. O'Neill, L. A. J. & Bowie, A. G. The family of five: TIR-domain-containing adaptors in Toll-like receptor signaling. *Nature Reviews Immunology* **7,** 353–364 (2007).
42. Nakajima, T. *et al.* Natural selection in the TLR-related genes in the course of primate evolution. *Immunogenetics* **60,** 727–735 (2008).
43. Fornarino, S. *et al.* Evolution of the TIR domain-containing adaptors in humans: swinging between constraint and adaptation. *Molecular Biology and Evolution* **28,** 3087–3097 (2011).
44. Mikami, T., Miyashita, H., Takatsuka, S., Kuroki, Y. & Matsushima, N. Molecular evolution of vertebrate Toll-like receptors: evolutionary rate difference between their leucine-rich repeats and their TIR domains. *Gene* **503,** 235–243 (2012).
45. Murphy, D. & Oshin, F. Reptile-associated salmonellosis in children aged under 5 years in South West England. *Archives of Disease in Childhood* **100**, 364-365 (2015).
46. Lukac, M., Pedersen, K. & Prukner-Radovcic, E. Prevalence of *Salmonella* in captive reptiles from Croatia. *Journal of Zoo and Wildlife Medicine* **46,** 234–240 (2015).
47. Warwick, C., Lambiris, A. J., Westwood, D. & Steedman, C. Reptile-related salmonellosis. *Journal of the Royal Society of Medicine* **94,** 124–126 (2001).
48. Whitten, T., Bender, J. B., Smith, K., Leano, F. & Scheftel, J. Reptile-associated salmonellosis in Minnesota, 1996–2011. *Zoonoses and Public Health* **62,** 199–208 (2015).
49. Pasmans, F. *et al.* Characterization of *Salmonella* isolates from captive lizards. *Veterinary Microbiology* **110,** 285–291 (2005).

50. Elhadad, D., McClelland, M., Rahav, G. & Gal-Mor, O. Feverlike Temperature is a virulence regulatory cue controlling the motility and host cell entry of typhoidal *Salmonella*. *Journal of Infectious Diseases* **212,** 147–156 (2015).
51. Yang, Y., Khoo, W. J., Zheng, Q., Chung, H.-J. & Yuk, H.-G. Growth temperature alters *Salmonella* Enteritidis heat/acid resistance, membrane lipid composition and stress/virulence related gene expression. *International Journal of Food Microbiology* **172,** 102–109 (2014).
52. Seixas, R. *et al.* Severe fibrinonecrotic enteritis caused by *Pseudomonas aeruginosa* in a captive monitor lizard (*Varanus niloticus*). *Journal of Zoo and Wildlife Medicine* **45,** 410–412 (2014).
53. Folkesson, A. *et al.* Adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis airway: an evolutionary perspective. *Nature Reviews Microbiology* **10,** 841–851 (2012).
54. Wlasiuk, G., Khan, S., Switzer, W. M. & Nachman, M. W. A history of recurrent positive selection at the Toll-like receptor 5 in primates. *Molecular Biology and Evolution* **26,** 937–949 (2009).
55. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948 (2007).
56. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30,** 2725–2729 (2013).
57. Xu, Y. *et al.* Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature* **408,** 111–115 (2000).
58. Van Asten, F. J., Hendriks, H. G., Koninkx, J. F., Van der Zeijst, B. A. & Gaastra, W. Inactivation of the flagellin gene of *Salmonella enterica* serotype enteritidis strongly reduces invasion into differentiated Caco-2 cells. *FEMS Microbiology Letters* **185,** 175–179 (2000).
59. Rodriguez, M. S., Thompson, J., Hay, R. T. & Dargemont, C. Nuclear retention of IkappaBalpha protects it from signal-induced degradation and inhibits nuclear factor kappaB transcriptional activation. *Journal of Biological Chemistry* **274,** 9108–15 (1999).

Supplement

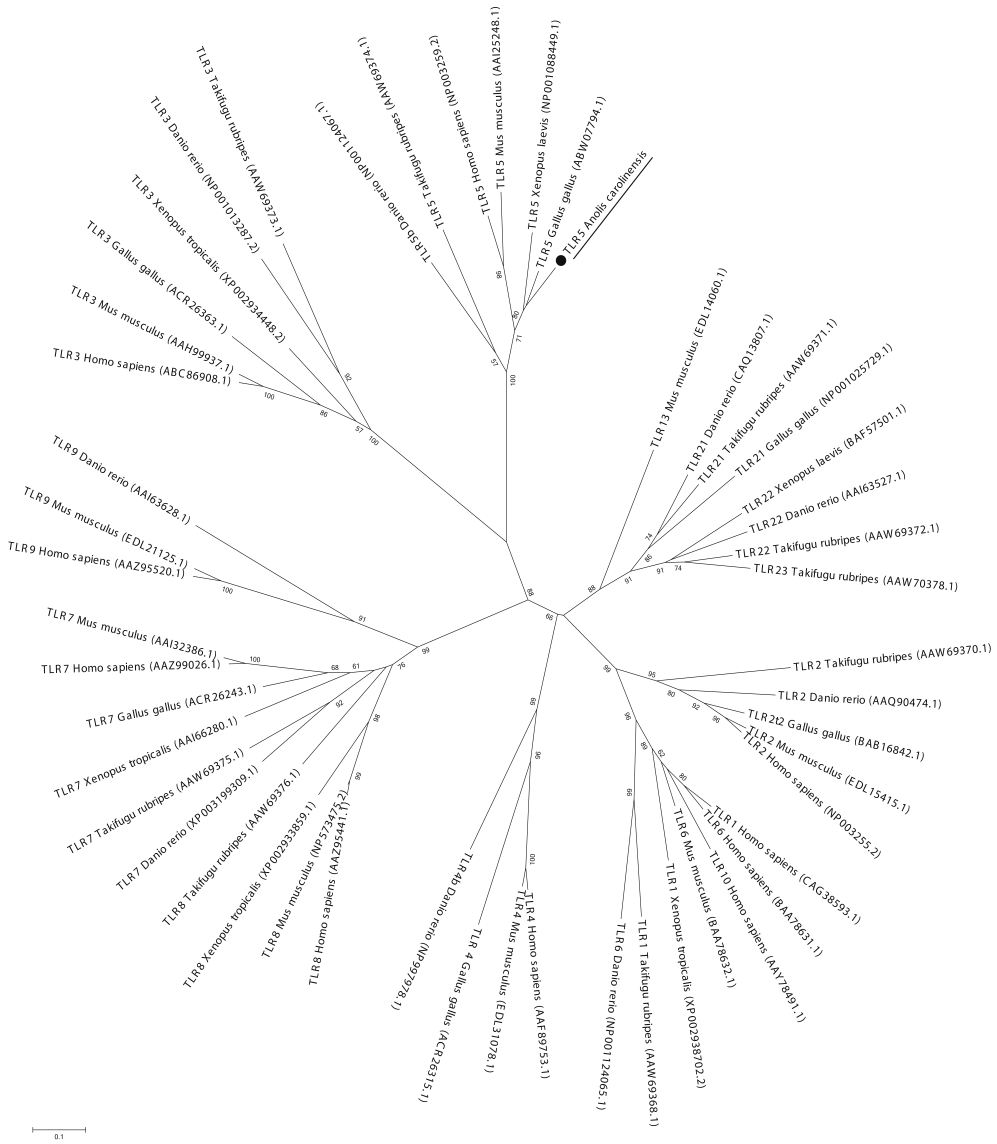
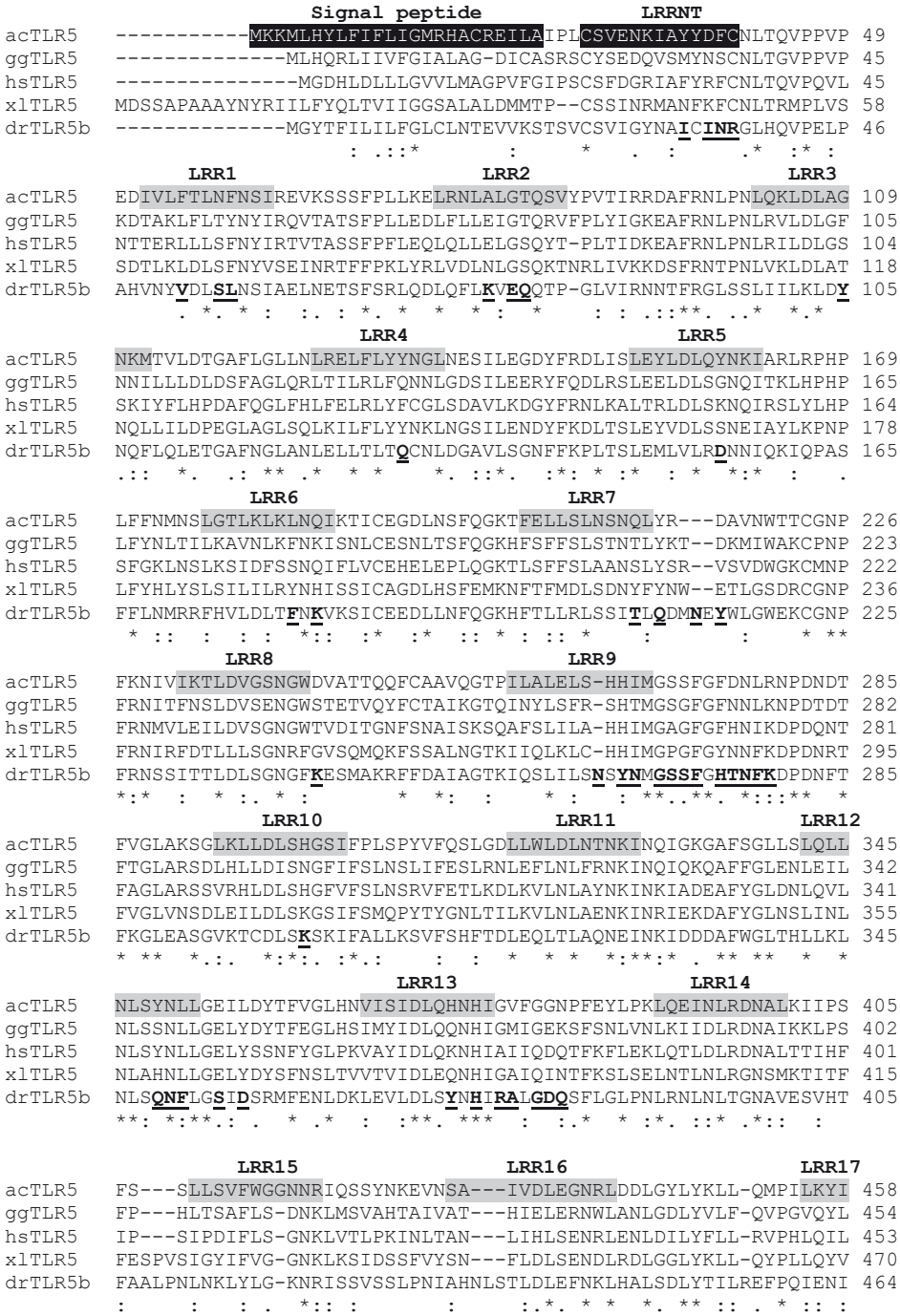


Figure S1. Phylogenetic position of *A. carolinensis* TLR5.
MEGA6 software[1] was used to construct a phylogenetic tree of different TLR families according to the Neighbour-Joining method with a bootstrap analysis of 10,000 iterations to indicate the relative support for each branch. acTLR5 is underlined and indicated with ●. GenBank accession numbers of the TLR protein sequences are given in brackets following the species name. Only bootstrap values of > 50 are shown. The tree is drawn to scale with the scale bar representing the evolutionary distance of 0.1 amino acid substitutions per site in the TLRs.

1. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30,** 2725–2729 (2013).


Signal peptide
LRRNT
LRR1
LRR2
LRR3
LRR4
LRR5
LRR6
LRR7
LRR8
LRR9
LRR10
LRR11
LRR12
LRR13
LRR14
LRR15
LRR16
LRR17
acTLR5
ggTLR5
hsTLR5
xlTLR5
drTLR5b

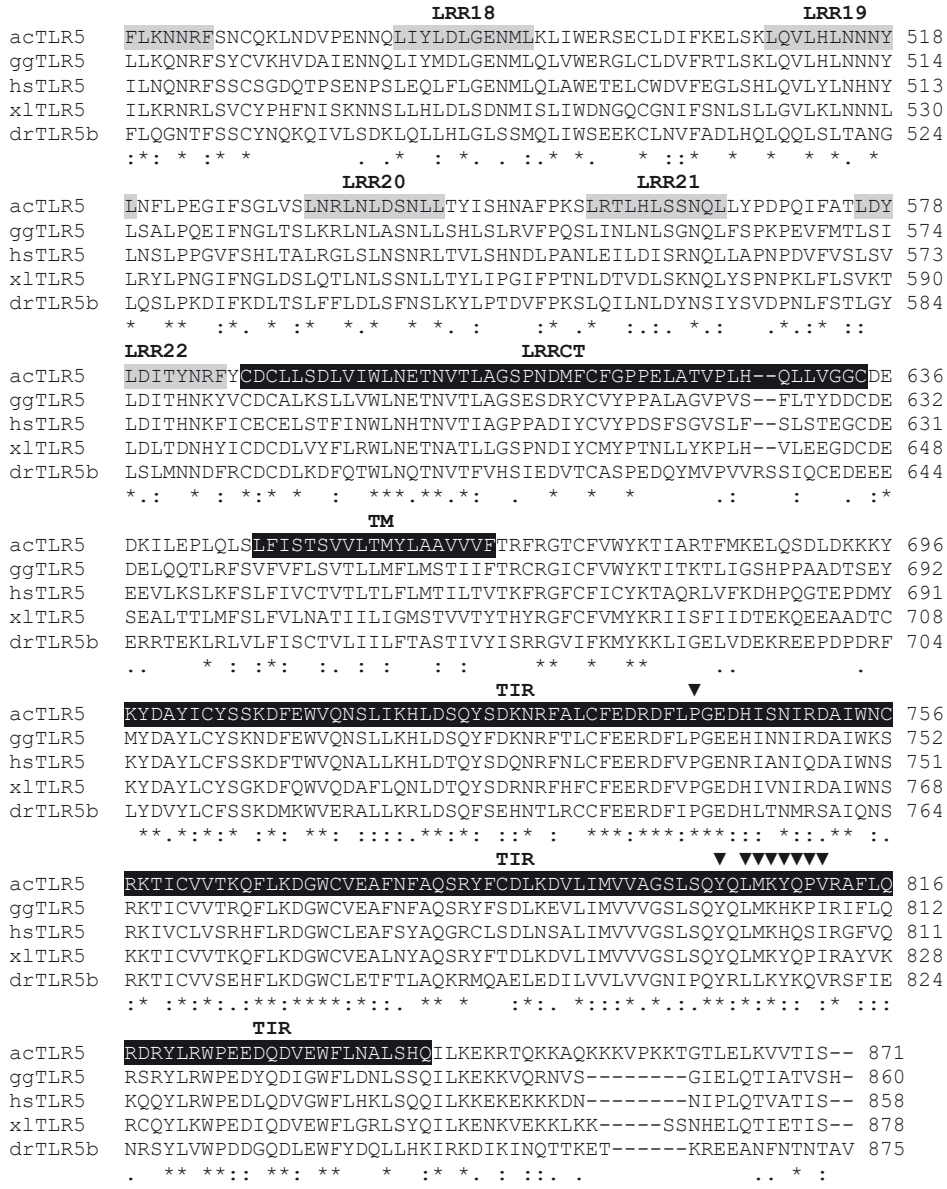
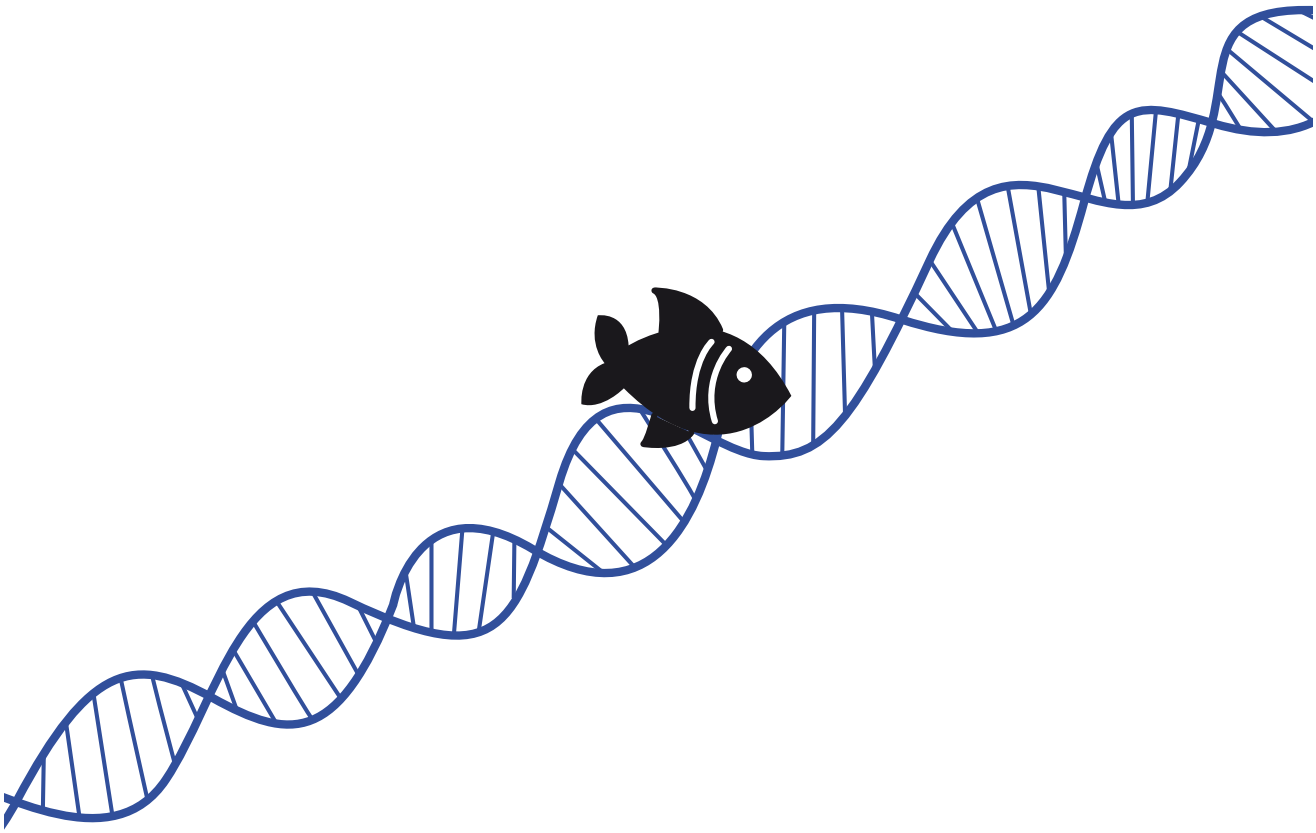


Figure S2. Multiple sequence alignment of vertebrate TLR5 proteins.

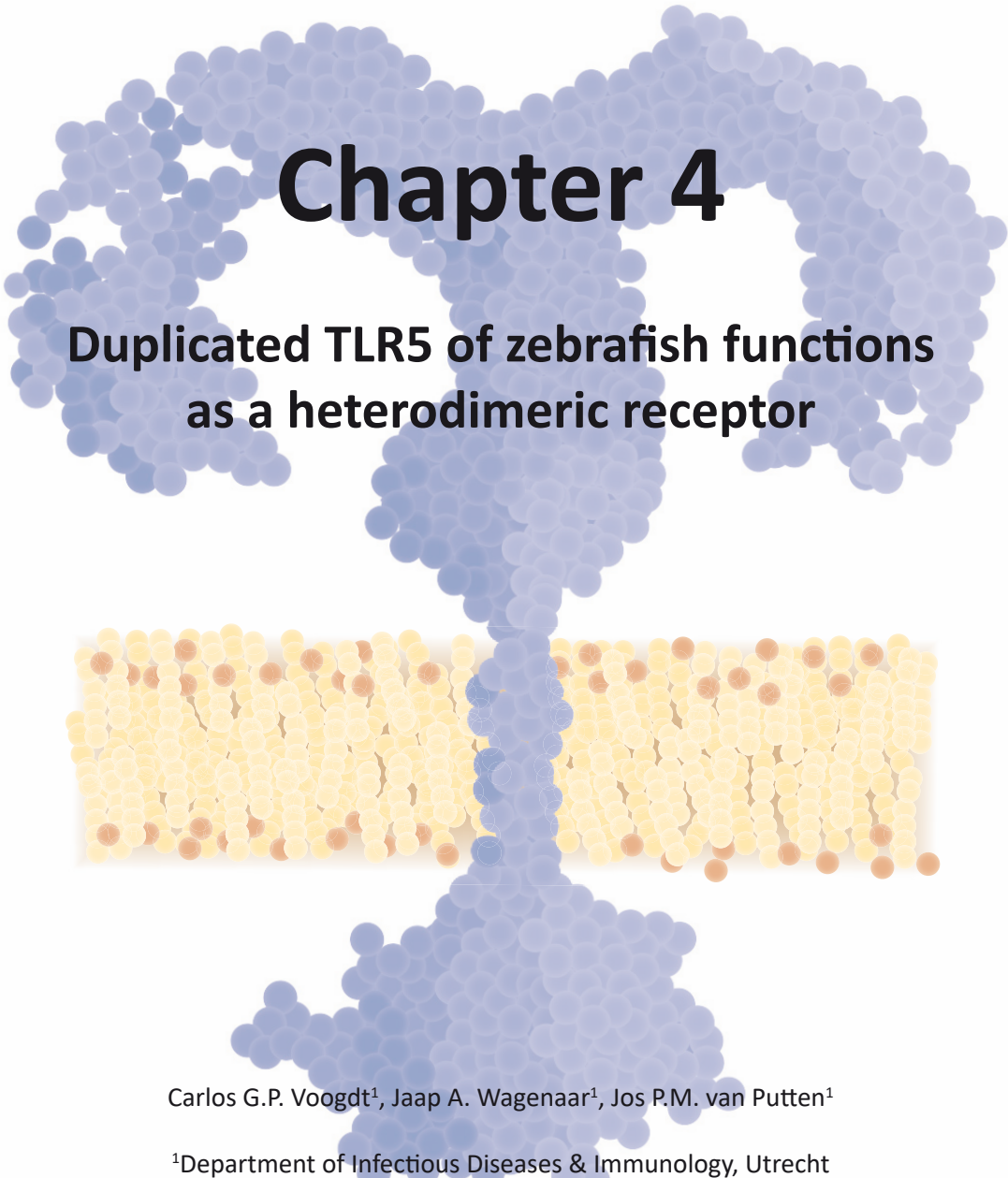
Alignment of TLR5 from *Anolis carolinensis* (ac), *Gallus gallus* (gg, chicken, NCBI reference sequence: ABW07794.1), *Homo sapiens* (hs, human, NP003259.2), *Xenopus laevis* (xl, African clawed frog, NP001088449.1) and *Danio rerio* (dr, zebrafish, NP001124067.1). The amino acid sequences were aligned using the ClustalW server with default settings. Asterisks (*) indicate identical residues in all sequences, double dots (:) indicate highly similar residues, single dots (.) indicate similar residues and bars (-) indicate gaps to complete the sequence alignment. acTLR5 leucine rich repeats (LRR 1 to 22) are shaded in gray while acTLR5 signal peptide, N-terminal LRR (LRRNT), C-terminal LRR (LRRCT), transmembrane domain (TM) and TIR domain are shaded in black. ▼ above residues in the TIR domain indicate conserved residues important for TLR5 signaling. Residues in bold and underlined in the drTLR5b sequence are the residues involved in flagellin binding by drTLR5b (adapted from Yoon et al.[1]).

1. Yoon, S. I. *et al.* Structural basis of TLR5-flagellin recognition and signaling. *Science* **335,** 859–64 (2012).



Chapter 4

Duplicated TLR5 of zebrafish functions as a heterodimeric receptor



Carlos G.P. Voogdt[1], Jaap A. Wagenaar[1], Jos P.M. van Putten[1]

[1]Department of Infectious Diseases & Immunology, Utrecht University, Utrecht, the Netherlands

Proceedings of the National Academy of Sciences of the United States of America (PNAS) Vol. 115, No. 14, E3221-E3229 2018

Abstract

Toll-like receptor 5 (TLR5) of mammals, birds and reptiles detects bacterial flagellin and signals as a homodimeric complex. Structural studies using truncated TLR5b of zebrafish confirm the homodimeric TLR5-flagellin interaction. Here we provide evidence that zebrafish (*Danio rerio*) TLR5 unexpectedly signals as a heterodimer composed of the duplicated gene products drTLR5b and drTLR5a. Flagellin induced signaling by the zebrafish TLR5 heterodimer increased in the presence of the TLR trafficking chaperone UNC93B1. Targeted exchange of drTLR5b and drTLR5a regions revealed that TLR5 activation needs a heterodimeric configuration of the receptor ectodomain and cytoplasmic domain, consistent with ligand-induced changes in receptor conformation. Structure-guided substitution of the presumed principle flagellin binding site in human TLR5 with corresponding zebrafish TLR5 residues abrogated human TLR5 activation, indicating species-specific TLR5-flagellin interaction. Our findings indicate that the duplicated TLR5 of zebrafish underwent sub-functionalization through concerted co-evolution to form a unique heterodimeric flagellin receptor that operates fundamentally different than TLR5 of other species.

Significance statement

Toll-like receptors (TLRs) are highly conserved innate receptors that form homo- or heterodimers to detect microbial danger signals and activate the immune system. TLR5 detects flagellin of bacteria and functions as a homodimeric receptor complex. A crystallized fragment of TLR5b of the zebrafish (*Danio rerio*) serves as a model structure for the homodimeric TLR5-flagellin interaction. Here we report that zebrafish TLR5 unexpectedly functions as a heterodimeric flagellin receptor composed of the duplicated gene products TLR5b and TLR5a. The unique heterodimeric nature of zebrafish TLR5 indicates important receptor differences between species, contributes to a deeper understanding of the activation mechanism of TLRs and provides an illustrative example of the functional co-evolution of duplicated genes.

Introduction

Toll-like receptors (TLRs) are evolutionarily highly conserved membrane bound receptors that initiate activation of the immune system upon detection of microbial ligands[1,2]. TLRs are composed of a cytoplasmic Toll-Interleukin-1 (TIR) signaling domain, a single transmembrane segment, and a leucine rich repeat (LRR) containing ectodomain[3]. Microbe driven diversification of the ectodomain has resulted in a family of distinct TLR receptors across species that recognize a wide variety of conserved microbial structures[4]. For example, TLR4 detects bacterial lipopolysaccharide (LPS)[5], TLR15 is activated by microbial proteases[6], while microbial nucleic acids are sensed by TLR3, 7-9 and 21[7-11]. Bacterial flagellin, the major constituent of the flagellum of motile bacteria including pathogens like *Salmonella* spp. and *Pseudomonas aeruginosa,* is recognized by TLR5[12,13]. Irrespective of the type of ligand, TLRs need to form receptor dimers to exert their function[14]. Whereas TLR1 and TLR6 function in a heterodimeric complex with TLR2, other TLRs including TLR3, TLR4, TLR7-9 and also TLR5 act as homodimers. In all cases, binding of ligand to the TLR complex induces a conformational change which positions the intracellular TIR domains of the dimer into close proximity, thereby forming a docking station for intracellular adapter molecules. Binding of adapter molecules initiates further signaling that activates transcription factors like NF-κB which translocate to the nucleus to start pro-inflammatory gene expression[14].

Structural analyses of crystallized TLR-ligand complexes have fundamentally contributed to detailed understanding of TLR-ligand interaction and dimerization. To date the ectodomains of many mammalian TLRs have been crystallized[15-18] with the exception of TLR5. For this receptor, only the protein structure of the N-terminal part of the ectodomain of TLR5b of zebrafish (*Danio rerio*) in complex with *Salmonella* flagellin has been resolved. The structure indicates that bacterial flagellin binds to homodimers of zebrafish TLR5b[19], consistent with the homodimeric nature of mammalian TLR5. However, direct evidence that binding of flagellin to zebrafish TLR5b initiates signaling is lacking.

Zebrafish is a widely used model animal and belongs to the teleostei or bony fish. Teleost fish have underwent an additional round of whole genome duplication. The duplication and reorganization of TLR genes has led to a higher diversity of TLR gene repertoires among teleosts compared to non-teleost vertebrates[20-24]. The common carp (*Cyprinus carpio*) has two paralogs of TLR2[20,25] and the Atlantic cod (*Gadus morhua*) has lost TLR5 but massively expanded TLR8 genes[22]. Zebrafish carries two paralogs of TLR4[26] and retained both CpG DNA receptors TLR9 (present

in mammals) and TLR21 (present in birds)[27]. In addition, zebrafish has duplicated TLR5 resulting in the two paralogs TLR5a and TLR5b. The functional consequences of these TLR duplications in teleosts are poorly understood.

In the present study we provide evidence that zebrafish TLR5b does not signal as a conventional TLR5 homodimer but instead co-operates with its paralog TLR5a to form a unique heterodimeric TLR5 that responds to flagellin. Structure-guided amino acid substitution and functional analyses of the N-terminal part of TLR5 indicate species-specific functional differences between zebrafish and human TLR5. Targeted exchange of multiple defined regions of zebrafish TLR5a and TLR5b indicates that a transregional heterodimeric configuration is required for receptor activation and signal initiation which is consistent with ligand-induced changes in receptor conformation.

Results

drTLR5a and drTLR5b are both required for functional flagellin recognition

Since the published crystal structure of the zebrafish TLR5b (drTLR5b) ectodomain indicates binding of flagellin to a receptor homodimer[19], we investigated whether stimulation of full length drTLR5b with bacterial flagellin activates the canonical TLR5 inducible NF-κB signaling pathway. The corresponding *tlr5b* gene was amplified from zebrafish and cloned in front of the C-terminal FLAG tag of an expression vector. The plasmid was transfected into human HeLa-57A cells that do not express endogenous TLRs but stably express an NF-κB inducible luciferase reporter gene. Stimulation of drTLR5b transfected cells with recombinant flagellin of *Salmonella enterica* serovar Enteritidis ($FliC^{SE}$) or *Pseudomonas aeruginosa* ($FliC^{PA}$)[28] failed to activate NF-κB (Fig. 1A), although confocal microscopy confirmed expression of FLAG tagged drTLR5b in transfected cells (Fig. 1B). Both $FliC^{SE}$ and $FliC^{PA}$ did activate NF-κB in HeLa-57A cells transfected with human TLR5 (hTLR5), indicating that TLR expression and signaling are functional in these cells and that the flagellins displayed TLR5 activating capacity (Fig. 1C). To test whether drTLR5b might recognize native instead of recombinantly produced flagellin, we stimulated transfected cells with whole cell lysates of highly motile *S.* Enteritidis, *P. aeruginosa* and of the fish pathogen *Aeromonas hydrophila*. None of the lysates activated cells in a drTLR5b-dependent fashion (Fig. 1D).

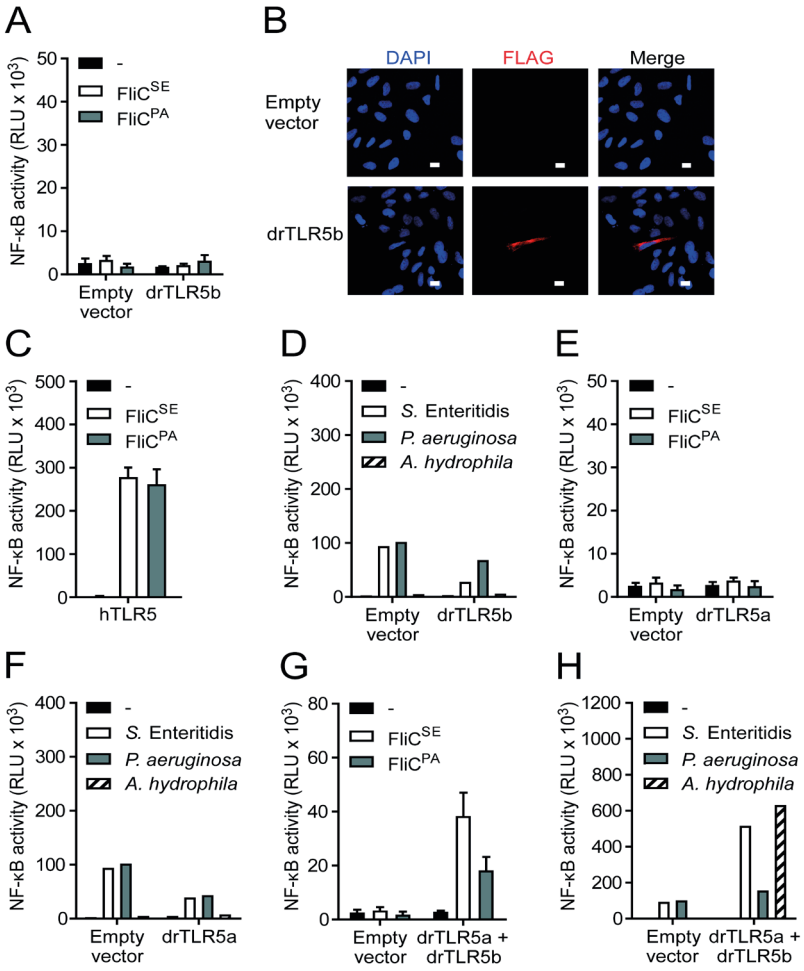


Figure 1. Induction of NF-κB by zebrafish TLR5b (drTLR5b) and drTLR5a upon stimulation with purified flagellins and bacterial lysates. **(A, C-H)** HeLa-57A cells were transfected with drTLR5b, drTLR5a, human TLR5 (hTLR5) or drTLR5b and drTLR5a combined as indicated. Control cells were transfected with empty vector. Cells were stimulated (5h) with vehicle (-) or 1 µg/ml of purified recombinant *Salmonella enterica* serovar Enteritidis (FliC^{SE}) or *Pseudomonas aeruginosa* (FliC^{PA}) flagellin **(A, C, E, G)** or 2 µg/ml of total protein from lysates of *S.* Enteritidis, *P. aeruginosa* or *Aeromonas hydrophila* **(D, F, H)**. Note that the NF-κB response of the drTLR5 heterodimer is higher when stimulated with bacterial lysate compared to purified recombinant flagellin. **(B)** Fluorescence microscopy of empty vector or FLAG tagged drTLR5b transfected HeLa-57A cells stained with M2 α-FLAG and DAPI for nuclear visualization. White scale bar is 10 µm. NF-κB activity is represented by luciferase activity in relative light units (RLU). Values are the mean ± SEM of three independent experiments **(A, C, E, G)** or a representative of three independent experiments **(D, F, H)** all performed in duplicate.

Since in zebrafish the *tlr5* gene is duplicated, we questioned whether its paralog, annotated as TLR5a (drTLR5a), could be the functional receptor to flagellin. dr*tlr5a* and dr*tlr5b* are positioned in tandem on chromosome 20. Both

genes translate into proteins of 881 amino acids that have characteristic TLR5 features including a signal peptide, N-terminal leucine rich repeat (NTLRR), 22 consecutive LRR's, C-terminal LRR (CTLRR), transmembrane domain (TM) and an intracellular Toll-Interleukin-1 receptor (TIR) signaling domain (Fig. S1). Overall, the protein sequences of drTLR5b and drTLR5a are 83% similar with highest similarity in the transmembrane and TIR domains and lowest similarity in the LRR ligand binding ectodomain, suggesting a possible different interaction of flagellin with drTLR5a than with drTLR5b. We cloned drTLR5a and expressed the receptor as described for drTLR5b. Stimulation of drTLR5a transfected cells with $FliC^{SE}$ and $FliC^{PA}$ failed to activate NF-κB (Fig. 1E). drTLR5a also did not respond to bacterial lysates (Fig. 1F). Since both receptors separately were not functional, we explored the possibility that duplicated drTLR5b and drTLR5a may have co-evolved to form a functional receptor. Stimulation of cells co-transfected with drTLR5a and drTLR5b indeed yielded robust activation of NF-κB when exposed to $FliC^{SE}$ and $FliC^{PA}$ (Fig. 1G). Likewise, drTLR5a and drTLR5b together activated NF-κB in response to lysate of motile bacteria, especially *A. hydrophila* (Fig. 1H). These data strongly suggest that drTLR5a and drTLR5b do not function as conventional homodimers but instead form functional TLR5 heterodimers.

Zebrafish UNC93B1 facilitates formation of functional drTLR5a and drTLR5b heterodimers

To gain additional evidence for drTLR5a and drTLR5b acting as heterodimeric receptors, we constructed drTLR5a with a C-terminal HA tag and visualized receptor localization in drTLR5a-HA and drTLR5b-FLAG transfected HeLa-57A cells. Confocal microscopy revealed overlapping expression of drTLR5a and drTLR5b throughout cells rather than at a specific cellular location (Fig. 2A). Previously, human TLR5 was found to localize at the cell surface only in the presence of the TLR trafficking chaperone UNC-93 homolog B1 (UNC93B1)[29]. UNC93B1 interacts with TLRs via acidic amino acids in the TLR extracellular juxtamembrane region (EJM)[29,30] and is required for TLR3, 7, 8 and 9 to traffic to endo-lysosomal compartments[31-33]. Inspection of the amino acid sequences of drTLR5a and drTLR5b revealed identical EJM's that contained multiple acidic residues (Fig. S1), suggesting that drTLR5a and drTLR5b may interact with UNC93B1. Co-transfection of the zebrafish homolog of UNC93B1 (drUNC93B1) resulted in relocation of drTLR5a-HA and drTLR5b-FLAG to the same vesicle like compartments (Fig. 2B). Separate co-transfection of FLAG-tagged drUNC93B1 with drTLR5a-HA or drTLR5b-HA showed independent interaction of drUNC93B1 with both receptors (Fig. S2A). Staining of the drTLR5a and drTLR5b containing

vesicles using antibodies against Early Endosome Antigen-1 (EEA-1) or Lysosomal Associated Membrane Protein-1 (LAMP-1) demonstrated stronger association of drTLR5a and drTLR5b with LAMP-1 than with EEA-1 positive vesicles (Fig. S2B).

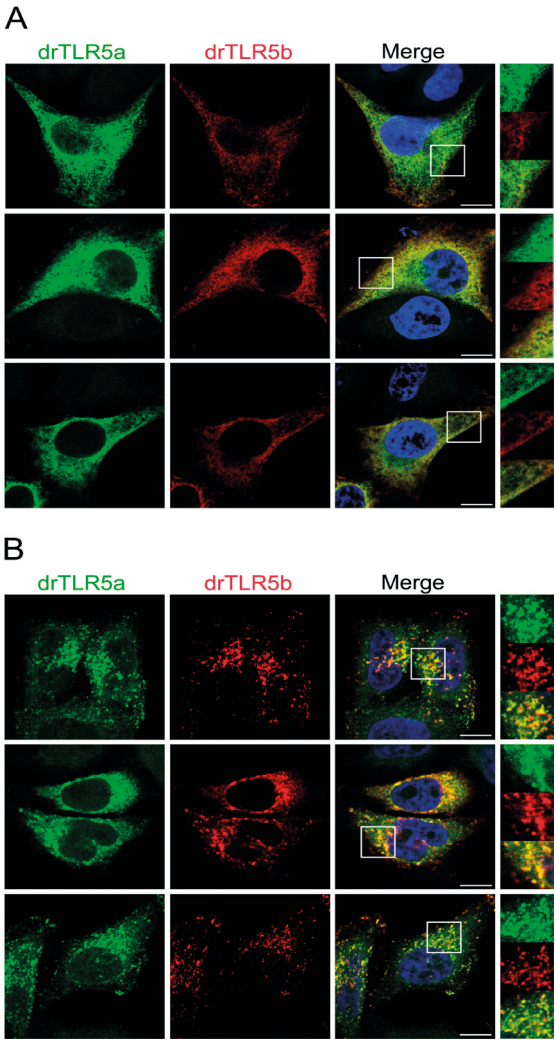


Figure 2. Effect of zebrafish UNC93B1 (drUNC93B1) on localization of drTLR5a and drTLR5b. Confocal microscopy on HeLa-57A cells transfected with; **(A)** drTLR5-HA and drTLR5b-FLAG or **(B)** drTLR5a-HA and drTLR5b-FLAG and untagged drUNC93B1. Merge images show nuclei stained with DAPI (blue). Images are selected from three independent experiments and three representative images are shown for each transfected group. Scale bar in merge images is 10 µm.

Flagellin stimulation of cells expressing drUNC93B1 and both drTLR5 receptors induced a potent (>10-fold) increase in NF-κB activity compared to cells lacking drUNC93B1 (Fig. 3). drUNC93B1 alone or in combination with either drTLR5a or drTLR5b did not induce NF-κB activity towards $FliC^{SE}$ (Fig. 3). Exposure of cells to $FliC^{SE}$ for three hours did not alter the localization of either drTLR5a or drTLR5b at LAMP-1 positive vesicles (Fig. S3). Together, these findings indicate that zebrafish UNC93B1 operates as a trafficking chaperone to both

drTLR5a and drTLR5b and that despite accumulative localization in vesicles, separate receptors cannot function as a typical TLR5 homodimer, whereas drUNC93B1-mediated relocation of both receptors facilitates the formation of functional drTLR5a and drTLR5b heterodimers.

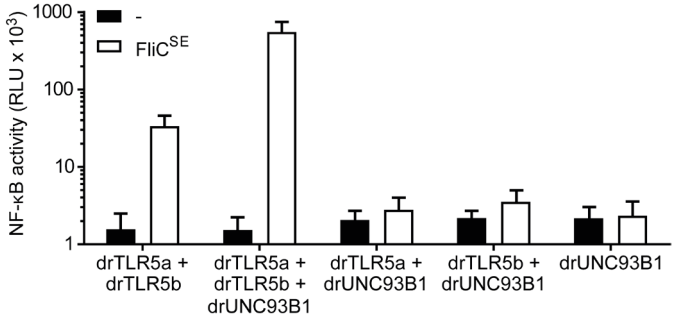


Figure 3. Effect of drUNC93B1 on drTLR5a and drTLR5b mediated NF-κB activation. HeLa-57A cells expressing drUNC93B1, drTLR5a and/or drTLR5b in the indicated combinations were stimulated (5h) with vehicle (-) or 1 μg/ml *Salmonella enterica* serovar Enteritidis flagellin ($FliC^{SE}$). Data shows NF-κB activity represented by luciferase activity in relative light units (RLU). Values are the mean ± SEM of three independent experiments performed in duplicate.

The N-terminal fragment of drTLR5b but not drTLR5a is functional as a homodimer

Our finding that full-length drTLR5b cannot signal as a homodimer was unexpected as the reported crystallized protein structure indicated that the N-terminal fragment (NTLRR-LRR14) of drTLR5b binds flagellin as a homodimeric complex[19]. To investigate this discrepancy between binding and signaling in response to flagellin, we first determined whether the N-terminal region of the drTLR5b ectodomain can function in homodimeric configuration when all other regions of the receptor complex are heterodimeric. For this, we replaced the N-terminal part of the ectodomain (NTLRR-LRR14) of drTLR5a for the corresponding drTLR5b sequence. The exchanged protein fragment matches exactly with the published crystallized structure of drTLR5b (Fig. S1). The resulting chimeric receptor was named "BA-A". In this terminology "BA" indicates the combination of the N-terminal part of the ectodomain (N-ECD) of drTLR5b and the C-terminal part of the ectodomain (C-ECD) of drTLR5a; "-" indicates the transmembrane region and; "A" the drTLR5a intracellular domain (ICD) (Fig. 4A). Exposure of cells co-transfected with chimera BA-A and wild type drTLR5b (BB-B) to $FliC^{SE}$ (Fig. 4B) or $FliC^{PA}$ (Fig. S4) strongly activated NF-κB. This suggests that flagellin can bind to the homodimeric drTLR5b N-ECD, as the crystal structure dictates, and is capable of inducing functional dimerization when other regions of the receptor are in a

heterodimeric configuration. To determine whether the N-ECD of drTLR5a can act in a similar way, we constructed the inverse chimera "AB-B" and co-transfected this chimera with wild type drTLR5a (AA-A). Interestingly, stimulation of this receptor combination with FliC^{SE} (Fig. 4B) or FliC^{PA} (Fig. S4) failed to activate NF-κB. To verify that the AB-B chimeric construct was expressed and functional, cells were co-transfected with chimeras AB-B and BA-A which would reinstate a fully complementary heterodimeric complex. FliC^{SE} stimulation of these cells yielded NF-κB activation (Fig. 4B). Cells transfected with only the BA-A or AB-B chimeras did not respond to FliC^{SE} (Fig. 4C). These results indicate that the interaction with flagellin differs between drTLR5a and drTLR5b and that heterodimerization is required for TLR5 signal transduction.

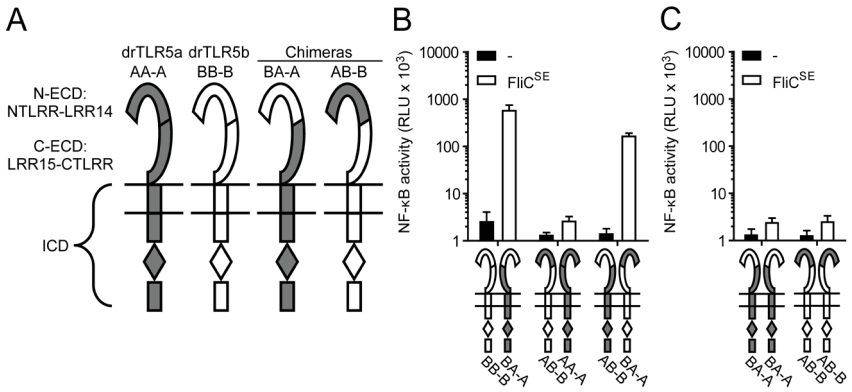


Figure 4. Functionality of the N-terminal part of the ectodomain of drTLR5b and drTLR5a. **(A)** Schematic representation of wild type and chimeric TLR5 constructs. N-ECD: N-terminal part of the ectodomain (ECD) ranging from the N-terminal leucine rich repeat (NTLRR) to LRR14. C-ECD: C-terminal part of the ECD ranging from LRR15 to the C-terminal LRR (CTLRR), ICD: intracellular domain. **(B-C)** HeLa-57A cells transfected with drUNC93B1 and the indicated receptor combinations were stimulated (5h) with vehicle (-) or 1 µg/ml *Salmonella enterica* serovar Enteritidis flagellin (FliC^{SE}). Data shows NF-κB activity represented by luciferase activity in relative light units (RLU). Values are the mean ± SEM of three independent experiments performed in duplicate.

Transfer of N-ECD amino acids from drTLR5b to drTLR5a and human TLR5

The N-ECD protein sequences of drTLR5a and drTLR5b vary in 108 out of 369 amino acids (excluding the signal peptide). To analyze the possible structural implications of this difference, we modelled the N-ECD sequence of drTLR5a onto the crystallized structure of the ligand-free (PDB ID: 3v44) and the flagellin bound drTLR5b N-ECD (PDB ID: 3v47). Superposition of the structures predicted that the drTLR5 N-ECDs are structurally highly similar both in the absence (Fig. 5A) or presence (Fig. S5A) of flagellin, except for a laterally protruding loop which

connects LRR9 and LRR10. This loop was found to change conformation upon binding of flagellin[19]. Yoon *et al*. marked this loop as the flagellin binding "hotspot" since its deletion severely impaired flagellin binding. The hotspot forms part of 45 residues in the drTLR5b N-ECD crystal structure that are involved in dimerization and flagellin binding. Twenty-one of the 45 residues in drTLR5b, including six residues in the hotspot, are different at the same position in drTLR5a (Fig. S5B). In an attempt to reconstitute a functional homodimeric drTLR5a N-ECD, we adapted chimera AB-B by replacing the deviant drTLR5a N-ECD residues for their corresponding drTLR5b residues. We first tested the effect of replacing 14 of the 21 residues that were least similar by constructing chimera A^{14}B-B. Although the chimera was expressed (Fig. S5C), co-transfection of cells with chimera A^{14}B-B and AA-A (drTLR5a) did not restore responsiveness to $FliC^{SE}$. When all 21 deviant residues were replaced for their drTLR5b counterparts, functionality was still not restored as shown by the lack of $FliC^{SE}$ induced activation of chimera A^{21}B-B co-expressed with AA-A. Co-transfection of chimeras A^{14}B-B or A^{21}B-B with chimera BA-A also did not result in a $FliC^{SE}$ responsive receptor complex (Fig. S5D). These results indicate that the transfer of amino acids involved in flagellin binding and dimerization in drTLR5b to drTLR5a is not sufficient to reconstitute a functional N-terminal domain.

Of the 45 interacting residues in drTLR5b, the hotspot is the only cluster of residues that is somewhat evolutionarily conserved in TLR5 of avian, reptilian and mammalian species, including human TLR5 (Fig. S6). Although the loop of hTLR5 contains different residues (Fig. 5B), modelling of the hTLR5 N-ECD onto drTLR5b predicted no structural difference of the LRR9 loop, neither in absence (Fig. 5C) or presence (Fig. S7A) of flagellin. To determine whether the predicted different loop conformation of drTLR5a versus drTLR5b and hTLR5 would affect receptor function, we replaced the loop in homodimeric hTLR5 for the structurally dissimilar drTLR5a hotspot loop (hTLR5-hs5a). hTLR5 carrying the structurally similar drTLR5b hotspot loop (hTLR5-hs5b) was constructed as a control. While receptors were expressed (Fig. S7B) and wild type hTLR5 activated NF-κB in response to $FliC^{SE}$, hTLR5-hs5a as well as hTLR5-hs5b failed to respond to $FliC^{SE}$ (Fig. 5D). These findings suggest that the LRR9 loop is crucial for hTLR5 activation and that, despite predicted structural similarity, this flagellin binding hotspot cannot be functionally exchanged between drTLR5b and hTLR5.

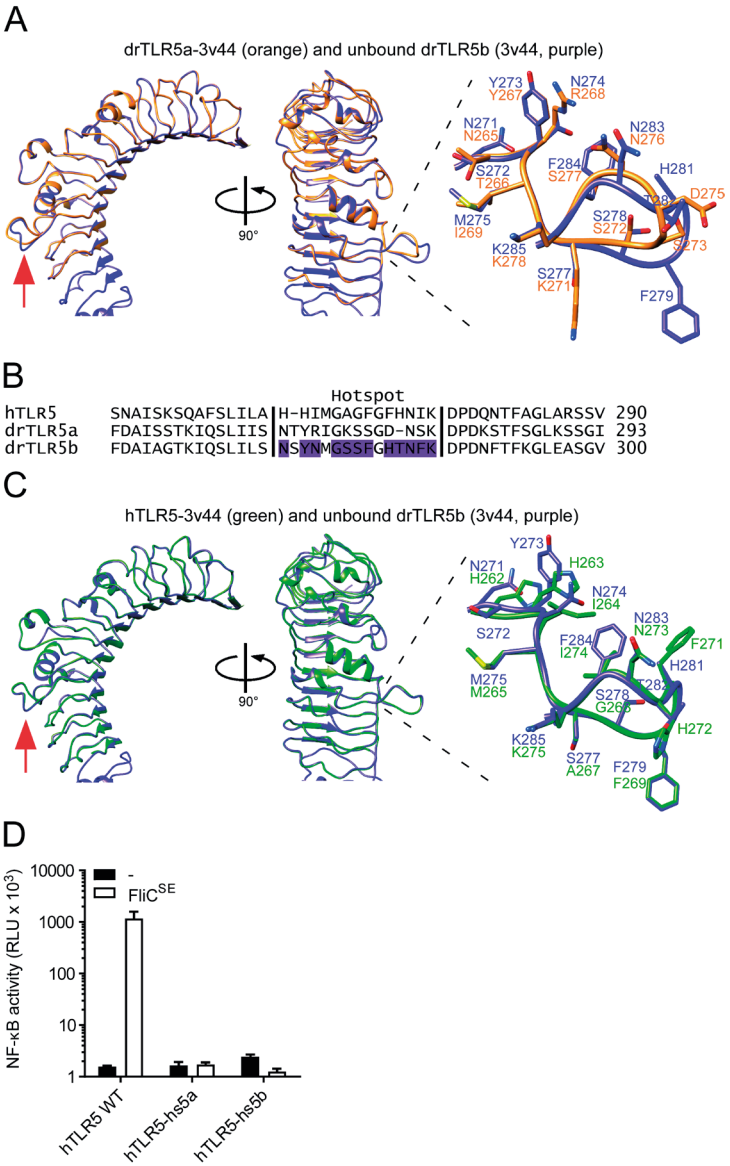


Figure 5. Structural modelling and functionality of the flagellin binding hotspot in human TLR5, drTLR5a and drTLR5b. **(A)** Superposition of unbound drTLR5b (PDB ID: 3v44, purple) and a model of drTLR5a based on 3v44 (drTLR5a-3v44, orange). Red arrow indicates the flagellin binding hotspot that forms a loop between LRR9 and LRR10. **(B)** Alignment of the putative flagellin binding hotspot of human TLR5 (hTLR5) and drTLR5a with drTLR5b. Purple colored residues in drTLR5b are involved in flagellin binding (see Yoon *et al*.[19]). **(C)** Superposition of unbound drTLR5b (PDB ID: 3v44, purple) and a model of human TLR5 based on 3v44 (hTLR5-3v44, green). **(D)** HeLa-57A cells transfected with wild type hTLR5 (hTLR5 WT), hTLR5 containing the hotspot of drTLR5a (hTLR5-hs5a) or the hotspot of drTLR5b (hTLR5-hs5b) were stimulated (5h) with vehicle (-) or 1 µg/ml *Salmonella enterica* serovar Enteritidis flagellin (FliC^{SE}). Data shows NF-κB activity represented by luciferase activity in relative light units (RLU). Values are the mean ± SEM of three independent experiments performed in duplicate.

Functional drTLR5 requires heterodimerization of either the N- or C-terminal ectodomain

To identify other (non N-ECD) region(s) in drTLR5a and drTLR5b that must operate in a heterodimeric configuration to establish a functional TLR5 dimer, we systematically exchanged domains or regions between drTLR5a and drTLR5b. Expression of different combinations of these chimeric and wild type receptors enabled us to determine the contribution of the different domains to the formation of functional heterodimeric TLR5. Reduced or loss-of-function upon homodimerization would indicate that the substituted domain must be in heterodimeric configuration to maintain functionality. Having established that the drTLR5b N-ECD can function in homodimeric configuration (see above), we next tested the effect of homodimerization of the full drTLR5b ectodomain while keeping the intracellular domain heterodimeric. For this, we constructed and expressed the combination of the chimeric receptor "BB-A" (consisting of the entire drTLR5b ectodomain linked to the drTLR5a ICD) (Fig. 6A) and wild type drTLR5b (BB-B). Exposure of the transfected cells to $FliC^{SE}$ did not result in NF-κB activity. Likewise, the reverse chimera AA-B in combination with wild type drTLR5a (AA-A) failed to respond to $FliC^{SE}$. Combined expression of BB-A and AA-B, which restored complete heterodimerization, did activate NF-κB upon $FliC^{SE}$ exposure ensuring that chimeras were expressed and functional (Fig. 6B). These findings indicate that the full ectodomains of drTLR5a and drTLR5b cannot function in homodimeric configuration.

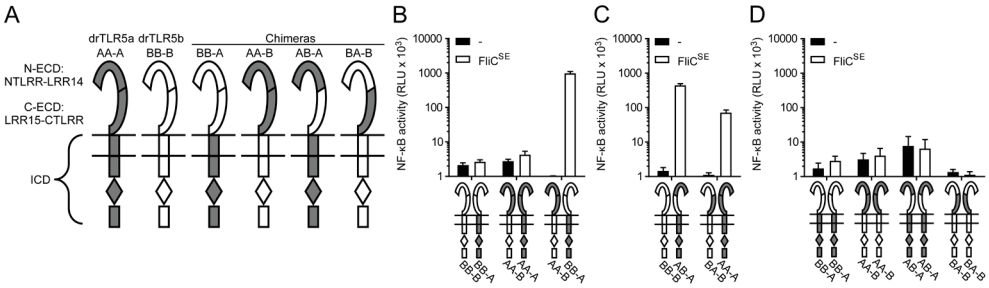


Figure 6. Heterodimerization of the ectodomain of drTLR5a and drTLR5b. **(A)** Schematic representation of wild type and chimeric TLR5 constructs. N-ECD: N-terminal part of the ectodomain (ECD) ranging from N-terminal leucine rich repeat (NTLRR) to LRR14. C-ECD: C-terminal part of the ECD ranging from LRR15 to the C-terminal LRR (CTLRR), ICD: intracellular domain. **(B-D)** HeLa-57A cells transfected with drUNC93B1 and the indicated receptor combinations were stimulated (5h) with vehicle (-) or 1 μg/ml *Salmonella enterica* serovar Enteritidis flagellin ($FliC^{SE}$). Data shows NF-κB activity represented by luciferase activity in relative light units (RLU). Values are the mean ± SEM of three independent experiments performed in duplicate.

As the N-terminal part of the drTLR5b ectodomain but not the full ectodomain did signal in homodimeric configuration (Fig. 4B), we anticipated that the remaining C-terminal part (LRR15-CTLRR) must be heterodimeric to allow dimer activation. To verify this, chimera AB-A was constructed (Fig. 6A) and co-transfected with wild type BB-B. Unexpectedly, stimulation of these cells with $FliC^{SE}$ still activated NF-κB, despite the homodimeric configuration of the drTLR5b C-ECD. Similarly, homodimerization of the drTLR5a C-ECD tested by expressing chimera BA-B in combination with wild type AA-A conferred NF-κB activation upon $FliC^{SE}$ stimulation (Fig. 6C). All ectodomain chimeras transfected individually did not respond to $FliC^{SE}$ stimulation (Fig. 6D). Taken together, these results show that the full length drTLR5a and drTLR5b ectodomains cannot function in homodimeric configuration, but that heterodimerization of either the N- or C-terminal LRR region is required to allow receptor activation.

Zebrafish TLR5 signaling requires heterodimerization across multiple regions in the intracellular domain

After delineating the extracellular needs for functional TLR5 heterodimerization, we determined the configurational requirements for the formation of functional intracellular receptor domains (ICDs). At the amino acid level, the ICDs of drTLR5a and drTLR5b are very similar (Fig. S1) suggesting that the ICDs of drTLR5a and drTLR5b may initiate signaling in homodimeric configuration. However, cells co-transfected with wild type BB-B and chimera AA-B, forming a heterodimeric ECD but homodimeric ICD of drTLR5b, did not respond to $FliC^{SE}$. Likewise, co-expression of wild type AA-A and chimera BB-A failed to activate NF-κB in response to $FliC^{SE}$ (Fig. 7B). To explore which distinct part of the ICD must be heterodimeric to allow signal initiation, we subdivided the ICD into four regions; transmembrane (TM), cytoplasmic juxtamembrane (CJM), TIR, and Tail region and constructed chimeras of drTLR5b in which these regions were substituted with their drTLR5a counterparts (Fig. 7A, Fig. S1). Co-expression of wild type drTLR5a and chimeric drTLR5b carrying the drTLR5a TM region still activated NF-κB, indicating that the TM region does not have to be heterodimeric in order to form a functional dimer. Interestingly, receptor combinations forming homodimeric CJM, TIR or Tail regions also still activated NF-κB when stimulated with $FliC^{SE}$ (Fig. 7C). As expected from their fully homodimeric configuration, the ICD chimeras transfected individually were unresponsive to $FliC^{SE}$ stimulation (Fig. 7D). These results indicate that while separate regions of the ICD can have identical sequences and still form a functional dimer, an entirely homodimeric ICD blocks the ability to initiate signaling. This suggests that heterodimeric interactions of multiple regions

between the TLR5 ICDs are required to constitute a functional zebrafish TLR5 receptor.

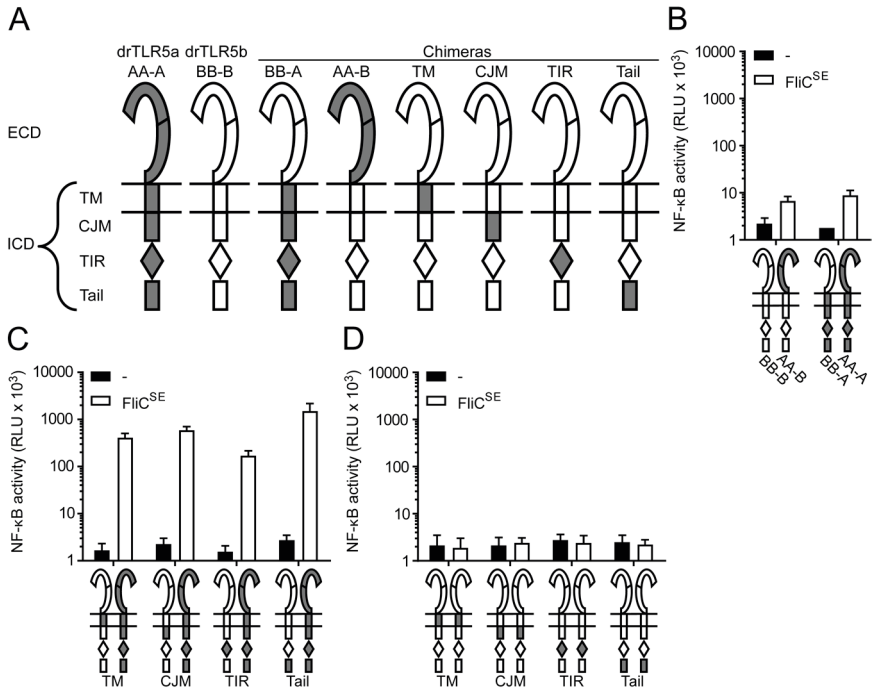


Figure 7. Heterodimerization of the intracellular domain of drTLR5a and drTLR5b. **(A)** Schematic representation of wild type and chimeric TLR5 constructs. ECD: ectodomain, TM: transmembrane, CJM: cytoplasmic juxtamembrane, TIR: Toll-Interleukin-1 receptor domain. **(B-D)** HeLa-57A cells transfected with drUNC93B1 and the indicated receptor combinations were stimulated (5h) with vehicle (-) or 1 µg/ml *Salmonella enterica* serovar Enteritidis flagellin ($FliC^{SE}$). Data shows NF-κB activity represented by luciferase activity in relative light units (RLU). Values are the mean ± SEM of three independent experiments performed in duplicate.

Discussion

Detection of flagellin by TLR5 is important to initiate immune responses against motile pathogenic bacteria[34-36]. Vertebrate animals including mammals, birds and reptiles have one TLR5 ortholog that detects flagellin as a homodimeric complex[12,13,37]. The crystal structure of drTLR5b in complex with flagellin provided structural support for the homodimeric mode of action of other TLR5 species. However, direct evidence that drTLR5b is functional as a homodimeric receptor has been lacking. Here we show that zebrafish TLR5 only responds to flagellin as a heterodimeric complex composed of the products of the duplicated *tlr5* genes; drTLR5a and drTLR5b. When expressed separately, drTLR5b and drTLR5a were

unable to activate NF-κB in response to purified flagellins or lysates of motile bacteria. Even after expression of drUNC93B1, which led to redistribution and more robust NF-κB activation by the heterodimeric receptor, individual drTLR5a and drTLR5b did not operate as homodimers. Additional evidence for the requirement to form a heterodimeric complex came from the use of chimeric receptors that were only functional in combination with other chimeras composed of swapped complementary regions (that allowed full heterodimerization). Given that TLRs universally function via dimerization[3,38] our data firmly indicate that drTLR5a and drTLR5b detect flagellin as a heterodimeric complex. Our results provide the molecular basis for the observation that knockdown of either dr*tlr5a* or dr*tlr5b* in zebrafish embryos hampers upregulation of pro-inflammatory genes after flagellin challenge[39,40].

Zebrafish UNC93B1 had a profound effect on the trafficking of drTLR5a and drTLR5b and signaling by the TLR5 heterodimer. drUNC93B1 enhanced relocalization of both receptors towards LAMP-1 specific compartments and this coincided with a strongly increased capacity of cells to respond to flagellin. UNC93B1 is a membrane protein that complexes in the ER to TLR3, 7, 8, 9 and TLR5 via acidic residues in the TLR EJM[29, 41, 42]. The UNC93B1-TLR complex then exits the ER to transit to the Golgi apparatus. Further trafficking of TLRs to endosomes, lysosomes or the cell surface involves other trafficking partners[41, 42]. In the EJM of both drTLR5a and drTLR5b multiple acidic residues have been conserved and the ability of drUNC93B1 to facilitate relocalization of drTLR5a and drTLR5b independently, identifies drUNC93B1 as a trafficking chaperone for both zebrafish TLR5 receptors.

The use of an extensive set of chimeric receptors revealed that the N-ECD of drTLR5b but not drTLR5a could act in homodimeric configuration to activate the receptor complex. In the drTLR5b N-ECD, 45 residues are involved in dimerization and flagellin binding[19]. Of these 45 residues 21 are different in the drTLR5a N-ECD and of these 21 residues 14 differ significantly. Our attempts to restore homodimeric functionality of the drTLR5a N-ECD by replacing these differing residues in the AB-B chimera for their drTLR5b counterparts and combining this chimera with wild type drTLR5a (AA-A), failed. In addition, replacement of the differing residues in an otherwise heterodimeric complex (by combining chimeras A^{14}B-B or A^{21}B-B with BA-A) abrogated function. These findings indicate that the 45 residues involved in flagellin binding and dimerization of the drTLR5b N-ECD are either not sufficient for the formation of a functional homodimeric drTLR5a N-ECD or that epistatic interactions within the chimeric N-ECD prevent the formation of a functional receptor. Information as to whether this difference between the drTLR5a and drTLR5b N-ECD is caused by a lack of interaction of drTLR5a with flagellin or

drTLR5a signaling defects awaits elucidation of the flagellin-drTLR5a N-ECD crystal structure.

Our finding that the N-ECD of drTLR5b has maintained the ability to function as a homodimer is fully in line with the homodimeric binding of flagellin in the drTLR5b N-ECD crystal structure and indicates that drTLR5b more closely resembles homodimeric TLR5 of other species. Nonetheless, the drTLR5b ectodomain was only functional if at least one of the ECD regions (N- or C-ECD) was in heterodimeric configuration with drTLR5a. The requirement for heterodimerization in the ECD was not specifically linked to the N- or C-ECD regions. This suggests that flagellin induced receptor complex activation depends on the ectodomains of drTLR5a and drTLR5b to have different conformations rather than on specific amino acid residues. Structural information on the C-ECD's of TLR5 is needed to elucidate such different conformations in the ectodomains.

Modeling of hTLR5 onto drTLR5b predicted very high structural similarity of the N-ECD, including at the loop between LRR9 and LRR10 which forms the flagellin binding hotspot. Yet, substitution of amino acids that form this loop in hTLR5 for those of drTLR5b completely abrogated hTLR5 activation. According to Yoon *et al*. more than 60% of interactions in the drTLR5b loop would be conserved in hTLR5 and the loop would thus also form a flagellin binding hotspot in hTLR5. Our findings confirm that the hotspot is critical for hTLR5 activation but also indicate that the hTLR5 and drTLR5b hotspots are functionally distinct. The flagellin binding hotspot forms a pocket which accommodates a highly conserved arginine residue in flagellin (R90). Flagellin binding and thus insertion of R90 in the hotspot pocket alters the conformation of the LRR9 loop[19]. Mutation of R90 preserves binding of flagellin to hTLR5 and mouse TLR5 (mTLR5) but reduces receptor activation[43, 44] suggesting that insertion of R90 into the hotspot pocket initiates a conformational change that is necessary for receptor activation[44]. Although the hotspot pocket of mTLR5 and hTLR5 differ only by three residues (Fig. S6), R90 mutations differentially affect mTLR5 and hTLR5 activation[43, 44]. In addition, substitution of the proline residue (P268) in the mTLR5 hotspot for the alanine residue (A267) in hTLR5 changed mTLR5 flagellin recognition specificity towards the pattern of hTLR5[43]. Homology models constructed of hTLR5 and mTLR5 based on drTLR5b predicted differential binding affinities between hTLR5 and mTLR5 to flagellin residue E93. Flagellin residue E93 interacts extensively with drTLR5b but mutation of E93 had no functional effect on both human and mouse TLR5[44]. In the present study, substitution of the hTLR5 hotspot with corresponding drTLR5b residues effectively changed eight and inserted one residue. This alteration was sufficient to completely preclude hTLR5 activation suggesting that the conformational change of the drTLR5b hotspot upon flagellin binding differs from

what is necessary to induce functional homodimerization of hTLR5. Our findings corroborate with the published work and provide additional evidence that TLR5-flagellin recognition is species-specific. This is important to consider when using TLR5 of non-target species as a model to define the exact mechanism of flagellin binding and flagellin induced changes required for TLR5 activation.

Our data demonstrate that the cytoplasmic domains of drTLR5a and drTLR5b initiated signaling in a hetero- but not homodimeric configuration. This was unexpected since the drTLR5a and drTLR5b cytoplasmic domains are highly similar. When tested separately, all of the defined cytoplasmic regions, including the TIR domain, could function in homodimeric configuration only if other regions were heterodimeric. Since a fully homodimeric cytoplasmic domain could not signal, these findings suggest functional interdependence of the distinct regions, i.e. a requirement for transregional heterodimeric configuration. This may be best explained by different, conformational changes occurring in hetero- versus homodimeric cytoplasmic domains. TLR activation involves ligand induced conformational changes in the ectodomain that transit through the transmembrane region to result in dimerization of the cytoplasmic TIR domains[14, 45]. The TLR TIR domain consists of five β-sheets surrounded by five α-helixes[46]. These secondary structures are connected by loops and the BB loops of two TIR domains have to align correctly to form a dimerization interface that is necessary for the TIR domains to recruit signaling adapters[14, 47]. Based on our findings we postulate that activation of the drTLR5 heterodimer starts with a flagellin binding induced conformational change in the ectodomain, probably in the LRR9 loop, that in turn induces a conformational change, e.g. rotational, across multiple regions in the heterodimeric cytoplasmic domain to correctly align the TIR BB loops to form an appropriate dimerization interface. Such a ligand induced rotational conformational change in drTLR5 could be similar to those occurring in other type I membrane receptors like the epidermal growth factor and thrombopoietin receptors[14, 48, 49] and may also be required for activation of other heterodimeric TLRs like the TLR1/2 and TLR2/6 combinations. Further research using structural and biophysical experiments is necessary to validate this scenario.

The drTLR5a and drTLR5b receptors respond to flagellin as a complementary heterodimer. Whereas in zebrafish (a teleost fish) the *tlr5a* and *tlr5b* paralogs have likely arisen through a tandem gene duplication, non-teleost vertebrates including mammals, birds, reptiles and also the spotted gar (*Lepisosteus occulatus*), a holostean fish[50], all have one *tlr5* ortholog. This implies that the common ancestor of teleosts and non-teleost vertebrates likely had one TLR5 ortholog that may have operated as a homodimer. Partitioning of the ancestral function, i.e. flagellin detection, over the drTLR5a and drTLR5b paralogs points to

sub-functionalization of drTLR5a and drTLR5b that may be explained by the duplication-degeneration-complementation (DDC) model[51, 52]. According to this model, duplicated genes randomly acquire different degenerative mutations to the extent that protein paralogs must complement each other to maintain the ancestral function. drTLR5a and drTLR5b form an interesting example of protein sub-functionalization since the receptors maintain flagellin detection by physical complementation. Although drTLR5a and drTLR5b clearly detect flagellin as a heterodimer, we cannot exclude that homodimeric drTLR5a or drTLR5b may detect ligands other than flagellin. However, such ligands will likely not include common and abundant molecules from Gram-negative bacteria since neither drTLR5a nor drTLR5b responded to lysates of such bacteria. To our knowledge zebrafish TLR5 is the first example of TLR sub-functionalization in teleost fish. It is likely that in other species at least some of the duplicated TLR genes also evolved through DDC to maintain ancestral functions. This may explain the conservation in mammals of the TLR1, TLR2 and TLR6 heterodimeric partners.

In conclusion, the duplicated TLR5 genes of the zebrafish model organism encode two functional receptors that have sub-functionalized via concerted co-evolution to detect bacterial flagellins as a unique TLR5 heterodimer. The zebrafish TLR5 heterodimer expands our understanding of the mechanism of ligand induced TLR activation and provides an illustrative example of the functional co-evolution of duplicated genes.

Methods and Materials

DNA isolation and cDNA synthesis

Zebrafish tissue was kindly provided by Dr. J. den Hertog (Hubrecht Institute, Utrecht, The Netherlands). Genomic DNA was isolated from zebrafish tissue using the high pure template kit (Roche) according to the manufacturer's instructions. RNA was extracted from tissue lysed with RLT buffer (1% β-mercaptoethanol) (Qiagen) in 1.4 mm Fastprep lysing matrix tubes (MPbio) in a Magna Lyser centrifuge (6,500 × g, 40 s, RT) (Roche). Total RNA was isolated using the RNeasy mini kit (Qiagen), treated with DNase I (1 U/mg RNA, Thermo Fisher Scientific) and reverse transcribed into cDNA using the RevertAid First Strand cDNA Synthesis kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

Recombinant DNA techniques

Phusion high-fidelity DNA polymerase, dNTP's, fast digest restriction endonucleases, T4 DNA ligase and primers were purchased from Thermo Fisher

Scientific. PCR products were extracted from agarose gel using the GeneJET Gel Extraction Kit (Thermo Fisher Scientific). Plasmids were isolated with Nucleobond Xtra Midi prep kit (Macherey-Nagel).

DNA constructs

Zebrafish *tlr5b* and *tlr5a* were amplified from genomic DNA with forward primers CV039 and CV095 (Table S1) resp. to add a BamHI restriction site and Kozak sequence at the start of the gene and reverse primers CV054 and CV096 resp. to add a NotI restriction site at the end of the gene. Both dr*tlr5b* and dr*tlr5a* PCR products were digested with BamHI and NotI and ligated into pTracer-CMV2ΔGFP/3×FLAG[11] yielding drTLR5b and drTLR5a carrying a C-terminal 3×FLAG tag. To create HA (hemagglutinin-epitope) tagged drTLR5a and drTLR5b, pTracer-CMV2ΔGFP/3×FLAG was modified by inserting a 3×HA sequence (YPYDVPDYA) ending with a stop codon in the 3×FLAG sequence using NotI and ClaI, yielding pTracer-CMV2ΔGFP/3×HA. dr*tlr5a* and dr*tlr5b* were cut from pTracer-CMV2ΔGFP/3×FLAG using BamHI and NotI and inserted in pTracer-CMV2ΔGFP/3×HA yielding drTLR5b and drTLR5a carrying a C-terminal 3×HA tag. Zebrafish *unc93b1* was amplified from cDNA with forward primer CV159 and reverse primer CV160 or reverse primer CV180 to maintain the stop codon. The PCR products were cloned into pTracer-CMV2ΔGFP/3×FLAG using KpnI and NotI restriction enzymes yielding drUNC93B1 with a C-terminal 3×FLAG tag or untagged drUNC93B1. Human TLR5 was cloned from pUNO-hTLR5-GFP (Invivogen) with primers CV257 and CV259, ligated into pTracer-CMV2ΔGFP/3×FLAG using BamHI and NotI to yield hTLR5 with a C-terminal 3×FLAG tag. Chimeric DNA constructs were created by standard overlap extension PCR technique. The DNA fragments encoding the N-terminal ECD in chimera's A^{14}B-B, A^{21}B-B, hTLR5-hs5a and hTLR5-hs5b were purchased from GeneArt (Thermo Fisher Scientific) and fused to the receptor genes by overlap extension PCR. Chimeric PCR products were ligated in pTracer-CMV2ΔGFP/3×FLAG using BamHI and NotI. Primer sequences used for chimera construction are available upon request. All DNA constructs were verified by sequencing (Macrogen). The sequences were deposited in GenBank with accession numbers; drTLR5b: MF983797, drTLR5a: MF983798, drUNC93B1: MF983799.

Protein sequence analysis and TLR5 modelling

Protein sequences were aligned using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/)[53] with default settings. Leucine rich repeats were identified by manual sequence inspection according to Matsushima *et*

al.[54] and with use of the Leucine Rich Repeat Finder (http://www.lrrfinder.com/lrrfinder.php)[55]. The transmembrane domain was predicted with TMHMM Server v 2.0 (http://www.cbs.dtu.dk/services/TMHMM/)[56]. The TIR domain was predicted by identifying five alternating β-sheets and α-helices[46] using the Proteus Protein Structure Prediction server (http://wks80920.ccis.ualberta.ca/proteus/)[57]. Structural models of the ECD N-terminal part of drTLR5a and hTLR5 were created using SWISS-MODEL (default settings) (https://swissmodel.expasy.org/)[58] with the drTLR5b crystal structure (PDB: 3v44) or drTLR5b-flagellin crystal structure (PDB: 3v47) as templates. UCSF Chimera (http://www.rbvi.ucsf.edu/chimera)[59] was used to view protein models.

Isolation of recombinant flagellins and bacterial lysate

Construction and purification of recombinant HIS-tagged FliC of *Salmonella enterica* serovar Enteritidis ($FliC^{SE}$)(strain 90-13-706 CVI, Lelystad, the Netherlands) and *Pseudomonas aeruginosa* ($FliC^{PA}$)(reptile clinical isolate, Utrecht University) has been described previously[13, 37]. Purified flagellins were stored in buffer; 4 M urea, 10 mM NaH_2PO_4, 100 mM TRIS (all Sigma) at -20 °C. For lysate isolation, *S.* Enteritidis, *P. aeruginosa* and *Aeromonas hydrophila* (fish clinical isolate, Utrecht University) were grown (24 h) on sheep blood agar plates (Biotrading) at 28 °C. Single colonies were grown (28 °C, 16 h) in 5 mL HI broth (Biotrading) at 160 rpm and placed on ice. Microscopic examination confirmed that all bacteria were highly motile. Cultures were normalized to an OD_{550} of 3, pelleted by centrifugation (30 min, 5000 × *g*, 4 °C), washed with 1 mL Dulbecco's phosphate buffered saline (DPBS, Sigma), vortexed, and collected by centrifugation (15 min, 5000 × *g*, 4 °C). Pellets were re-dissolved in 2 mL DPBS and placed at 70 °C for 1h. Heat killed bacteria were sonicated (6 × 15 s, Vibra-cell, Sonics, USA) and centrifuged (14,000 × *g*, 40 min, 4 °C). Supernatants were concentrated with Pierce protein concentrators 10 kDa (Thermo Fisher Scientific) for 15 min at 5000 × *g*, RT. Total protein concentration of lysates was determined by BCA assay (Thermo Fisher Scientific). Lysate supernatants were stored at − 20 °C until use.

Cell culture and transient transfection

Human HeLa-57A cells (that stably carry a NF-κB luciferase reporter construct[60]) were cultured in DMEM (Thermo Fisher Scientific) with 5% fetal calf serum (Bodinco) at 37 °C and 10% CO_2. For transfection, cells were grown to 80% confluency in a 12 wells plate and transfected with 666,66 ng of each plasmid using Fugene HD (Promega) at a DNA to Fugene ratio of 1:3 following the manufacturer's

instructions. Empty pTracer 3×FLAG plasmid was used to equalize the total amount (2 μg) of transfected plasmid added to each well.

Luciferase assay

Twenty-four hours after transfection cells were redistributed in a 96-well plate. After 24 h cells were washed twice with medium without FCS and stimulated with purified flagellin or bacterial lysate in 100 μL medium without FCS. After 5 h at 37 °C cells were washed twice with DPBS (Sigma) and lysed with reporter lysis buffer (50 μL, Promega) at − 80 °C for at least 1 h. After thawing 15 μL lysate was mixed with 37 μL luciferase reagent (Promega) and luciferase activity was measured with a TriStar2 luminometer (Berthold). Results were expressed in relative light units (RLU). Receptors were assigned to be unresponsive when luciferase activity after flagellin stimulation did not exceed 10 RLU ($\times 10^3$).

Confocal laser microscopy

Twenty-four hours after transfection cells were seeded onto glass coverslips. After 24 h cells were prepared for confocal microscopy. Cells were washed with TBS and fixed (30 min) with TBS/1.5% paraformaldehyde (Affimetrix). Cells were permeabilized and blocked (30 min) with TBS containing 0.1% Saponin and 0.2% BSA (both Sigma). Next, cells were incubated (1 h) with mouse M2-α-FLAG (F3165, Sigma), mouse α-EEA-1 (610456, BD Bioscience) or rabbit α-LAMP-1 (24170, Abcam), washed with TBS, incubated (1 h) with Alexa Fluor-568 goat-α-mouse IgG (A11031, Thermo Fisher Scientific) or Alexa Fluor-568 goat-α-rabbit IgG (A11036, Thermo Fished Scientific), washed with TBS and incubated (1 h) with Alexa Fluor-488 mouse-α-HA (A21287, Thermo Fisher Scientific). After antibody staining cells were washed with TBS and milliQ and embedded in Prolong diamond mounting solution (Thermo Fisher Scientific). Cells were viewed on a Leica SPE-II or Nikon A1R/STORM laser confocal microscope and images were processed using Leica LAS AF or Nikon NIS software.

Acknowledgements

We kindly thank Dr. R. Wubbolts (Center for Cell Imaging, Department of Biochemistry and Cell Biology, Utrecht University, Utrecht, the Netherlands) for acquisition of confocal microscope data and Dr. R.T. Hay (Institute of Biomolecular Sciences, School of Biomedical Sciences, University of St. Andrews, Scotland, UK) for providing the HeLa-57A cell-line.

Author contributions

C.V., J.W. and J.v.P. designed research, C.V. performed experiments and analyzed all data, C.V. wrote the paper with support from J.W. and J.v.P.

References

1. Medzhitov, R. Toll-like receptors and innate immunity. *Nature Reviews Immunology* **1,** 135–45 (2001)
2. Leulier, F. & Lemaitre, B. Toll-like receptors--taking an evolutionary approach. *Nature Reviews Genetics* **9**, 165–78 (2008).
3. Gay, N. J. & Gangloff, M. Structure and function of Toll receptors and their ligands. *Annual Review of Biochemistry* **76,** 141–165 (2007).
4. Roach, J. C. *et al.* The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 9577–9582 (2005).
5. Poltorak, A. *et al.* Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* **282,** 2085–2088 (1998).
6. Zoete, M. R. de, Bouwman, L. I., Keestra, A. M. & Putten, J. P. M. van. Cleavage and activation of a Toll-like receptor by microbial proteases. *Proceedings of the National Academy of Sciences of the United States of America* **108,** 4968–4973 (2011).
7. Alexopoulou, L., Holt, A. C., Medzhitov, R. & Flavell, R. A. Recognition of double-stranded RNA and activation of NF-κB by Toll-like receptor 3. *Nature* **413**, 732–738 (2001).
8. Hemmi, H. *et al.* Small anti-viral compounds activate immune cells via the TLR7 MyD88–dependent signaling pathway. *Nature Immunology* **3**, 196–200 (2002).
9. Jurk, M. *et al.* Human TLR7 or TLR8 independently confer responsiveness to the antiviral compound R-848. *Nature Immunology* **3**, 499–499 (2002).
10. Hemmi, H. *et al.* A Toll-like receptor recognizes bacterial DNA. *Nature* **408**, 740–745 (2000).
11. Keestra, A. M., Zoete, M. R. de, Bouwman, L. I. & Putten, J. P. M. van. Chicken TLR21 is an innate CpG DNA receptor distinct from mammalian TLR9. *Journal of Immunology* **185**, 460–467 (2010).
12. Hayashi, F. *et al.* The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* **410**, 1099–1103 (2001).
13. Voogdt, C. G. P., Bouwman, L. I., Kik, M. J. L., Wagenaar, J. A. & Putten, J. P. M. van. Reptile Toll-like receptor 5 unveils adaptive evolution of bacterial flagellin recognition. *Scientific Reports* **6**, 19046 (2016).
14. Gay, N. J., Symmons, M. F., Gangloff, M. & Bryant, C. E. Assembly and localization of Toll-like receptor signalling complexes. *Nature Reviews Immunology* **14**, 546–558 (2014).
15. Botos, I., Segal, D. M. & Davies, D. R. The structural biology of Toll-like receptors. *Structure* **19**, 447–459 (2011).
16. Tanji, H., Ohto, U., Shibata, T., Miyake, K. & Shimizu, T. Structural reorganization of the Toll-Like receptor 8 dimer induced by agonistic ligands. *Science* **339**, 1426–1429 (2013).
17. Ohto, U. *et al.* Structural basis of CpG and inhibitory DNA recognition by Toll-like receptor 9. *Nature* **520**, 702–705 (2015).
18. Zhang, Z. *et al.* Structural analysis reveals that Toll-like receptor 7 is a dual receptor for guanosine and single-stranded RNA. *Immunity* **45**, 737–748 (2016).
19. Yoon, S. I. *et al.* Structural basis of TLR5-flagellin recognition and signaling. *Science* **335,** 859–64 (2012).
20. Palti, Y. Toll-like receptors in bony fish: From genomics to function. *Developmental & Comparative Immunology* **35**, 1263–1272 (2011).

21. Christoffels, A. *et al.* Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. *Molecular Biology and Evolution* **21**, 1146–1151 (2004).
22. Solbakken, M. H. *et al.* Evolutionary redesign of the Atlantic cod (*Gadus morhua* L.) Toll-like receptor repertoire by gene losses and expansions. *Scientific Reports* **6**, 25211 (2016).
23. Jaillon, O. *et al.* Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* **431**, 946–957 (2004).
24. Van de Peer, Y., Maere, S. & Meyer, A. The evolutionary significance of ancient genome duplications. *Nature Reviews Genetics* **10**, 725–732 (2009).
25. Fink, I. R. *et al.* Molecular and functional characterization of Toll-like receptor (Tlr)1 and Tlr2 in common carp (*Cyprinus carpio*). *Fish & Shellfish Immunology* **56**, 70–83 (2016).
26. Sullivan, C. *et al.* The gene history of zebrafish tlr4a and tlr4b is predictive of their divergent functions. *Journal of Immunology* **183**, 5896–5908 (2009).
27. Yeh, D.-W. *et al.* Toll-like receptor 9 and 21 have different ligand recognition profiles and cooperatively mediate activity of CpG-oligodeoxynucleotides in zebrafish. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 20711–20716 (2013).
28. Song, W. S. & Yoon, S. Crystal structure of FliC flagellin from *Pseudomonas aeruginosa* and its implication in TLR5 binding and formation of the flagellar filament. *Biochemical and Biophysical Research Communications* **444**, 109–115 (2014).
29. Huh, J.-W. *et al.* UNC93B1 is essential for the plasma membrane localization and signaling of Toll-like receptor 5. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 7072-7077 (2014).
30. Kim, J. *et al.* Acidic amino acid residues in the juxtamembrane region of the nucleotide-sensing TLRs are important for UNC93B1 binding and signaling. *Journal of Immunology* **190**, 5287–5295 (2013).
31. Kim, Y.-M., Brinkmann, M. M., Paquet, M.-E. & Ploegh, H. L. UNC93B1 delivers nucleotide-sensing toll-like receptors to endolysosomes. *Nature* **452**, 234–238 (2008).
32. Itoh, H. *et al.* UNC93B1 physically associates with human TLR8 and regulates TLR8-mediated signaling. *PLOS ONE* **6**, e28500 (2011).
33. Qi, R., Singh, D. & Kao, C. C. Proteolytic processing regulates Toll-like receptor 3 stability and endosomal localization. *Journal of Biological Chemistry* **287**, 32617–32629 (2012).
34. Carvalho, F. A. *et al.* Transient inability to manage Proteobacteria promotes chronic gut inflammation in TLR5-deficient mice. *Cell Host & Microbe* **12**, 139–152 (2012).
35. Vijay-Kumar, M. *et al.* Deletion of TLR5 results in spontaneous colitis in mice. *Journal of Clinical Investigation* **117**, 3909–3921 (2007).
36. Andersen-Nissen, E. *et al.* Cutting edge: Tlr5−/− mice are more susceptible to *Escherichia coli* urinary tract infection. *Journal of Immunology* **178**, 4717–4720 (2007).
37. Keestra, A. M., de Zoete, M. R., van Aubel, R. A. M. H. & Putten, J. P. M. van. Functional characterization of chicken TLR5 reveals species-specific recognition of flagellin. *Molecular Immunology* **45**, 1298–1307 (2008).
38. Bryant, C. E. *et al.* Advances in Toll-like receptor biology: Modes of activation by diverse stimuli. *Critical Reviews in Biochemistry and Molecular Biology* **50**, 359–379 (2015).
39. Stockhammer, O. W., Zakrzewska, A., Hegedûs, Z., Spaink, H. P. & Meijer, A. H. Transcriptome profiling and functional analyses of the zebrafish embryonic innate immune response to *Salmonella* infection. *Journal of Immunology* **182**, 5641–5653 (2009).
40. Yang, S., Marín-Juez, R., Meijer, A. H. & Spaink, H. P. Common and specific downstream signaling targets controlled by Tlr2 and Tlr5 innate immune signaling in zebrafish. *BMC Genomics* **16**, (2015).
41. Lee, B. L. *et al.* UNC93B1 mediates differential trafficking of endosomal TLRs. *eLife* **2**, e00291 (2013).

42. Pelka, K., Shibata, T., Miyake, K. & Latz, E. Nucleic acid-sensing TLRs and autoimmunity: novel insights from structural and cell biology. *Immunology Reviews* **269**, 60–75 (2016).
43. Andersen-Nissen, E., Smith, K. D., Bonneau, R., Strong, R. K. & Aderem, A. A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. *Journal of Experimental Medicine* **204**, 393–403 (2007).
44. Forstnerič, V., Ivičak-Kocjan, K., Ljubetič, A., Jerala, R. & Benčina, M. Distinctive recognition of flagellin by human and mouse Toll-like receptor 5. *PLOS ONE* **11**, e0158894 (2016).
45. Latz, E. *et al.* Ligand-induced conformational changes allosterically activate Toll-like receptor 9. *Nature Immunology* **8**, 772–779 (2007).
46. Xu, Y. *et al.* Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature* **408**, 111–115 (2000).
47. Nyman, T. *et al.* The crystal structure of the human Toll-like receptor 10 cytoplasmic domain reveals a putative signaling dimer. *Journal of Biological Chemistry* **283**, 11861–11865 (2008).
48. Endres, N. F. *et al.* Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* **152**, 543–556 (2013).
49. Matthews, E. E. *et al.* Thrombopoietin receptor activation: transmembrane helix dimerization, rotation, and allosteric modulation. *FASEB Journal* **25**, 2234–2244 (2011).
50. Braasch, I. *et al.* The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. *Nature Genetics* **48**, 427–437 (2016).
51. Force, A. *et al.* Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531–1545 (1999).
52. Innan, H. & Kondrashov, F. The evolution of gene duplications: classifying and distinguishing between models. *Nature Reviews Genetics* **11**, 97–108 (2010).
53. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**, 539 (2011).
54. Matsushima, N. *et al.* Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors. *BMC Genomics* **8**, 124 (2007).
55. Offord, V., Coffey, T. J. & Werling, D. LRRfinder: A web application for the identification of leucine-rich repeats and an integrative Toll-like receptor database. *Developmental & Comparative Immunology* **34**, 1035–1041 (2010).
56. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology* **305**, 567–580 (2001).
57. Montgomerie, S., Sundararaj, S., Gallin, W. J. & Wishart, D. S. Improving the accuracy of protein secondary structure prediction using structural alignment. *BMC Bioinformatics* **7**, 301 (2006).
58. Biasini, M. *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research* **42**, W252–W258 (2014).
59. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605–1612 (2004).
60. Rodriguez, M. S., Thompson, J., Hay, R. T. & Dargemont, C. Nuclear retention of IkappaBalpha protects it from signal-induced degradation and inhibits nuclear factor kappaB transcriptional activation. *Journal of Biological Chemistry* **274**, 9108–15 (1999).

Supplement

Figure S1. Alignment of zebrafish drTLR5a and drTLR5b. The amino acid sequences were aligned using the Clustal Omega server with default settings. Asterisks (*) indicate identical residues, double dots (:) highly similar residues, single dots (.) somewhat similar residues and bars (-) indicate gaps to complete the sequence alignment. Signal peptide, N-terminal leucine rich repeat (NTLRR), 22 consecutive LRR's and C-terminal LRR (CTLRR) are shaded in grey. Grey arrows indicate start and end of specified regions; N-ECD: N-terminal part of ectodomain domain (ECD), C-ECD: C-terminal part of ECD, EJM: extracellular juxta membrane, TM: transmembrane, CJM: cytoplasmic juxta membrane, TIR: Toll-IL-1 receptor and Tail region.

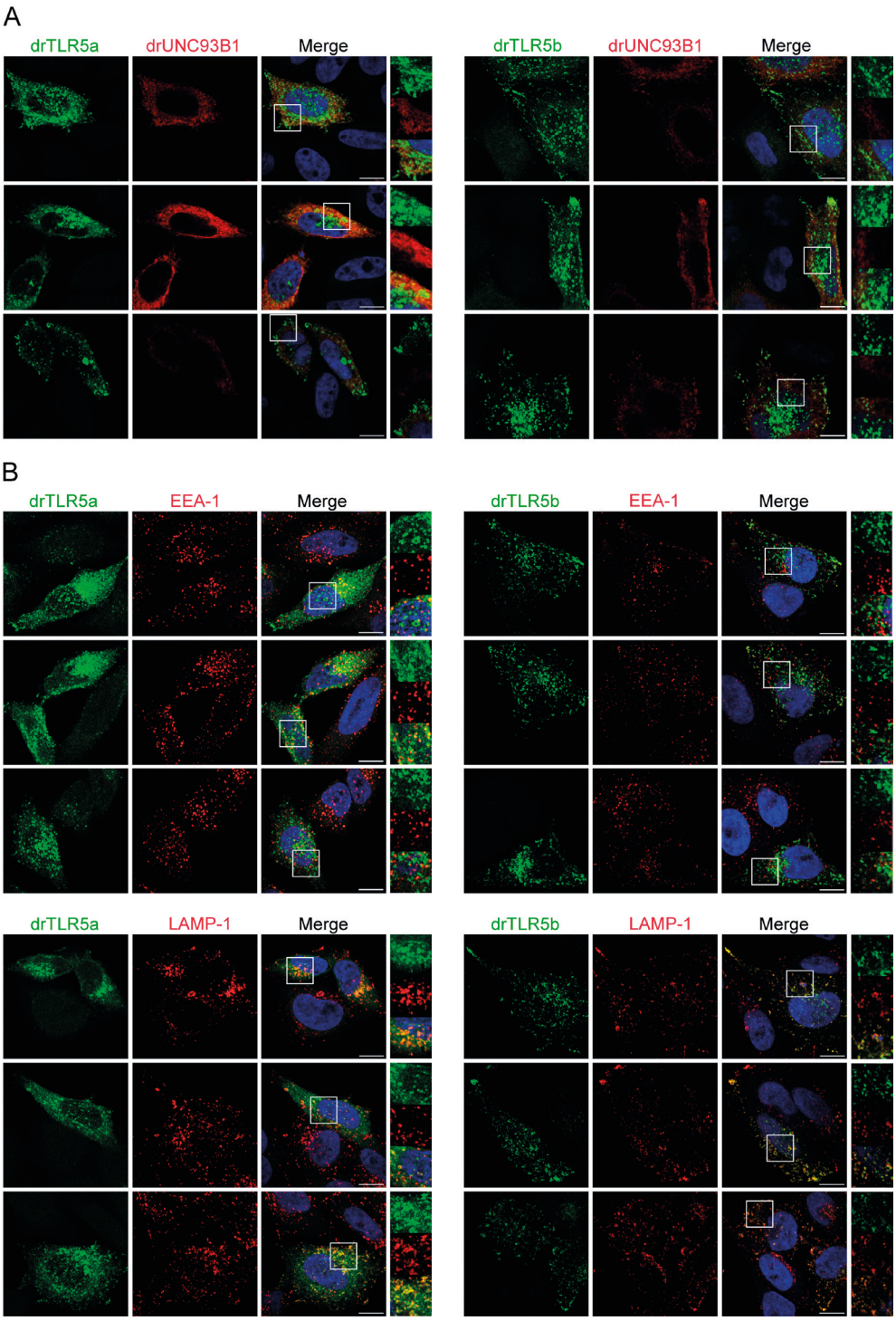

A
drTLR5a
drUNC93B1
Merge
drTLR5b
drUNC93B1
Merge
B
drTLR5a
EEA-1
Merge
drTLR5b
EEA-1
Merge
drTLR5a
LAMP-1
Merge
drTLR5b
LAMP-1
Merge

Figure S2 on previous page. Vesicular localization of drTLR5a and drTLR5b in the presence of drUNC93B1. Confocal microscopy on HeLa-57A cells expressing; **(A)** drTLR5a-HA or drTLR5b-HA and drUNC93B1-FLAG or **(B)** drTLR5a-HA or drTLR5b-HA and untagged drUNC93B1 co-stained for EEA-1 or LAMP-1. Merge images show nuclei stained with DAPI (blue). Images are selected from three independent experiments and three representative images are shown for each transfected group. Scale bar in merge images is 10 µm.

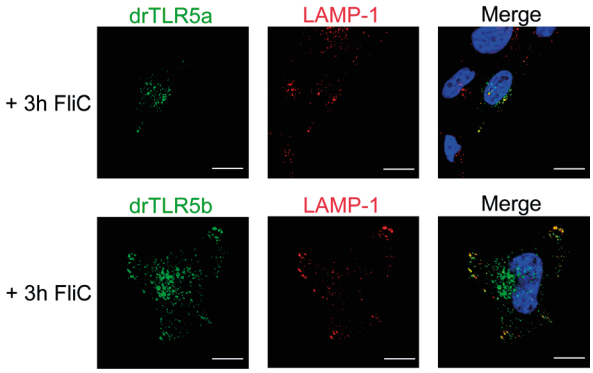


Figure S3. Localization of zebrafish drTLR5a and drTLR5b during flagellin stimulation. HeLa-57A cells expressing drTLR5a-HA or drTLR5b-HA and untagged drUNC93B1 were stimulated for three hours with 1 µg/ml $FliC^{SE}$ and co-stained for LAMP-1. Merge images show nucleus stained with DAPI (blue). Images are representative of three independent experiments. Scale bar is 10 µm.

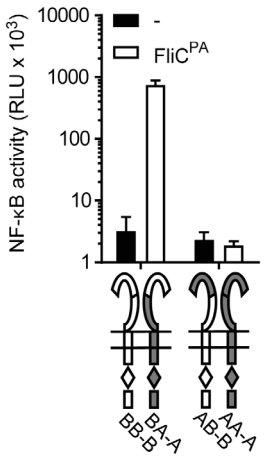


Figure S4. Stimulation of homodimeric N-ECD of drTLR5a and drTLR5b with *P. aeruginosa* flagellin ($FliC^{PA}$). HeLa-57A cells transfected with drUNC93B1 and the indicated receptor combinations were stimulated (5h) with vehicle (-) or 1 µg/ml *P. aeruginosa* flagellin ($FliC^{PA}$). Data shows NF-κB activity represented by luciferase activity in relative light units (RLU). Values are the mean ± SEM of two independent experiments performed in duplicate.

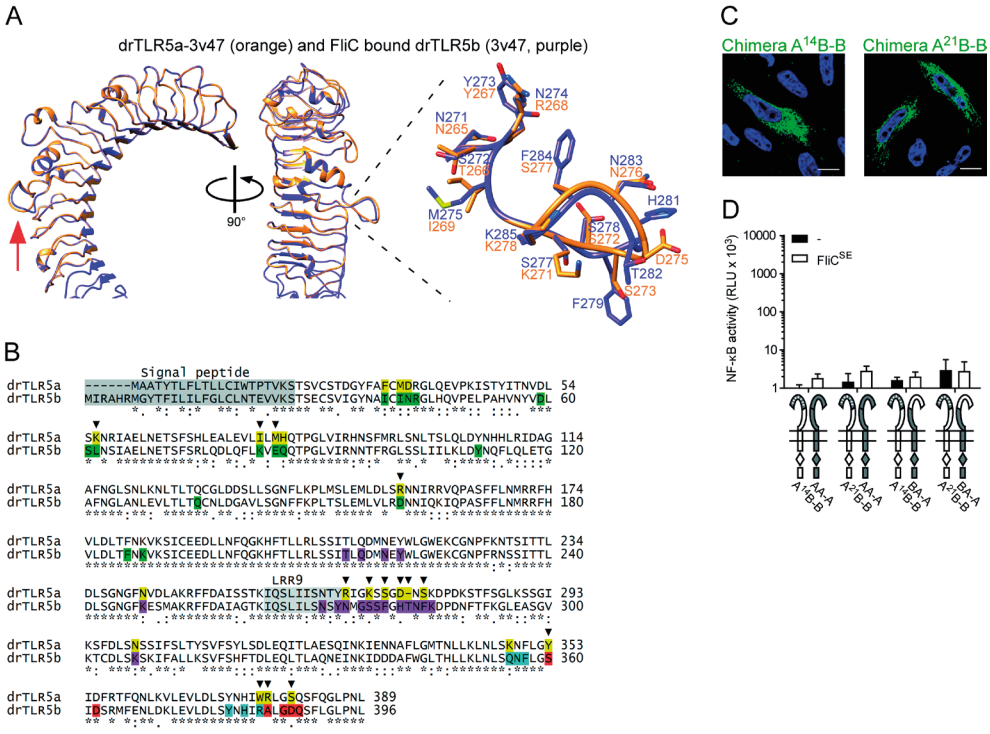


Figure S5. (A) Superposition of FliC bound drTLR5b (PDB ID: 3v47, purple) and a model of drTLR5a based on 3v47 (drTLR5a-3v47, orange). Red arrow indicates the flagellin binding hotspot that forms a loop between LRR9 and LRR10. **(B)** Sequence alignment of the drTLR5a and drTLR5b N-ECD. The amino acid sequences were aligned using the Clustal Omega server with default settings. Asterisks (*) indicate identical residues, double dots (:) highly similar residues, single dots (.) somewhat similar residues and bars (-) indicate gaps to complete the sequence alignment. The 45 residues in the drTLR5b crystal model responsible for flagellin binding and dimerization are color-coded according to Yoon *et al.*: green; primary interface-A, purple; primary interface-B, cyan; dimerization interface-β, red; dimerization interface-α. Green and purple residues interact with flagellin, red residues interact with the second flagellin molecule in the 2:2 stoichiometric homodimeric TLR5-flagellin crystal model, cyan residues interact with the second drTLR5b molecule, the bold phenylalanine (F) at position 284 interacts with flagellin and with the second drTLR5b molecule, see also Yoon *et al.*[19]. Yellow-colored amino acids in the drTLR5a sequence indicate the 21 different binding or dimerization residues that were replaced in chimera AB-B to produce chimera A^{21}B-B. The arrowheads (▼) indicate 14 of these 21 most deviating residues that were replaced to produce chimera A^{14}B-B. **(C)** HeLa-57A cells were transfected with FLAG tagged chimeras A^{14}B-B or A^{21}B-B and stained with α-FLAG (green) and DAPI (blue) for nuclear visualization. Scale bar is 10 µm. **(D)** HeLa-57A cells transfected with drUNC93B1 and the indicated receptor combinations were stimulated (5h) with vehicle (-) or 1 µg/ml *Salmonella enterica* serovar Enteritidis flagellin ($FliC^{SE}$). Data shows NF-κB activity represented by luciferase activity in relative light units (RLU). Values are the mean ± SEM of three independent experiments performed in duplicate.

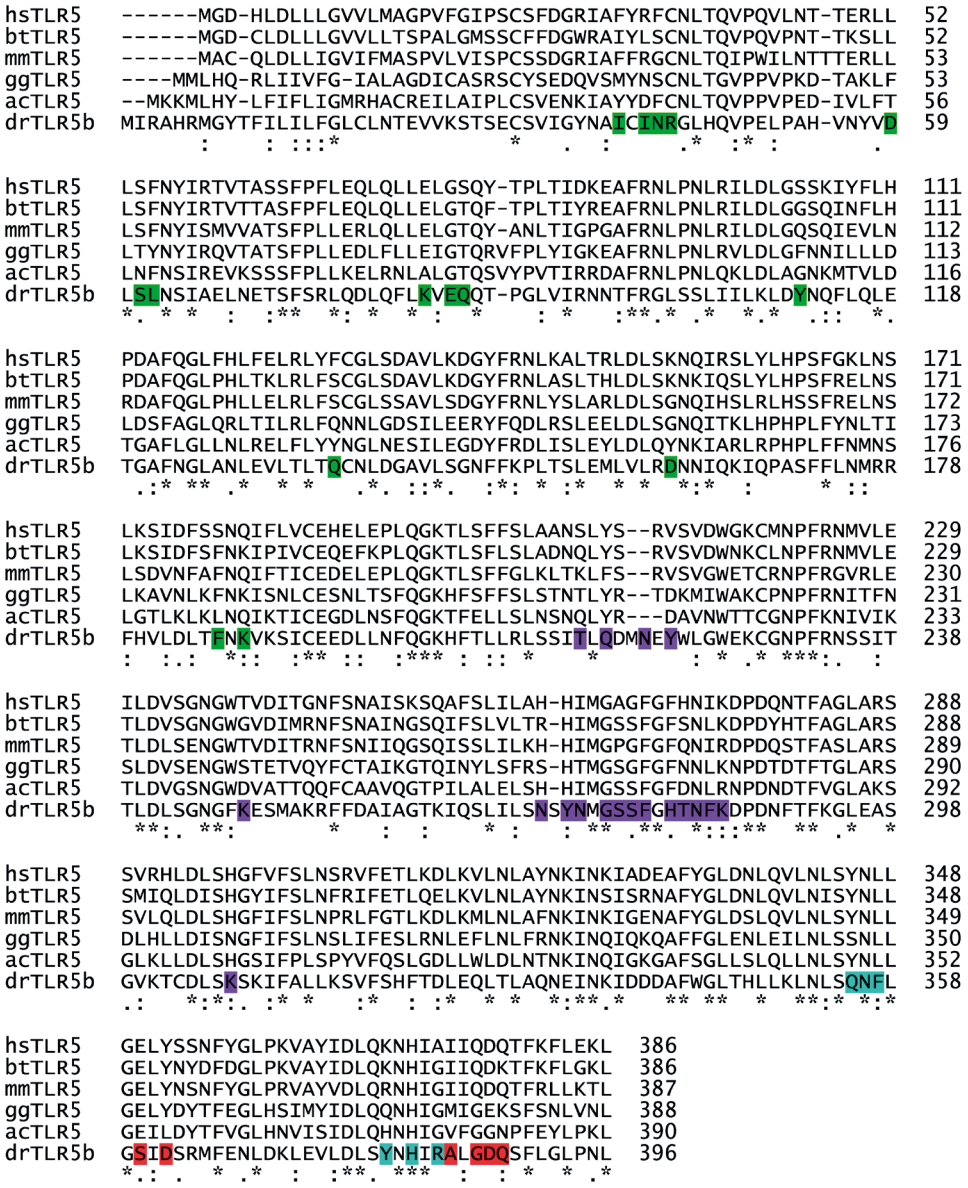


Figure S6. Sequence alignment of the TLR5 N-ECD of different vertebrates. Alignment of the TLR5 N-ECD of *Homo sapiens* (hs, human, NCBI reference sequence: AAI09119.1), *Bos taurus* (bt, bovine, ABC68311.1), *mus musculus* (mm, mouse, AAI25248.1), *Gallus gallus* (gg, chicken, ACR26275.1), *Anolis carolinensis* (ac, green anole lizard, ALT10445.1) and zebrafish. Alignment was constructed with the Clustal Omega server with default settings. Asterisks (*) indicate identical residues, double dots (:) highly similar residues, single dots (.) somewhat similar residues and bars (-) indicate gaps to complete the sequence alignment. Residues in the drTLR5b crystal model responsible for flagellin binding and dimerization are color-coded according to Yoon *et al.*[19]: green; primary interface-A, purple; primary interface-B, cyan; dimerization interface-β, red; dimerization interface-α. Green and purple residues interact with flagellin, red residues interact with the second flagellin molecule in the 2:2 stoichiometric homodimeric TLR5-flagellin crystal model, cyan residues interact with the second drTLR5b molecule, see also Yoon *et al.*

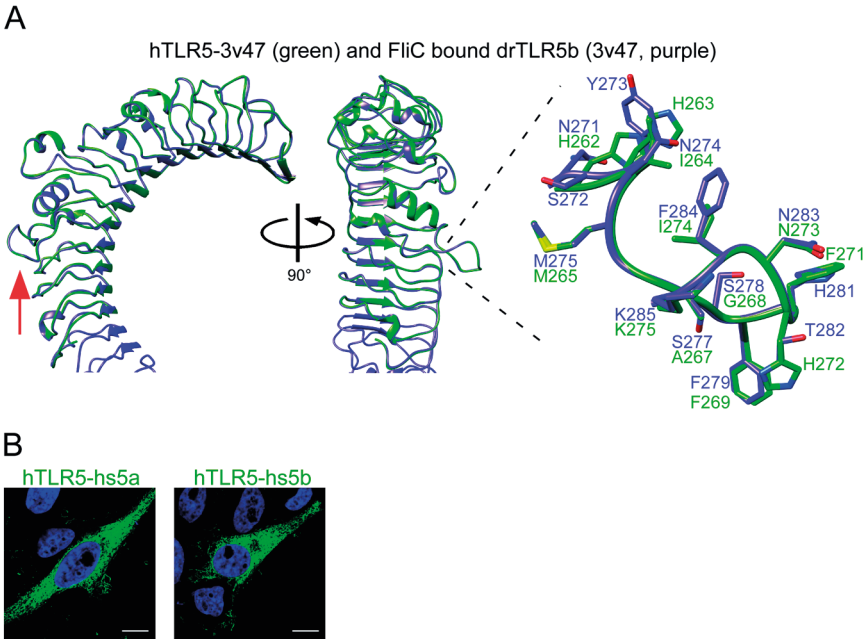


Figure S7. (A) Structural modelling of human TLR5 onto drTLR5b. Superposition of FliC bound drTLR5b (PDB ID: 3v47, purple) and a model of hTLR5 based on 3v47 (hTLR5-3v47, green). Red arrow indicates the flagellin binding hotspot that forms a loop between LRR9 and LRR10. **(B)** Expression of hTLR5 chimeric constructs. HeLa-57A cells were transfected with FLAG tagged hTLR5-hs5a or hTLR5-hs5b and stained with α-FLAG (green) and DAPI (blue) for nuclear visualization. Images are representative of two independent experiments. Scale bar is 10 μm.

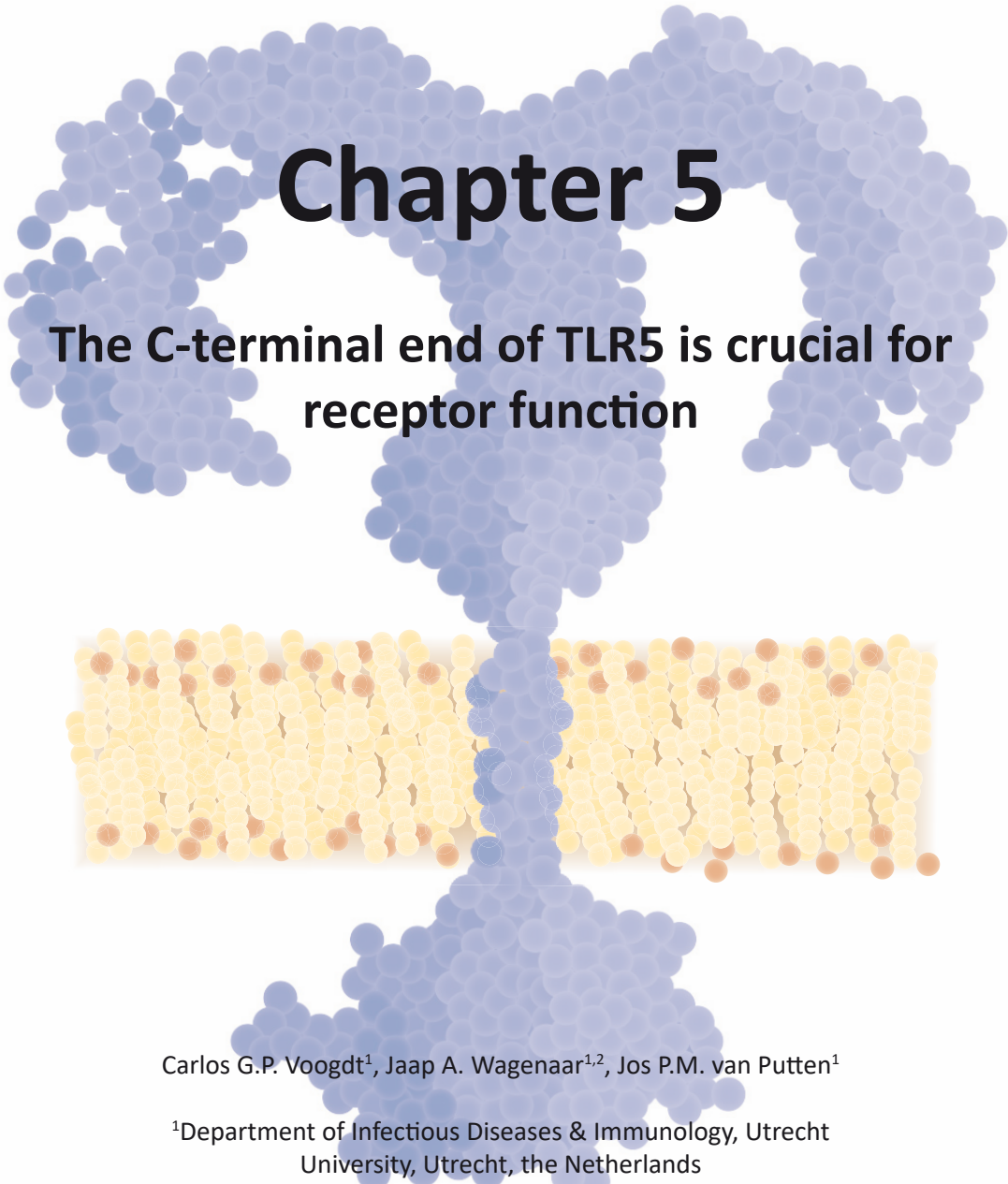
Table S1. Primers used in this study

Product	Primer	Sequence (5'-3')
BamHI-Kozak-drTLR5b-**NotI**	CV039 F	CC**GGATCC**GCCACCATGATCCGTGCTCACAGAATGG
	CV054 R	CC**GCGGCCGC**TTACTGCTGTGTTGGTGTTGAAATTG
BamHI-Kozak-drTLR5a-**NotI**	CV095 F	CC**GGATCC**GCCACCATGGCAGCTACATACACTTTATTTC
	CV096 R	CC**GCGGCCGC**TAACTGCAGTGTCTGCTTGAAC
KpnI-Kozak-drUNC93B1-**NotI**	CV159 F	CC**GGTACC**GCCACCATGGCAGCACTGATCGC
	CV160 R	CC**GCGGCCGC**AGTGTTGGACGTAGTCATCTC
KpnI-Kozak-drUNC93B1-**NotI**-stop codon	CV159 F	CC**GGTACC**GCCACCATGGCAGCACTGATCGC
	CV180 R	CC**GCGGCCGC**TTAGTGTTGGACGTAGTCATCTC
BamHI-Kozak-hTLR5-**NotI**	CV257 F	CC**GGATCC**GCCACCATGGGAGA
	CV259 R	CC**GCGGCCGC**TGGAGATGGTTGCTACAGTTTG

Restriction enzyme sequences are in bold, the Kozak sequence in forward primers is underlined

Chapter 5

The C-terminal end of TLR5 is crucial for receptor function



Carlos G.P. Voogdt[1], Jaap A. Wagenaar[1,2], Jos P.M. van Putten[1]

[1]Department of Infectious Diseases & Immunology, Utrecht University, Utrecht, the Netherlands
[2]Wageningen Bioveterinary Research, Lelystad, the Netherlands

Abstract

Toll-like receptors (TLRs) are evolutionarily conserved innate immune receptors that detect signals during infection or tissue damage. TLRs are transmembrane proteins that consist of distinct regions which are required for different actions. The ectodomain binds ligands and initiates receptor dimerization. The extracellular juxta membrane region of some TLRs interacts with the chaperone UNC93B1, whereas the cytoplasmic TIR domain is necessary for ligand induced downstream signaling. At its ultimate C-terminal end, TLRs contain a receptor specific tail region. The function of this tail region is still unknown. Here we report that deletion of the C-terminal amino acids of human TLR5 prevents its cell surface localization and abrogates flagellin mediated receptor activation. Targeted mutagenesis of the TLR5 tail revealed that the order of the amino acids, rather than evolutionarily conserved charged residues, determines receptor trafficking and function. Substitution of the amino acids in the tail of surface localized human TLR5 by the corresponding residues of lysosomal localized zebrafish TLR5b still enabled this chimeric receptor to localize at the cell surface. The tail regions of both human and zebrafish TLR5b are predicted to be phosphorylated at a threonine residue and this is dependent on the order of the tail sequence. These findings show for the first time that the tail region of TLR5 is essential for receptor trafficking in a possible threonine phosphorylation dependent manner.

Introduction

Toll-like receptors (TLRs) form a family of proteins that are highly conserved throughout animal evolution. The receptors play a central role in the immune system by detecting microbe or damage associated molecular patterns. Upon detection of such patterns, TLRs initiate signaling to activate transcription factors that induce the expression of defense mechanisms. TLRs are type I membrane proteins that have a highly conserved structure with distinct regions. The TLR ectodomain (ECD) is involved in ligand-binding, receptor dimerization and binding of co-receptors[1–3]. The single transmembrane (TM) region is necessary for embedment in endolysosomal or plasma membranes, receptor dimerization and signal transduction[4,5]. The cytoplasmic part of the protein contains a Toll/Interleukin-1 (TIR) domain which is highly conserved across TLRs within and between species. Ligand-induced receptor dimerization connects two TIR domains which activates a signaling cascade that results in activation of transcription factors like Nuclear factor κB (NF-κB)[6,7]. Insights in the contribution of the distinct TLR regions to receptor functioning are indispensible to understand TLR biology and develop strategies to therapeutically target TLRs.

TLR function critically depends on the appropriate subcellular localization of the receptor and this differs among TLR subfamily members. In general, TLR1, 2, 4, and 6 function from the plasma membrane, while TLR3, 7, 8 and 9 function from endolysosomes[6]. The endolysosomal receptors require the TLR chaperone UNC93B1 for intracellular trafficking and interact with UNC93B1 via acidic amino acids present in the receptor's extracellular juxta membrane (EJM) region[8,9]. The EJM of TLR5, a receptor for bacterial flagellin, also contains acidic residues that interact with UNC93B1 but in contrast to TLR3, 7, 8 and 9, human TLR5 is localized at the plasma membrane[10]. Intriguingly, we previously described that a TLR5 ortholog of the zebrafish contains acidic residues in the EJM but traffics to lysosomes when co-expressed with zebrafish UNC93B1 in human cells[11]. The exact role of UNC93B1 in the different receptor localization is still unclear[12,13] and the mentioned studies suggest that regions other than the EJM may be involved in receptor localization and hence function.

The TLR receptor ends at its C-terminus beyond the TIR domain with a TLR specific sequence termed the tail region. The amino acids of the tail region are somewhat conserved among TLR orthologs in terms of length and composition. Yet, the contribution of these amino acids to receptor trafficking and function is unknown. Here we investigated the potential significance of the TLR tail sequence for the proper localization and function of human TLR5.

Results

Comparison of the C-terminal tail sequences of human TLRs

The TIR domain of TLRs adopts a globular fold that consists of five alternating β-sheets and α-helices. The TIR domain ends with an α-helix that spans approximately seven amino acids following a highly conserved phenylalanine[14]. We defined the C-terminal tail region of TLRs as all amino acids positioned at the carboxyl end following the last α-helix of the TIR domain (Fig. 1). Sequence comparison of a part of the TIR domain and the tail region of all human TLRs showed high heterogeneity in tail length (Fig. 1). The composition of the TLR tails was also highly diverse and showed no evolutionary conservation of particular residues across all TLRs. All TLR tails did however contain one or multiple charged amino acids and especially TLR5 contained a remarkable sequence of consecutive charged residues dominated by lysines, possibly pointing towards a charge-based role for the tail reminiscent of the function of the charged residues in the receptor's EJM (Fig. 1).

The C-terminal tail of TLR5 is required for receptor trafficking and functioning

To learn more about the potential physiological significance of the TLR tail region, we first assessed whether the TLR5 tail plays a role in receptor trafficking. Hereto, we cloned full length wild type human TLR5 (hsTLR5-WT) and a TLR5 lacking the tail region (hsTLR5-Δtail) in an expression vector in fusion with a C-terminal HA tag sequence. The plasmids were transfected together with plasmids encoding human UNC93B1 and a plasma membrane localized red fluorescent protein (pmRFP) into HEK293 cells which had been shown previously to express TLR5 at the plasma membrane[10]. Confocal microscopic examination of transfected HEK293 cells confirmed that hsTLR5-WT co-localized with RFP at the plasma membrane (Fig. 2A). Strikingly, removal of the tail affected receptor trafficking as the hsTLR5-Δtail was no longer observed at the plasma membrane and instead resided intracellularly (Fig. 2A). To determine whether the tail deletion also affected TLR5 function, HeLa-57A cells were transfected with the different TLR5 constructs. HEK293 cells were not suitable to test the function of hsTLR5-WT and hsTLR5-Δtail as these cells express endogenous TLR5 and respond to flagellin (data not shown). HeLa-57A cells lack endogenous TLR5 and stably carry a NF-κB-luciferase reporter system to monitor TLR5 activation. Stimulation of HeLa-57A cells expressing hsTLR5-WT with flagellin of *Salmonella enterica* serovar Enteritidis (FliC) resulted in strong activation of NF-κB (Fig. 2B). In contrast, the addition of FliC to HeLa-57A cells expressing hsTLR5-Δtail failed to induce NF-κB activity (Fig. 2B). Detection of the HA tag via confocal microscopy ensured that both receptors were expressed in

HeLa-57A cells (Fig. 2C). These results are in line with the observed mislocalization of hsTLR5-Δtail in HEK293 cells and suggest that the tail region is required for TLR5 function.

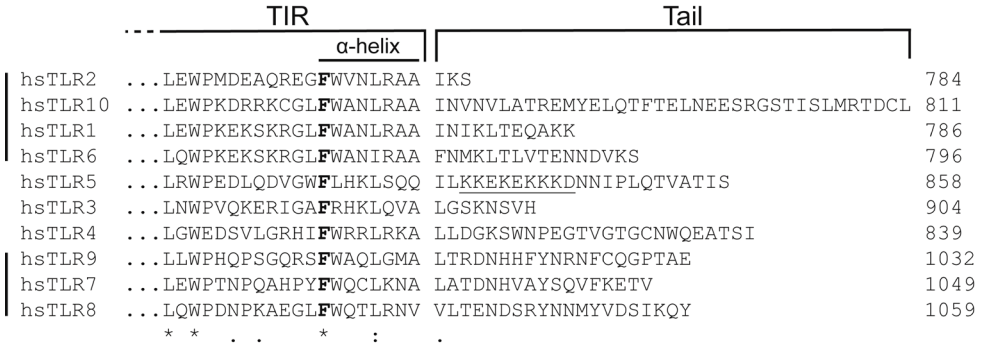


Figure 1. Sequence comparison of human (*Homo sapiens*, hs) TLR tails. Sequences comprising the last part of the TIR domain and the tail of all human TLRs were aligned using Clustal Omega. Asterisks (*) indicate identical residues, double dot (:) indicates residue with strongly similar properties (e.g. FYW), single dots (.) indicate residues with weakly similar properties (e.g. CSA). TLR2, 10, 1 and 6 and TLR9, 7 and 8 form subfamilies indicated by the vertical bars on the left. The conserved phenylalanine at the start of the last α-helix is shown in bold. A short sequence of consecutive charged residues in TLR5 is underlined.

To exclude that the inability to function of the hsTLR5-Δtail was caused by the HA tag sequence directly following the TIR domain at the receptor C-terminus, we constructed full length and tail truncated receptors in which the HA tag was placed between the signal peptide and the first leucine rich repeat at the receptor N-terminus, leaving the TIR domain untagged. Flagellin stimulation of HeLa-57A cells transfected with the N-terminally tagged full length hsTLR5-WT resulted in NF-κB activation, although the response was less potent than noted for the C-terminally tagged hsTLR5-WT (Fig. 2D). Importantly, cells transfected with N-terminally tagged hsTLR5-Δtail still failed to respond to FliC (Fig. 2D). These findings corroborate that the tail of TLR5, comprising the last 23 amino acids of the protein, plays an important role in receptor trafficking and function.

The order of the TLR5 tail sequence is important

Comparison of the tail of human TLR5 with those of TLR5 from other vertebrates revealed that the occurrence of negative and positive charged residues, mainly lysines, is a conserved feature among TLR5 of diverse species (Fig. 3A). As charged residues may influence protein interactions, we investigated whether the charge or specifically the lysines in the tail of TLR5 were involved in receptor trafficking and function. Hereto, we constructed a TLR5 in which all charged residues in the tail

were substituted with neutral alanines or glycines (hsTLR5-Neu) and a TLR5 in which the lysines in the tail were replaced with arginines (hsTLR5-KtoR)(Fig. 3B). Both mutated receptors were tagged C-terminally with an HA tag. Confocal microscopy showed that both hsTLR5-Neu and hsTLR5-KtoR still localized at the plasma membrane of HEK293 cells (Fig. 3C). In addition, expression and FliC stimulation of hsTLR5-Neu yielded similar NF-κB activity as observed for hsTLR5-WT. FliC stimulation of hsTLR5-KtoR resulted in even stronger NF-κB activation than the hsTLR5-WT (Fig. 3D). These results imply that the amino acid charge of the TLR5 tail does not critically contribute to the TLR5 cellular localization and flagellin mediated ability to signal.

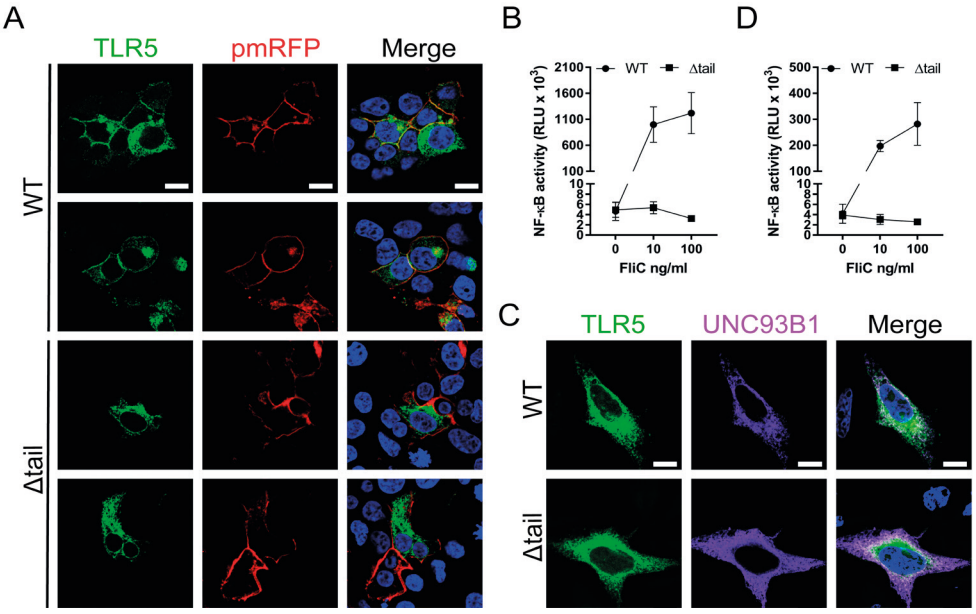


Figure 2. Deletion of the TLR5 tail prevents receptor plasma membrane localization and receptor mediated NF-κB activation. **(A)** Confocal microscopy images of HEK293 cells transfected with plasmids encoding untagged UNC93B1, a plasma membrane localizing red fluorescent protein (pmRFP) and C-terminally HA-tagged hsTLR5-WT or hsTLR5-Δtail. Merge images show nuclei in blue. Scale bar is 10 µm and applies to all images. Two representative images per transfected group, taken from two independent experiments, are shown. **(B & D)** HeLa-57A cells were transfected with UNC93B1-FLAG and C-terminally HA-tagged TLR5 **(B)** or N-terminally HA-tagged TLR5 **(D)** and stimulated with flagellin (FliC). NF-κB activity is represented by luciferase activity measured in relative light units (RLU). Values are mean ± SEM of three **(B)** or two **(D)** independent experiments. **(C)** Confocal microscopy images of HeLa-57A cells expressing C-terminally HA-tagged hsTLR5-WT or hsTLR5-Δtail and FLAG-tagged UNC93B1. Merge images show nuclei in blue. Scale bar is 10 µm and applies to all images. Images are representatives of two independent experiments.

In further search for the tail feature that influences receptor trafficking, we investigated whether perhaps certain amino acid motifs in the tail were involved. To test this, we randomly scrambled the 23 amino acids of the tail sequence yielding hsTLR5-scram (Fig. 3B). Strikingly, HEK293 cells expressing hsTLR5-scram failed to show localization of the receptor at the plasma membrane (Fig. 3C). Scrambling of the tail sequence also completely blocked receptor activation as hsTLR5-scram transfected HeLa-57A cells no longer responded to FliC stimulation (Fig. 3D). These findings suggest that TLR5 localization at the plasma membrane and its responsiveness to flagellin relies on the order of the amino acids in the receptor tail.

Human TLR5 with a zebrafish tail is positioned at the cell surface

To learn more about a putative amino acid motif in the C-terminal tail that may be involved in TLR5 trafficking and function, we made use of the TLR5 ortholog of the zebrafish which localizes to intracellular vesicles rather than the plasma membrane[11]. We constructed a chimeric human TLR5 in which the 23 C-terminal amino acids were replaced by the tail sequence of zebrafish TLR5b (hsTLR5-drtail)(Fig. 3A). Control HEK293 cells transfected with HA-tagged wild type zebrafish TLR5b (drTLR5b) and zebrafish UNC93B1 showed receptor localization at lysosome associated membrane protein 1 (LAMP-1) positive vesicles, consistent with previous findings in HeLa-57A cells (Fig 4A)[11]. Interestingly, expression of the chimeric hsTLR5-drtail together with human UNC93B1 in HEK293 cells resulted in receptor localization at the plasma membrane, like hsTLR5-WT (Fig. 4B). Furthermore, HeLa-57A cells expressing hsTLR5-drtail showed activation of NF-κB in response to FliC exposure (Fig. 4C). These results suggest that despite considerable differences in composition, the zebrafish TLR5b tail may contain a similar, order-dependent, feature as the human TLR5 tail that is required for initiation of correct receptor trafficking and function.

A

```
                    TIR                    Tail
                     α-helix
hsTLR5   ...EDLQDVGWFLHKLSQQ ILKKEKEKKKDNNIPLQTVATIS          858
mmTLR5   ...EDLQDVGWFLDKLSGC ILKEEKGKKRSSSIQLRTIATIS          859
btTLR5   ...EDLQDVDWFLNKLSQW LPKKKKERKKDMWIPLQNVPTIS          858
ggTLR5   ...EDYQDIGWFLDNLSSQ ILKEKKVQRNVSGIELQTIATVSH         861
acTLR5   ...EEDQDVEWFLNALSHQ ILKEKRTQKKAQKKKVPKKTGTLELKVVTIS 871
drTLR5b  ...DDGQDLEWFYDQLLHK IRKDIKINQTTKETKREEANFNTNTAV      881
            :: **: ** . *    : *. : .:
```

B

```
WT      ...ILKKEKEKKKDNNIPLQTVATIS
Neu     ...ILAAGGAGAGGNNIPLQTVATIS
KtoR    ...ILRRERERRRDNNIPLQTVATIS
Scram   ...NKKSKQVIPKIKLDIETLNTEAK
```

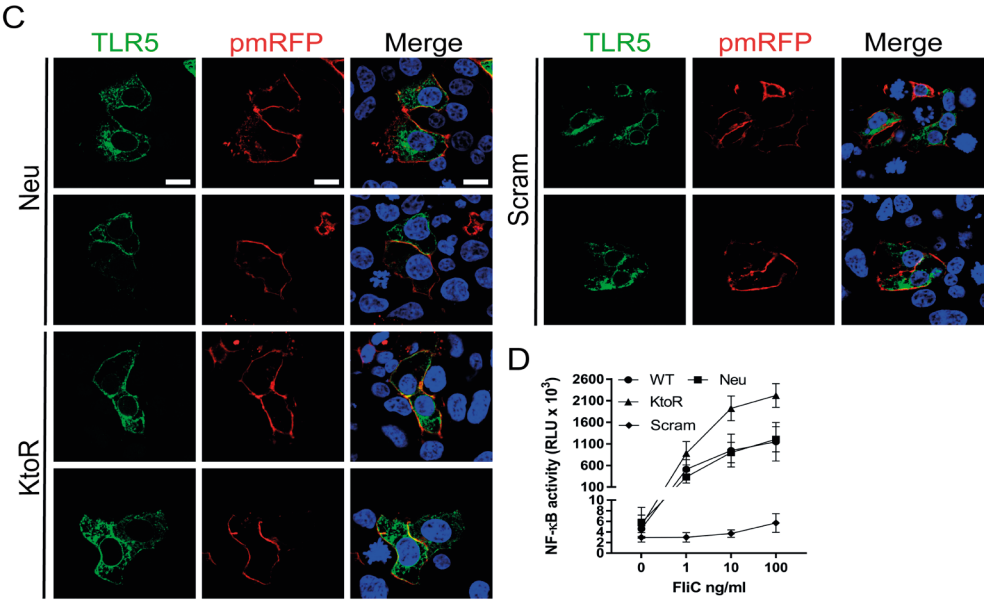


Figure 3. The sequence order of the TLR5 tail is important for receptor localization and activity. **(A)** Sequence alignment of the tail of TLR5 from human (*Homo sapiens*, hs), mouse (*Mus musculus*, mm), bovine (*Bos taurus*, bt), chicken (*Gallus gallus*, gg), Anolis lizard (*Anolis carolinensis*, ac) and zebrafish (*Danio rerio*, dr). Sequences were aligned using Clustal Omega. Asterisks (*) indicate identical residues, double dots (:) indicate residues with strongly similar properties (e.g. FYW), single dots (.) indicate residues with weakly similar properties (e.g. CSA). **(B)** Sequences of the altered hsTLR5 tails: wild type (WT), neutralized (Neu), lysines replaced with arginines (KtoR) and randomly scrambled (Scram). **(C)** HEK293 cells transfected with untagged UNC93B1, plasma membrane localizing red fluorescent protein (pmRFP) and C-terminally HA-tagged hsTLR5-Neu, hsTLR5-KtoR or hsTLR5-Scram. Merge images show nuclei in blue. Scale bar is 10 µm and applies to all images. Two representative images per transfected group, taken from two independent experiments, are shown. **(D)** HeLa-57A cells transfected with human UNC93B1-FLAG and hsTLR5-WT, hsTLR5-Neu, hsTLR5-KtoR or hsTLR5-Scram were stimulated with flagellin (FliC). NF-κB activity is represented by luciferase activity measured in relative light units (RLU). Values are mean ± SEM of three independent experiments.

Prediction of putative threonine phosphorylation in the TLR5 tail

During the search for a common motif in the tails of human TLR5 and its orthologs including zebrafish TLR5b, we noticed a shared conservation of threonine residues (Fig. 3A). Threonine can be phosphorylated but the accessibility of threonine to kinases depends on the tertiary structure adopted by the surrounding amino acid sequence. Bioinformatics analysis of the human and zebrafish TLR5 tails predicted phosphorylation of T853 in human TLR5 and T864 and T868 in zebrafish TLR5b. Interestingly, the confidence of predicted phosphorylation in the scrambled TLR5 tail was much lower, suggesting that rearrangement of the residues in the scrambled tail could negatively affect threonine phosphorylation (Table 1). The substitution of charged amino acids for neutral or arginines for lysines was predicted not to influence the phosphorylation event, which is consistent with our experimental results. Overall, and bearing the limitations of the bioinformatics analysis in mind, our results point towards threonine phosphorylation in the TLR5 tail as a possible requirement for appropriate tail-mediated receptor localization and function.

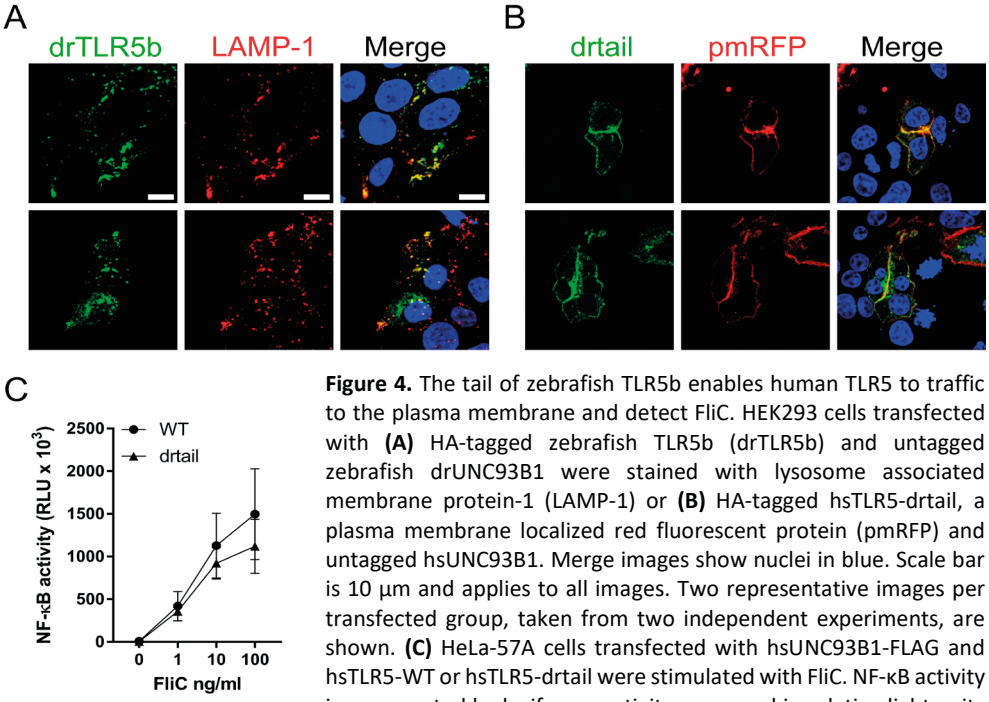


Figure 4. The tail of zebrafish TLR5b enables human TLR5 to traffic to the plasma membrane and detect FliC. HEK293 cells transfected with **(A)** HA-tagged zebrafish TLR5b (drTLR5b) and untagged zebrafish drUNC93B1 were stained with lysosome associated membrane protein-1 (LAMP-1) or **(B)** HA-tagged hsTLR5-drtail, a plasma membrane localized red fluorescent protein (pmRFP) and untagged hsUNC93B1. Merge images show nuclei in blue. Scale bar is 10 µm and applies to all images. Two representative images per transfected group, taken from two independent experiments, are shown. **(C)** HeLa-57A cells transfected with hsUNC93B1-FLAG and hsTLR5-WT or hsTLR5-drtail were stimulated with FliC. NF-κB activity is represented by luciferase activity measured in relative light units (RLU). Values are mean ± SEM of three independent experiments.

Table 1. Predicted phosphorylation sites in TLR5 tail by NetPhos 3.1

TLR	Residue	Position	Score (0-1)	Kinase
hsTLR5-WT	T	853	0,724	PKC
drTLR5b	T	864	0,898	PKC
	T	868	0,968	unsp*
hsTLR5-Scram	S	839	0,573	PKG
	T	852	0,504	CKII

* unsp: unspecific

Discussion

TLRs have a highly conserved overall structure with distinct regions being involved in different aspects of receptor processing and function. However, the role of the C-terminal amino acids forming the tail following the intracellular TIR domain has thus far remained largely obscured. In the present study, we provide evidence that the tail region of TLR5 is essential for receptor trafficking and hence ligand induced signaling. As the tail regions of TLR5 orthologs are rather conserved, our findings may be instrumental in further unraveling the function of the TLR tail sequence in TLR biology.

A strong first indication of the relevance of the TLR tail region was the mislocalization and inability to respond to flagellin of a truncated human TLR5 lacking the tail region. As microscopy indicated that the truncated receptor was expressed, the results at first sight pointed towards a role of the tail region in receptor trafficking. Sequence comparison of human TLR5 with other TLR5 orthologs showed that the presence of multiple, mainly positively, charged amino acids is a conserved feature of the TLR5 tail. Evolutionary conservation of the charged residues is suggestive of a function which may e.g. involve the regulation of charge-based interactions with cytosolic or other membrane proteins[15,16]. Yet, a constructed human TLR5 carrying neutral alanines and glycines instead of charged residues or carrying arginines instead of lysines still induced FliC mediated NF-κB activation in HeLa-57A cells and these receptors were still transported to the plasma membrane in HEK293 cells. While the charged residues in the tail region may be involved in other receptor actions, these findings indicate that the charge of the TLR5 tail is not required for proper TLR5 trafficking and flagellin-mediated signaling to activate NF-κB.

Since the charged residues in the tail were not essential to TLR5, we probed the possibility that other membrane sorting signals might be present in the TLR5 tail. Intracellular transport of membrane proteins requires trafficking chaperones such as

the adapter protein (AP) complexes. These typically recognize cargo proteins by di-leucine-based ([DE]xxxL[LI]) or tyrosine-based (YxxØ, in which Ø is a hydrophobic residue) sorting signals in the cargo's cytoplasmic domain[17], and human TLR7 was found to interact with AP-4 through a tyrosine motif at the start of the TIR domain[12]. However, the tail region of any of the TLR5 orthologs studied here does not contain a di-leucine or tyrosine based sorting motif (Fig. 3A). Furthermore, whereas wild type zebrafish drTLR5b is transported to intracellular LAMP-1 positive vesicles[11], our chimeric human TLR5 carrying the tail of drTLR5b still localized at the plasma membrane. On the other hand, random rearrangement of the human TLR5 tail sequence (yielding hsTLR5-scram) rendered the receptor incapable of localizing at the plasma membrane, just as we observed after complete deletion of the tail sequence. Combined, these findings suggest that the TLR5 tail lacks a putative non-canonical sorting motif that directly dictates transport to the plasma membrane but does contain an evolutionarily conserved sequence that is required for initiating appropriate receptor trafficking.

By inspection of the tail of different TLR5 orthologs we noticed that several threonine residues are evolutionarily conserved. Further bioinformatics analysis predicted that phosphorylation of threonine in the tail of wild type human TLR5 is more likely than phosphorylation of the rearranged threonines in the receptor carrying the scrambled tail sequence, consistent with the defective localization and signaling by hsTLR5-scram. Although predictive, these analyses may suggest that threonine phosphorylation in the TLR5 tail could be necessary for initiating receptor trafficking. Further support for this hypothesis comes from the correct localization and ligand detection by the chimeric human TLR5 carrying the zebrafish TLR5b tail, which despite significant sequence deviation is also predicted to be phosphorylated at threonines with high confidence (Table 1). Phosphorylation of the TLR intracellular domain is well described but is mostly limited to tyrosine residues in the TIR domain. These tyrosine phosphorylation events are involved in signal transduction[6,18,19]. One study reported phosphorylation of a tyrosine residue in the tail of TLR8 which is involved in signaling, although an effect on trafficking was not excluded[20]. Future efforts using receptor mutants and kinase inhibitors are necessary to elucidate whether threonine phosphorylation in the tail occurs and contributes to TLR5 processing. In addition, large variation in length and composition among the tail regions of other human TLRs could provide specificity in the potential role(s) that the tail regions play. As TLR subfamily members share a highly conserved structure but are distinct in terms of regulation, trafficking, dimerization and ligand-binding, the highly diverse tail region seems as a logical candidate region that may contribute to at least some of this specificity. The use of orthology-based chimeric receptors, composed of evolutionarily relevant receptor

sequences, can be instrumental in deciphering what specific receptor features are involved in intracellular transport.

Overall, the current study shows that the tail of human TLR5 is important to establish a functional receptor at the plasma membrane. This function of the tail appears to be different from the interaction of the receptor's EJM region with the trafficking chaperone UNC93B1 as well as the events that dictate the lysosomal localization of zebrafish TLR5b.

Methods and materials

RNA isolation and cDNA synthesis

HeLa-57A cells were grown to confluence in a T25 cell culture flask. RNA was extracted using RNA-Bee (Amsbio) according to the manufacturer's instructions. RNA was treated with DNAse (Thermo Fisher Scientific) and reverse transcribed into cDNA using the RevertAid First Strand cDNA Synthesis kit (Thermo Fisher Scientific) following the manufacturer's instructions.

Plasmid constructs

Phusion high-fidelity DNA polymerase, dNTPs, fast digest restriction endonucleases, T4 DNA ligase, and primers were purchased from Thermo Fisher Scientific. Full length and tail truncated human *tlr5* were amplified from a pUNO hsTLR5-GFP plasmid (Invivogen) by PCR with primers listed in Table 2. The product was digested with BamHI and NotI restriction enzymes and ligated in a pTracer-CMV2ΔGFP/3 × HA[11], yielding hsTLR5 or hsTLR5-Δtail with a C-terminal 3 × HA-tag. To construct N-terminally tagged receptors, full length hsTLR5 and hsTLR5-Δtail were amplified by PCR with a forward primer containing an overlap to the HA coding sequence and a reverse primer containing a stop codon. PCR products were purified, mixed with a synthetic sequence (GeneArt, Thermo Fisher Scientific) containing a hsTLR5 signal peptide coding sequence and a 3 × HA coding sequence, and fused by overlap PCR using forward and reverse primers containing a BamHI and NotI site, respectively. PCR products were digested and ligated into pTracer-CMV2ΔGFP/3 × HA to yield hsTLR5 or hsTLR5-Δtail with a N-terminal 3 × HA-tag between the signal peptide and first leucine rich repeat. hsTLR5 tail mutants were amplified with reverse primers containing the modified sequence and inserted into pTracer-CMV2ΔGFP/3 × HA. The chimeric receptor consisting of hsTLR5 carrying the tail of drTLR5b was constructed by standard overlap extension PCR technique and ligated in pTracer-CMV2ΔGFP/3 × HA following BamHI and NotI restriction to yield hsTLR5-drtail

with a C-terminal 3 × HA tag. Construction of TLR5b and UNC93B1 of zebrafish has been described previously[11]. Human UNC93B1 was amplified by PCR from HeLa-57A cell derived cDNA with a forward primer containing a HindIII restriction site and reverse primer with or without stop codon and EcoRV restriction site. PCR products were digested and ligated in pTracer-CMV2ΔGFP/3 × FLAG[21] to yield untagged and C-terminally 3 × FLAG-tagged UNC93B1. Constructs were verified by sequencing (Macrogen).

Table 2. Primers used in this study

Product	Direction	5' to 3' sequence
hsTLR5-WT	Forward	CCGGATCCGCCACCATGGGAGA
	Reverse	CCGCGGCCGCTGGAGATGGTTGCTACAGTTTG
hsTLR5-Δtail	Forward	CCGGATCCGCCACCATGGGAGA
	Reverse	CCGCGGCCGCGCCAGCCAACATCCTGGAGATC
hsTLR5-WT N-terminal HA	Forward	TGACTACGCTATTCCTTCCTGCTCCTTTGATG
	Reverse	CCGCGGCCGCTCAGGAGATGGTTGCTACAGTTTG
hsTLR5-Δtail N-terminal HA	Forward	TGACTACGCTATTCCTTCCTGCTCCTTTGATG
	Reverse	CCGCGGCCGCTCACTGTTGAGAGAGTTTATGAAGAAACC
hsTLR5-signalpeptide-3xHAtag	Forward	ATCGTACCGGATCCGCCACCATGGGAGACCACCTGGACCTTCTCCTAGG AGTGGTGCTCATGGCCGGTCCTGTGTTTGGATACCCATATGACGTTCCA GACTACGCGTATCCGTACGACGTTCCGGATTACGCTTACCCTTACGATGT ACCTGACTACGCTATTCCTTCCTGCTCCTTTGATGGC
hsUNC93B1 untagged	Forward	CCAAGCTTGCCACCATGGAGGCGGAGCCGCC
	Reverse	CCGATATCCTGCTCCTCCGGCCCGTC
hsUNC93B1 FLAG-tagged	Forward	CCAAGCTTGCCACCATGGAGGCGGAGCCGCC
	Reverse	CCGATATCTCACTGCTCCTCCGGCCCGTCTCCC

Cell culture and transient transfection

HeLa-57A cells stably transfected with an NF-κB-luciferase reporter and HEK293 cells were cultured in DMEM (Thermo Fisher Scientific) supplemented with 5% FCS at 37°C and 10% CO_2. Cells were seeded in a 12 well-plate and after 24 hours cells were transiently transfected at 70% confluence. HeLa-57A cells were transfected with 800 ng of hsTLR5 containing plasmid and 200 ng of human UNC93B1-FLAG containing plasmid. HEK293 cells were transfected with plasmids containing hsTLR5 (400 ng), plasma membrane localized RFP (400 ng)(kind gift from Dr. K. Strijbis) and untagged human UNC93B1 (200 ng) or drTLR5b (400 ng), empty vector (400 ng) and untagged zebrafish UNC93B1 (200 ng). Plasmids were transfected using Fugene HD (Promega) at a DNA to Fugene ratio of 1:3 according to the manufacturer's instructions.

Luciferase NF-κB reporter assay

Twenty-four hours after transfection HeLa-57A cells were redistributed into a 96-well plate. After 24-hours cells were washed twice with DMEM without FCS and stimulated with purified *Salmonella enterica* serovar Enteritidis flagellin (FliC)[22]. After 5 h at 37°C cells were lysed in 50 µl reporter lysis buffer (Promega) at -80°C for 24-hours. After thawing the lysate was mixed with luciferase reagent (Promega) and luciferase activity was measured in a TriStar2 luminometer (Berthold). NF-κB activity is represented by luciferase activity which was measured in relative light units (RLU).

Confocal microscopy

Glass cover slips were coated overnight with 0.02% poly-L-lysine in phosphate buffered saline (PBS)(both Sigma) at room temperature (RT). Twenty-four hours after transfection cells were redistributed onto coated cover slips and cultured in DMEM with 5% FCS. After twenty-four hours cells were washed once with TRIS buffered saline (TBS) and fixed with 1.5% paraformaldehyde (Affimetrix) in TBS. Cells were blocked and permeabilized with 0.2% Bovine serum albumin (BSA) and 0.1% saponin (both Sigma) in TBS for 30 min at RT. Next, cells were incubated (1 h) with M2 anti-FLAG antibody (F3165, Sigma, 1:500) or rabbit anti-human LAMP-1 (24170, Abcam, 1:200) and DAPI (Molecular Probes, 1:2500), washed with TBS, incubated (1 h) with Alexa Fluor-568 goat-anti-mouse IgG (A11031, 1:200) or Alexa Fluor-568 goat-anti-rabbit IgG (A11036, 1:200)(both Thermo Fisher Scientific), washed with TBS and incubated (1 h) with Alexa Fluor-488 mouse-α-HA (A21287, Thermo Fisher Scientific, 1:300). Antibody incubations were in blocking buffer (TBS, BSA and saponin). After staining, cells were washed three times with TBS and once with MilliQ and embedded in Prolong Diamond mounting solution (Thermo Fisher Scientific). Cells were imaged on a Leica SPE-II laser confocal microscope, and images were processed using Leica LAS AF software.

Bioinformatics analyses

Sequences were aligned with Clustal Omega https://www.ebi.ac.uk/Tools/msa/clustalo/. Phosphorylation was predicted with NetPhos 3.1 http://www.cbs.dtu.dk/services/NetPhos/. GenBank accession numbers used in the TLR tail alignments were: human TLR1 NP_003254.2, TLR2 NP_003255.2, TLR3 NP_003256.1, TLR4 NP_612567.1, TLR5 NP_003259.2, TLR6 NP_006059.2, TLR7 NP_057646.1, TLR8 NP_057694.2, TLR9 NP_059138.1, TLR10 NP_112218.2, *Mus musculus* TLR5 NP_058624.2, *Bos*

taurus TLR5 NP_001035591.1, *Gallus gallus* TLR5 NP_001019757.1, *Anolis carolinensis* TLR5 ALT10445.1 and *Danio rerio* TLR5b AVQ55078.1.

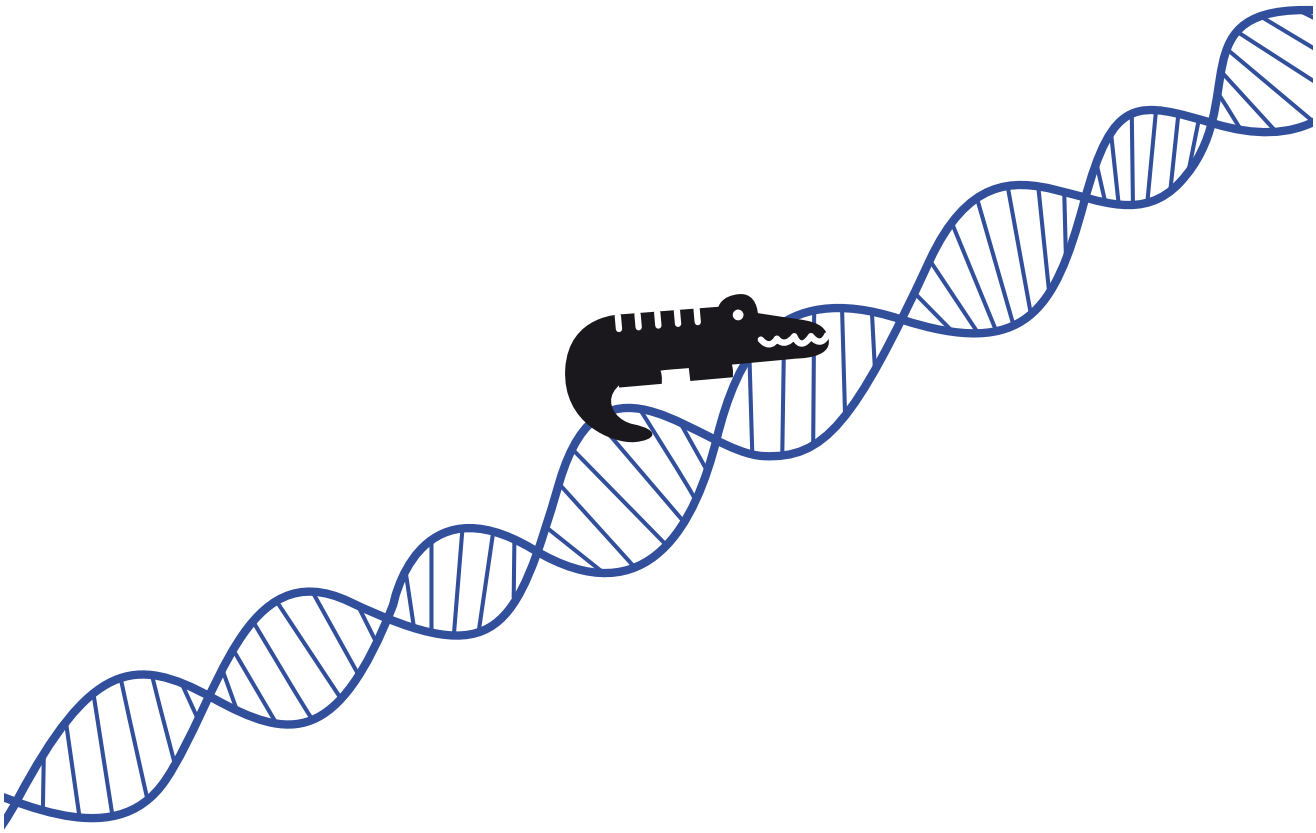
Acknowledgments

We kindly thank Caren Leemberg for performing pilot experiments.

References

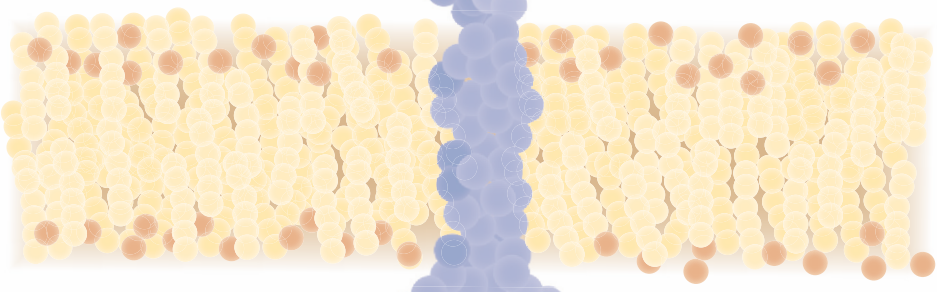
1. Gao, D. & Li, W. Structures and recognition modes of toll-like receptors. *Proteins: Structure, Function, and Bioinformatics* **85**, 3–9 (2017).
2. Jin, M. S. *et al.* Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* **130**, 1071–1082 (2007).
3. Park, B. S. *et al.* The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex. *Nature* **458**, 1191–1195 (2009).
4. Reuven, E. M., Fink, A. & Shai, Y. Regulation of innate immune responses by transmembrane interactions: Lessons from the TLR family. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1838**, 1586–1593 (2014).
5. Ruysschaert, J.-M. & Lonez, C. Role of lipid microdomains in TLR-mediated signalling. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1848**, 1860–1867 (2015).
6. Gay, N. J., Symmons, M. F., Gangloff, M. & Bryant, C. E. Assembly and localization of Toll-like receptor signalling complexes. *Nature Reviews Immunology* **14**, 546–558 (2014).
7. O'Neill, L. A. J. & Bowie, A. G. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nature Reviews Immunology* **7**, 353–364 (2007).
8. Kim, Y.-M., Brinkmann, M. M., Paquet, M.-E. & Ploegh, H. L. UNC93B1 delivers nucleotide-sensing toll-like receptors to endolysosomes. *Nature* **452**, 234–238 (2008).
9. Kim, J. *et al.* Acidic amino acid residues in the juxtamembrane region of the nucleotide-sensing TLRs are important for UNC93B1 binding and signaling. *Journal of Immunology* **190**, 5287–5295 (2013).
10. Huh, J.-W. *et al.* UNC93B1 is essential for the plasma membrane localization and signaling of Toll-like receptor 5. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 7072–7077 (2014).
11. Voogdt, C. G. P., Wagenaar, J. A. & Putten, J. P. M. van. Duplicated TLR5 of zebrafish functions as a heterodimeric receptor. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E3221–E3229 (2018).
12. Lee, B. L. *et al.* UNC93B1 mediates differential trafficking of endosomal TLRs. *eLife* **2**, e00291 (2013).
13. Pelka, K. *et al.* The chaperone UNC93B1 regulates Toll-like receptor stability independently of endosomal TLR transport. *Immunity* **48**, 911-922.e7 (2018).
14. Xu, Y. *et al.* Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature* **408**, 111–115 (2000).
15. Xu, B., Stippec, S., Robinson, F. L. & Cobb, M. H. Hydrophobic as well as charged residues in both MEK1 and ERK2 are important for their proper docking. *Journal of Biological Chemistry* **276**, 26509–26515 (2001).
16. Zhao, N., Pang, B., Shyu, C.-R. & Korkin, D. Charged residues at protein interaction interfaces: unexpected conservation and orchestrated divergence. *Protein Science* **20**, 1275–1284 (2011).

17. Park, S. Y. & Guo, X. Adaptor protein complexes and intracellular transport. *Bioscience Reports* **34**, (2014).
18. Chattopadhyay, S. & Sen, G. C. Tyrosine phosphorylation in Toll-like receptor signaling. *Cytokine & Growth Factor Reviews* **25**, 533–541 (2014).
19. Leifer, C. A. & Medvedev, A. E. Molecular mechanisms of regulation of Toll-like receptor signaling. *Journal of Leukocyte Biology* **100**, 927–941 (2016).
20. Rajagopal, R., Waller, A. S., Mendoza, J. D. & Wightman, P. D. The covalent modification and regulation of TLR8 in HEK-293 cells stimulated with imidazoquinoline agonists. *Biochemical Journal* **409**, 275–287 (2008).
21. Keestra, A. M., Zoete, M. R. de, Bouwman, L. I. & Putten, J. P. M. van. Chicken TLR21 Is an Innate CpG DNA Receptor Distinct from Mammalian TLR9. *Journal of Immunology* **185**, 460–467 (2010).
22. Voogdt, C. G. P., Bouwman, L. I., Kik, M. J. L., Wagenaar, J. A. & van Putten, J. P. M. Reptile Toll-like receptor 5 unveils adaptive evolution of bacterial flagellin recognition. *Scientific Reports* **6**, 19046 (2016).



Chapter 6

Evolutionary regression and species-specific codon usage of TLR15



Carlos G.P. Voogdt[1], Mark E. Merchant[2], Jaap A. Wagenaar[1,3], Jos P.M. van Putten[1]

[1]Department of Infectious Diseases & Immunology, Utrecht University, Utrecht, the Netherlands
[2]Department of Chemistry, McNeese State University, Lake Charles, LA, United States
[3]Wageningen Bioveterinary Research, Lelystad, the Netherlands

Frontiers in Immunology Vol. 9, Article 2626 2018

Abstract

Toll-like receptors (TLRs) form an ancient family of innate immune receptors that detect microbial structures and activate the host immune response. Most subfamilies of TLRs (including TLR3, TLR5 and TLR7) are highly conserved among vertebrate species. In contrast, TLR15, a member of the TLR1 subfamily, appears to be unique to birds and reptiles. We investigated the functional evolution of TLR15. Phylogenetic and synteny analyses revealed putative TLR15 orthologs in bird species, several reptilian species and also in a shark species, pointing to an unprecedented date of origin of TLR15 as well as large scale reciprocal loss of this TLR in most other vertebrates. Cloning and functional analysis of TLR15 of the green anole lizard (*Anolis carolinensis*), salt water crocodile (*Crocodylus porosus*), American alligator (*Alligator mississippiensis*) and chicken (*Gallus gallus*) showed for all species TLR15 specific protease-induced activation of NF-κB, despite highly variable TLR15 protein expression levels. The variable TLR15 expression was consistent in both human and reptilian cells and could be attributed to species-specific differences in TLR15 codon usage. The species-specific codon bias was not or barely noted for more evolutionarily conserved TLRs (e.g. TLR3). Overall, our results indicate that TLR15 originates before the divergence of chondrichthyes fish and tetrapods and that TLR15 of both avian and reptilian species has a conserved function as protease activated receptor. The species-specific codon usage and large scale loss of TLR15 in most vertebrates suggest evolutionary regression of this ancient TLR.

Introduction

Toll-like receptors (TLRs) are innate immune receptors that have a critical role in the early detection of infection[1]. The general architecture of TLRs consists of a ligand-binding extracellular domain containing multiple leucine rich repeats (LRR), a single transmembrane domain and an intracellular Toll-interleukin-1 (TIR) signaling domain[2]. Ligand-induced TLR signaling activates immune-related transcription factors (e.g. nuclear factor κB, NF-κB) which induce expression of pro-inflammatory genes. The importance of TLRs in the immune system is underlined by the strong evolutionary conservation of this family of receptors. The prototypical Toll receptor originates at the base of metazoa approximately 600 million years ago[3]. Subsequently, extensive gene duplication and gene loss events have resulted in 10 different TLRs in some mammals (including humans) to more than 20 TLRs in teleost fish[4–6]. Evolutionary diversification of the TLR ligand-binding domain to detect diverse types of microbial structures has resulted in distinct TLR subfamilies.

One of the TLR subfamilies that has evolved highly dynamically is the TLR1 subfamily that comprises TLR1, TLR2, TLR6 and TLR10. Members of this subfamily typically function as heterodimeric receptors. Heterodimers of TLR2 and TLR1 or TLR6 respond to microbial lipopeptides such as Pam_3CSK_4 (TLR2/TLR1) or FSL-1 (TLR2/TLR6)[7–9]. In mammals TLR1 and TLR6 arose by tandem duplication and are limited in divergence due to gene conversion[10]. For reasons unknown, TLR10 has been preserved in some mammals (including humans) and has been lost in other vertebrates. Among teleost fish, the common carp (*Cyprinus carpio*) duplicated its TLR2 gene[11], whereas the Atlantic cod (*Gadus morhua*) lost TLR1, TLR2, TLR6 and TLR10 altogether[12]. In birds, duplications of TLR1 and TLR2 are abundant[6,13] which has left chicken (*Gallus gallus*) with the paralogs TLR2A, TLR2B, TLR1A (also known as TLR16) and TLR1B. Chicken TLR2B/TLR1A heterodimers show dual recognition of the Pam_3CSK_4 and FSL-1 ligands[14]. Interestingly, one gene duplicate within the TLR1 subfamily that appears to have evolved independently of other TLR1 subfamily members is TLR15. This TLR functions as a homodimer rather than as a heterodimer and signals upon proteolytic cleavage of its extracellular domain by microbial proteases[15,16]. TLR15 is absent in mammals and was first described in chicken[17]. A partial related sequence has also been identified in the genome of the reptile *Anolis carolinensis* suggesting that TLR15 may be unique to the reptilian lineage[16].

Reptiles can be broadly subdivided in lepidosauria (lizards, snakes, amphisbaenians and tuatara) and archosauria (crocodiles and birds). The position of turtles among reptiles is still debated but molecular analyses tend to group turtles

within archosauria[18]. Reptiles were the first vertebrates that could permanently colonize terrestrial habitats and thereby came into contact with prehistoric terrestrial microbiota which shaped the immune system of reptiles and descending animals. Despite its central role in vertebrate evolution little is known about the reptilian immune system (but see [19–22]), especially at the level of reptile-microbe interactions. Previously, we unveiled adaptive evolution of TLR5 of the *Anolis carolinensis* lizard indicating different sensitivity of lizard and human TLR5 to bacterial flagellins[23]. Given the dynamic evolution of TLR1 subfamily members and the recent increase in available whole genome sequences of reptiles and other non-mammals, we here aimed to define the extent of genomic and functional evolutionary conservation of TLR15 in non-avian reptiles.

Bioinformatics analyses uncovered the presence of TLR15 outside the reptilian lineage, as well as loss of TLR15 within the reptilian lineage. Functional activation assays with recombinant lepidosaurian (*Anolis carolinensis*) and archosaurian (*Crocodylus porosus* and *Alligator mississippiensis*) TLR15 revealed conservation of function among reptilian and chicken TLR15 orthologs. Markedly variable expression efficiency of different reptilian TLR15s in both human and reptilian cells could be experimentally attributed to species-specific codon usage of the respective TLR15 genes. Finally, interspecies variability of codon usage in TLR15 was higher compared to TLRs which evolved more stably within vertebrates.

Results

Identification of TLR15 in vertebrates

In order to identify potential TLR15 sequences in vertebrate genomes we investigated the evolutionary relationship among TLR1 subfamily members from a diverse set of species (Table S1) using a maximum likelihood based phylogenetic tree (Fig. 1). Predicted TLR1 and TLR2 sequences of the uro-chordate *Ciona intestinalis*, an invertebrate, were used to root the tree. Analysis of the tree revealed four separate branches within the TLR1 subfamily; (i) TLR2, (ii) TLR1/6/10, (iii) a second group of TLR2 present only in fish, amphibians and non-avian reptiles, and (iv) a group containing chicken TLR15 (Fig. 1). An additional sequence more distantly related to the TLR2 precursor was found in the Australian ghost shark, spotted gar and medaka but not in other vertebrates. Due to low sequence homology, the TLR15 branch did not contain TLR sequences of teleost fish, coelacanth, amphibians or mammals. On the contrary, several sequences of birds as well as reptiles annotated in the database as TLR1 or TLR2 showed highest homology to chicken TLR15 and thus grouped in the TLR15 phylogenetic branch.

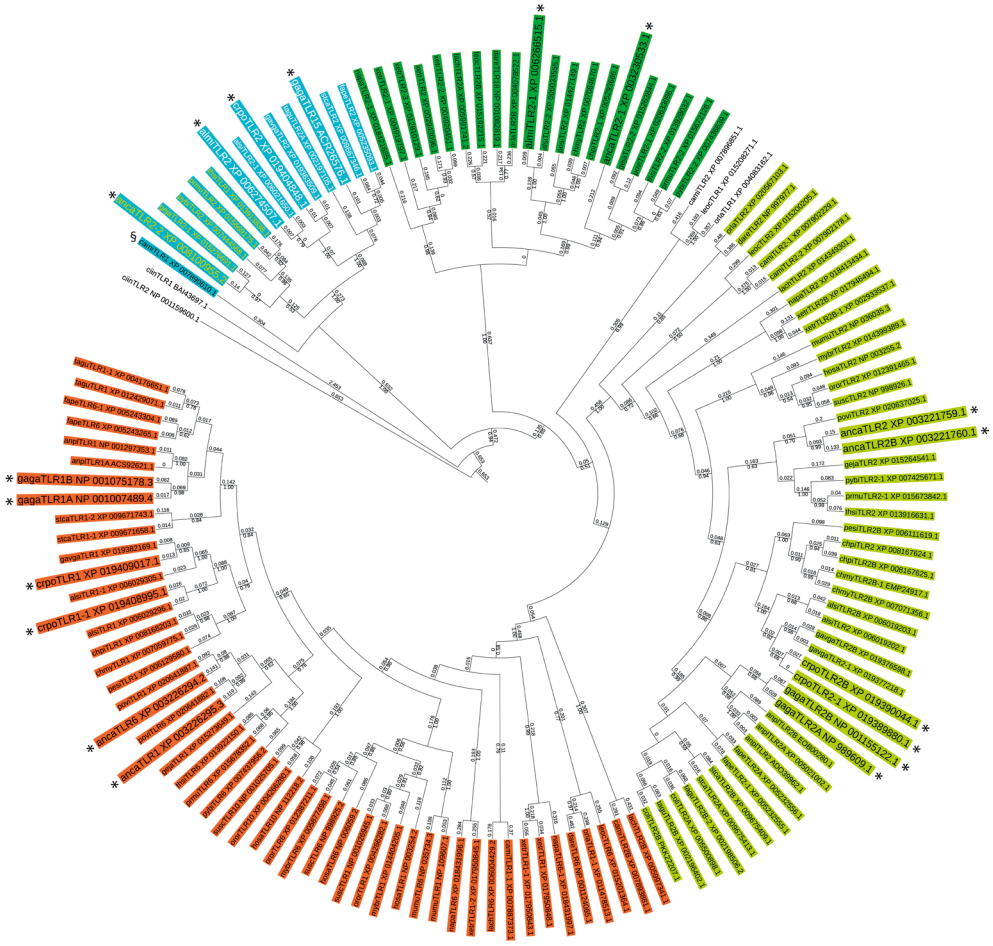


Figure 1. Phylogenetic tree of the vertebrate TLR1 subfamily. The evolutionary history of the vertebrate TLR1 subfamily was inferred by using the Maximum Likelihood method based on the JTT matrix-based model (50). The tree with the highest log likelihood (-28950.33) is shown. For each TLR sequence the GenBank accession number is indicated. For species abbreviations see Table S1. Bootstrap analysis was performed with 250 iterations and the fraction of trees in which the associated taxa clustered together is shown below the branches (only bootstrap values greater than 0.5 are shown). Branch lengths indicate the number of substitutions per site and are shown above the branches. The analysis involved 136 full length amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 252 positions in the final dataset. Evolutionary analyses were conducted in MEGA7[41]. The tree was customized using iTol[42]. Four separate groups within the TLR1 subfamily are identified; (i) TLR2 (yellow), (ii) TLR1/6/10 (orange), (iii) a group of TLR2 present only in fish, amphibians and non-avian reptiles (green) and (iv) the group containing chicken TLR15 (blue). In the TLR15 branch, lepidosaurians are show in yellow letters and archosaurians are shown in white letters. The predicted TLR1 of the ghost shark (*Callorhinchus milii*, camiTLR1) is indicated with the § symbol. All TLR1 subfamily members of *Anolis carolinensis* (anca), *Crocodylus porosus* (crpo), *Alligator mississippiensis* (almi) and *Gallus gallus* (gaga) are shown enlarged and are indicated with an asterisk.

This supports the notion that TLR15 is unique to the reptilian lineage. Unexpectedly however, a predicted TLR1 sequence of the Australian ghost shark (*Callorhinchus milii*) also clustered with high bootstrap support within the chicken TLR15 branch. Apart from the shark sequence, the TLR15 phylogeny recapitulates with high support the division of reptilia into lepidosauria and archosauria (Fig. 1). Yet, annotated TLR1, 2 or 6 sequences of the bearded dragon (*Pogona vitticeps*, a lizard) and all three species of turtles could be placed somewhere in the tree but none of the TLR sequences of these reptiles clustered within the TLR15-containing branch. This suggests that TLR15 has been lost from most non-reptilian lineages as well as from specific reptilian lineages after the divergence of lepido- and archosaurians.

To gain additional evidence for reciprocal loss of TLR15 in teleost fish, amphibians, mammals, turtles and the bearded dragon and to confirm the conservation of putative TLR15 in the other reptiles, we collected the genomic region surrounding *tlr15* from the NCBI Gene database and compared the gene synteny in this region between chicken and other species. This showed that chicken *tlr15* is flanked by *psme4*, *erlec1*, *gpr75*, *chac2* and *asb3*. These genes are absolutely conserved and arranged in this order in all species investigated here, except for zebrafish in which *chac2* and *asb3* are replaced by *agpat4* and *map3k4* (Fig. 2). While all five genes surrounding chicken *tlr15* are conserved in the same order in the bearded dragon, no gene was identified between *erlec1* and *gpr75* in this reptilian species. The same was true for the Chinese softshell turtle. In the green sea turtle a predicted TLR2 pseudogene with high homology to chicken TLR15 is situated between *erlec1* and *gpr75.* In the genome of the painted turtle, a 214 amino acid coding sequence is conserved between *erlec1* and *gpr75* that has high homology to the TIR domain of chicken TLR15 and thus may represent a remnant of TLR15. Teleost fish, coelacanth and amphibians carry no genes between *erlec1* and *gpr75.* In humans a microRNA encoding sequence is present at this position and in mice a pseudogene is predicted at this location but the residual protein sequence lacks leucine rich repeats, transmembrane regions or a TIR domain. Conversely, all of the predicted TLR1 or TLR2 sequences of species that grouped in the TLR15 branch of the phylogenetic tree, including the predicted TLR1 of the ghost shark (Fig. 1), mapped between *erlec1* and *gpr75* (Fig. 2).

Together, the phylogenetic and synteny analyses strongly suggest that the precursor of TLR15 is an ancient gene duplicate of the TLR2/1/6/10 precursor dating back at least to the common ancestor of chondrichthyes fish and tetrapods and that TLR15 has been reciprocally lost from the teleost fish, coelacanth, amphibian, mammalian and even specific reptilian lineages.

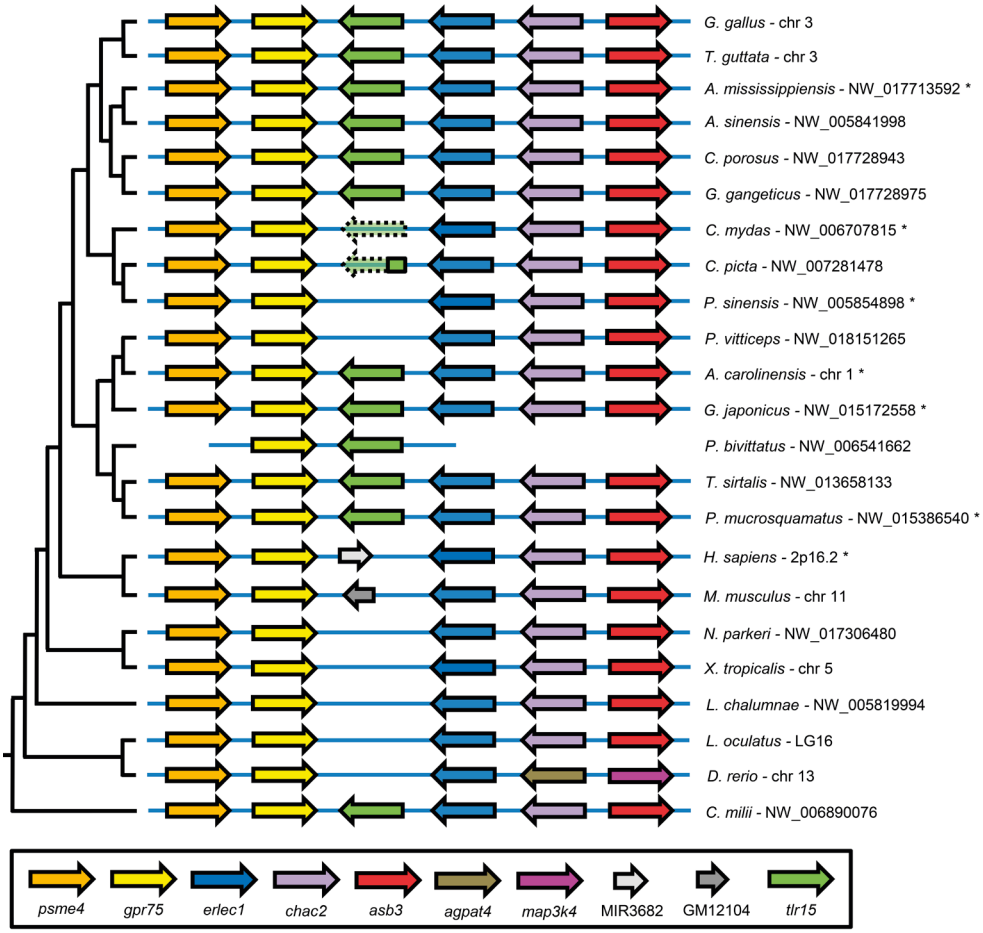


Figure 2. TLR15 gene synteny comparison among vertebrates. The genomic region containing chicken (*G. gallus*) *tlr15* compared to the same genomic region of the indicated species shows very high conservation of gene synteny. Genomic regions were collected from the NCBI Gene database and ordered according to the species phylogeny, shown on the left. Following the scientific name of each species is the NCBI Gene database identifier for this genomic region. *Tlr15* is shown in green and putative *tlr15* pseudogenes in the genomes of the *C. mydas* and *C. picta* turtles are shown in transparent color with a dashed line. Genes in this conserved genomic region are: proteasome activator complex subunit 4 (psme4), G-protein coupled receptor 75 (gpr75), endoplasmic reticulum lectin 1 (erlec1), ChaC cation transport regulator homolog 2 (chac2), ankyrin repeat and SOCS box containing 3 (asb3), 1-acylglycerol-3-phosphate O-acyltransferase 4 (agpat4), mitogen-activated protein kinase kinase kinase 4 (map3k4), microRNA 3682 (MIR3682). GM12104 in the mouse genome is a pseudogene without TLR features. Genomic regions of species indicated with an asterisk are annotated in NCBI's Gene bank database in reverse order.

Cloning and characteristics of reptilian TLR15

Gene evolution is largely driven by selection on function. We therefore investigated whether the putative TLR15 genes found in lepidosaurian and archosaurian reptiles

still encode a functional receptor. Hereto, putative *tlr15* genes were amplified from DNA of the lepidosaurian *Anolis carolinensis* (ancaTLR15) and archosaurians *Crocodylus porosus* and *Alligator mississippiensis* (crpoTLR15 and almiTLR15 resp.). Genes were cloned upstream of a C-terminal hemagglutinin (HA) or FLAG-tag sequence in an expression vector. Reptilian *tlr15* genes comprise a single exon and encode proteins of 823 (ancaTLR15), 877 (crpoTLR15) and 875 (almiTLR15) amino acids in length. Comparison of the putative TLR15 protein sequences of reptiles and chicken (gagaTLR15, 868 amino acids) showed that both full length crocodilian TLR15 proteins are more similar to chicken (69%) than to anolis TLR15 (59%) (Table S2). All proteins had a similar architecture consisting of an extracellular domain (ECD) with 20 (gagaTLR15, crpoTLR15 and almiTLR15) or 18 (ancaTLR15) leucine rich repeats (LRRs), a C-terminal LRR (CTLRR), a single transmembrane region and a highly conserved intracellular TIR domain. Like TLR1, TLR6 and TLR10 of other species, all TLR15 sequences lack a cysteine containing N-terminal LRR (NTLRR) and like TLR2, the CTLRR of TLR15 is characterized by a CxCx24Cx20C cysteine motif[24] (Fig. S1).

Activation of reptilian TLR15 by proteases

Functionality of the cloned reptilian TLR15s was assessed after transfection of the plasmids encoding ancaTLR15, crpoTLR15 or almiTLR15, together with an NF-κB-luciferase reporter, into human HEK293 cells. Stimulation of reptilian TLR15-transfected cells with Proteinase K resulted in increased NF-κB activity (Fig. 3), as was observed for the positive control cells transfected with gagaTLR15[15]. Reptilian and chicken TLR15 did not respond to the canonical TLR2/1 or TLR2/6 ligands Pam_3CSK_4 or FSL-1 resp. while these ligands were able to activate NF-κB in control cells that expressed the chicken TLR2B/TLR1A heterodimer[14]. Previously, we and others found that gagaTLR15 is activated by secreted proteases of fungi pathogenic to poultry[15,16]. *Chrysosporium* anamorph of *Nannizziopsis vriesii* (CANV) is a pathogenic fungus which can cause a fatal condition called yellow fungus disease in infected reptiles[25]. Use of culture supernatant of a clinical CANV isolate in our stimulation assay potently activated NF-κB in cells expressing reptilian or chicken TLR15 but not chicken TLR2B/TLR1A. Addition of the serine protease inhibitor PMSF to the CANV culture supernatant strongly reduced its TLR15 activating capacity, confirming the response of TLR15 to proteolytic activity (Fig. 3). The responsiveness of the newly identified TLR15 (previously erroneously annotated as TLR2) of *A. carolinensis*, *C. porosus* and *A. mississippiensis* indicates that these receptors are indeed still functional and share functional characteristics with chicken TLR15 and not chicken TLR2 or TLR1.

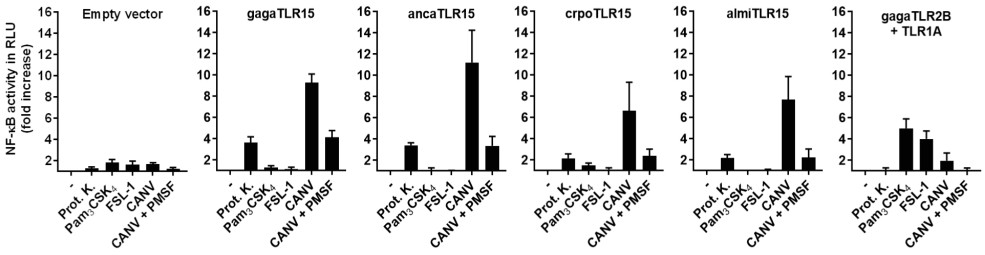


Figure 3. NF-κB activation by reptilian TLR15. HEK293 cells transiently transfected with an NF-κB luciferase reporter plasmid and empty vector, chicken (gaga), anolis (anca), crocodile (cpro), alligator (almi) TLR15 or the gagaTLR2B and gagaTLR1A plasmids were stimulated (5 h) with Proteinase K (100 ng/ml), Pam_3CSK_4 (100 ng/ml), FSL-1 (100 ng/ml), 10 ul of *Chrysosporium* anamorph of *Nannizziopsis vriesii* (CANV) sterile culture supernatant or 10 ul of CANV supernatant pre-treated (30 min) with 1 mM PMSF. Values are the mean ± SEM fold increase of NF-κB activity, represented by luciferase activity in Relative Light Units (RLU), in stimulated cells over unstimulated control cells from three independent experiments performed in duplicate.

Proteolytic cleavage and variable expression of TLR15

To ensure that the protease treatment of cells transfected with reptilian TLR15 resulted in proteolytic cleavage of the receptor, cells expressing a C-terminal FLAG-tagged TLR15 were incubated with Proteinase K, lysed and subjected to Western blotting using a FLAG-specific antibody. Proteinase K cleaved ancaTLR15 to form a similarly sized product as gagaTLR15 (slightly higher than 70 kDa)[15], yet the efficiency of cleavage of ancaTLR15 was substantially less than noted for gagaTLR15 (Fig. 4A). Although both crpoTLR15 and almiTLR15 have a similar molecular size as gagaTLR15 and ancaTLR15, no cleaved forms of these receptors were detected; however, this may be due to the generally low level of expression of these receptors in whole cell lysates (Fig. 4A). Detection of the various TLR15 receptors using confocal microscopy showed that HEK293 cells transfected with gagaTLR15 strongly expressed this receptor at the cell surface, in line with previous findings in different cell-lines[15]. Detection of the ancaTLR15 also indicated strong expression but this TLR resided mostly intracellularly (Fig. 4B). Interestingly, despite a higher protein similarity to gagaTLR15 than to ancaTLR15, both crocodilian TLR15s localized mostly intracellularly but with low signal intensity (Fig. 4B), consistent with the observed low protein expression levels for these receptors (Fig. 4A). In an attempt to improve expression of the crocodilian receptors we transfected plasmids into viper heart (VH-2) reptilian cells rather than human HEK293 cells. In the reptilian cells the crpoTLR15 and almiTLR15 proteins still could not be detected by Western blotting in contrast to ancaTLR15 and gagaTLR15 (Fig. 4C). These findings suggest that the different TLR15s display a species-specific difference in protein expression efficiency.

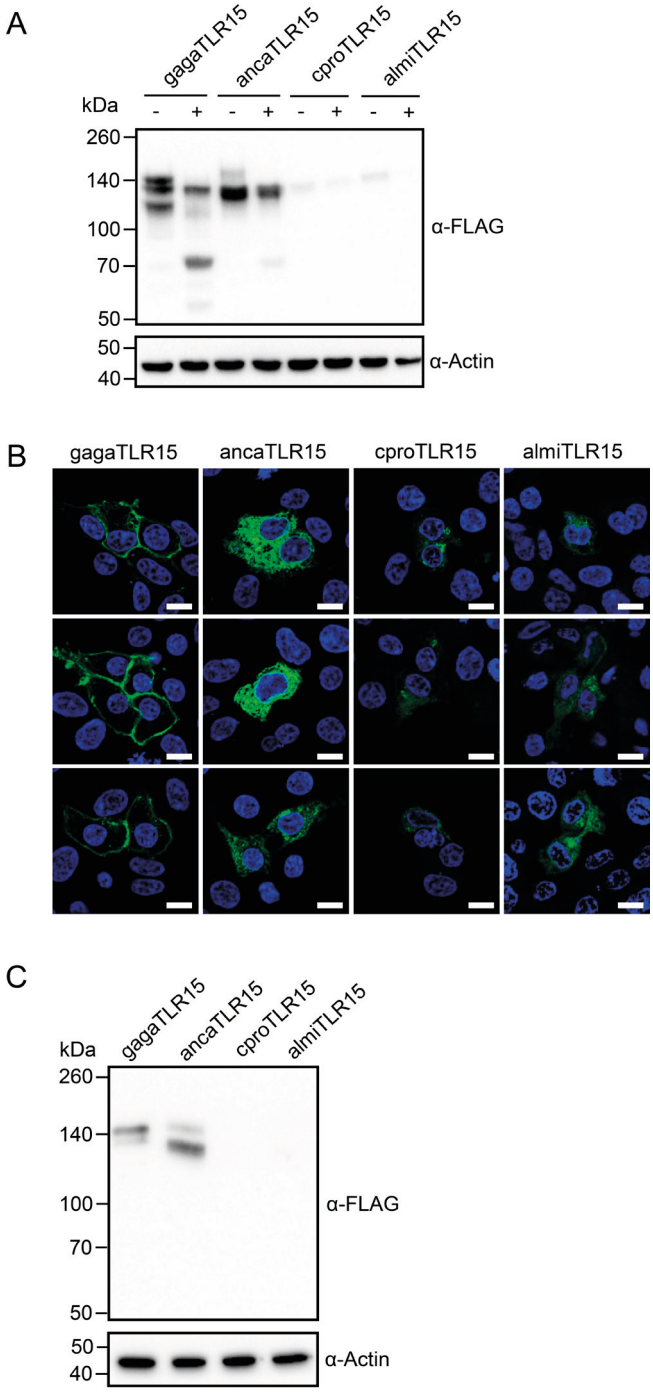


Figure 4. Proteolytic cleavage and expression of reptilian TLR15. **(A)** Immunoblot analysis of HEK293 cells expressing C-terminally FLAG-tagged chicken (gaga), anolis (anca), crocodile (cpro) or alligator (almi) TLR15 left untreated (-) or stimulated (+) (1 h) with 250 ng/ml Proteinase K. Mature TLR15 is approximately 140 kDa. Treatment with Proteinase K results in cleavage of gagaTLR15 and ancaTLR15 to form a cleaved receptor fragment that is slightly higher than 70 kDa. Note that crpoTLR15 and almiTLR15 are poorly expressed compared to ancaTLR15 and gagaTLR15. Beta-actin was detected to confirm equal loading of total protein onto SDS-PAGE gel. **(B)** Confocal microscopy on HEK293 cells expressing C-terminally HA-tagged TLR15 (green). Note that crpoTLR15 and almiTLR15 show lower expression compared to ancaTLR15 and gagaTLR15. All images were produced with the same microscopy settings. Nuclei are stained with DAPI (blue). White scale bar is 10 µm. Three representative images from two independent experiments are shown for each transfected group. **(C)** Immunoblot analysis of reptilian viper heart (VH-2) cells transfected with the different FLAG-tagged TLR15s. The rabbit α-human Beta actin antibody cross reacts with a specific protein in VH-2 cell lysate which was used to confirm equal loading of total protein onto SDS-PAGE gel. For (A) and (C); results are representative of three independent experiments.

Species-specific codon bias of TLR15

In search for the molecular basis of the variable expression levels of the different TLR15s in human and reptilian cells, we compared the codon usage of the different *tlr15* genes. Codon usage bias, the organism-specific use of different synonymous codons to encode the same amino acid, is well known for its effect on heterologous protein expression efficiency. To investigate whether the limited expression of crpoTLR15 and almiTLR15 protein in human cells could be due to codon bias, we analyzed codon usage of the four *tlr15* genes in relation to the genome-wide codon usage in human (Table S3). While gaga*tlr15* and anca*tlr15* contain more frequently than infrequently used codons, the opposite was found for crpo*tlr15* and almi*tlr15* (Fig. 5A). The higher number of infrequent codons in crocodile and alligator *tlr15* transcripts may reduce translation efficiency of these receptors resulting in lower protein expression levels compared to chicken and anolis TLR15.

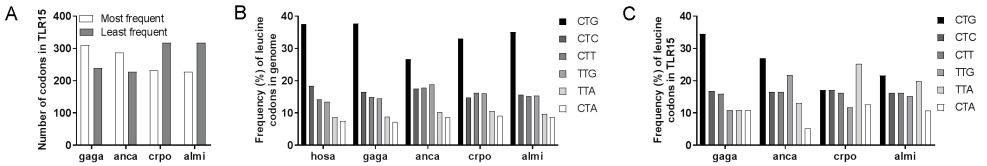


Figure 5. Species-specific bias in TLR15 codon usage. **(A)** The number of most frequent or least frequent codons in chicken (gaga), anolis (anca), crocodile (cpro) and alligator (almi) TLR15 according to codon prevalence in the human genome (see also Table S3). **(B)** Frequency of the six codons encoding the amino acid leucine in the genomes of *Homo sapiens* (hosa), *Gallus gallus* (gaga), *Anolis carolinensis* (anca), *Crocodylus porosus* (crpo) and *Alligator mississippiensis* (almi). **(C)** Frequency of leucine codons in TLR15 of the indicated species.

As codon usage may differ between species, we next compared human, chicken, anolis, crocodile and alligator genome-wide usage of leucine codons. We focused on leucine as this amino acid can be encoded by six codons and is the most abundant amino acid in TLRs, including TLR15. Results showed that genome-wide leucine codon usage is conserved among human, chicken, anolis, crocodile and alligator and that in all species the CTG leucine codon is most abundant and CTA and TTA are least abundant (Fig. 5B). In clear contrast to genome-wide usage of leucine codons, leucine codon usage in *tlr15* genes is markedly different between species. For example, CTG codon usage in gagaTLR15 is 34% (41/119) versus only 17% in crpoTLR15 (19/111) while TTA codon usage is just 11% in gagaTLR15 (13/119) but 25% in crpoTLR15 (28/111). The latter is more than twice the average of TTA codon usage in both the chicken and crocodile genome. (Fig. 5C). Additional analysis of the other TLR15 sequences identified within the phylogenetic tree also showed extensive variation in leucine codon usage among lepido- and archosaurians

despite similar genome-wide leucine codon usage in these species (Fig. 6A). Interestingly, the same analysis of TLR3, TLR5 and TLR7, which unlike TLR15 are highly conserved among vertebrate species, revealed a more conserved pattern of leucine codon usage among the same set of species, especially in the case of TLR3 (Fig. 6B). These findings indicate that TLR15, which has been lost in most vertebrates, shows a species-specific bias of leucine codon usage with greater interspecies variability than TLRs that have been conserved across most vertebrates

To verify that the identified species-specific codon usage in *tlr15* genes is a major cause of the observed variable expression levels of TLR15, we transfected HEK293 and VH-2 cells with synthetic alligator and crocodile *tlr15* genes that had been codon optimized according to human codon usage. For leucine residues in almiTLR15 and crpoTLR15 the optimization resulted in ≤ 5% of leucines being encoded by the CTC codon and ≥ 95% being encoded by the most frequently used CTG codon. Transfection of the codon optimized genes resulted in very high expression of both TLR15 proteins in human HEK293 cells as well as in reptilian VH-2 cells (Fig. 7). This clearly indicates that gene-specific evolutionary changes of codon usage have a major impact on relative protein expression efficiency, including the expression of reptilian TLR15.

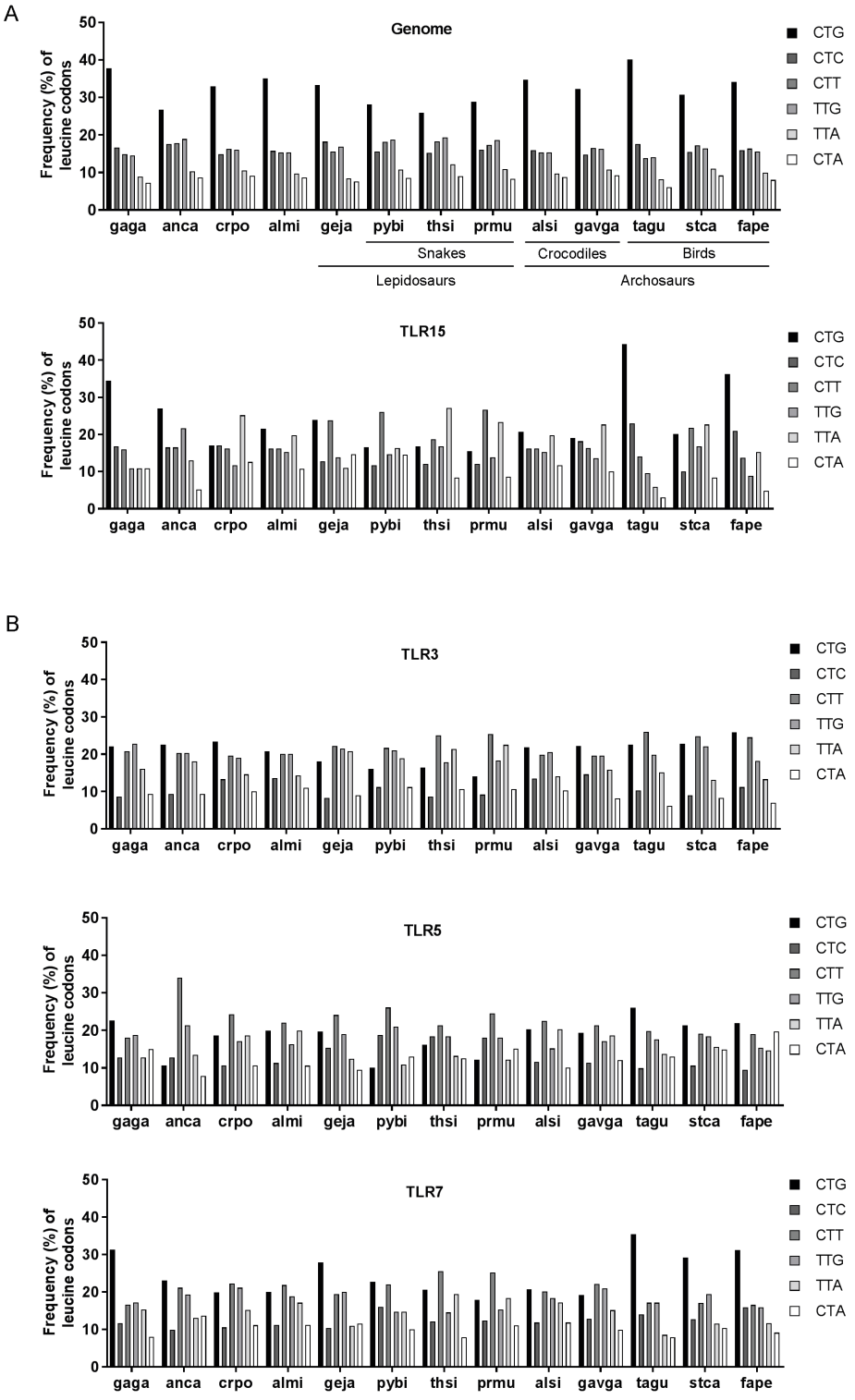

A
Genome
Frequency (%) of leucine codons
CTG
CTC
CTT
TTG
TTA
CTA
gaga anca crpo almi geja pybi thsi prmu alsi gavga tagu stca fape
Snakes
Crocodiles
Birds
Lepidosaurs
Archosaurs
TLR15
B
TLR3
TLR5
TLR7

Figure 6 on previous page. Leucine codon usage in TLRs of lepido- and archosaurians. **(A)** Frequency (percentage) of the six codons encoding leucine across genome-wide coding sequences (upper panel) or in TLR15 (lower panel) in lepidosaurians and archosaurians. **(B)** Frequency of leucine codons in TLR3, TLR5 and TLR7 in lepidosaurians and archosaurians. Species abbreviations: gaga (*Gallus gallus*, bird), anca (*Anolis carolinensis*, lizard), crpo (*Crocodylus porosus*, crocodile), almi (*Alligator mississippiensis*, crocodile), geja (*Gekko japonicus*, lizard), pybi (*Python bivittatus*, snake), thsi (*Thamnophis sirtalis*, snake), prmu (*Protobothrops mucrosquamatus*, snake), alsi (*Alligator sinensis*, crocodile), gavga (*Gavialis gangeticus*, crocodile), tagu (*Taeniopygia guttata*, bird), stca (*Struthio camelus australis*, bird), fape (*Falco peregrinus*, bird).

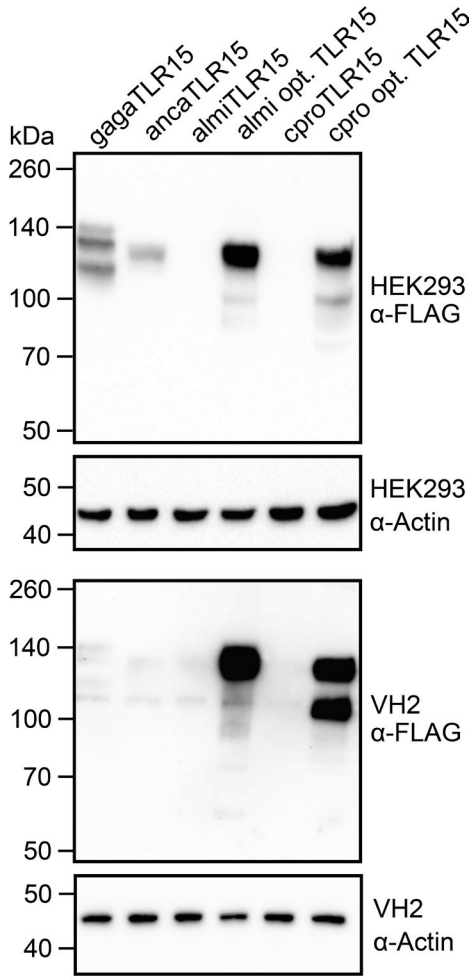


Figure 7. Expression of codon optimized alligator and crocodile TLR15. Immunoblot analyses of HEK293 cells and VH-2 cells expressing C-terminal FLAG-tagged chicken (gaga), anolis (anca), alligator (almi),codon optimized alligator (almi opt.), crocodile (cpro) or codon optimized crocodile (crpo opt.) TLR15. Mature TLR15 is approximately 140 kDa. Beta-actin was detected to confirm equal loading of HEK293 and VH-2 total protein onto SDS-PAGE gel. Results are representative of at least two independent experiments.

Discussion

Throughout evolution, duplications and losses of TLR genes have resulted in varying repertoires of TLRs among animal lineages. While some TLRs are highly conserved in nearly all vertebrates, other TLRs, most notably members of the TLR1 subfamily, evolved much more dynamically and are occasionally only found in specific vertebrate lineages. In the present work we provide evidence that (i) TLR15 is evolutionarily older than expected, (ii) *tlr15* genes display species-specific codon usage, (iii) the *tlr15* gene underwent evolutionary regression in most vertebrates including certain reptilian lineages, and (iv) that the activation of reptilian TLR15 by external proteases is a conserved feature that functionally distinguishes TLR15 from other TLR1 subfamily members.

TLR15 was originally identified in chickens[17] and later studies found additional TLR15 orthologs in other avian and also four reptilian species (green anole lizard, Burmese python, Chinese alligator and American alligator)[6,16,26]. By combining phylogenetic and synteny analyses results, we identified TLR15 orthologs in even more reptilian species. All of the newly identified *tlr15* genes in lepidosaurians and archosaurians (including avian *tlr15* genes) are erroneously annotated in the database as TLR1 or TLR2 sequences. However, based on their phylogeny as well as highly conserved position in the genome and, most convincingly, the activation of anolis (lepidosaurian), salt water crocodile and alligator (archosaurians) TLR15 by proteolytic cleavage, we consider these sequences as bonafide TLR15 orthologs.

Unexpectedly, our bioinformatics search of a wide range of vertebrate genomes also led to the identification of a TLR sequence in the genome of the Australian ghost shark that has high homology to, and the same synteny as, avian and reptilian TLR15. The presence of this putative TLR15 ortholog in a shark species suggests that TLR15 did not arise in the sauropsid lineage but instead originates before the divergence of chondrichthyes fish and tetrapods. This would date the origin of the ancestral TLR15 to at least 465 million years ago[27] while reptiles share a common ancestor with birds roughly 284 million years ago[27,28].

The successful expression of recombinant reptilian TLR15s in human cells allowed us to perform functional studies. NF-κB reporter assays with TLR15-transfected cells clearly showed that the cloned reptilian TLR15 was proteolytically cleaved and activated by fungal proteases, as has previously been reported for chicken TLR15[15,16]. Chicken and reptilian TLR15 are predicted to share many structural characteristics including a highly conserved region in LRR11 which may be involved in the TLR15 activation process[26].

Given that both crocodilian TLR15 protein sequences are more similar to chicken than to anolis TLR15, we were surprised to find highly variable expression levels between the different TLR15s. Western blot analysis and confocal microscopy indicated high expression of gagaTLR15 and ancaTLR15 in a human and reptilian cell-line while crpoTLR15 and almiTLR15 protein levels were much lower. Investigation of *tlr15* codon usage pointed to a potential molecular basis for this difference in protein expression. The analysis revealed that both crocodilian *tlr15* genes are biased more towards using unpreferred codons than the *tlr15* genes of chicken and anolis. This was especially true for codons for leucine, the most abundant amino acid in TLR15. Codon optimization of both crocodile *tlr15* genes resulted in strongly increased protein levels, indicating that codon bias is an important determinant of TLR15 expression. Practically, these findings demonstrate that codon usage is a significant factor to consider when designing experiments to study TLRs in heterologous expression systems

The biological rationale for the use of unpreferred codons by crocodilian and some other reptilian TLR15s remains to be clarified. We found higher variability in leucine codon usage among TLR15s compared to TLRs that are more conserved among vertebrates such as TLR3, indicating that variation in leucine codon bias is not a general feature of TLR evolution following speciation, but is perhaps more related to the biological role of TLR15. Zhong et al. first described that most human *tlr* genes are not enriched with preferred codons and that this considerably limits TLR expression[29]. Codon bias in mammalian *tlr7* leading to low cytosine-guanine (CG) content was shown to limit *tlr7* transcription and this has been proposed to form a regulatory mechanism to prevent over-expression of TLR7 which can lead to auto-immune disease[30]. Thus, some TLRs may have become biased in codon usage under a selective pressure to maintain suboptimal codons that limit their expression efficiency. However, we identified variation in the biased usage of leucine codons in TLR15 among reptiles. While crocodilian and snake TLR15s are biased to contain more unpreferred leucine codons, most bird and lizard TLR15s contain predominantly preferred leucine codons. Additionally, the biased leucine codon usage among the three relatively close related snake species is inconsistent (Fig. 6A). Given this diversity in TLR15 leucine codon usage, which is not in line with the evolutionary relations among these species, it seems likely that the species-specific TLR15 codon bias is more the result of neutral mutation and drift than of selection. Evolution of codon bias through neutral mutation and drift is common for most genes in higher eukaryotes[31–34].

From the perspective of immune system evolution, it is noteworthy that the variable leucine codon usage among TLR15s coincides with large scale loss of TLR15 from the teleost fish, amphibian and mammalian lineages. Perhaps even more

striking are the identified TLR15-like remnants or complete absence of a TLR15-like sequence in turtles which are genetically closely related to crocodiles and birds[18]. These independent gene loss events in different animal lineages and even among reptiles, suggest multiple moments of redundancy of TLR15 throughout vertebrate evolution. It is possible that in species that lost TLR15, other receptors for microbial proteases have taken over its role[35–37]. In other words, it can be speculated that species-specific codon usage and the persistence of unpreferred codons are part of the onset to gradual functional redundancy and eventually disappearance of TLR15 from a genome, but this awaits detailed analysis of TLR codon usage in relation to the evolutionary history of TLRs across various vertebrate lineages.

Methods and materials

Isolation of reptilian DNA and ethics statement

Whole blood from *Crocodylus porosus* and *Alligator mississippiensis* was collected via the spinal vein[38] with an 18 ga needle and a 3 mL syringe. Blood was immediately transferred to a 4 mL heparin Vaccutainer™, and 200 µL was centrifuged at 2500 × *g* for 5 min. DNA was isolated from the resulting cell pellet using a Qiamp® DSP DNA kit (Qiagen). Genomic DNA was precipitated with 3 M NaOAc in the presence of 70% isopropyl alcohol. Precipitated DNA was washed with 70% EtOH and resuspended in nuclease free water. All procedures related to the handling of crocodilians were conducted as approved by the McNeese State University Animal Care and Use Committee. Genomic DNA of *Anolis carolinensis* was isolated as described[23]. The procedure was approved by the Animal Ethics Committee of Utrecht University (study number 2014.II.04.031).

Plasmid constructs

Phusion high-fidelity DNA polymerase, dNTPs, fast digest restriction endonucleases, T4 DNA ligase, and primers were purchased from Thermo Fisher Scientific. The TLR15 gene of *Anolis carolinensis* was amplified from genomic DNA by touchdown PCR with gene-specific primers listed in Table S4. The purified ancaTLR15 gene was next amplified to add a KpnI restriction and kozak site at the 5' end of the gene and an overlap on a 3 × Hemagglutinin-epitope (HA) sequence on the 3' end of the gene. A 3 × HA sequence was amplified from a pTracer-CMV2ΔGFP/3 × HA vector[39] to add an overlap to ancaTLR15 at the start of the HA sequence and a PacI restriction site at the end of the sequence. The ancaTLR15 gene and 3 × HA sequence were subsequently fused by standard overlap PCR, digested with KpnI and PacI and ligated in pTracer-CMV2ΔGFP to yield ancaTLR15

carrying a C-terminal 3 × HA-tag. TLR15 genes of *Crocodylus porosus* and *Alligator mississippiensis* were amplified from genomic DNA by touchdown PCR (for primers see Table S4), digested with KpnI and NotI and ligated into pTracer-CMV2ΔGFP/3 × HA (from which ancaTLR15 was removed) or a vector with a 3 × FLAG epitope tag (pTracer-CMV2ΔGFP/3×FLAG[40]) to yield crpoTLR15 or almiTLR15 with C-terminal 3 × HA-tag or 3 × FLAG-tag. The gagaTLR15 gene was cut from gagaTLR15-pTracer[15] with KpnI and NotI and ligated into pTracer-CMV2ΔGFP/3 × HA or pTracer-CMV2ΔGFP/3 × FLAG to yield gagaTLR15 with C-terminal 3 × HA-tag or 3 × FLAG-tag. The codon optimized *A. mississippiensis* and *C. porosus* TLR15 genes were synthesized by GeneArt (Thermo Fisher Scientific) and subcloned using KpnI and NotI restriction sites into pTracer-CMV2ΔGFP/3 × FLAG. All constructs were verified by sequencing (Macrogen). TLR15 sequences were deposited in GenBank with the following accession numbers: anca*tlr15* (MH395322), crpo*tlr15* (MH395323), almi*tlr15* (MH395324).

Phylogenetic and synteny analysis

Protein sequences of the TLR1 subfamily of multiple vertebrate species including actinopterygii, sarcopterygii, chondrichtyes, amphibia, mammalia, and reptilia (lepidosaurs and archosaurs, including aves) (Table S1), were identified by BLASTp on the species' ref_seq database of the NCBI using the TIR domain of TLRs from reference species (anole, chicken, human, zebrafish, xenopus) as queries. Full length sequences were collected in FASTA format. When multiple copies of the same annotated TLR (i.e. duplications) were found in the one species, receptors were denoted as TLR2, TLR2-1, TLR2-2 etc. Sequences were aligned with the MUltiple Sequence Comparison by Log- Expectation (MUSCLE) sequence aligner of the EMBL-EBI (https://www.ebi.ac.uk/Tools/msa/muscle/). MEGA7 software[41] was used to construct a phylogenetic tree from the aligned TLR sequences by Maximum Likelihood analysis. The analysis involved 136 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 252 positions in the final dataset. The best fitting substitution model was JTT+G+I (AICc: 65408,13971; BIC: 67720,59789; Gamma: 1,275251655, Invariant: 0,033209398). Two-hundred and fifty bootstrap iterations were run and the tree with the highest log likelihood was exported in Newick format and customized in iTol (https://itol.embl.de)[42]. The genomic region surrounding the *tlr15* locus was inspected for neighboring genes by searching the NCBI Gene database with the gagaTLR15 gene ID or the gene ID of predicted TLR1 or 2 sequences of species clustering with gagaTLR15 in the phylogenetic tree. For species in which no TLR15 ortholog was found, the region between *erlec1* and *gpr75* was analyzed by using

BLASTx on NCBIs non-redundant protein sequence database (NR) of *Gallus gallus* (taxid: 9031). Database searches were performed in April 2018.

Protein sequence and codon analyses

Protein sequences were aligned using the Clustal Omega sequence aligner (https://www.ebi.ac.uk/Tools/msa/clustalo/)[43]. To identify the TIR domain[44] the secondary structure was predicted using Jpred4 (http://www.compbio.dundee.ac.uk/jpred4/index_up.html)[45]. Transmembrane region was predicted with http://www.cbs.dtu.dk/services/TMHMM/[46]. LRRs were identified by manual sequence inspection according to[24] and with use of the Leucine rich repeat finder web tool http://www.lrrfinder.com/[47]. Signal peptides were predicted with http://www.cbs.dtu.dk/services/SignalP/[48]. Codon usage tables of the different species were retrieved from https://hive.biochemistry.gwu.edu/dna.cgi?cmd=refseq_processor&id=563612[49].

Cell culture and transient transfection

HEK293 cells were cultured in DMEM (Thermo Fisher Scientific) supplemented with 5% FCS (Bodinco) at 37°C and 10% CO_2. Viper heart (VH-2) cells from a Russell's viper (*Daboia russelli*) were cultured in M199 (Thermo Fisher Scientific) with Hank salts and 10% FCS at 31°C in air. Cells were transiently transfected at 70% confluency with Fugene HD (Promega) at a DNA to Fugene ratio of 1:3 according to the manufacturer's instructions.

Fungal supernatant

Chrysosporium anamorph of *Nannizziopsis vriesii* (CANV) isolated from an agama (lizard) patient was kindly provided by the Veterinary Microbiological Diagnostics Center (VMDC) of the Utrecht University. CANV was grown in 25 ml M199 liquid medium for 7 days at 26°C. Supernatant was collected by centrifugation (3,000 × g, 5 min, room temperature (RT)) and sterilized by passaging through a 0.2 µm filter. Supernatant was stored at 4°C until use (within 24-hours). Prior to addition to cells, CANV supernatant was treated with 1 mM of phenylmethane sulfonyl fluoride (PMSF) for 30 min at RT.

Luciferase NF-κB reporter assay

Cells were transfected in a 12-well plate with 50 ng of an NF-κB-luciferase reporter plasmid and 450 ng of HA-tagged TLR15 plasmid or 225 ng of gagaTLR2B and 225 ng of gagaTLR1A plasmid[14]. Twenty-four hours after transfection cells were redistributed into a 96-well plate. After 24-hours cells were washed twice with

DMEM without FCS and stimulated with: 100 ng/ml of Proteinase K (Sigma), Pam_3CSK_4, FSL-1 (both Invivogen) or 10 µl PMSF treated or untreated CANV culture supernatant in a total of 100 µl DMEM without FCS. After 5 h at 37°C cells were lysed in 50 µl reporter lysis buffer (Promega) at -80°C for 24-hours. After thawing lysate was mixed with luciferase reagent (Promega) and luciferase activity was measured in a TriStar2 luminometer (Berthold). NF-κB activity is represented by luciferase activity in Relative Light Units (RLU). Results were expressed as fold increase in NF-κB activity of stimulated over unstimulated cells.

Confocal microscopy

Cells were transfected in a 12-well plate with 500 ng of HA-tagged TLR15 plasmid. Glass coverslips were coated overnight with 0.02% Poly-L-lysine (Sigma) at RT. Coverslips were washed three times with PBS (Sigma) and 24-hours after transfection cells were seeded onto coated coverslips. Twenty-four hours after seeding onto coverslips cells were washed once with TRIS-buffered saline (TBS) and fixed with TBS/1.5% paraformaldehyde (Affimetrix). Cells were permeabilized and blocked (30 min) with TBS containing 0.1% saponin and 0.2% BSA (both Sigma). Next, cells were incubated (1 h) with Alexa Fluor-488 conjugated mouse α-HA antibody (A21287; Thermo Fisher Scientific) and DAPI (Molecular Probe). After staining cells were washed with TBS and MilliQ and embedded in Prolong Diamond mounting solution (Thermo Fisher Scientific). Cells were imaged on a Leica SPE-II laser confocal microscope and images were processed with Leica LAS AF software.

TLR15 expression and cleavage

HEK293 and VH-2 cells were transfected in a 6-well plate with 1,000 ng of FLAG-tagged TLR15 plasmids. After 48-hours HEK293 cells were washed with DMEM and incubated (1 h, 37°C) with 250 ng/ml Proteinase K in DMEM without FCS. HEK293 and VH-2 cells were lysed in lysis buffer (25 mM TRIS, 150 mM NaCl, 0.5% NP-40, 1 mM EDTA, 5% glycerol, 1 cOmplete protease inhibitor tablet (Roche)), centrifuged (3,000 × g, 3 min, RT) and total protein concentration in the supernatant was measured by BCA assay (Thermo Fisher Scientific). Samples were equalized and run on SDS-PAGE gel and blotted onto PVDF membrane. Membranes were blocked with 5% nonfat milk in TBS-Tween (0.1%) and incubated (1 h, RT) with M2 mouse α-FLAG antibody (F3165; Sigma) or rabbit α-human Beta actin (bs-0061R; Bioss) followed by incubation (1 h, RT) with goat α-mouse (A2304; Sigma) or goat α-rabbit (A4914; Sigma) HRP conjugated antibody. HRP chemiluminescence was detected with Clarity western ECL (Bio-rad). The rabbit α-

human Beta actin cross reacts with a specific protein in VH-2 cell lysate which is likely actin of *D. russelli* due to very high evolutionary conservation of actin.

Acknowledgments

We kindly thank the Veterinary Microbiological Diagnostics Center (VMDC) for providing the CANV isolate, Berend Snel (Theoretical Biology and Bioinformatics, Department of Biology, Science Faculty, Utrecht University, Utrecht, The Netherlands) for help with conceptualizing the bioinformatics and Chris Deiter of Crocodile Encounters (Angleton, Texas, USA) for allowing the collection of blood from crocodilian species.

Author contributions

CV, JW and JvP designed research, MM and CV isolated reptilian DNA, CV performed research, CV analyzed data, CV wrote manuscript with support from all authors.

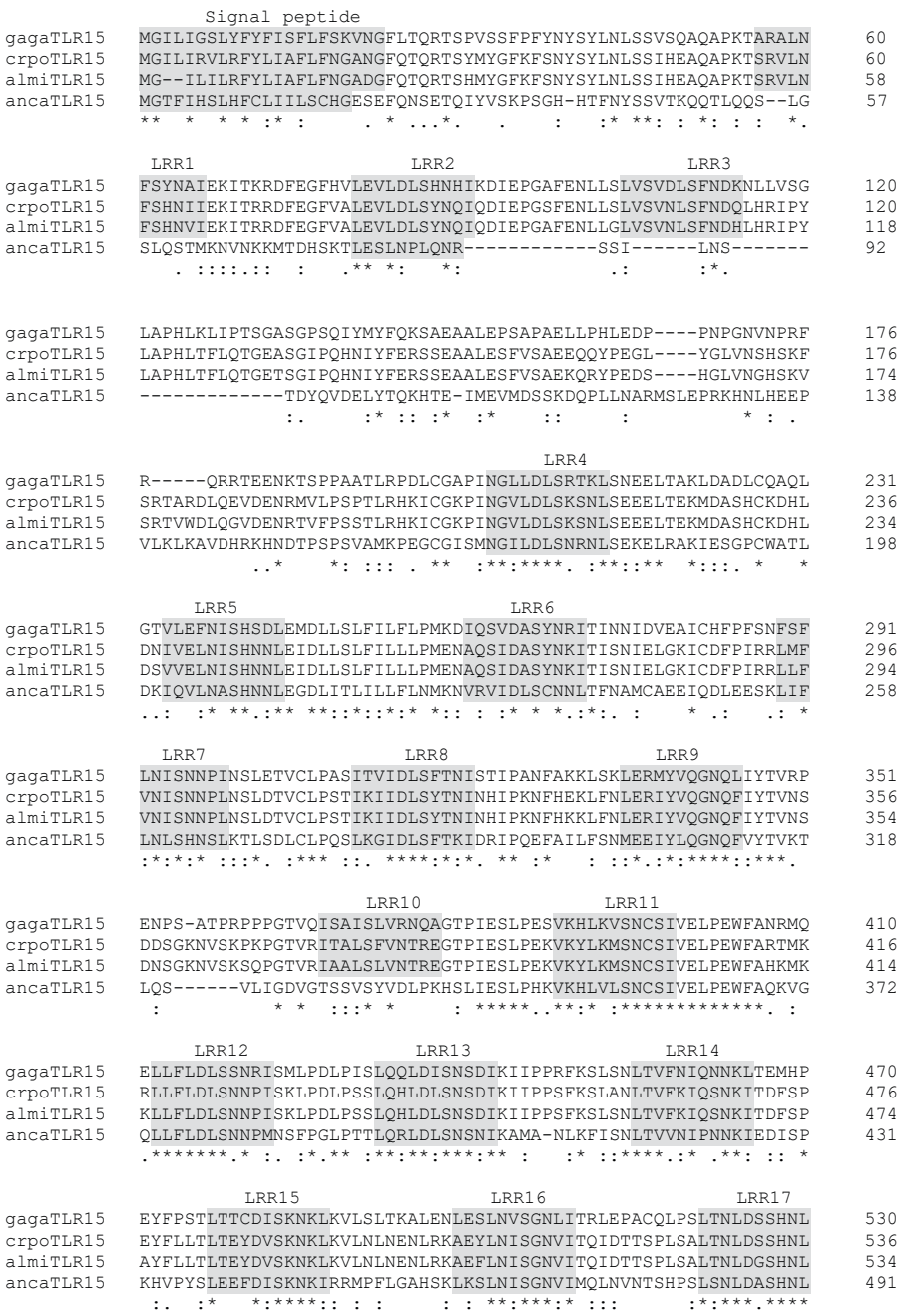
References

1. Medzhitov, R. Toll-like receptors and innate immunity. *Nature Reviews Immunology* **1,** 135–45 (2001).
2. Gay, N. J. & Gangloff, M. Structure and function of Toll receptors and their ligands. *Annual Review of Biochemistry* **76,** 141–165 (2007).
3. Leulier, F. & Lemaitre, B. Toll-like receptors--taking an evolutionary approach. *Nature Reviews Genetics* **9**, 165–78 (2008).
4. Roach, J. C. *et al.* The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 9577–9582 (2005).
5. Rauta, P. R., Samanta, M., Dash, H. R., Nayak, B. & Das, S. Toll-like receptors (TLRs) in aquatic animals: Signaling pathways, expressions and immune responses. *Immunology Letters* **158**, 14–24 (2014).
6. Alcaide, M. & Edwards, S. V. Molecular evolution of the Toll-like receptor multigene family in birds. *Molecular Biology and Evolution* **28**, 1703–1715 (2011).
7. Jin, M. S. *et al.* Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* **130**, 1071–1082 (2007).
8. Kang, J. Y. *et al.* Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. *Immunity* **31**, 873–884 (2009).
9. Kiura, K. *et al.* The synthetic analogue of mycoplasmal lipoprotein FSL-1 induces dendritic cell maturation through Toll-like receptor 2. *FEMS Immunology & Medical Microbiology* **46**, 78–84 (2006).
10. Kruithof, E. K., Satta, N., Liu, J. W., Dunoyer-Geindre, S. & Fish, R. J. Gene conversion limits divergence of mammalian TLR1 and TLR6. *BMC Evolutionary Biology* **7**, 148 (2007).

11. Fink, I. R. *et al.* Molecular and functional characterization of Toll-like receptor (Tlr)1 and Tlr2 in common carp (*Cyprinus carpio*). *Fish & Shellfish Immunology* **56**, 70–83 (2016).
12. Solbakken, M. H. *et al.* Evolutionary redesign of the Atlantic cod (*Gadus morhua* L.) Toll-like receptor repertoire by gene losses and expansions. *Scientific Reports* **6**, 25211 (2016).
13. Velová, H., Gutowska-Ding, M. W., Burt, D. W., Vinkler, M. & Yeager, M. Toll-like receptor evolution in birds: gene duplication, pseudogenisation and diversifying selection. *Molecular Biology and Evolution* **35**, 2170-2184, (2018).
14. Keestra, A. M., de Zoete, M. R., van Aubel, R. A. & Putten, J. P. M. van. The central leucine-rich repeat region of chicken TLR16 dictates unique ligand specificity and species-specific interaction with TLR2. *Journal of Immunology* **178**, 7110–9 (2007).
15. Zoete, M. R. de, Bouwman, L. I., Keestra, A. M. & Putten, J. P. M. van. Cleavage and activation of a Toll-like receptor by microbial proteases. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 4968–4973 (2011).
16. Boyd, A. C. *et al.* TLR15 Is unique to avian and reptilian lineages and recognizes a yeast-derived agonist. *Journal of Immunology* **189**, 4930–4938 (2012).
17. Higgs, R. *et al.* Induction of a novel chicken Toll-like receptor following *Salmonella enterica* serovar Typhimurium infection. *Infection and Immunity.* **74**, 1692–1698 (2006).
18. Wang, Z. *et al.* The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. *Nature Genetics* **45**, 701–706 (2013).
19. Zimmerman, L. M., Vogel, L. A. & Bowden, R. M. Understanding the vertebrate immune system: insights from the reptilian perspective. *Journal of Experimental Biology* **213**, 661–671 (2010).
20. Merchant, M., Fleury, L., Rutherford, R. & Paulissen, M. Effects of bacterial lipopolysaccharide on thermoregulation in green anole lizards (*Anolis carolinensis*). *Veterinary Immunology and Immunopathology* **125**, 176–181 (2008).
21. Merchant, M., Williams, S., Trosclair III, P. L., Elsey, R. M. & Mills, K. Febrile response to infection in the American alligator (*Alligator mississippiensis*). *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **148**, 921–925 (2007).
22. Merchant, M. E. *et al.* Comparisons of innate immune activity of all known living crocodylian species. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **143**, 133–137 (2006).
23. Voogdt, C. G. P., Bouwman, L. I., Kik, M. J. L., Wagenaar, J. A. & Putten, J. P. M. van. Reptile Toll-like receptor 5 unveils adaptive evolution of bacterial flagellin recognition. *Scientific Reports* **6**, 19046 (2016).
24. Matsushima, N. *et al.* Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors. *BMC Genomics* **8**, 124 (2007).
25. Johnson, R. S. P., Sangster, C. R., Sigler, L., Hambleton, S. & Paré, J. A. Deep fungal dermatitis caused by the *Chrysosporium* anamorph of *Nannizziopsis vriesii* in captive coastal bearded dragons (*Pogona barbata*). *Australian Veterinary Journal* **89**, 515–519 (2011).
26. Wang, J., Zhang, Z., Chang, F. & Yin, D. Bioinformatics analysis of the structural and evolutionary characteristics for toll-like receptor 15. *Peer Journal* **4**, e2079 (2016).
27. Hedges, S. B., Marin, J., Suleski, M., Paymer, M. & Kumar, S. Tree of life reveals clock-like speciation and diversification. *Molecular Biology and Evolution* **32**, 835–845 (2015).
28. Benton, M. J. & Donoghue, P. C. J. Paleontological evidence to date the tree of life. *Molecular Biology and Evolution* **24**, 26–53 (2007).
29. Zhong, F. *et al.* Deviation from major codons in the Toll-like receptor genes is associated with low Toll-like receptor expression. *Immunology* **114**, 83–93 (2005).
30. Newman, Z. R., Young, J. M., Ingolia, N. T. & Barton, G. M. Differences in codon bias and GC content contribute to the balanced expression of TLR7 and TLR9. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E1362-1371 (2016).
31. Chamary, J. V., Parmley, J. L. & Hurst, L. D. Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nature Reviews Genetics* **7**, 98–108 (2006).

32. Hershberg, R. & Petrov, D. A. Selection on codon bias. *Annual Review of Genetics* **42**, 287–299 (2008).
33. Plotkin, J. B. & Kudla, G. Synonymous but not the same: the causes and consequences of codon bias. *Nature Reviews Genetics* **12**, 32–42 (2011).
34. Reis, M. dos, Savva, R. & Wernisch, L. Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Research* **32**, 5036–5044 (2004).
35. Ossovskaya, V. S. & Bunnett, N. W. Protease-activated receptors: Contribution to physiology and disease. *Physiological Reviews* **84**, 579–621 (2004).
36. Chavarría-Smith, J., Mitchell, P. S., Ho, A. M., Daugherty, M. D. & Vance, R. E. Functional and evolutionary analyses identify proteolysis as a general mechanism for NLRP1 inflammasome activation. *PLOS Pathogens* **12**, e1006052 (2016).
37. LaRock, C. N. *et al.* IL-1β is an innate immune sensor of microbial proteolysis. *Science Immunology* **1**, (2016).
38. Zippel, K. C., Lillywhite, H. B. & Mladinich, C. R. J. Anatomy of the crocodilian spinal vein. *Journal of Morphology* **258**, 327–335 (2003).
39. Voogdt, C. G. P., Wagenaar, J. A. & Putten, J. P. M. van. Duplicated TLR5 of zebrafish functions as a heterodimeric receptor. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E3221-E3229 (2018).
40. Keestra, A. M., Zoete, M. R. de, Bouwman, L. I. & Putten, J. P. M. van. Chicken TLR21 is an innate CpG DNA receptor distinct from mammalian TLR9. *Journal of Immunology* **185**, 460–467 (2010).
41. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**, 1870–1874 (2016).
42. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research* **44**, W242–W245 (2016).
43. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**, 539 (2011).
44. Xu, Y. *et al.* Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature* **408**, 111–115 (2000).
45. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Research* **43**, W389–W394 (2015).
46. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology* **305**, 567–580 (2001).
47. Offord, V., Coffey, T. J. & Werling, D. LRRfinder: A web application for the identification of leucine-rich repeats and an integrative Toll-like receptor database. *Developmental & Comparative Immunology* **34**, 1035–1041 (2010).
48. Petersen, T. N., Brunak, S., Heijne, G. von & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **8**, 785–786 (2011).
49. Athey, J. *et al.* A new and updated resource for codon usage tables. *BMC Bioinformatics* **18**, 391 (2017).
50. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences* **8**, 275–282 (1992).

Supplement


Signal peptide
LRR1
LRR2
LRR3
LRR4
LRR5
LRR6
LRR7
LRR8
LRR9
LRR10
LRR11
LRR12
LRR13
LRR14
LRR15
LRR16
LRR17

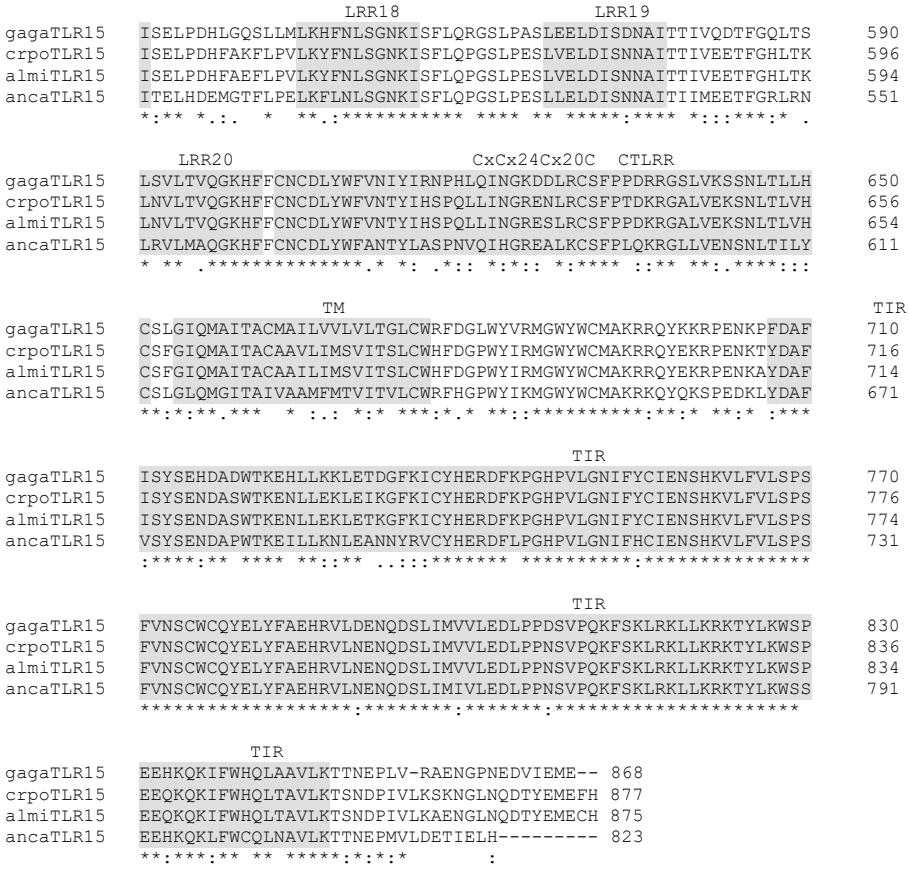


Figure S1. Comparison of TLR15 protein sequences. Amino acid sequences of chicken (gaga), crocodile (cpro), alligator (almi) and anolis (anca) TLR15 were aligned using the Clustal Omega server with default settings. Asterisks (*) indicate identical residues, double square dots (:) indicate highly similar residues, single square dots (.) indicate somewhat similar residues, and bars (–) indicate gaps to complete the sequence alignment. Signal peptide, LRRs, CTLRR, TM and TIR domain are shaded in gray.

Table S1. Species of which TLR1 subfamily protein sequences were collected from GenBank

Species	Abbreviation	Taxonomic group	Common name
Danio rerio	dare	actinopterygii (post TS-WGD*)	Zebrafish
Oryzias latipes	orla	actinopterygii (post TS-WGD)	Medaka
Lepisosteus oculatus	leoc	actinopterygii (pre TS-WGD)	Spotted gar
Nanorana parkeri	napa	amphibia	High Himalaya frog
Xenopus tropicalis	xetr	amphibia	Western clawed frog
Callorhinchus milii	cami	chondrichtyes	Australian ghost shark
Homo sapiens	hosa	mammalia	Human
Mus musculus	mumu	mammalia	Mouse
Sus scrofa	susc	mammalia	Pig
Myotis brandtii	mybr	mammalia	Brandt's bat
Orcinus orca	oror	mammalia	Killer whale
Chelonia mydass	chmy	reptilia	Green sea turtle
Chrysemys picta bellii	chpi	reptilia	Painted turtle
Pelodiscus sinensis	pesi	reptilia	Chinese softshell turtle
Alligator mississippiensis	almi	reptilia (archosaur)	American alligator
Alligator sinensis	alsi	reptilia (archosaur)	Chinese alligator
Crocodylus porosus	crpo	reptilia (archosaur)	Salt water crocodile
Gavialis gangeticus	gavga	reptilia (archosaur)	Indian Gharial
Anas platyrhynchos	anpl	reptilia (archosaur, aves)	Mallard duck
Columba livia	coli	reptilia (archosaur, aves)	Rock pigeon
Falco peregrinus	fape	reptilia (archosaur, aves)	Peregrine falcon
Gallus gallus	gaga	reptilia (archosaur, aves)	Chicken
Struthio camelus australis	stca	reptilia (archosaur, aves)	South African ostrich
Taeniopygia guttata	tagu	reptilia (archosaur, aves)	Zebrafinch
Anolis carolinensis	anca	reptilia (lepidosaur)	Green anole lizard
Gekko japonicus	geja	reptilia (lepidosaur)	Schlegel's Japanese gecko
Pogona vitticeps	povi	reptilia (lepidosaur)	Bearded dragon
Protobothrops mucrosquamatus	prmu	reptilia (lepidosaur)	Brown spotted pit viper
Python bivittatus	pybi	reptilia (lepidosaur)	Burmese python
Thamnophis sirtalis	thsi	reptilia (lepidosaur)	Common garter snake
Latimeria chalumnae	lach	sarcopterygii	Coelacanth
Ciona intestinalis	ciin	urochordate (non-vertebrate)	Sea squirt

*TS-WGD: teleost specific whole genome duplication

Table S2. Identity matrix of full length TLR15 protein sequences

Identity (%)	gagaTLR15	crpoTLR15	almiTLR15	ancaTLR15
gagaTLR15	100	69	70	59
crpoTLR15		100	95	52
almiTLR15			100	61
ancaTLR15				100

Table S3. Percentage (%) and absolute number (Counts) of codons per amino acid in the TLR15 genes of the indicated species in relation to the percentage of codon usage in the human genome. Per amino acid, the most frequent codon in the human genome is color coded in blue and the least frequent codon is color coded in red.

			TLR15							
			G. gallus		*A. carolinensis*		*C. porosus*		*A. mississippiensis*	
Amino acid	codon	% in human genome	%	Counts	%	Counts	%	Counts	%	Counts
ALA	GCA	24,776	45,238	19	25,000	7	41,935	13	39,394	13
	GCG	9,203	0,000	0	3,571	1	0,000	0	0,000	0
	GCC	38,471	23,810	10	39,286	11	12,903	4	21,212	7
	GCT	27,550	30,952	13	32,143	9	45,161	14	39,394	13
ARG	AGA	23,320	48,485	16	39,130	9	48,387	15	42,857	12
	AGG	21,662	27,273	9	17,391	4	25,806	8	32,143	9
	CGA	11,339	0,000	0	13,043	3	3,226	1	7,143	2
	CGG	19,415	9,091	3	13,043	3	9,677	3	7,143	2
	CGC	16,170	6,061	2	4,348	1	3,226	1	3,571	1
	CGT	8,094	9,091	3	13,043	3	9,677	3	7,143	2
ASN	AAC	50,301	49,153	29	56,923	37	31,429	22	32,353	22
	AAT	49,699	50,847	30	43,077	28	68,571	48	67,647	46
ASP	GAC	50,851	46,341	19	53,125	17	48,649	18	42,105	16
	GAT	49,149	53,659	22	46,875	15	51,351	19	57,895	22
CYS	TGC	51,452	55,556	10	42,105	8	25,000	4	35,294	6
	TGT	48,548	44,444	8	57,895	11	75,000	12	64,706	11
GLN	CAA	28,135	41,935	13	55,882	19	51,724	15	50,000	14
	CAG	71,865	58,065	18	44,118	15	48,276	14	50,000	14
GLU	GAA	45,289	63,265	31	67,308	35	73,684	42	76,364	42
	GAG	54,711	36,735	18	32,692	17	26,316	15	23,636	13
GLY	GGA	26,775	32,258	10	23,333	7	41,935	13	41,176	14
	GGG	24,445	25,806	8	30,000	9	19,355	6	17,647	6
	GGC	31,859	25,806	8	30,000	9	19,355	6	23,529	8
	GGT	16,921	16,129	5	16,667	5	19,355	6	17,647	6
HIS	CAC	55,719	52,174	12	48,387	15	38,462	10	33,333	10
	CAT	44,281	47,826	11	51,613	16	61,538	16	66,667	20
ILE	ATA	18,421	27,273	15	22,000	11	35,938	23	38,710	24
	ATC	43,837	29,091	16	36,000	18	21,875	14	17,742	11
	ATT	37,742	43,636	24	42,000	21	42,188	27	43,548	27
LEU	CTA	7,531	10,924	13	5,217	6	12,613	14	10,811	12
	CTG	37,520	34,454	41	26,957	31	17,117	19	21,622	24
	CTC	18,439	16,807	20	16,522	19	17,117	19	16,216	18
	CTT	14,224	15,966	19	16,522	19	16,216	18	16,216	18
	TTA	8,724	10,924	13	13,043	15	25,225	28	19,820	22
	TTG	13,561	10,924	13	21,739	25	11,712	13	15,315	17
LYS	AAA	45,960	60,417	29	60,714	34	75,439	43	77,586	45
	AAG	54,040	39,583	19	39,286	22	24,561	14	22,414	13
MET	ATG	100,000	100,000	15	100,000	25	100,000	14	100,000	12
PHE	TTC	51,173	45,652	21	48,571	17	26,531	13	25,000	12
	TTT	48,827	54,348	25	51,429	18	73,469	36	75,000	36
PRO	CCA	29,335	37,500	21	48,649	18	42,857	18	46,341	19
	CCG	10,118	8,929	5	0,000	0	2,381	1	0,000	0

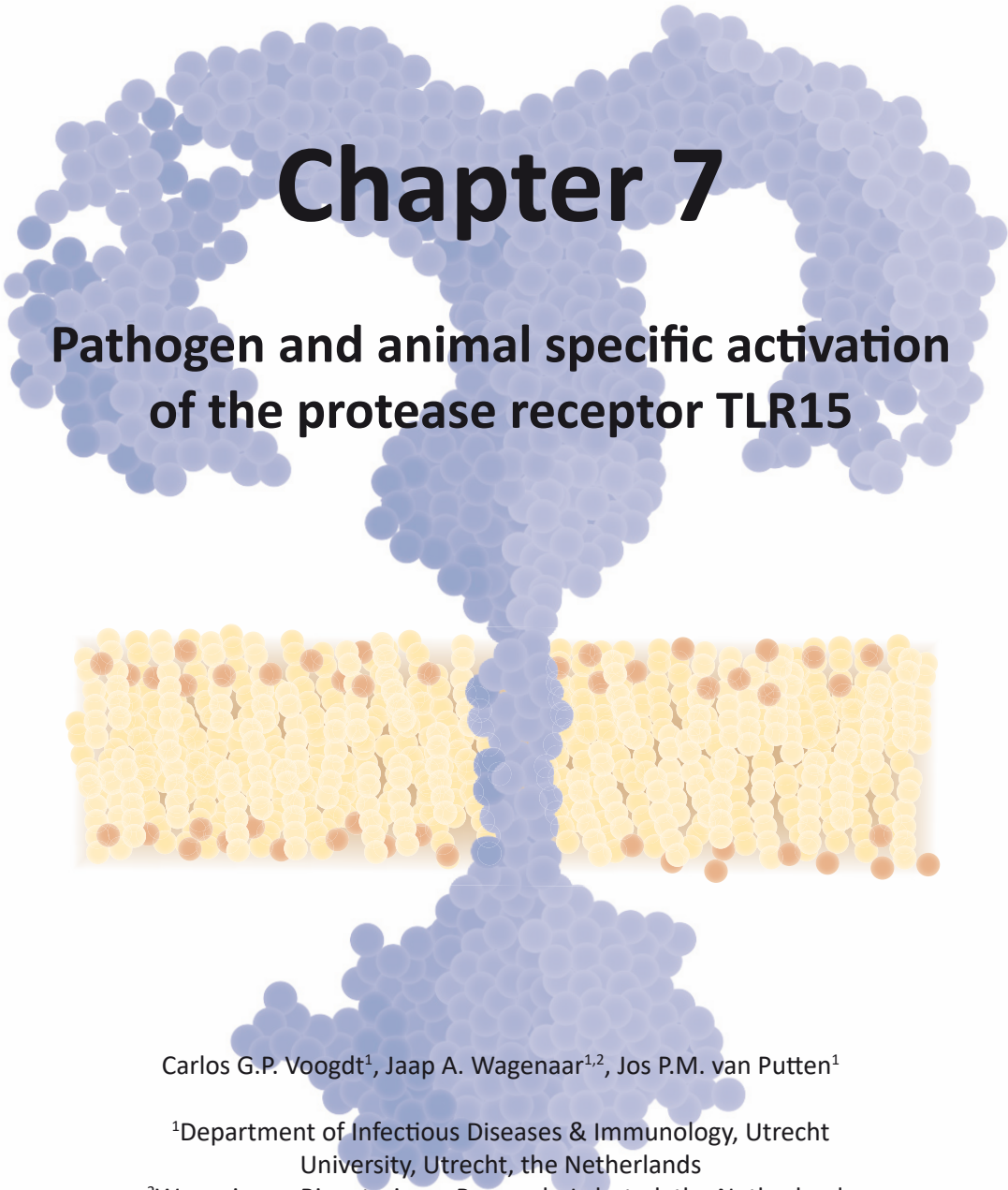
	CCC	30,528	23,214	13	24,324	9	26,190	11	24,390	10
	CCT	30,019	30,357	17	27,027	10	28,571	12	29,268	12
SER	AGC	23,091	28,571	24	24,359	19	25,610	21	21,687	18
	AGT	16,029	14,286	12	19,231	15	17,073	14	19,277	16
	TCA	16,228	15,476	13	19,231	15	24,390	20	22,892	19
	TCG	4,887	2,381	2	7,692	6	0,000	0	2,410	2
	TCC	20,391	17,857	15	11,538	9	8,537	7	9,639	8
	TCT	19,373	21,429	18	17,949	14	24,390	20	24,096	20
THR	ACA	30,155	36,585	15	42,105	16	44,186	19	45,238	19
	ACG	10,569	9,756	4	0,000	0	9,302	4	9,524	4
	ACC	33,186	24,390	10	21,053	8	18,605	8	14,286	6
	ACT	26,090	29,268	12	36,842	14	27,907	12	30,952	13
TRP	TGG	100,000	100,000	10	100,000	11	100,000	10	100,000	11
TYR	TAC	53,245	52,174	12	52,381	11	40,625	13	44,828	13
	TAT	46,755	47,826	11	47,619	10	59,375	19	55,172	16
VAL	GTA	12,803	29,545	13	23,256	10	35,556	16	31,915	15
	GTG	44,514	29,545	13	32,558	14	24,444	11	25,532	12
	GTC	23,046	15,909	7	18,605	8	6,667	3	6,383	3
	GTT	19,637	25,000	11	25,581	11	33,333	15	36,170	17
STOP	TAA	27,439	0,000	0	100,000	1	0,000	0	0,000	0
	TAG	21,951	0,000	0	0,000	0	100,000	1	100,000	1
	TGA	50,610	100,000	1	0,000	0	0,000	0	0,000	0
Total				869		824		878		876
Sum of most frequent codons			35,67	310	34,83	287	26,42	232	25,91	227
Sum of least frequent codons			27,62	240	27,67	228	36,22	318	36,30	318

Table S4. Primers used for cloning reptilian TLR15

Product	Primer	Sequence (5' – 3')
ancaTLR15	Forward	ATGGGAACTTTCATCCACAGTCTGCA
	Reverse	ATGCAGCTCGATTGTTTCATCC
KpnI-kozak-ancaTLR15-3xHA-overlap	Forward	CCGGTACCGCCACCATGGGAACTTTCATCCACAGTCTGCA
	Reverse	CATATGGGTAGCGGCCGCTATGCAGCTC
ancaTLR15-overlap-3xHA-PacI	Forward	GCATAGCGGCCGCTACCCATATGACGTTCCAG
	Reverse	CCTTAATTAATCAAGCGTAGTCAGGTACATCG
KpnI-kozak-crpoTLR15-NotI	Forward	CCGGTACCGCCACCATGGGTATCCTCATCAGAGTTC
	Reverse	CCGCGGCCGCGGTGAAATTCCATCTCATATGTATCTTG
KpnI-kozak-almiTLR15-NotI	Forward	CCGGTACCGCCACCATGGGTATCCTCATTCTTCG
	Reverse	CCGCGGCCGCGGTGGCATTCCATCTCATATG

Chapter 7

Pathogen and animal specific activation of the protease receptor TLR15



Carlos G.P. Voogdt[1], Jaap A. Wagenaar[1,2], Jos P.M. van Putten[1]

[1]Department of Infectious Diseases & Immunology, Utrecht University, Utrecht, the Netherlands
[2]Wageningen Bioveterinary Research, Lelystad, the Netherlands

Abstract

During infection diverse micro-organisms use proteases to shape their niche and to manipulate host immune responses. As a countermeasure, hosts have evolved receptors that directly detect the presence of microbial proteases. Toll-like receptor 15 (TLR15), which has been lost from most vertebrates but is conserved in birds and most reptiles, is activated by microbial proteolytic activity. The specificity of this event however, is still largely unknown. Here we compared the activation of TLR15 from chicken, green anole lizard, Australian salt water crocodile, and American alligator in transfected HEK293 cells after exposure of the cells to bacterial and fungal pathogens. TLR15 activation was found to be strongly species-specific, in terms of both the receptors and pathogens tested. Exposure to *Pseudomonas aeruginosa*, but not fungal pathogens, revealed a strong difference in TLR15 activation even between two highly homologous TLR15 orthologs. Functional analysis of chimeric receptors constructed of these highly homologous orthologs indicated that the N-terminal part of the receptor ectodomain, which is most variable, is not responsible for the species-specific response. Testing of several *P. aeruginosa* mutants lacking distinct proteases for their ability to activate TLR15, identified one specific secreted protease as the predominant TLR15 agonist. Our findings indicate that throughout the evolution of birds and reptiles, TLR15 evolved a species-specific sensitivity to detect diverse microbes including a virulence factor of the opportunistic pathogen *P. aeruginosa*.

Introduction

Pathogenic micro-organisms including bacteria, fungi and parasites often use proteases as virulence factors to evade, block, or alter host proteins that are necessary for inducing an appropriate immune response. Numerous strategies of microbial protease-mediated manipulation of host responses have been identified. Reported strategies include degradation of immunoglobulins by bacterial IgA1 proteases[1], degradation of monomeric flagellin from *Pseudomonas aeruginosa* by its alkaline protease to avoid detection of flagellin by host receptors[2], and degradation of host complement factors by Alp1 alkaline protease of *Aspergillus fumigates*[3]. The host can detect microbial proteases by the activation of signaling pathways following proteolytic cleavage of different sensory proteins. For example, pro-IL-1β, the immature form of the IL-1β pro-inflammatory cytokine, is cleaved by a cysteine protease of Group A *Streptococcus* (GAS) and thus functions as a direct sensor of microbial proteolytic activity[4]. In addition, protease activated receptors (PARs), a family of four highly conserved G-protein coupled receptors that are involved in several physiological processes[5,6], can initiate signaling upon cleavage of their N-terminus by proteases of multiple bacterial species including *P. aeruginosa* and *Porphyromonas gingivalis*[7–10].

Another receptor that is activated by proteases is Toll-like receptor 15 (TLR15). TLR15 is part of the TLR family of innate immune receptors that directly sense microbe associated molecular patterns and activate pro-inflammatory transcription factors such as Nuclear factor κB (NF-κB) that initiate an immune response[11]. Structurally, TLRs are composed of a Toll/interleukin-1 (TIR) signaling domain, a single transmembrane domain, and an ectodomain (ECD) involved in ligand interaction. By evolutionary diversification of the ECD, distinct TLR subfamilies have been formed that together detect a wide variety of microbial structures. These structures are often highly conserved and essential for microbial viability. Well known TLR ligands are lipopolysaccharides (TLR4), proteins (flagellin, TLR5), cell wall components (TLR1, 2 and 6) and nucleic acids (TLR3, 7-9)[12,13]. The importance of detecting these types of microbial ligands is exemplified by the strong conservation of their respective TLRs throughout vertebrate evolution[11]. Some TLRs evolved more dynamically and are now restricted to certain clades of vertebrate species. The protease receptor TLR15, for example, has been lost from teleost fish, amphibians and mammals but has been retained in birds and reptiles. Among these animals the TLR15 gene sequence displays a large species-specific variation in codon usage bias[14]. Yet, despite these signs of evolutionary regression the TLR15 gene of birds and reptiles still encodes a functional receptor

that is activated by proteolytic cleavage. Whether TLR15 of birds and reptiles can be activated in a protease- and species-specific manner and thus respond to different microbes is unknown.

Previously we and others found that chicken and reptilian TLR15 can activate NF-κB upon exposure to pathogenic fungi[14–16]. Here we report that chicken and reptilian TLR15s are activated by additional microbial species and searched for possible species-specific receptor activation.

Results

Activation of TLR15 is species-specific

In search for novel microbial TLR15 agonists that may reveal species-specific activation of TLR15, we first investigated whether TLR15 of chicken and different reptilians can be activated by distinct microbial species. Hereto, HEK293 cells were transfected with an NF-κB luciferase reporter plasmid and an expression plasmid containing the gene encoding either chicken (*Gallus gallus*, gaga), green anole lizard (*Anolis carolinensis*, anca), Australian salt water crocodile (*Crocodylus porosus*, crpo) or American alligator (*Alligator mississippiensis*, almi) TLR15. Control cells were transfected with the reporter plasmid and an empty expression vector. Transfected cells were exposed to cell-free culture supernatant of different micro-organisms including *P. aeruginosa* strain PAO1 and clinical isolates of *Aeromonas hydrophila* (bacterium), *Candida guilliermondii* (yeast) and *Chrysosporium* anamorph of *Nannizziopsis vriesii* (CANV, fungus). CANV supernatant was used as a positive control as it can activate all four species of TLR15 [14]. Empty vector transfected control cells showed no activation of NF-κB against any of the microbial supernatants. As expected, supernatant of CANV induced a strong and similar level of NF-κB activity among the different TLR15 expressing cells (Fig. 1). Supernatants of the other microbial species also induced NF-κB in TLR15 transfected cells although the level of NF-κB activation varied depending on which TLR15 species was expressed. The *C. guilliermondii* supernatant potently induced NF-κB in gagaTLR15, crpoTLR15 and almiTLR15 expressing cells but induction was much lower in ancaTLR15 expressing cells. The supernatant of *A. hydrophila* did not stimulate gagaTLR15 and ancaTLR15 expressing cells, but was able to activate NF-κB in cells expressing crpoTLR15 and almiTLR15, be it marginally. The response to *P. aeruginosa* supernatant was most contrasting as gagaTLR15 and almiTLR15 expressing cells showed strong induction of NF-κB compared to the poor NF-κB activation in cells expressing crpoTLR15 and ancaTLR15 (Fig. 1). These results indicate that TLR15 of reptiles and birds can be activated by secreted

factors of different prokaryotic and eukaryotic micro-organisms but also that the activation of TLR15 displays species-specificity.

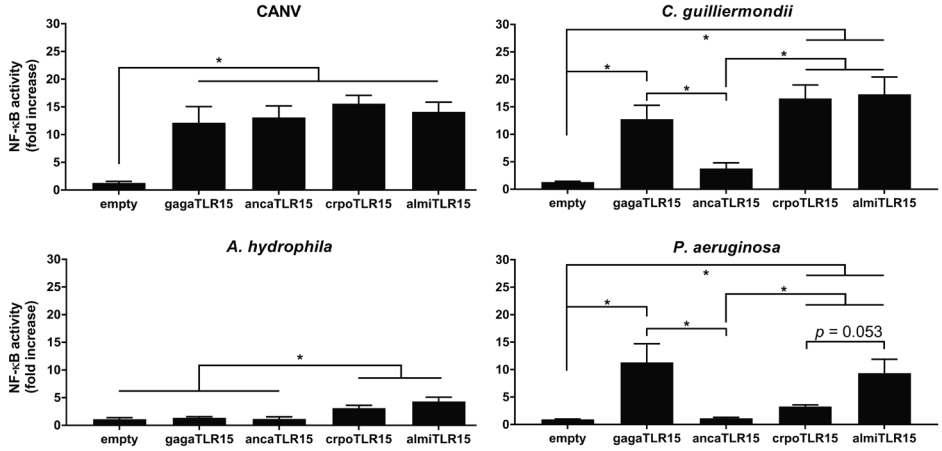


Figure 1. Species-specificity in TLR15 activation. HEK293 cells were transfected with an NF-κB luciferase reporter construct and expression plasmids containing either chicken (gaga), anolis (anca), crocodile (crpo) or alligator (almi) TLR15. Cells were exposed (5 h) to cell-free culture supernatant (10%) of *Chrysosporium* anamorph of *Nannizziopsis vriesii* (CANV), *Candida guilliermondii*, *Aeromonas hydrophila* and *Pseudomonas aeruginosa*. Values are the mean ± SEM fold increase of NF-κB activity in stimulated cells over unstimulated control cells from four independent experiments performed in duplicate. Differences with $p < 0.05$ by unpaired Student's t-test are indicated with an asterisk.

Mapping of the TLR15 region involved in species-specific activation

To learn more about the molecular basis of the species-specific TLR15 activation, we focused on the strong difference in *P. aeruginosa* induced NF-κB activity between cells expressing the relatively closely related crpoTLR15 and almiTLR15. Crocodiles and alligators have diverged approximated 87 million years ago[17] and evolve relatively slowly[18]. A sequence alignment of their TLR15 proteins, which are 877 (crocodile) and 875 (alligator) amino acids in length, indicated that the receptors are 95% identical with an overall similarity of even 97% (Fig. 2). The majority of the 48 amino acids that differ between crpoTLR15 and almiTLR15 are located in the N-terminal part of the ECD (27/48). To determine whether this region of the receptors is responsible for the differential sensitivity to *P. aeruginosa* supernatant, we constructed chimeric receptors in which the N-terminal part of the ECD of one species was fused to the ECD C-terminal part of the other species. This yielded the chimeras al-crTLR15 and cr-alTLR15. The chimera al-crTLR15 carried the N-terminal ECD of alligator and the C-terminal ECD, transmembrane and TIR domain of crocodile TLR15. The cr-alTLR15 chimera had the opposite configuration (Fig. 3A). Exposure of natural (wild type) and chimeric TLR15 transfected cells to CANV

supernatant showed that both chimeric receptors yielded potent NF-κB responses, similar to the wild type receptors (Fig. 3B). In contrast, for *P. aeruginosa* supernatant we unexpectedly observed that chimera cr-alTLR15 responded similar as wild type almiTLR15, while chimera al-crTLR15 mimicked the response of wild type crpoTLR15. This finding indicates that not the (most variable) N-terminal part of the ECD but the 21 different amino acids in the C-terminal part of the ECD, the transmembrane and the signaling domain dictate the species-specific response between crocodile and alligator TLR15 to *P. aeruginosa* supernatant.

The LasB protease of *P. aeruginosa* activates TLR15

As *P. aeruginosa* is a well characterized opportunistic pathogens with a broad host range including birds and reptiles, we attempted to identify the bacterial protease(s) responsible for the observed strong species-specific activation of TLR15. *P. aeruginosa* secretes multiple proteases into the environment including LasA, LasB (elastase), Protease IV, AprA (alkaline protease) and a probable aminopeptidase. To identify the TLR15 activating protease(s), we measured NF-κB activity in cells transfected with empty vector or the different TLR15 orthologs after exposure to the supernatants from wild type *P. aeruginosa* (WT) and isogenic mutants strains lacking *lasA*, *lasB*, *protease IV*, *aprA* and the probable aminopeptidase. As expected, exposure of cells transfected with empty vector or ancaTLR15 to the different supernatants did not result in activation of NF-κB. In contrast, addition of supernatant to gagaTLR15, almiTLR15 and crpoTLR15 expressing cells showed increased NF-κB activation for all mutant strains except for the mutant lacking LasB (elastase). This strongly suggests that the LasB elastase rather than other proteases secreted by *P. aeruginosa*, is the predominant agonist of this bacterium for bird and reptilian TLR15.

Figure 2. Sequence alignment of alligator (almi) and crocodile (crpo) TLR15. Sequences were aligned using Clustal Omega[19]. Asterisks (*) indicate identical residues, double dots (:) indicate residues with strongly similar properties (e.g. FYW), single dots (.) indicate residues with weakly similar properties (e.g. CSA) and bars (-) indicate gaps to complete the sequence alignment. Signal peptide, leucine rich repeats (LRR) 1-20, C-terminal LRR (CTLRR), transmembrane domain (TM), and Toll/interleukine-1 (TIR) domain were identified using[20,21]. The N-terminal (N-term) part, indicated between arrows, of one species was fused to the C-terminal part (C-term) of the other species to create chimeric receptors. Differences between almi and crpoTLR15 are highlighted in yellow.

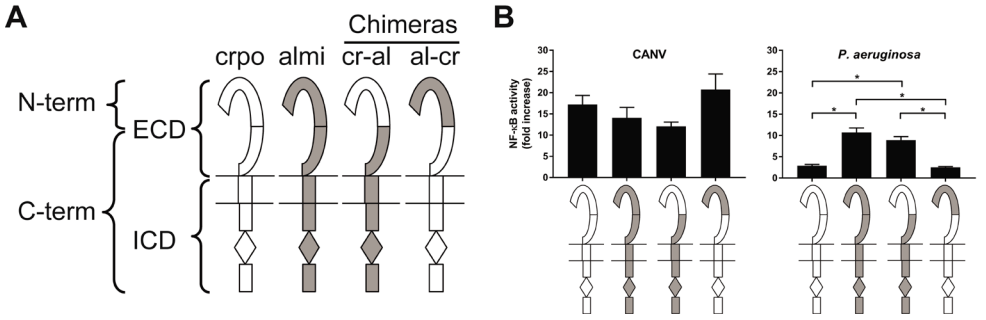


Figure 3. The N-terminus of TLR15 is not involved in species-specific activation by *P. aeruginosa*. (**A**) Schematic representation of wild type crocodile (crpo), alligator (almi), and chimeric TLR15 constructs. ECD: ectodomain, ICD: intracellular domain, N-term: N-terminus of the ECD, C-term: C-terminal part of ECD and ICD. (**B**) HEK293 cells transfected with the indicated receptors were exposed (5 h) to cell-free culture supernatant (10%) of *Chrysosporium* anamorph of *Nannizziopsis vriesii* (CANV) or *P. aeruginosa*. Values are the mean ± SEM fold increase of NF-κB activity in stimulated cells over unstimulated control cells from four independent experiments performed in duplicate. Differences with $p < 0.05$ by unpaired Student's t-test are indicated with *.

Discussion

Numerous microbial pathogens use proteases to attack host proteins in order to avoid or dampen immune responses directed against them. As a countermeasure, hosts have evolved receptors that can directly sense the presence of microbial proteases. In the present study we provide evidence that the TLR15 protease receptor of birds and reptiles is activated by distinct bacterial and fungal pathogens and that the receptor responses to these pathogens are highly species-specific. In addition, we identified the protease LasB as a predominant TLR15 activator secreted by *P. aeruginosa*.

The activation of TLR15 by culture supernatant of two bacterial and two fungal pathogens suggests that TLR15 may function as a broad sensor of microbial proteolytic activity. Our results indicate that besides chicken TLR15[15,16], reptilian TLR15 in general have conserved this broad sensory function. However, despite its conservation as a protease sensor, the activation of the different TLR15s by the different microbes was highly species-specific. For example, almiTLR15 was activated by all microbes tested while ancaTLR15 was only activated by CANV. Furthermore, while gagaTLR15 and crpoTLR15 were both strongly activated by *C. guilliermondii*, the activation of crpoTLR15 was substantially lower than that of gagaTLR15 upon exposure to *P. aeruginosa*. The similar NF-κB activation by all TLR15 orthologs in response to CANV but ortholog-specific induction in response to the other microbes, eliminates differences in receptor expression as an explanation

for the observed species-specificity. Our data therefore suggests that sensitivity of TLR15 to certain microbes has been gained or lost during specific host-microbe co-evolution. In line with this is the observation that both crocodile and alligator TLR15 functionally more resemble chicken than anolis TLR15, which corroborates with crocodiles being more closely related to birds than to lizards. Differences in TLR-ligand interaction between species provide a functional basis to uncover features that are important for receptor functioning[22–26]. In this regard, our finding of a strong difference in the activation of crocodile and alligator TLR15 upon exposure to *P. aeruginosa* is of particular interest. almiTLR15 and crpoTLR15 proteins differ only at 5% of their residues and chimeric constructs revealed that the N-terminus, which harbors most of the variable residues, is not involved in the species-specific response. This indicates that the functional difference between crpoTLR15 and almiTLR15 against *P. aeruginosa* relies on very few, possibly even one, amino acid most likely located in the C-terminal region of the receptor ectodomain. Future studies using additional chimeric constructs and single residue mutants are required to identify these amino acids. This information may then be used to predict sensitivity of TLR15 from other species to *P. aeruginosa* and to investigate the evolutionary conservation of the sensitive sequence among other TLRs. In addition, it may explain the apparent resistance of other TLRs to proteolytic enzymes[16].

The discovery that the secreted protease LasB of *P. aeruginosa* activated TLR15, provides the first identification of a specific TLR15 agonist. The contribution of LasB to *P. aeruginosa* mediated TLR15 activation became evident when a mutant in which the *lasB* gene was deleted could no longer activate TLR15 while deletion of other protease encoding genes still resulted in activation of TLR15. However, the residual response of almiTLR15 to the LasB mutant supernatant suggests that almiTLR15 can also be activated by other proteases produced by *P. aeruginosa*. The finding that crpoTLR15 and gagaTLR15 did not respond to the LasB mutant strongly suggests that LasB of *P. aeruginosa* is the sole activator of these TLR15 orthologs. *P. aeruginosa* is an opportunistic pathogen that can infect reptiles, birds and mammals[27–29]. LasB is a versatile protease that is used as a virulence factor to aid the bacterium in immune evasion[2], tissue destruction[30,31], and immune silencing[32]. The advantage for a host of detecting LasB by TLR15 therefore seems obvious. For example, LasB can cleave thrombin to form an anti-inflammatory peptide that blocks TLR4 function[32]. Direct detection of LasB by TLR15 would counteract the pathogen's attempt of evading a TLR4-mediated immune response.

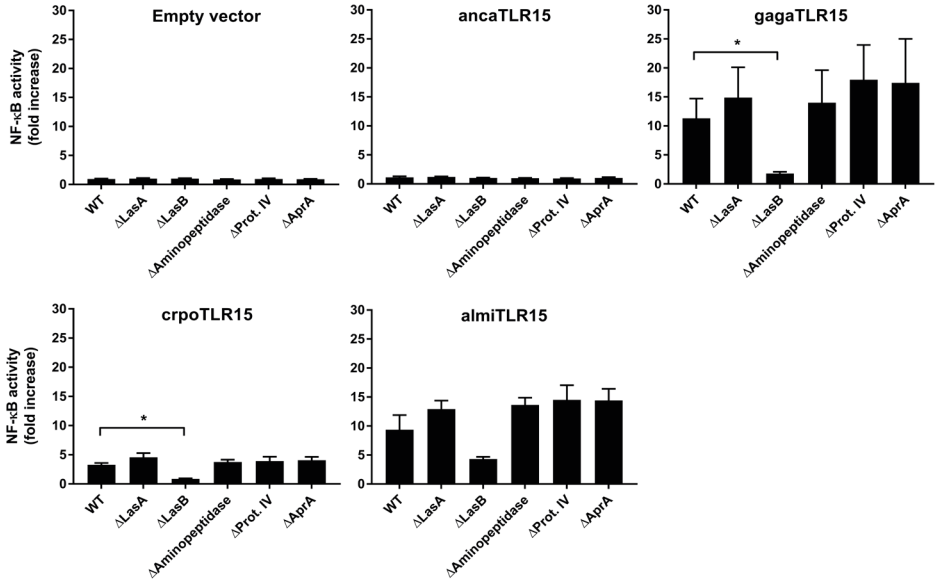


Figure 4. LasB (elastase) of *P. aeruginosa* is required for potent TLR15 activation. HEK293 cells were transfected with the indicated constructs and exposed (5 h) to cell-free culture supernatant of wild type (WT) or protease mutant (Δ) *P. aeruginosa* strains. Values are the mean ± SEM fold increase of NF-κB activity in stimulated cells over unstimulated control cells from four independent experiments performed in duplicate. Differences with $p < 0.05$ by unpaired Student's t-test are indicated with an asterisk.

Yet, despite this theoretically useful role in immunity, TLR15 has been lost multiple independent times throughout vertebrate evolution[14]. In addition, TLR15 of the anole lizard showed a limited microbial detection range compared to the other TLR15 orthologs and did not even respond to *P. aeruginosa*. Unlike other TLRs which detect highly conserved structures that are often essential for microbial viability, TLR15 detects proteases that are used as virulence factors. Virulence factors may evolve more dynamically than e.g. cell-wall structures and indeed there are pathogenic *P. aeruginosa* isolates that naturally lack LasB[32]. It can be imagined that a dynamic occurrence of proteases even among strains of the same microbial species during host-microbe co-evolution may underlie the strong species-specificity in TLR15 activation. In other words, co-evolution of different hosts with microbial strains that may or may not have certain proteases will likely result in species-specific differences in receptor sensitivity due to unique selective pressures exerted during host-microbe interactions. Furthermore, the dynamic evolution of its ligand may even be one of the reasons for the evolutionary regression of TLR15 as it may not provide sufficient and/or consistent selective pressure to force conservation of the receptor in all vertebrate species.

Methods and materials

Plasmid constructs

Construction of plasmids containing gagaTLR15, ancaTLR15 and codon optimized versions of crpoTLR15 and almiTLR15 has been described previously[14]. Phusion high-fidelity DNA polymerase, dNTPs, fast digest restriction endonucleases, T4 DNA ligase, and primers were purchased from Thermo Fisher Scientific. To construct chimeric receptors cr-alTLR15 and al-crTLR15, DNA fragments corresponding to the N-ECD and C-terminal part of the receptors (C-term) were amplified from plasmids containing almiTLR15 and crpoTLR15[14] by PCR with primers (Table 1) to yield almiTLR15 N-ECD, almiTLR15 C-term, crpoTLR15 N-ECD and crpoTLR15 C-term. PCR products were purified, mixed and fused by overlap PCR followed by amplification with forward and reverse primers containing a KpnI and NotI restriction site, respectively. Full-length chimeric TLR15 genes were cut with KpnI and NotI and ligated in a pTracer-CMV2ΔGFP/3 × FLAG vector to yield cr-alTLR15 and al-crTLR15 with a C-terminal 3 × FLAG-tag. Constructs were verified by sequencing (Macrogen).

Table 1. Primers used in this study

Product	Direction	5' to 3' sequence
almiTLR15 N-ECD	Forward	CCTAGTGGTACCGCCACCATGGGAATCCTGATCCTGCG
	Reverse	GTCTGCAGAAAGGTCAGGTGAGGGGCCAGATAAGGGAT
almiTLR15 C-term	Forward	CTGGATCTGTCCAAGAGCAACCTGAGCGAGGAAGAACTG
	Reverse	GTAGCGGCCGCGGTGGCATTCCATCTCGTAGGTGTC
crpoTLR15 N-ECD	Forward	CCTAGTGGTACCGCCACCATGGGCATCCTGATCCGGGTGCTGAG
	Reverse	GTCTGCAGAAAGGTCAGGTGAGGGGCCAGATAAGGGAT
crpoTLR15 C-term	Forward	CTGGATCTGTCCAAGAGCAACCTGAGCGAGGAAGAACTG
	Reverse	ATAGCGGCCGCGGTGAAATTCCATCTCGTAGGTATC
al-crTLR15	Forward	CCTAGTGGTACCGCCACCATGGGAATCCTGATCCTGCG
	Reverse	ATAGCGGCCGCGGTGAAATTCCATCTCGTAGGTATC
cr-alTLR15	Forward	CCTAGTGGTACCGCCACCATGGGCATCCTGATCCGGGTGCTGAG
	Reverse	GTAGCGGCCGCGGTGGCATTCCATCTCGTAGGTGTC

Microbial strains

Clinical isolates of *Aeromonas hydrophila* (turtle), *Candida guilliermondii* (chameleon) and *Chrysosporium* anamorph of *Nannizziopsis vriesii* (CANV, agama) were kindly provided by the Veterinary Microbiological Diagnostics Center (VMDC) of the Utrecht University. *Pseudomonas aeruginosa* reference strain PAO1

protease mutants *lasA* (PA1871), *lasB* (elastase, PA3724), *protease IV* (PA4175), *aprA* (alkaline protease, PA1249) and a probable aminopeptidase (PA2939) were generated by Jacobs et al.[33] (http://www.gs.washington.edu/labs/manoil/libraries.htm). Mutants and the isogenic PAO1 wild type strain were kindly provided by Dr. B. Bardoel of the University Medical Center Utrecht.

Preparation of microbial supernatants

Single colonies of *P. aeruginosa* wild type and mutant strains and *A. hydrophila* were grown in Luria-Bertani (LB) broth (Biotrading) and incubated for 17 h at 31°C. *C. guilliermondii* was grown in 10 ml LB broth for 48 hours at 30°C. CANV was grown in 25 ml M199 liquid medium (Thermo Fisher Scientific) for 7 days at 26°C. Supernatants were collected by centrifugation (3,000 × g, 5 min, room temperature (RT)) and sterilized by passaging through a 0.2 μm filter. Supernatants were stored at 4°C until use (within 24-hours).

Cell culture and transient transfection

HEK293 cells were cultured in DMEM (Thermo Fisher Scientific) supplemented with 5% FCS (Bodinco) at 37°C and 10% CO_2. Cells were seeded in a 12 well-plate and after 24 hours at 70% confluence transiently transfected with 450 ng TLR15 containing plasmid and 50 ng of an NF-κB-luciferase reporter plasmid using Fugene HD (Promega) at a DNA to Fugene ratio of 1:3 according to the manufacturer's instructions.

Luciferase NF-κB reporter assay

Twenty-four hours after transfection cells were redistributed into a 96-well plate. After an additional 24-hours of incubation, cells were washed twice with DMEM without FCS and stimulated with 10 μl of cell-free microbial culture supernatant in a total of 100 μl DMEM without FCS. After 5 h at 37°C cells were lysed in 50 μl reporter lysis buffer (Promega) at -80°C for 24-hours. After thawing lysate was mixed with luciferase reagent (Promega) and luciferase activity was measured in a TriStar2 luminometer (Berthold). NF-κB activity is represented by luciferase activity which was measured in relative light units. Results were expressed as fold increase in NF-κB activity of cells stimulated with microbial supernatants over cells that received the equivalent amount of cell-free LB broth.

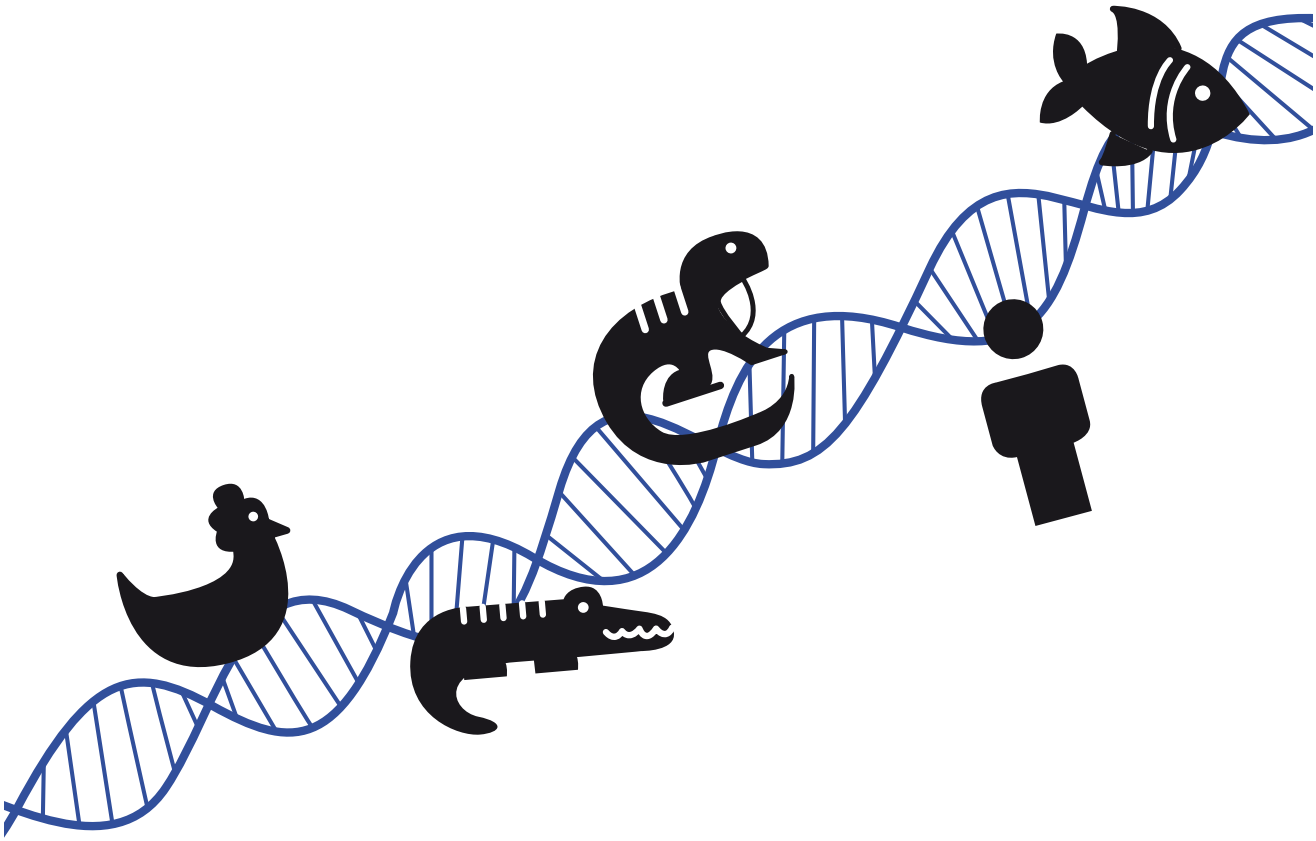
Statistics

Groups were compared using an unpaired Student's *t*-test. Differences with a $p < 0.05$ were considered significant and are indicated in graphs by an asterisk. Analyses were performed with Graphpad Prism 7 software.

References

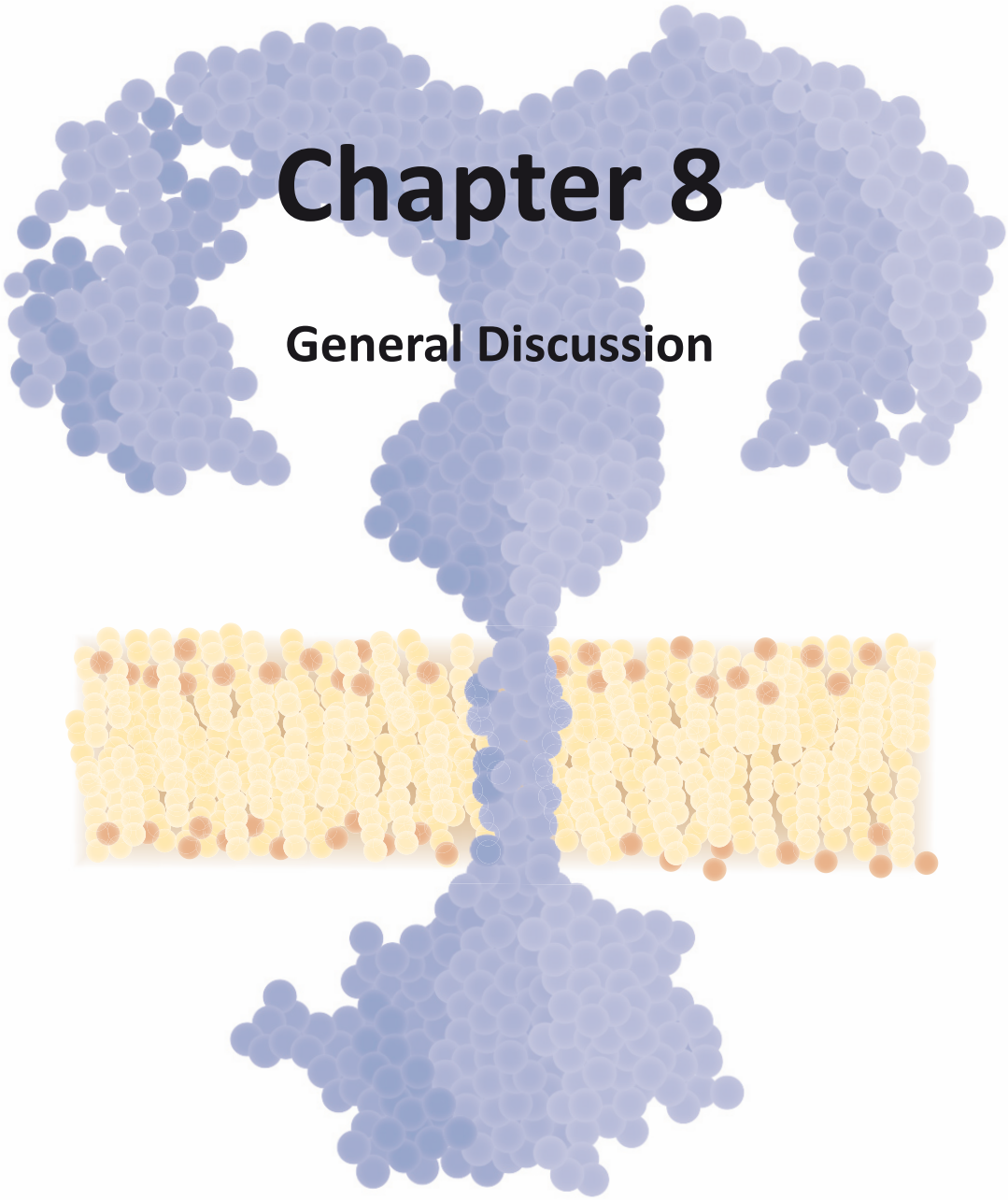
1. Senior, B. W. & Woof, J. M. Sites in the CH3 Domain of human IgA1 that influence sensitivity to bacterial IgA1 proteases. *The Journal of Immunology* **177**, 3913–3919 (2006).
2. Bardoel, B. W. *et al. Pseudomonas* evades immune recognition of flagellin in both mammals and plants. *PLOS Pathogens* **7**, e1002206 (2011).
3. Behnsen, J. *et al.* Secreted *Aspergillus fumigatus* protease Alp1 degrades human complement proteins C3, C4, and C5. *Infection and Immunity* **78**, 3585–3594 (2010).
4. LaRock, C. N. *et al.* IL-1β is an innate immune sensor of microbial proteolysis. *Science Immunology* **1**, eaah3539–eaah3539 (2016).
5. Rothmeier, A. S. & Ruf, W. Protease-activated receptor 2 signaling in inflammation. *Seminars in Immunopathology* **34**, 133–149 (2012).
6. Ossovskaya, V. S. & Bunnett, N. W. Protease-activated receptors: contribution to physiology and disease. *Physiological Reviews* **84**, 579–621 (2004).
7. Kida, Y., Higashimoto, Y., Inoue, H., Shimizu, T. & Kuwano, K. A novel secreted protease from *Pseudomonas aeruginosa* activates NF-κB through protease-activated receptors. *Cellular Microbiology* **10**, 1491–1504 (2008).
8. Kida, Y., Taira, J., Yamamoto, T., Higashimoto, Y. & Kuwano, K. EprS, an autotransporter protein of *Pseudomonas aeruginosa*, possessing serine protease activity induces inflammatory responses through protease-activated receptors. *Cellular Microbiology* **15**, 1168–1181 (2013).
9. Chung, W. O., Hansen, S. R., Rao, D. & Dale, B. A. Protease-activated receptor signaling increases epithelial antimicrobial peptide expression. *The Journal of Immunology* **173**, 5165–5170 (2004).
10. Lourbakos, A. *et al.* Activation of protease-activated receptors by gingipains from *Porphyromonas gingivalis* leads to platelet aggregation: a new trait in microbial pathogenicity. *Blood* **97**, 3790–3797 (2001).
11. Leulier, F. & Lemaitre, B. Toll-like receptors — taking an evolutionary approach. *Nature Reviews Genetics* **9**, 165–178 (2008).
12. Kawai, T. & Akira, S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature Immunology* **11**, 373–384 (2010).
13. De Nardo, D. Toll-like receptors: Activation, signalling and transcriptional modulation. *Cytokine* **74**, 181–189 (2015).
14. Voogdt, C. G. P., Merchant, M. E., Wagenaar, J. A. & van Putten, J. P. M. Evolutionary regression and species-specific codon usage of TLR15. *Frontiers in Immunology.* **9**, (2018).
15. Boyd, A. C. *et al.* TLR15 is unique to avian and reptilian lineages and recognizes a yeast-derived agonist. *The Journal of Immunology* **189**, 4930–4938 (2012).
16. Zoete, M. R. de, Bouwman, L. I., Keestra, A. M. & Putten, J. P. M. van. Cleavage and activation of a Toll-like receptor by microbial proteases. *Proceedings of the National Academy of Science of the United States of America* **108**, 4968–4973 (2011).
17. Kumar, S., Stecher, G., Suleski, M. & Hedges, S. B. TimeTree: a resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution* **34**, 1812–1819 (2017).
18. Green, R. E. *et al.* Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. *Science* **346**, 1254449 (2014).

19. Sievers, F. & Higgins, D. G. Clustal Omega for making accurate alignments of many protein sequences. *Protein Science* **27**, 135–145 (2018).
20. Offord, V. & Werling, D. LRRfinder2.0: A webserver for the prediction of leucine-rich repeats. *Innate Immunity* **19**, 398–402 (2013).
21. Matsushima, N. *et al.* Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors. *BMC Genomics* **8**, 124 (2007).
22. Anwar, M. A., Panneerselvam, S., Shah, M. & Choi, S. Insights into the species-specific TLR4 signaling mechanism in response to *Rhodobacter sphaeroides* lipid A detection. *Scientific Reports* **5**, 7657 (2015).
23. Ishida, H., Ohto, U., Shibata, T., Miyake, K. & Shimizu, T. Structural basis for species-specific activation of mouse Toll-like receptor 9. *FEBS Letters* **592**, 2636–2646 (2018).
24. Werling, D., Jann, O. C., Offord, V., Glass, E. J. & Coffey, T. J. Variation matters: TLR structure and species-specific pathogen recognition. *Trends in Immunology* **30**, 124–130 (2009).
25. Voogdt, C. G. P., Wagenaar, J. A. & Putten, J. P. M. van. Duplicated TLR5 of zebrafish functions as a heterodimeric receptor. *Proceedings of the National Academy of Science of the United States of America* **115**, E3221–E3229 (2018).
26. Andersen-Nissen, E., Smith, K. D., Bonneau, R., Strong, R. K. & Aderem, A. A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. *Journal of Experimental Medicine* **204**, 393–403 (2007).
27. Seixas, R. *et al.* Severe fibrinonecrotic enteritis caused by *Pseudomonas aeruginosa* in a captive monitor lizard (*Varanus niloticus*). *Journal of Zoo and Wildlife Medicine* **45**, 410–412 (2014).
28. Walker, S. E., Sander, J. E., Cline, J. L. & Helton, J. S. Characterization of *Pseudomonas aeruginosa* isolates associated with mortality in broiler chicks. *Avian Diseases* **46**, 1045–1050 (2002).
29. Folkesson, A. *et al.* Adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis airway: an evolutionary perspective. *Nature Reviews Microbiology* **10**, 841–851 (2012).
30. Beaufort, N. *et al.* The thermolysin-like metalloproteinase and virulence factor LasB from pathogenic *Pseudomonas aeruginosa* induces anoikis of human vascular cells: Pseudomonal LasB induces vascular cell anoikis. *Cellular Microbiology* **13**, 1149–1167 (2011).
31. Golovkine, G. *et al.* VE-Cadherin cleavage by LasB protease from *Pseudomonas aeruginosa* facilitates type III secretion system toxicity in endothelial cells. *PLOS Pathogens* **10**, e1003939 (2014).
32. Plas, M. J. A. van der *et al. Pseudomonas aeruginosa* elastase cleaves a C-terminal peptide from human thrombin that inhibits host inflammatory responses. *Nature Communications* **7**, 11567 (2016).
33. Jacobs, M. A. *et al.* Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Science of the United States of America* **100**, 14339–14344 (2003).



Chapter 8

General Discussion



General Discussion

Toll-like receptors (TLRs) form an intricate part of an animals' immune system. The primary task of these receptors is to detect danger signals. The evolutionary conservation of TLRs throughout the animal kingdom indicates that the receptors have remained crucial ever since their emergence more than 500 million years ago. While the overall TLR structure has not changed throughout its long evolutionary history, gene duplications and environmental selective pressure have driven diversification of TLRs. This has eventually led to a family of structurally highly similar receptors of which each member detects a distinct set of molecules. These molecules include danger signals released during sterile tissues damage or comprise ligands derived from invading microorganisms. By directly detecting microbial ligands, TLRs are involved in shaping species-specific host-microbe interactions. In addition, by recognizing microbial as well as endogenous ligands, TLRs are increasingly acknowledged to play an important role in infectious and non-infectious diseases. This has fueled research aimed at modulating TLR function through therapeutic interventions. Yet, after two decades of intense research the factors that shape TLR evolution and determine their function are still not fully understood. This holds true for key processes such as receptor expression, intracellular trafficking, and species-specific ligand interactions. Possibly as a result of this, the exact role of TLRs in many diseases is still unclear and therapeutics designed to target TLRs often fail to reach the clinic. This situation demands a better understanding of fundamental aspects of TLR biology. In this thesis, several key aspects of TLR biology have been dissected using a species-comparative approach. This approach was taken based on the conviction that evolution-driven differences and similarities between TLRs (or any other protein) of diverse species will reveal features that are important for receptor function. This Chapter provides an overview on the current understanding of TLR biology from an evolutionary and cell biological perspective. The evolutionary history, life cycle and species-specificity of TLRs are discussed with special emphasis on the novel findings obtained from the receptors that are central in this thesis: TLR5 and TLR15.

TLR evolution

Throughout evolution, interacting species continuously exert a selective pressure on each other that forces adaptation and counter-adaptation. Selection for the most advantageous adaptations drives molecular evolution and this is strongest among the interactions between hosts and parasitic microorganisms as they have an intimate yet

adversary relationship[1,2]. Gene products that operate at the interface of host-parasite interactions hence show some of the fastest rates of molecular evolution[3–6]. TLRs are such operators at the host-parasite frontier by directly sensing microbial products with their ectodomain and initiating signaling events with their intracellular TIR domain. Because of this dual function (sensing and signaling) TLRs undergo distinct intramolecular selective modes of evolution. The TLR ectodomains that sense microbe specific ligands (TLR1, 2, 4, 5 and 6 in humans) must be able to adapt to new or altered microbial ligands and thus evolve under positive (diversifying) selection (which increases the frequency of an advantageous allele in a population). In contrast, the ectodomains of nucleic acid sensing TLRs (TLR3, 7, 8 and 9) as well as the TIR domain of all TLRs do not allow mutations that may increase the chance of sensing endogenous nucleic acids or affect signaling and therefore evolve under negative (purifying) selection (which expels detrimental alleles from a population)[7–10]. Throughout this thesis evidence for both modes of selection acting on TLRs has been obtained. The species-specific TLR responses reported in **Chapters 3 & 7** indicate diversification of the TLR ectodomain following species-specific host-microbe co-evolution (discussed later in this Chapter), while the ligand-induced signaling via bird, reptilian and even fish TLRs in human cells, despite more than 430 million years of independent evolution, exemplify the extreme purification of the TIR signaling domain (**Chapters 3, 4 & 6**). Findings like these provide a functional basis to interpret the relevance of signatures of selection that may be encountered in future evolutionary analyses on TLRs. Each TLR co-evolved with its specific ligand(s) and therefore has a unique evolutionary history. Below, some aspects are discussed concerning the evolutionary history of TLR5 and TLR15.

Evolution of TLR5

TLR5 is conserved in all classes of vertebrates including fish, amphibians, reptiles, birds and mammals. Mammalian TLR5 is critically involved in correctly shaping the enteric microbiota and maintaining homeostasis[11–13]. As hosts and microbes co-evolve, the evolution of TLR5 is linked to the evolution of its ligand which is the bacterial protein flagellin. Flagellin is the major building block of the flagella fiber that allows many (pathogenic) bacteria to be motile and to colonize and penetrate host mucosal barriers. Bacterial flagellin proteins are composed of multiple domains of which the D0 and D1 domains adopt an elongated α-helix shape that is essential for the polymerization of multiple flagellin subunits to construct the flagellum[14]. Due to its function in polymerization, the D1 domain tolerates little structural variation and is thus highly conserved among diverse bacterial clades[15]. This feature has marked the flagellin D1 domain as a target for the innate immune system. In this thesis, we showed that *Salmonella* flagellin activates TLR5 of the green anole lizard

(**Chapter 3**), thereby expanding the repertoire of functional TLR5 as a flagellin sensor to reptilian species. The detection of the *Salmonella* flagellin D1 domain by TLR5 of both reptiles and humans illustrates that this domain has remained an excellent, conserved target of TLR5 for more than 240 million years (**Chapter 3**). However, while flagellins of Gammaproteobacteria, including *Salmonella*, activate TLR5, flagellins of Epsilonproteobacteria, such as *Campylobacter*, do not activate reptilian and mammalian TLR5 (**Chapter 3**)[16]. Flagellins of the Epsilonproteobacteria have several different amino acids in the D1 domain that still enable flagellin polymerization and thus bacterial motility, but these different residues somehow adopt a conformation that precludes activation of TLR5. For example, gamma-flagellins contain a positively charged arginine at position 90 which is crucial for TLR5 activation[16–18], whereas epsilon-flagellins carry at this position a threonine residue. The smaller size and lack of charge of threonine compared to arginine may be insufficient for TLR5 to bind or become activated upon binding of flagellin types containing threonine at position 90. The fact that epsilon-flagellins still enable bacterial motility but are not detected by TLR5, portrays an evolutionary scenario that Epsilonproteobacteria may have actively adapted their flagellin D1 domain to evade detection by host TLR5[16]. However, the class of Epsilonproteobacteria also contains many species of free-living motile bacteria that have D1 domain residues that are identical or very similar to the residues in e.g. *Campylobacter* flagellin but that are not associated with any host[19] (Fig. 1).

```
γ Salmonella enterica                83 EINNNLQRVRELAVQSAN 100
ε Campylobacter jejuni               83 EQLKILDTIKTKATQAAQ 100
ε Caminibacter mediatlanticus        83 EQVKILNTIKVKATQAAQ 100
ε Sulfurimonas autotrophica          83 EYVKIINSVRTKAIQAAS 100
ε Arcobacter nitrofigilis            81 EQSNILDTVKAKLIQANT 98
```

Figure 1. Amino acid sequence alignment of the TLR5-interacting region in flagellins of Gamma- and Epsilon-proteobacteria. This region of residues is situated in the D1 domain of flagellin. Grey shaded residues in flagellin of *Salmonella enterica* are important for activation of TLR5[16-18]. This region in flagellin of Epsilonproteobacteria is different from the Gammaproteobacteria (*S. enterica*), and in particular the threonine (T) and lysine (K) at positions 90 and 94 respectively are significantly different from the arginine (R) and leucine (L) in flagellins of Gammaproteobacteria. These deviant residues are conserved among host-associated Epsilonproteobacteria such as *C. jejuni* and free-living bacteria that inhabit deep-sea vents such as *C. mediatlanticus* and *S. autotrophica* or salt marshes such as *A. nitrofigilis*[19]. GenBank accession numbers of flagellins are: NP_460912.1 (*S. enterica*), YP_002344727.1 (*C. jejuni*), WP_007474650.1 (*C. mediatlanticus*), WP_013325936.1 (*S. autotrophica*) and WP_013136714.1 (*A. nitrofigilis*). Sequences were aligned with Clustal Omega.

The conservation of these "TLR5-evading" residues in flagellin of both host-associated and free-living Epsilonproteobacteria rather suggests the alternative evolutionary scenario that epsilon-flagellins evolved independently of any selective

pressure exerted by a host. If so, it is TLR5, rather than flagellin, that underwent evolutionary adaptation and became specialized to sense gamma-flagellins but not epsilon-flagellins. Since the residues in gamma- and epsilon-flagellins relevant for TLR5 are quite distinct, the adaptation of TLR5 to detect one type of flagellin may exclude the possibility to detect the other type. Perhaps Gammaproteobacteria exerted a relatively stronger selective pressure compared to Epsilonproteobacteria during host-microbe co-evolution which forced TLR5 to specialize on the detection of gamma-flagellins.

One factor that may have allowed TLR5 to specialize on a subset of bacterial flagellins, is a partial takeover of flagellin sensing by the NAIP5-NLRC4 inflammasome. NAIP5 [(Nucleotide binding domain leucine rich repeat, NLR)(NLR family apoptosis inhibitory protein 5)] is a cytosolic receptor that unlike TLR5 binds the D0 domain of flagellin. This initiates multimerization of NLRC4 [NLR family CARD (caspase activation and recruitment domain) containing-4] to induce an interleukin-1β mediated inflammatory response[20–23]. Although it remains to be seen whether NAIP5 can detect the D0 domain of epsilon-flagellins, the sensing of conserved but distinct domains in bacterial flagellin by TLR5 and NAIP5-NLRC4 enable the immune detection of a broad array of flagellated bacteria.

Although TLR5 has remained highly conserved among vertebrates, there are some exceptions to the stability of its evolution. In some teleost fish including the grass carp (*Ctenopharyngodon idella*)[24] and the zebrafish, the *tlr5* gene has been duplicated in tandem. Gene duplications are a prime source for molecular innovation by allowing neofunctionalization, although most often one of the duplicated paralogs rapidly undergoes pseudogenization[25]. Duplicated genes can also undergo subfunctionalization. In the studies described in **Chapter 4** we discovered that the two zebrafish TLR5 paralogs subfunctionalized to form a unique complementary heterodimeric flagellin receptor. This unexpected evolutionary path of a TLR5 ortholog is possibly the result of the duplication, degeneration, complementation mode of duplicated gene evolution[26]. Besides this remarkable subfunctionalization through physical complementation of the duplicated genes to detect a ligand, the zebrafish TLR5 paralogs may have also subfunctionalized their expression (by diversifying their regulatory elements) as it has been found that in adult zebrafish, TLR5a is constitutively expressed while expression of TLR5b is induced upon infection with (non-flagellated) *Mycobacterium marinum*[27]. Although we showed that the zebrafish TLR5 paralogs clearly function as a heterodimer to detect flagellin (**Chapter 4**), we cannot exclude that due to their differential expression each of the paralogs also has unique functions, perhaps as homodimers or heterodimers with other molecules. Alternatively, the subfunctionalization of TLR5 paralog expression

may coincidentally provide the zebrafish an advantage by forming an additional regulatory mechanism to control heterodimer formation and thus limit the risk of auto-activation or detection of endogenous ligands. In contrast to the duplication of TLR5 is some fish species, TLR5 underwent pseudogenization in some populations of birds and mammals, including humans[28,29]. This may hint at some redundancy of TLR5 function (perhaps compensated partially by NAIP5-NLRC4) although humans harboring the non-functional TLR5 allele are more susceptible to infections with *Legionella pneumophila* and less susceptible to develop the autoimmune disease systemic lupus erythematosus[30,31]. This implies that some trade-off between the detection of infection and prevention of autoimmunity may have partially shaped TLR5 evolution, at least in mammals. Given the pronounced role of TLR5 in shaping of the microbiota it would be interesting to compare microbiota development between individuals with functional versus individuals with non-functional TLR5 alleles.

Evolution of TLR15

While TLR5 has remained highly conserved among vertebrates, TLR15 has been lost multiple independent times throughout vertebrate evolution. We came to this conclusion by identifying TLR15 orthologs in birds, reptiles and a shark species whilst finding no orthologs in fish, amphibians and mammals (**Chapter 6**). The presence of putative TLR15 in a shark reverses the view on the evolutionary history of this receptor as it proved that TLR15 did not originate in the reptilian clade. Such findings, which expand the understanding of a gene's evolutionary history through extended species comparative analysis, have the potential to significantly change how we view and value the function of a gene. In the case of TLR15 our results changed the view that the receptor is a novel invention in birds and reptiles that served a species-specific need to detect microbial proteases, to the new awareness that TLR15 is a receptor that has been largely redundant throughout vertebrate evolution and may only still serve a purpose in a very select group of animals.

The evolutionary regression of TLR15 in most vertebrate lineages may well relate to the nature of its ligands which we showed to be microbial proteases (**Chapter 6 & 7**). Unlike canonical TLR ligands such as DNA, cell-wall components or flagellin, secreted microbial proteases are not directly essential for microbe viability. These enzymes may thus escape evolutionary fixation and indeed the presence of protease as virulence factors can be species and even strain specific[32,33]. For example, LasB or elastase of *Pseudomonas aeruginosa* is a versatile protease contributing to *P. aeruginosa* pathogenicity and is likely the primary protease of this bacterium that activates TLR15 (**Chapter 7**). Yet, LasB is not conserved in all *P.*

aeruginosa strains including clinical isolates[34]. Such dynamic evolution of microbial proteases may not have exerted sufficient and/or continuous selective pressure on TLR15 to force its conservation in all vertebrates during species-specific co-evolution with virulent microorganisms. Nonetheless, the detection of microbial proteases that contribute to virulence is important for hosts and is a conserved trait from vertebrates to flies. During an infection in the fruit fly (*Drosophila melanogaster*), microbial proteases that enter the fly's hemolymph cleave a circulating protein called Persephone at a specific bait region. This initial cleavage of Persephone results in further processing by an endogenous cathepsin to result in activated Persephone, which is itself a protease. Persephone then cleaves the cytokine pro-spätzle to form active spätzle which is the ligand for and activates Toll-1 in the fruit fly. Activation of Toll-1 induces the expression of antimicrobial peptides that will attack the source of infection[35]. This multistep protease detection system in fruit flies is somewhat reminiscent of the complex hemostatic system in vertebrates in which multiple proteases consecutively activate others to eventually activate e.g. platelets or the protease activated receptors (PARs). PARs form a family of four highly conserved G-protein coupled receptors that are expressed on the surface of many different cells types in mammals. PARs are critically involved in diverse physiological systems and responses including hemostasis and inflammation[36,37]. Like TLRs, PARs can initiate pro-inflammatory signaling. PARs are activated upon direct cleavage of their N-termini by microbial proteases or when microbial proteases (once present in the bloodstream during bacteremia) cleave components of the hematopoietic system such as thrombin which in turn forms an activated endogenous protease that can activate PARs[37,38]. By directly and indirectly sensing microbial proteases and initiating immune responses, PARs form an elaborate system that detects microbial proteolytic activity and may have caused for TLR15 to become redundant in most vertebrates.

An additional factor that may underlie the loss of TLR15 from different vertebrate clades is the direct activation of TLR15 by proteolytic activity. This may pose a risk that the receptor becomes activated by endogenous proteases. For TLR15 to function as a broad sensor of (infectious) microbes (**Chapter 7**), the receptor must be promiscuous to different types of microbial proteases. Maintaining broad microbial protease detection may conflict with avoiding receptor activation by endogenous proteases such as cell-surface expressed ADAMs (a disintegrin and metalloproteinase) or secreted neutrophil elastase. Accidental activation of TLR15 by endogenous proteases may lead to auto-inflammation and preventing this could, in some species, have formed a driving force to erase the *tlr15* gene. It can be speculated that those animals that have retained TLR15 (at least birds and several reptilian groups) perhaps circumvented such conflict by tightly regulating *tlr15* gene

expression. Indeed, multiple bird species show none to very low basal tissue expression of *tlr15* transcripts but strongly upregulate the gene during infection[39–41].

Beyond the extensive gene loss across vertebrates, we discovered that TLR15 orthologs also display a species-specific bias in the usage of leucine codons that strongly affects protein expression (**Chapter 6**). The actual driving force behind codon usage bias remains elusive but is generally thought to consist of either non-randomness in mutations or natural selection, or possibly a combination of these processes[42]. Bias towards the use of optimal codons (codons that occur most frequently throughout the genome and for which the levels of their cognate tRNA are relatively high) is likely the result of a selective pressure to enhance the translational efficiency and accuracy of highly expressed genes. Conversely, bias towards the use of sub-optimal codons may result from a selective pressure to actively dampen the efficiency of protein translation. This has been suggested as an explanation for the predominant sub-optimal codon composition of all human TLR genes (except TLR9 which is highly codon optimized) as high TLR protein levels can result in detrimental auto-activation[43–45]. Some TLR15 gene orthologs such as those of crocodiles and snakes are also biased for sub-optimal codon usage (**Chapter 6**). Yet in the case of TLR15 gene orthologs, it is less likely that the codon usage pattern is shaped by natural selection because it is variable between relatively closely related species and does not adhere to species phylogeny. The TLR15 species-specific codon bias, and its limitations on protein expression, is therefore better explained by mutational drift due to a lack of selective pressure (at least in some species).

Whether TLR15 provides an advantage in clearing infections, whether it can be activated by endogenous proteases and, if so, whether this has restrained TLR15 evolution are interesting questions of which the answers will help to better understand the exceptional evolutionary history of this TLR.

Do non-immune functions shape TLR evolution?

As gene evolution is directed by function the intriguing question arises whether not only host-microbe interactions, but also non-immune functions of TLRs affect their evolution. In **Chapter 2** we briefly discussed that TLRs can detect endogenous (damage associated) ligands and thus are involved in activities other than immune responses against pathogens. This is not surprising given the evolutionary and structural relationship of TLRs with the protostomian Toll receptors which, besides immunity, are involved in e.g. embryogenesis and neuronal development. What if the necessity to detect (or avoid) endogenous ligands also exerts a selective pressure on TLR evolution? The direction of the receptor's molecular evolution may then

come to depend on the balance between microbial vs. endogenous ligand detection. When receptor adaptations for microbial ligand detection conflict with the detection or avoidance of endogenous ligands (or vice versa) the evolutionary "freedom" of a TLR may be restraint. Perhaps such restraints on TLR evolution are the underlying reason why microbes still find opportunities to occasionally evade TLRs by altering their ligands and why receptors are unable to evolve complete avoidance of endogenous ligands to prevent autoimmune activation. Better understanding of the interplay between microbial and endogenous ligand detection by TLRs and how this has shaped receptor evolution, requires precise knowledge on the receptor binding sites for both ligand types and how adaptations to one may affect recognition of the other. Studying this in a species-comparative setting can further delineate how external vs. internal ligand detection impacted TLR evolution.

TLR expression

Comparative studies not only yield insights into the evolution of TLR structure and function but are also valuable to decipher cell biological aspects of the TLR life cycle, such as TLR expression and maturation, TLR trafficking, TLR-ligand interaction, and receptor activation. The life cycle of TLRs starts with receptor gene expression (Fig. 2-A). In all vertebrates thus far examined, TLRs are expressed in tissues as diverse as the spleen, brain, kidneys and even reproductive tissues (**Chapter 3**)[46–48]. The presence of TLRs on different cell types in immune and non-immune related tissues ensures the scanning for microbial and damage associated ligands throughout the body. While most immune cells express the majority of TLRs[49,50], receptor expression can be much more specific among subsets of immune cells[51]. Similarly, most epithelial cells, especially those interacting with the external environment such as intestinal or airway epithelial cells, express multiple TLRs[52], but paneth cells, located in the crypts of the intestinal villi and responsible for antimicrobial peptide secretion, specifically only express TLR5[53]. The highly selective TLR expression within and between cell types indicates that the transcriptional regulation of TLR expression involves receptor specific regulatory elements and diverse transcription factors (Fig. 2-A). Furthermore, TLRs can negatively regulate their own expression via cognate ligand detection[54] and they can cross-regulate the expression of other TLRs[55–58]. For example, activation of TLR5 by flagellin in mouse intestinal epithelial cells promotes gene expression of TLR2 and TLR4 while activation of TLR4 by lipopolysaccharide (LPS) limits the expression of TLR5[58]. Such differential self- and cross-translational regulation of TLRs indicates that the activation of different receptors leads to the formation of

unique activating and inhibitory transcriptional regulatory complexes that determine the expression of multiple TLR genes.

Currently, the mapping of regulatory elements and identification of transcription factors that are involved in the expression of each TLR is incomplete. In addition, the expression of TLR3, for example, differs between mouse and human cells[59,60] indicating that there is species-specificity in TLR expression and thus evolutionary diversification of regulatory elements upstream of TLR genes[61]. A formidable challenge lies in determining how receptor induced regulatory complexes operate in concert during the contact of cells with microbes or damaged tissues, when multiple ligands are present in a time and space dependent fashion. What would be the hierarchy in the formation of transcriptional regulatory complexes following activation of multiple receptors? Is there competition for the limited amount of transcription factors to form complexes? What are the consequences for the expression of virus detecting receptors (e.g. TLR7) when bacteria detecting receptors (e.g. TLR5) have been activated (and vice versa)? And how do transcriptional dynamics differ between species and how does that affect the translationality of (model) species? Answering these questions will help tremendously in understanding the regulation of TLR gene expression and its role in maintaining homeostasis and the initiation of appropriate immune responses.

In **Chapter 6** of this thesis we identified an additional layer of regulation of TLR expression at the protein level (Fig. 2-B). Using the species-comparative approach, we expressed *tlr15* genes from different species in a human and snake cell background and noticed large species-specific variation in the expression of TLR15 protein. Despite a relatively high amino acid similarity between chicken and crocodile TLR15, the crocodile TLR15 protein was expressed at much lower levels than the chicken TLR15 protein. This variation could be attributed to a difference in codon usage. Codon usage influences cellular protein levels by determining the efficiency of protein translation (Fig. 2-B1). In addition, codon usage, which we discussed earlier in an evolutionary context, is a notable factor that can also influence gene transcription through the level of guanine-cytosine (GC) bias. Genes with a high GC level are transcribed more efficiently and the resulting mRNA is more stable, leading to higher levels of proteins[62–64]. Hence, the low GC level of the mouse TLR7 gene limits its transcription and is thereby responsible for relatively low expression of TLR7[44]. The effects of codon usage bias on gene transcription and translation into protein are of great practical significance. As we have seen in **Chapter 6**, codon optimization of two crocodile *tlr15* genes (that naturally contain many sub-optimal codons) greatly increased protein production in both human and snake cells which allowed us to make inferences about the role of codon bias in

TLR15 evolution (see above). As most TLR genes are biased towards sub-optimal codon usage, codon optimization can greatly help to overcome some of the technical challenges related to the generally low protein expression of TLRs. Optimizing TLR gene codons should thus be considered for example for immunoprecipitation or structural studies on TLRs where large quantities of protein are required.

TLR maturation

During translation of mRNA by ribosomes bound to the endoplasmic reticulum (ER), membrane localized glycoproteins, such as TLRs, co-translationally enter the secretory pathway at the ER lumen for further processing (Fig. 2-B). Proteins residential in the ER lumen such as glycoprotein 96 (gp96 or GPR94) and protein associated with TLR4 (PRAT4A), have been identified as chaperones that are necessary for proper maturation of the TLR protein (Fig. 2-C). How these chaperones exactly contribute to TLR maturation and whether all TLRs require their service is not known although knock down of gp96 or PRAT4A prevents proper localization and signaling by TLR1, 2, 4, 5, 7 and 9[65–68]. As gp96 is a versatile chaperone involved also in the maturation of other leucine rich repeat proteins[69], it is surprising that TLR3 appears not to depend on gp96 (and PRAT4A) for its processing[65,68]. This suggests that TLR3 contains a highly specific structural feature compared to other TLRs and may depend on assistance of a different, still unknown chaperone. Further down the processing route, another ER residential protein called leucine rich repeat containing protein 59 (LRRC59) aids the chaperone UNC93B1 (discussed in the following section) with preparing antrograde transport of receptors to the Golgi complex by loading specifically the nucleic acid sensing TLRs into coat protein II (COPII) vesicles[70]. Such selectivity of the TLR-chaperone interactions at this basal level of receptor processing is intriguing. It suggests that subtle but highly specific receptor features are involved and strongly implies that additional factors contribute to TLR maturation. While passing through the ER lumen, TLRs are modified by the (extensive) acquisition of glycans on accessible asparagine residues in N-x-S/T sequons of the polypeptide chain (N: asparagine, S: serine, T: threonine, x: any amino acid except proline)(Fig. 2-C). The glycan moieties may be adapted whilst the receptor passes the Golgi system (Fig. 2-D). However, the exact glycan composition on a TLR has not yet been determined. N-glycosylation is vital for the function of many membrane anchored receptors as the addition of glycans can affect a proteins' tertiary structure and provide additional interfaces for interaction with other molecules. This influences receptor expression, mobility, endocytosis, and interaction with ligands or co-receptors[71–75].

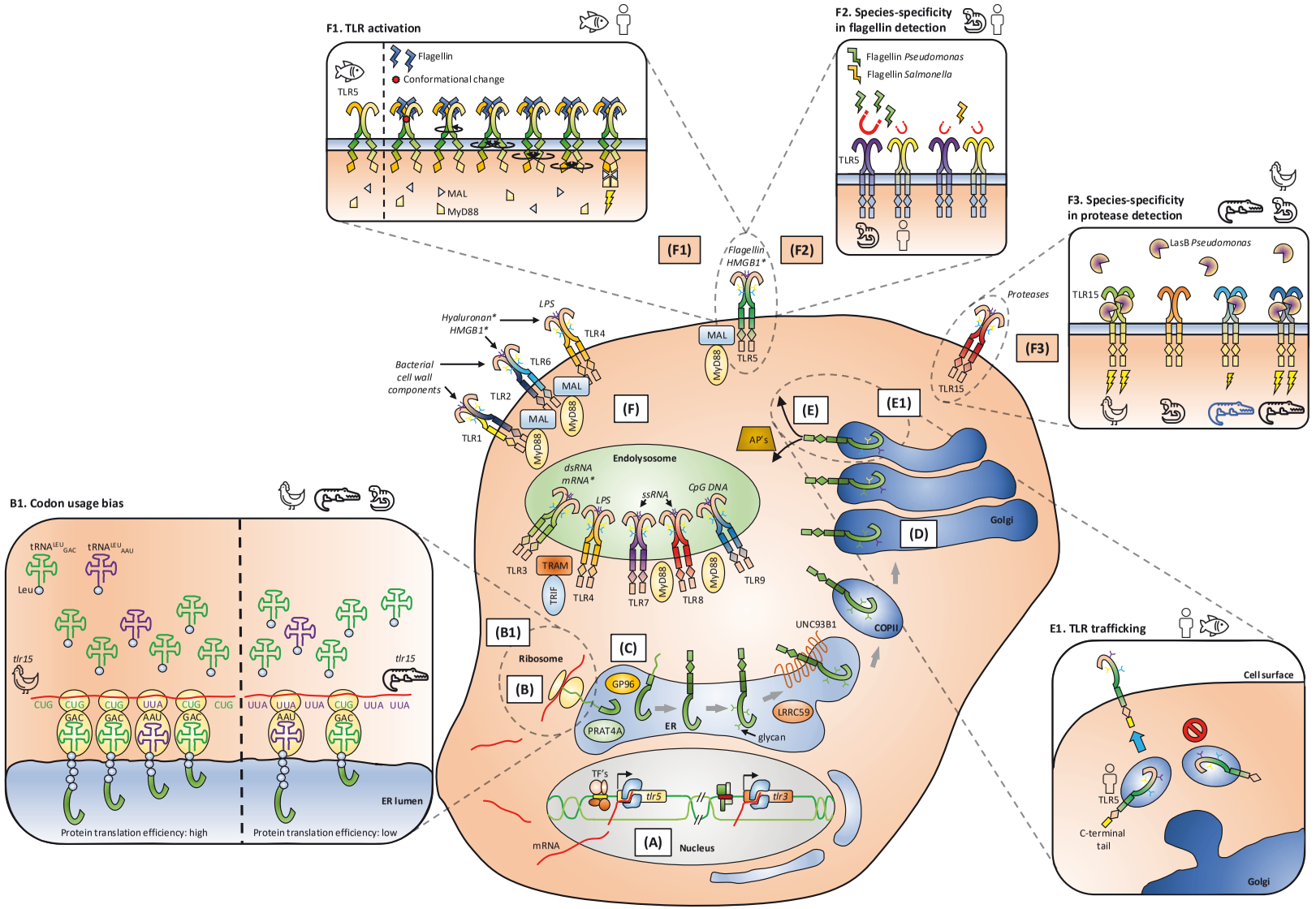

F1. TLR activation
Flagellin
Conformational change
TLR5
MAL
MyD88
F2. Species-specificity in flagellin detection
Flagellin Pseudomonas
Flagellin Salmonella
TLR5
F3. Species-specificity in protease detection
LasB Pseudomonas
TLR15
Proteases
(F1)
(F2)
(F3)
Flagellin HMGB1*
MAL
MyD88
TLR5
TLR15
LPS
Hyaluronan* HMGB1*
Bacterial cell wall components
TLR4
TLR6
TLR2
TLR1
MAL
MyD88
(F)
(E)
(E1)
AP's
Endolysosome
dsRNA mRNA*
LPS
ssRNA
CpG DNA
TLR3
TRAM
TRIF
TLR4
TLR7
TLR8
TLR9
MyD88
Golgi
(D)
COPII
UNC93B1
(C)
GP96
PRAT4A
ER
LRRC59
glycan
(B1)
Ribosome
(B)
TF's
tlr5
tlr3
mRNA
(A)
Nucleus
B1. Codon usage bias
tRNA LEU GAC
tRNA LEU AAU
Leu
tlr15
tlr15
CUG CUG CUG UUA CUG CUG
UUA UUA UUA CUG UUA UUA
GAC GAC AAU GAC
AAU GAC
ER lumen
Protein translation efficiency: high
Protein translation efficiency: low
E1. TLR trafficking
Cell surface
TLR5
C-terminal tail
Golgi

Figure 2 on previous page. A graphical overview of the TLR life cycle and functional TLR evolution, emphasizing in the separate windows several of the major findings described in this thesis. The animal icons represent a human, chicken, green anole lizard, crocodile/alligator and zebrafish. Icons inside the windows indicate the receptor species. Icons on the top right of the windows indicate which species were used for studying that particular aspect of TLR biology. **(A)** The TLR life cycle starts with transcription of the *tlr* gene in the nucleus. *tlr* genes (e.g. *tlr5* and *tlr3*) have unique promoter sequences to which different transcription factors (TF's) bind and initiate transcription of the receptor gene by RNA polymerase II (blue). **(B)** The transcribed TLR messenger RNA (mRNA) which exited the nucleus is translated by ribosomes into TLR protein. **(B1)** Codon usage bias can affect the efficiency of translation and thus the levels of protein expression. In the *tlr15* mRNA of the chicken, the high frequency of CUG codons for leucine (Leu) enables many ribosomes to produce chicken TLR15 protein because there are plenty of transfer RNA's available that match the CUG codon with their GAC anticodon ($tRNA^{LEU}_{GAC}$). The production of alligator TLR15 protein on the other hand is limited since its mRNA is biased to contain many UUA codons and the cognate tRNA ($tRNA^{LEU}_{AAU}$) for this codon is minimally available. **(C)** During translation the TLR proteins enter the lumen of the endoplasmic reticulum (ER). Here, ER resident proteins including glycoprotein 96 (GP96) and protein associated with TLR4 (PRAT4A) likely function as chaperones to properly fold the TLRs. During their transition through the ER, TLRs are glycosylated and eventually loaded into coat protein II (COPII) vesicles for transport to the Golgi, a process that is likely mediated by the chaperones leucine rich repeat containing protein 59 (LRRC59) and Unc-homolog B1 (UNC93B1). **(D)** In the Golgi, TLRs acquire different glycan groups. **(E)** With the aid of adapter proteins (AP's) TLRs are transported from the Golgi to either endolysosomes or the plasma membrane. **(E1)** For human TLR5 to localize at the plasma membrane, the presence of the receptor C-terminal tail is crucial. **(F)** At the proper location TLRs wait as inactive dimers for an encounter with ligands (written in italics, endogenous ligands are indicated with an *). **(F1)** Binding of a ligand to the receptor dimer induces conformational changes in the ectodomains. This may initiate the transmission of subsequent rotational conformational changes across multiple receptor regions. As a result, the intracellular TIR signaling domains of the receptors are aligned and can recruit adapter molecules like Myeloid differentiation primary response 88 (MyD88) and MyD88-adapter-like (MAL) or TIR-domain-containing adapter-inducing interferon-β (TRIF) and TRIF-related adaptor molecule (TRAM) that start pro-inflammatory signaling. **(F2)** Due to species-specific adaptations during host-microbe co-evolution, the ectodomain of lizard TLR5 has become more sensitive than human TLR5 to detect flagellin of *Pseudomonas* but not flagellin of *Salmonella*. Receptor sensitivity is illustrated with a magnet symbol. **(F3)** Evolution driven diversification of the ectodomain of TLR15 has also resulted in differences in the sensitivity to its ligand which are microbial proteases such as LasB of *Pseudomonas*. The response of chicken and alligator (black) TLR15 to LasB is stronger than the response of crocodile (blue) TLR15 while TLR15 of the green anole lizard does not respond to LasB at all. Not illustrated in this figure are the pro-inflammatory signaling cascades initiated by the TLRs upon ligand detection and receptor recycling or degradation after ligand interaction. Animal icons were downloaded from www.flaticon.com.

Indeed, glycosylation is required for appropriate localization and ligand detection as shown for TLR2, 3 and 4[76–80]. All TLRs have multiple N-x-S/T sequons in their ectodomain and collectively these sites can gain a substantial amount of glycan structures. In **Chapter 6** for example, we observed that the mature form of TLR15 of birds and reptiles has a mass of roughly 140 kilo Daltons (kDa) when expressed in both human and snake cells whereas the naked TLR15 protein is predicted to have a mass of only 100 kDa. Certainly this large bulk of glycans on these (and likely other) TLRs affects receptor processing or function but how exactly awaits clarification. Considering the evolution of TLR glycosylation, the number of N-x-S/T sequons predicted to be glycosylated is relatively well conserved among vertebrate TLR orthologs (e.g. TLR5 of anole lizard, human and zebrafish have respectively 6, 7 and 8 putative glycosylation sites). Yet, the position of these sites in the receptor is not strongly conserved which may relate to subtle differences in the overall 3D architecture between orthologs. Since glycans can affect ligand interaction and protein dimerization among others, the potential differential spatial organization of (diverse) glycan structures on TLRs from different species allows for the exciting possibility that glycans can contribute to species-specificity in TLR functioning. As the glycosylation machinery can differ among species and among tissues, the potential diversity in TLR biology that this entails will be staggering. Glycosylation is continuously recognized as a key factor that influences virtually all aspects of cell biology[81] and glycosylation will therefore surely be one of the main contributors to the mechanisms that underlie TLR function. In fact, the intriguing finding that a lectin (glycan binding protein) from a pathogenic fungus can activate TLR2 through the receptor glycans may open a whole new path towards non-canonical glycan mediated TLR activation[82,83].

Intracellular transport of TLRs

After acquisition of the correct glycans in the Golgi system, TLRs are transported to either endolysosomal compartments or the plasma membrane (Fig. 2-E). The plasma membrane harbors TLR1, 2, 4, 5, 6, 10 and 15, while TLR3, 7, 8 and 9 end up in endolysosomal vesicles. The reason for this spatial distinction makes sense when considering the receptor ligands (Fig. 2-F). All the plasma membrane localizing receptors detect ligands that are specific to microorganisms such as lipopolysaccharide (TLR4), peptidoglycan (TLR2) or flagellin (TLR5). Placement of these receptors on the plasma membrane of diverse cell types allows scanning for the presence of microbe specific molecules in the direct extracellular environment. The receptors that localize at endolysosomal vesicles on the other hand, detect

microbial nucleic acid motifs such as viral double (TLR3) or single stranded RNA (TLR7 & 8) or unmethylated CpG-DNA (TLR9). As these microbial derived ligands of genetic nature are in essence very similar to the genetic material of the host, compartmentalization of their receptors forms a safety mechanism to limit unwanted interaction of TLRs with endogenous nucleic acids[84].

To reach the plasma membrane or endolysosomal vesicles, TLRs must be assisted by proteins dedicated to the regulation of intracellular transport. Initially, Unc-homolog B1 (UNC93B1), a membrane bound protein that physically associates with TLRs that carry acidic amino acids in the extracellular juxtamembrane region (TLR3, 7, 8 and 9), was identified as a trafficking chaperone that determined receptor localization towards endolysosomes[85,86]. However, this view changed when UNC93B1 was found to also interact with the plasma membrane localizing human TLR5[87](**Chapter 5**). Furthermore, the interaction between UNC93B1 and TLRs beyond the ER seems not to be necessary for proper receptor localization and function[88]. UNC93B1 may therefore be considered more as an ER chaperone required for maturation of a subset of TLRs, possibly by mediating loading of TLRs into COPII vesicles, rather than directing TLR localization. UNC93B1 is a highly conserved protein[89] and its function in TLR processing is conserved at least throughout vertebrate evolution as reported in **Chapter 4.** We demonstrated that the addition of zebrafish UNC93B1 to zebrafish TLR5 aids in intracellular transport and facilitated functional heterodimerization of TLR5a and TLR5b. Interestingly though, while human TLR5 is transported to the plasma membrane in HEK293 cells (**Chapter 5**), zebrafish TLR5b, in the presence of UNC93B1, is transported to lysosomes in both HEK293 and HeLa-57A cells (**Chapter 4 & 5**). Obviously, the lysosomal localization of zebrafish receptors in human cells may not reflect the situation *in vivo* due to incompatibility of zebrafish receptors with one or multiple human factors. But as human and zebrafish TLR5 are relatively similar (human TLR5 is more similar to zebrafish TLR5b than to other human TLRs), their differential trafficking to defined locations in the same cell-type strongly suggests that differences between specific receptor regions are involved. Such regions have been identified in a subset of TLRs[90–92] but the mechanism by which they direct receptor localization remains unclear. For example, through a tyrosine-based motif in the TIR domain, TLR7 specifically interacts with adapter protein 4 (AP-4), and not other APs. This interaction is suggested to directly translocate TLR7 from the Golgi to endolysosomes[93]. While different APs transport cargo to distinct organelles, several APs recognize their clients by the same sorting signals such as the tyrosine-based motif[94]. The selectivity of the TLR7-AP-4 interaction can therefore not be explained by the tyrosine-based motif alone and implies that additional receptor

features co-determine the specific interaction with trafficking chaperones. This scenario is supported by the fact that plasma membrane localized TLR2 and TLR4 have the exact same tyrosine-based motif in their TIR domain as intracellular TLR7 and that the tyrosine of this motif in TLR9 was suggested to be involved in receptor maturation (although AP interaction was not studied)[95].

As APs are crucial for intracellular transport and necessary for some TLRs to respond to ligands[96,97], it is very likely that APs and other transportation complexes such as the endosomal sorting complexes required for transport (ESCRT) system[98], critically regulate the localization of all TLRs. Yet, the required information to reach a certain location by selective interactions with the trafficking machinery must be derived from receptor specific features present in the TLR sequence/structure. Our understanding of these features is still rudimentary[84] and it is currently not clear whether these regions truly direct trafficking or are required for receptor maturation. It is important to be able to make the distinction between trafficking and maturation as in both processes unique factors will contribute to a proper TLR lifecycle and hence disruption of any of these factors may induce (cell-type specific) TLR dysfunction. A promising novel receptor region to be explored for better understanding receptor maturation and trafficking is the TLR C-terminal tail (Fig. 2-E1). In **Chapter 5** we provided the first evidence that the tail of a TLR plays an important role. Disrupting the sequence order in the tail of human TLR5 prevented receptor localization at the plasma membrane and flagellin-induced signaling. One of our findings, by use of a chimeric human-zebrafish TLR5, suggests that the TLR5 tail is not directly involved in trafficking but may function in an earlier (maturation) process, possibly through threonine phosphorylation. As the tail region is highly diverse among TLR homologs within one species but evolutionarily conserved among TLR orthologs it may confer some sort of specificity in TLR processing. The discovery that the tail is involved in TLR processing provides a new opportunity to discover critical TLR partners, e.g. specific kinases or phosphatases, that may be linked to distinct TLR function and dysfunction and could therefore form novel candidates for therapeutic modulation during TLR-mediated diseases.

TLR-ligand interaction and receptor activation

Once the TLRs have been transported and embedded in the endolysosomal or cell surface membranes they are ready to interact with ligands via the receptor ectodomain. Available crystallized structures of TLR ectodomains in complex with their ligands have revealed two basic principles of TLR-ligand binding; I) TLRs adopt a curved structure which is formed by tight packing of the β-sheets that are in

between the consecutive leucine rich repeats (LRR), and II) TLRs bind ligand as a dimeric complex in which often the ligand acts as a cross linker to connect the two TLR monomers. The interaction between the two TLR monomers and their ligand is direct except in the case of TLR4 that needs the co-receptor MD-2 to bind its ligand LPS[99,100]. Collectively, TLRs can detect an impressive chemical diversity in ligands such as lipopolysaccharide (TLR4-MD-2), RNA (TLR3, 7 and 8) and DNA (TLR9) but also small lipopeptides of 1.5 kDa (TLR2/1 or TLR2/6) and large proteins of over 50 kDa such as flagellin (TLR5). In addition to these canonical microbial ligands, TLRs have been reported to detect endogenous ligands[101] like, mRNA (TLR3[102]), the extracellular matrix glycosaminoglycan hyaluronan (TLR2 and 4[103]), the nuclear protein high mobility group box 1 (HMGB1, TLR2, 4 and 5[104]), and even *de novo* synthesized small molecules that look nothing like the classical microbial agonists (TLR4, 7 and 8[105–107]). As discussed above, these unique binding capacities of the different TLRs are the result of divergent evolution of ancestral TLR duplicates. While the LRR domains remained conserved to form the curved shape that provides a large surface area, the regions in between the LRRs evolved under a selective (microbial) pressure to form homolog specific microstructures at the lateral sides of each receptor. This has ultimately resulted in a collection of receptors that share a LRR-based backbone but have unique structural binding features such as a hydrophobic pocket in TLR2 to accommodate lipopeptides[108], or a lateral protruding loop on TLR5 to bind hydrophilic flagellin residues[109]. Evolution has thus provided a relatively small set of receptors that detect widely different structures with high inter-receptor specificity.

After ligand binding TLRs become activated and initiate signaling. The current model of ligand-induced TLR activation dictates that, in the absence of ligand, TLRs are embedded in the membrane as loosely attached inactive dimers. Binding of ligand is accompanied by conformational changes and tighter dimerization of the ectodomains of the involved receptors. As a result, the intracellular TIR domains get closer, dimerize, and form a complex to which signal transducing adapter molecules can bind[110,111]. Our findings described in **Chapter 4**, which we obtained with chimeric receptors composed of zebrafish drTLR5a and drTLR5b, strongly imply that an additional force is required for TLR activation (Fig. 2-F1). We found that deliberate homodimerization of a single receptor region in an otherwise heterodimeric zebrafish TLR5 complex still allowed receptor activation by flagellin. However, when multiple regions of the extra- or intra-cellular domains were situated in a homodimeric configuration, receptor activation was completely blocked. This suggests that ligand-induced conformational changes in the ectodomain initiate interactions between multiple regions (extra- and intra-cellular) of the two monomers that are all required for signal transduction that ultimately

allows proper alignment and dimerization of the TIR domains. This scenario is based on our finding that a chimeric TLR5 complex that contained a heterodimeric ectodomain and homodimeric intracellular domain was not functional (**Chapter 4**). When we assume that the heterodimeric ectodomain could still bind flagellin, then the homodimeric TIR domains in this non-functional chimeric complex were likely at the same distance as in a functional complex composed of wild type receptors. Yet although in close proximity, the TIR domains in the chimeric complex were unable to adopt the right configuration to initiate signaling. As the zebrafish TLR5a and TLR5b receptors are highly similar and equally sized, we propose that a ligand-binding induced rotational twist, generated along the axes of multiple regions, is required to properly align the TIR domains in the functional configuration (Fig. 2-F1). This ligand-induced rotational mode of TLR activation would be preferable over ligand-induced activation by simply increasing the proximity of monomers. Rotation based activation of a dimer is energetically favorable as receptors would have minimal movement in the membrane and it also avoids the risk of inappropriate activation by accidental collision of monomers[112]. To become activated, zebrafish TLR5 likely relies on different, complementary rotational changes in TLR5a and TLR5b upon flagellin binding. Other heterodimeric TLRs such as TLR2/1 and TLR2/6 may depend on similar conformational changes. Interestingly, human TLR5, which functions as a homodimer, was found by electron microscopy to form asymmetric dimers in the absence of flagellin[113]. Possibly, human TLR5 also requires different rotational changes in the two monomers to align its TIR domains but it is more difficult to envision this (compared to a heterodimer) as the human TLR5 monomers are identical. Perhaps human TLR5 activation is opposite to e.g. activation of the epidermal growth factor receptor (EGFR) which is a symmetric homodimer of which the intracellular kinase domains are activated by asymmetric conformational changes following binding of its growth factor ligand[110,112,114]. Currently, the precise activational mechanism of TLRs remains elusive and our proposition of rotational activation must be confirmed with structural data. The only available high resolution TLR5 crystal structure, that of a homodimeric fragment of zebrafish TLR5b[109], is of little value to explain either heterodimeric zebrafish or homodimeric human TLR5 activation. In general, understanding the exact ligand-induced transregional conformational changes that result in TLR activation is precluded due to the lack of juxtamembrane and transmembrane regions in TLR crystal structures. Therefore, cryo-electron microscopy, for its rapidly increasing resolution and suitability for studying membrane associated structures[115], will likely be the key technique to unravel the precise structural mechanism of TLR activation.

After ligand induced receptor activation and correct positioning of the TIR domains, TLRs will interact with TIR domain-containing adaptor proteins (**Chapter 2**). The adaptor proteins connect with the TLR via TIR-TIR interactions and subsequently recruit kinases. The kinases will then propagate the signal that was started by the TLR via multiple intermediates to ultimately result in the activation of several transcription factors[110]. Once TLRs have initiated the signaling cascade, the receptors are translocated to the Golgi system for recycling or to lysosomes for degradation[116–118]. Meanwhile, the activated transcription factors may have already started the expression of a new generation of receptors thus completing the TLR lifecycle.

Species-specificity in ligand interaction

Bearing the TLR lifecycle and receptor-ligand interactions in mind, the question arises as to how evolution of TLR-ligand interaction results in species-specificity and what implications this entails. The adaptation of different animals to specific environments and the encountered pathogens has not only resulted in diversification of the TLRs within a single species (discussed above) but also resulted in differences among TLRs between species. Functional differences between TLR orthologs therefore reflect species-specific adaptations to those selective pressures that each animal faced during its co-evolution with microbes. This concept is strongly illustrated by our finding that TLR5 of the green anole lizard is more sensitive than human TLR5 to flagellin of *Pseudomonas aeruginosa* (**Chapter 3**)(Fig. 2-F2). *P. aeruginosa* is an opportunistic pathogen with the ability to thrive under diverse environmental circumstances and therefore infects a wide range of hosts, including fish, birds, mammals and reptiles[119–123]. Due to their poikilothermic nature and less efficient antibody response, reptiles probably rely more heavily on their innate immune system than endothermic birds and mammals[124]. As a result of these factors, *P. aeruginosa* may have had a stronger impact on the evolution of the reptilian immune system than on that of mammals which necessitated the anole lizard to evolve a flagellin binding site in its TLR5 that is more sensitive to *P. aeruginosa* flagellin. The equal sensitivity of anole and human TLR5 to flagellin from *Salmonella enterica* suggests that *S. enterica* did not exert such a differential selective pressure during the evolution of human and anole TLR5. Adaptation of the TLR5 flagellin binding site under specific selective pressures resulting in species-specific detection of flagellin is common across vertebrates[17,18,125,126]. Interestingly, TLR5 as well as TLR4 of animals are often found to be more sensitive to flagellin and LPS respectively than the human receptor variants[18,125,127–130]. Could this

differential sensitivity in TLR-ligand interaction, which may be based on only a few different amino acids in the ligand binding site, reflect a divergent evolutionary relationship between animals versus human with infectious agents in which the use of hygiene, antiseptics and medicine since ancient times possibly influenced human-microbe co-evolution? Evolutionary adaptation may preferentially occur at the level of receptor recognition and early signaling events as more downstream events are shared with other TLRs and adaptation of more downstream operating proteins would thus lead to less ligand specific, more drastic changes in the innate immune response.

Similar adaptive forces that drove TLR5-flagellin interaction likely also drove animal species-specific evolution of the ligand binding site in other TLRs such as TLR4[127-130], TLR9[131,132] and TLR15. TLR15 actually forms quite an extreme case of species-specific TLR activation as the receptors of chicken and three reptile species (which are relatively closely related) all show a unique profile of activation by bacterial, yeast and fungal pathogens (**Chapter 7**)(Fig. 2-F3). This differential sensitivity among TLR15 orthologs likely results from very specific host-microbe co-evolutionary interactions. As discussed at the beginning of this Chapter, these interactions may be strongly influenced by the dynamic evolution of the microbial proteases that activate TLR15 and possibly by potential receptor redundancy due to the presence of other protease activated receptors. In particular, the differential sensitivity to *P. aeruginosa* by the highly similar crocodile and alligator TLR15 indicates very specific adaptations during the co-evolution of these species. To answer whether crocodile TLR15 lost or alligator TLR15 gained sensitivity to *P. aeruginosa* requires investigation of TLR15 of additional crocodile species. While the 23 extant species of crocodile inhabit diverse tropical environments, their morphology, physiology and behavior are relatively similar and their rate of genome evolution is among the slowest of all tetrapods (amphibians, reptiles (including birds) and mammals)[133,134]. These features make crocodiles an attractive group to investigate how microbial (environmental) selective pressures shaped the innate immune system.

Animal evolution is strongly driven by co-evolving microbes and it is therefore safe to assume that the functioning of TLRs as well as many other immune receptors and effectors in each animal will have some level of specificity. Finding species-specific TLR functionality thus contributes to unraveling the evolutionary history of species. In addition, species-specific TLR responses can be fundamental in understanding the differences in susceptibility or resistance of hosts to (shared) pathogens. A striking example of this is the detection of LPS by mouse and human TLR4. Both TLR4 orthologs can detect hexa-acylated LPS with equal sensitivity but only mouse TLR4 can also detect tetra-acylated LPS. Switching from hexa- to tetra-

acylated LPS is a virulence strategy of *Yersinia pestis*, the agent causing bubonic plague, which enables this pathogen to evade detection by human but not mouse TLR4[129,135,136]. Mice are therefore less susceptible to succumb to *Y. pestis* infection and can form a reservoir for this bacterium with serious risks to public health. Knowing species-specific TLR responses is therefore highly valuable for uncovering the molecular mechanisms underlying zoonotic diseases. Species-specific TLR responses may even eventually be integrated with structural and sequence data in algorithms that could predict the sensitivity of new receptors and thus inform about an animal's susceptibility to disease which is useful for example in breeding and reintroduction programs.

Species-specific host-microbe co-evolution is also a major factor that influences the translational value of "model" species. The zebrafish is a well-known model species in developmental biology but is also increasingly popular in studies on immunology and infection biology[137]. In addition, a crystal structure of homodimeric zebrafish TLR5b was rapidly embraced as an important model for understanding the precise interactions between TLR5 and flagellin[18,109,138]. However, our unexpected finding that zebrafish TLR5 uniquely functions as a heterodimer rather than a homodimer as in all other vertebrates has direct consequences for the utilization of the homodimeric TLR5b crystal structure as a model (**Chapter 4**). Not only did we find that homodimeric zebrafish TLR5b is not functional, we also found that the amino acids interacting with flagellin according to the crystal structure are poorly conserved and not functionally exchangeable with residues in human TLR5. Crystal structures of TLRs from different species are instrumental for the exact determination of those amino acids that confer species-specific ligand interactions and for predicting interspecies sensitivity to agonists or antagonists[129,139]. Yet, for purposes such as the structure-guided rational design of TLR targeting pharmaceuticals, the structure must be of the target species or a close relative to avoid heavily confounding evolutionary influences. In this regard, it is thus understandable that TLRs of the zebrafish, which shared its last common ancestor with humans over 430 million years ago, make poor representatives for precisely knowing human TLR-ligand interactions.

Finally, as awareness is growing that species-specific TLR actions reflect evolutionarily relevant adaptations to receptor function, differential TLR actions between (preferably closely related) species can be used to molecularly dissect and pinpoint which amino acids play a significant role in TLR function. Based on the differential trafficking of zebrafish and human TLR5 in the same cell background, we used this comparative TLR approach in **Chapter 5** which resulted in the proposition that the tail of TLR5 is not a feature directly involved in trafficking but

may critically contribute to receptor processing by phosphorylation of a conserved threonine residue. An even stronger example of this approach is illustrated in **Chapter 7** with the differential response of crocodile and alligator TLR15 to *P. aeruginosa*. As these TLR15 orthologs are almost identical (95%) their differential sensitivity is mediated by only very few amino acids. With the use of chimeric receptors we found that this difference is not caused by amino acid variation in the receptor N-terminus, narrowing the difference of TLR15 sensitivity to *P. aeruginosa* down to 21 on a total of 875. Thus, species-specific TLR functioning contains an evolutionarily fueled discriminatory power that can greatly aid in the identification of functionally important sites.

Future perspectives and concluding remarks

Being primary sensors for environmental cues, TLRs play a crucial role in maintaining immune homeostasis. Understanding TLR biology is therefore fundamental to comprehend the mechanisms underlying health and disease. Because of the evolutionarily conserved nature of TLRs, the principles of their biology can be learned from and apply to both humans and animals. In this thesis we utilized a species-comparative approach that contributed valuable insights into several aspects of TLR biology including receptor evolution, intracellular transport, receptor activation, and ligand detection. Yet, TLR biology is far from fully understood and the receptors will continue to be the focus of intense future research. As discussed in this Chapter, multiple facets of TLR biology deserve additional attention.

The transcriptional regulation of TLRs is vastly complex with receptor specific regulatory elements and transcription factors driving self- and cross-regulation following receptor activation. Complicated by cell-type- and species-specificity, understanding transcriptional regulation of TLRs will take immense efforts but is ultimately essential for knowing the receptor dynamics that are needed for controlling infection or that erroneously contribute to detrimental auto-immune responses.

The ligand binding ectodomains of TLRs can be extensively decorated with glycan moieties which can affect almost every step in the TLR life cycle. Mapping the receptor sites involved and the exact glycan structures bound to it will help to uncover the extensive contribution of glycans to receptor functioning and promises to unveil the exciting possibility of non-canonical receptor activation by novel agonists such as microbial lectins.

TLRs vary roughly in length from 750 to 1050 amino acids and the structure that TLRs adopt contains unique signatures that direct receptor specific maturation,

intracellular transport, ligand binding and signal initiation. Precisely charting these signature regions will elucidate what regions participate in what processing steps and how structurally conserved receptors are handled with great specificity by the cell. Our novel findings described in **Chapter 5** set the trend for particular interest in the receptor specific C-terminal tail region.

Unique interactions throughout host-microbe co-evolution have resulted in intra- and inter-species variation among TLRs. Uncovering functional species-specificity in receptor-ligand interactions, which we have done in several Chapters in this thesis, contributes greatly to the understanding of the evolutionary history of both the receptor and its ligands. This information is vital for deciphering the molecular mechanisms underlying an animal's susceptibility or resistance to disease and hence the basis of zoonotic diseases. Species-specificity in TLR actions is furthermore crucial in evaluating the translationality of model organisms and can also be used as a tool to identify functionally important receptor regions, as we have shown in several Chapters using orthology based chimeric receptors.

The various aspects of TLR biology should not only be studied in relation to infectious diseases, i.e. with microbial ligands. It becomes increasingly clear that TLRs also take part in non-immune related processes. It is thus important to also investigate e.g. transcriptional regulation or species-specificity in the context of endogenous ligand detection during development or sterile inflammation. It goes without saying that investigating TLRs at such diverse levels promises to identify receptor regions or interacting partners that may make high potential targets for pharmaceuticals that modulate the TLR system in both human and veterinary medicine.

Obviously, every aspect of TLR biology, whether it is regulation, gene duplication or ligand binding, evolved from some ancestral state to become what it is now. Studying molecular evolution can therefore be fruitful in uncovering the mechanisms underlying function. Evolution is represented in species diversity and speciation (the evolutionary process of the formation of new species) is driven by adaptations to diverging environmental challenges. Comparing species thus holds power to identify important adaptations that were necessary for overcoming a certain challenge. When applied to receptor studies, this means that comparing TLRs of different species enables identification of receptor sites that are important for function. The current thesis is proof that such a species comparative approach, including the use of orthology based chimeric receptors, can be instrumental in uncovering fundamental principles of TLR biology.

In conclusion, Toll-like receptors are an important and fascinating family of proteins from an evolutionary, functional and health perspective. Their complexity demands that much work must still be done before we fully comprehend these receptors. Yet, the current thesis has shed new light on several aspects of TLR biology that together with the species comparative approach hopefully inspires and serves others in expanding and applying the knowledge on TLR biology.

References

1. Woolhouse, M. E. J., Webster, J. P., Domingo, E., Charlesworth, B. & Levin, B. R. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics* **32**, 569–577 (2002).
2. Paterson, S. *et al.* Antagonistic coevolution accelerates molecular evolution. *Nature* **464**, 275–278 (2010).
3. Drosophila 12 Genomes Consortium. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* **450**, 203–218 (2007).
4. Obbard, D. J., Jiggins, F. M., Halligan, D. L. & Little, T. J. Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Current Biology* **16**, 580–585 (2006).
5. McTaggart, S. J., Obbard, D. J., Conlon, C. & Little, T. J. Immune genes undergo more adaptive evolution than non-immune system genes in *Daphnia pulex*. *BMC Evolutionary Biology* **12**, 63 (2012).
6. Fumagalli, M. *et al.* Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLOS Genetics.* **7**, e1002355 (2011).
7. Barreiro, L. B. *et al.* Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. *PLOS Genetics.* **5**, e1000562 (2009).
8. Tschirren, B., Råberg, L. & Westerdahl, H. Signatures of selection acting on the innate immunity gene Toll-like receptor 2 (TLR2) during the evolutionary history of rodents. *Journal of Evolutionary Biology* **24**, 1232–1240 (2011).
9. Mikami, T., Miyashita, H., Takatsuka, S., Kuroki, Y. & Matsushima, N. Molecular evolution of vertebrate Toll-like receptors: evolutionary rate difference between their leucine-rich repeats and their TIR domains. *Gene* **503**, 235–243 (2012).
10. Velová, H., Gutowska-Ding, M. W., Burt, D. W. & Vinkler, M. Toll-like receptor evolution in birds: gene duplication, pseudogenization, and diversifying selection. *Molecular Biology and Evolution.* **35**, 2170–2184 (2018).
11. Vijay-Kumar, M. *et al.* Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. *Science* **328**, 228–231 (2010).
12. Kinnebrew, M. A. *et al.* Interleukin 23 production by intestinal CD103+CD11b+ dendritic cells in response to bacterial flagellin enhances mucosal innate immune defense. *Immunity* **36**, 276–287 (2012).
13. Carvalho, F. A. *et al.* Transient inability to manage Proteobacteria promotes chronic gut inflammation in TLR5-deficient mice. *Cell Host Microbe* **12**, 139–152 (2012).
14. Yonekura, K., Maki-Yonekura, S. & Namba, K. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424**, 643–650 (2003).
15. Beatson, S. A., Minamino, T. & Pallen, M. J. Variation in bacterial flagellins: from sequence to structure. *Trends Microbiology.* **14**, 151–155 (2006).
16. Andersen-Nissen, E. *et al.* Evasion of Toll-like receptor 5 by flagellated bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 9247–9252 (2005).
17. Andersen-Nissen, E., Smith, K. D., Bonneau, R., Strong, R. K. & Aderem, A. A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. *Journal of Experimental Medicine* **204**, 393–403 (2007).
18. Forstnerič, V., Ivičak-Kocjan, K., Ljubetič, A., Jerala, R. & Benčina, M. Distinctive recognition of flagellin by human and mouse Toll-like receptor 5. *PLOS One* **11**, e0158894 (2016).
19. Waite, D. W. *et al.* Comparative genomic analysis of the class Epsilonproteobacteria and proposed reclassification to Epsilonbacteraeota (phyl. nov.). *Frontiers in Microbiology* **8**, (2017).
20. Lightfield, K. L. *et al.* Critical function for Naip5 in inflammasome activation by a conserved carboxy-terminal domain of flagellin. *Nature Immunology* **9**, 1171–1178 (2008).

21. Zhao, Y. *et al.* The NLRC4 inflammasome receptors for bacterial flagellin and type III secretion apparatus. *Nature* **477**, 596–600 (2011).
22. Kofoed, E. M. & Vance, R. E. Innate immune recognition of bacterial ligands by NAIPs determines inflammasome specificity. *Nature* **477**, 592–595 (2011).
23. Tenthorey, J. L. *et al.* The structural basis of flagellin detection by NAIP5: A strategy to limit pathogen immune evasion. *Science* **358**, 888–893 (2017).
24. Jiang, Y. *et al.* Isolation and expression of grass carp toll-like receptor 5a (CiTLR5a) and 5b (CiTLR5b) gene involved in the response to flagellin stimulation and grass carp reovirus infection. *Fish and Shellfish Immunology* **44**, 88–99 (2015).
25. Innan, H. & Kondrashov, F. The evolution of gene duplications: classifying and distinguishing between models. *Nature Reviews Genetics* **11**, 97–108 (2010).
26. Force, A. *et al.* Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531–1545 (1999).
27. Meijer, A. H. *et al.* Expression analysis of the Toll-like receptor and TIR domain adaptor families of zebrafish. *Molecular Immunology* **40**, 773–783 (2004).
28. Wlasiuk, G., Khan, S., Switzer, W. M. & Nachman, M. W. A history of recurrent positive selection at the Toll-like receptor 5 in primates. *Molecular Biology and Evolution* **26**, 937–949 (2009).
29. Bainová, H. *et al.* First evidence of independent pseudogenization of Toll-like receptor 5 in passerine birds. *Developmental and Comparative Immunology* **45**, 151–155 (2014).
30. Hawn, T. R. *et al.* A common dominant TLR5 stop codon polymorphism abolishes flagellin signaling and is associated with susceptibility to Legionnaires' disease. *Journal of Experimental Medicine* **198**, 1563–1572 (2003).
31. Hawn, T. R. *et al.* A stop codon polymorphism of Toll-like receptor 5 is associated with resistance to systemic lupus erythematosus. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10593–10597 (2005).
32. Andrade, F. B. *et al.* Distribution of serine protease autotransporters of Enterobacteriaceae in typical and atypical enteroaggregative *Escherichia coli*. *Infection Genetics and Evolution* **50**, 83–86 (2017).
33. Kareem, S. M., Al-Kadmy, I. M. S., Al-Kaabi, M. H., Aziz, S. N. & Ahmad, M. *Acinetobacter baumannii* virulence is enhanced by the combined presence of virulence factors genes phospholipase C (plcN) and elastase (lasB). *Microbial Pathogensis* **110**, 568–572 (2017).
34. van der Plas, M. J. A. *et al. Pseudomonas aeruginosa* elastase cleaves a C-terminal peptide from human thrombin that inhibits host inflammatory responses. *Nature Communications* **7**, 11567 (2016).
35. Issa, N. *et al.* The circulating protease Persephone is an immune sensor for microbial proteolytic activities upstream of the *Drosophila* Toll pathway. *Molecular Cell* **69**, 539-550.e6 (2018).
36. Ossovskaya, V. S. & Bunnett, N. W. Protease-activated receptors: contribution to physiology and disease. *Physiological Reviews* **84**, 579–621 (2004).
37. Wojtukiewicz, M. Z., Hempel, D., Sierko, E., Tucker, S. C. & Honn, K. V. Protease-activated receptors (PARs)—biology and role in cancer invasion and metastasis. *Cancer Metastasis Reviews* **34**, 775–796 (2015).
38. Chang, A. K. *et al. Vibrio vulnificus* secretes a broad-specificity metalloprotease capable of interfering with blood homeostasis through prothrombin activation and fibrinolysis. *Journal of Bacteriology* **187**, 6909–6916 (2005).
39. Li, R. *et al.* Expression of immune-related genes of ducks infected with avian pathogenic *Escherichia coli* (APEC). *Frontiers in Microbiology* **7**, (2016).
40. Ying, S. *et al.* Time course effect of lipopolysaccharide on Toll-like receptors expression and steroidogenesis in the Chinese goose ovary. *Reproduction* **153**, 509–518 (2017).
41. Beaudet, J. *et al.* Transcriptional profiling of the chicken tracheal response to virulent *Mycoplasma gallisepticum* strain rlow. *Infection and Immunity* **85**, e00343-17 (2017).

42. Hershberg, R. & Petrov, D. A. Selection on Codon Bias. *Annual Reviews of Genetics* **42**, 287–299 (2008).
43. Deane, J. A. *et al.* Control of Toll-like receptor 7 expression is essential to restrict autoimmunity and dendritic cell proliferation. *Immunity* **27**, 801–810 (2007).
44. Newman, Z. R., Young, J. M., Ingolia, N. T. & Barton, G. M. Differences in codon bias and GC content contribute to the balanced expression of TLR7 and TLR9. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E1362–E1371 (2016).
45. Zhong, F. *et al.* Deviation from major codons in the Toll-like receptor genes is associated with low Toll-like receptor expression. *Immunology* **114**, 83–93 (2005).
46. Iqbal, M., Philbin, V. J. & Smith, A. L. Expression patterns of chicken Toll-like receptor mRNA in tissues, immune cell subsets and cell lines. *Veterinary Immunology and Immunopathology* **104**, 117–127 (2005).
47. Ishii, A., Kawasaki, M., Matsumoto, M., Tochinai, S. & Seya, T. Phylogenetic and expression analysis of amphibian *Xenopus* Toll-like receptors. *Immunogenetics* **59**, 281–293 (2007).
48. Zhang, X.-T. *et al.* Expression analysis of nine Toll-like receptors in yellow catfish (*Pelteobagrus fulvidraco*) responding to *Aeromonas hydrophila* challenge. *Fish and Shellfish Immunology.* **63**, 384–393 (2017).
49. Kigerl, K. A., de Rivero Vaccari, J. P., Dietrich, W. D., Popovich, P. G. & Keane, R. W. Pattern recognition receptors and central nervous system repair. *Experimental. Neurology.* **258**, 5–16 (2014).
50. Vijay, K. Toll-like receptors in immunity and inflammatory diseases: Past, present, and future. *International Immunopharmacology* **59**, 391–412 (2018).
51. Schreibelt, G. *et al.* Toll-like receptor expression and function in human dendritic cell subsets: implications for dendritic cell-based anti-cancer immunotherapy. *Cancer Immunology Immunotherapy* **59**, 1573–1582 (2010).
52. McClure, R. & Massari, P. TLR-dependent human mucosal epithelial cell responses to microbial pathogens. *Frontiers in Immunology* **5**, (2014).
53. Price, A. E. *et al.* A map of Toll-like receptor expression in the intestinal epithelium reveals distinct spatial, cell type-specific, and temporal patterns. *Immunity* **49**, 560-575.e6 (2018).
54. Takeshita, F. *et al.* Transcriptional regulation of the human TLR9 gene. *Journal of Immunology* **173**, 2552–2561 (2004).
55. Marshall, J. D., Heeke, D. S., Gesner, M. L., Livingston, B. & Nest, G. V. Negative regulation of TLR9-mediated IFN-α induction by a small-molecule, synthetic TLR7 ligand. *Journal of Leukocyte Biology* **82**, 497–508 (2007).
56. Dalpke, A. H., Lehner, M. D., Hartung, T. & Heeg, K. Differential effects of CpG-DNA in Toll-like receptor-2/-4/-9 tolerance and cross-tolerance. *Immunology* **116**, 203–212 (2005).
57. Peroval, M. Y., Boyd, A. C., Young, J. R. & Smith, A. L. A critical role for MAPK signalling pathways in the transcriptional regulation of Toll like receptors. *PLOS ONE* **8**, e51243 (2013).
58. van Aubel, R. A. M. H., Keestra, A. M., Krooshoop, D. J. E. B., van Eden, W. & van Putten, J. P. M. Ligand-induced differential cross-regulation of Toll-like receptors 2, 4 and 5 in intestinal epithelial cells. *Molecular Immunology* **44**, 3702–3714 (2007).
59. Rehli, M. Of mice and men: species variations of Toll-like receptor expression. *Trends in Immunology* **23**, 375–378 (2002).
60. Heinz, S. *et al.* Species-specific regulation of Toll-like receptor 3 genes in men and mice. *Journal of Biological Chemistry* **278**, 21502–21509 (2003).
61. Hagai, T. *et al.* Gene expression variability across cells and species shapes innate immunity. *Nature* 197-202 (2018).
62. Kudla, G., Lipinski, L., Caffin, F., Helwak, A. & Zylicz, M. High guanine and cytosine content increases mRNA levels in mammalian cells. *PLOS Biology.* **4**, e180 (2006).
63. Bauer, A. P. *et al.* The impact of intragenic CpG content on gene expression. *Nucleic Acids Research* **38**, 3891–3908 (2010).

64. Presnyak, V. *et al.* Codon optimality is a major determinant of mRNA stability. *Cell* **160**, 1111–1124 (2015).
65. Yang, Y. *et al.* Heat shock protein gp96 is a master chaperone for Toll-like receptors and is important in the innate function of macrophages. *Immunity* **26**, 215–226 (2007).
66. Liu, B. *et al.* Folding of Toll-like receptors by the HSP90 paralogue gp96 requires a substrate-specific cochaperone. *Nature Communications* **1**, 79 (2010).
67. Liu, W. *et al.* Interaction of Toll-like receptors with the molecular chaperone Gp96 is essential for its activation of cytotoxic T lymphocyte response. *PLOS ONE* **11**, e0155202 (2016).
68. Takahashi, K. *et al.* A protein associated with Toll-like receptor (TLR) 4 (PRAT4A) is required for TLR-dependent immune responses. *Journal of Experimental Medicine* **204**, 2963–2976 (2007).
69. Weekes, M. P. *et al.* Proteomic plasma membrane profiling reveals an essential role for gp96 in the cell surface expression of LDLR family members, including the LDL receptor and LRP6. *Journal of Proteome Research.* **11**, 1475–1484 (2012).
70. Tatematsu, M. *et al.* LRRC59 regulates trafficking of nucleic acid–sensing TLRs from the endoplasmic reticulum via association with UNC93B1. *Journal of Immunology* **195**, 4933–4942 (2015).
71. Kreer, C. *et al.* N-glycosylation converts non-glycoproteins into mannose receptor ligands and reveals antigen-specific T cell responses in vivo. *Oncotarget* **8**, 6857–6872 (2016).
72. Torreno-Pina, J. A. *et al.* Enhanced receptor–clathrin interactions induced by N-glycan–mediated membrane micropatterning. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 11037–11042 (2014).
73. Markkanen, P. M. H. & Petäjä-Repo, U. E. N-Glycan-mediated quality control in the endoplasmic reticulum is required for the expression of correctly folded δ-opioid receptors at the cell surface. *Journal of Biological Chemistry* **283**, 29086–29098 (2008).
74. Dennis, J. W., Lau, K. S., Demetriou, M. & Nabi, I. R. Adaptive regulation at the cell surface by N-glycosylation. *Traffic* **10**, 1569–1578 (2009).
75. Lee, H. S., Qi, Y. & Im, W. Effects of N-glycosylation on protein conformation and dynamics: protein data bank analysis and molecular dynamics simulation study. *Scientific Reports.* **5**, (2015).
76. Correia, J. da S. & Ulevitch, R. J. MD-2 and TLR4 N-Linked glycosylations are important for a functional lipopolysaccharide receptor. *Journal of Biological Chemistry* **277**, 1845–1854 (2002).
77. Kataoka, H. *et al.* Roles of N-linked glycans in the recognition of microbial lipopeptides and lipoproteins by TLR2. *Cellular Microbiology* **8**, 1199–1209 (2006).
78. Ohnishi, T., Muroi, M. & Tanamoto, K. N-Linked glycosylations at Asn26 and Asn114 of human MD-2 are required for Toll-like receptor 4-mediated activation of NF-κB by lipopolysaccharide. *Journal of Immunology* **167**, 3354–3359 (2001).
79. Sun, J. *et al.* Structural and functional analyses of the human Toll-like receptor 3 ROLE OF GLYCOSYLATION. *Journal of Biological Chemistry* **281**, 11144–11151 (2006).
80. Weber, A. N. R., Morse, M. A. & Gay, N. J. Four N-linked glycosylation sites in human Toll-like receptor 2 cooperate to direct efficient biosynthesis and secretion. *Journal of Biological Chemistry* **279**, 34589–34594 (2004).
81. Varki, A. Biological roles of glycans. *Glycobiology* **27**, 3–49 (2017).
82. Alegre-Maller, A. C. P. *et al.* Therapeutic administration of recombinant paracoccin confers protection against *Paracoccidioides brasiliensis* infection: involvement of TLRs. *PLOS Neglected Tropical Diseases* **8**, e3317 (2014).
83. Ricci-Azevedo, R., Roque-Barreira, M.-C. & Gay, N. J. Targeting and recognition of Toll-like receptors by plant and pathogen lectins. *Frontiers in Immunology* **8**, (2017).
84. Majer, O., Liu, B. & Barton, G. M. Nucleic acid-sensing TLRs: trafficking and regulation. *Current Opinion in Immunology* **44**, 26–33 (2017).

85. Kim, J. *et al.* Acidic amino acid residues in the juxtamembrane region of the nucleotide-sensing TLRs are important for UNC93B1 binding and signaling. *Journal of Immunology* **190**, 5287–5295 (2013).
86. Kim, Y.-M., Brinkmann, M. M., Paquet, M.-E. & Ploegh, H. L. UNC93B1 delivers nucleotide-sensing toll-like receptors to endolysosomes. *Nature* **452**, 234–238 (2008).
87. Huh, J.-W. *et al.* UNC93B1 is essential for the plasma membrane localization and signaling of Toll-like receptor 5. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 7072–7077 (2014).
88. Pelka, K. *et al.* The chaperone UNC93B1 regulates Toll-like receptor stability independently of endosomal TLR transport. *Immunity* **48**, 911-922.e7 (2018).
89. Kashuba, V. I. *et al.* hUNC93B1: a novel human gene representing a new gene family and encoding an unc-93-like protein. *Gene* **283**, 209–217 (2002).
90. Barton, G. M., Kagan, J. C. & Medzhitov, R. Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA. *Nature Immunology* **7**, 49–56 (2006).
91. Hart, B. E. & Tapping, R. I. Cell surface trafficking of TLR1 is differentially regulated by the chaperones PRAT4A and PRAT4B. *Journal of Biological Chemistry* **287**, 16550–16562 (2012).
92. Nishiya, T., Kajita, E., Miwa, S. & DeFranco, A. L. TLR3 and TLR7 are targeted to the same intracellular compartments by distinct regulatory elements. *Journal of Biological Chemistry* **280**, 37107–37117 (2005).
93. Lee, B. L. *et al.* UNC93B1 mediates differential trafficking of endosomal TLRs. *eLife* **2**, e00291 (2013).
94. Park, S. Y. & Guo, X. Adaptor protein complexes and intracellular transport. *Bioscience Reports* **34**, (2014).
95. Biswas, C. *et al.* Tyrosine 870 of TLR9 is critical for receptor maturation rather than phosphorylation-dependent ligand-induced signaling. *PLOS ONE* **13**, e0200913 (2018).
96. Blasius, A. L. *et al.* Slc15a4, AP-3, and Hermansky-Pudlak syndrome proteins are required for Toll-like receptor signaling in plasmacytoid dendritic cells. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 19973–19978 (2010).
97. Sasai, M., Linehan, M. M. & Iwasaki, A. Bifurcation of Toll-like receptor 9 signaling by adaptor protein 3. *Science* **329**, 1530–1534 (2010).
98. Chiang, C. *et al.* Cofactors required for TLR7- and TLR9-dependent innate immune responses. *Cell Host Microbe* **11**, 306–318 (2012).
99. Kim, H. M. *et al.* Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell* **130**, 906–917 (2007).
100. Park, B. S. *et al.* The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex. *Nature* **458**, 1191–1195 (2009).
101. Yu, L., Wang, L. & Chen, S. Endogenous toll-like receptor ligands and their biological significance. *Journal of Cellular and Molecular Medicine* **14**, 2592–2603 (2010).
102. Karikó, K., Ni, H., Capodici, J., Lamphier, M. & Weissman, D. mRNA is an endogenous ligand for Toll-like receptor 3. *Journal of Biological Chemistry* **279**, 12542–12550 (2004).
103. Jiang, D. *et al.* Regulation of lung injury and repair by Toll-like receptors and hyaluronan. *Nature Medicine* **11**, 1173–1179 (2005).
104. Das, N. *et al.* HMGB1 activates proinflammatory signaling via TLR5 leading to allodynia. *Cell Reports* **17**, 1128–1140 (2016).
105. Tanji, H., Ohto, U., Shibata, T., Miyake, K. & Shimizu, T. Structural reorganization of the Toll-like receptor 8 dimer induced by agonistic ligands. *Science* **339**, 1426–1429 (2013).
106. Zhang, Z. *et al.* Structural analysis reveals that Toll-like receptor 7 is a dual receptor for guanosine and single-stranded RNA. *Immunity* **45**, 737–748 (2016).

107. Wang, Y. *et al.* TLR4/MD-2 activation by a synthetic agonist with no similarity to LPS. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E884–E893 (2016).
108. Jin, M. S. *et al.* Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* **130**, 1071–1082 (2007).
109. Yoon, S. *et al.* Structural basis of TLR5-flagellin recognition and signaling. *Science* **335**, 859–864 (2012).
110. Gay, N. J., Symmons, M. F., Gangloff, M. & Bryant, C. E. Assembly and localization of Toll-like receptor signalling complexes. *Nature Reviews Immunology* **14**, 546–558 (2014).
111. Latz, E. *et al.* Ligand-induced conformational changes allosterically activate Toll-like receptor 9. *Nature Immunology* **8**, 772–779 (2007).
112. Maruyama, I. N. Activation of transmembrane cell-surface receptors via a common mechanism? The "rotation model". *BioEssays* **37**, 959–967 (2015).
113. Zhou, K., Kanai, R., Lee, P., Wang, H.-W. & Modis, Y. Toll-like receptor 5 forms asymmetric dimers in the absence of flagellin. *Journal of Structural Biology* **177**, 402–409 (2012).
114. Endres, N. F. *et al.* Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* **152**, 543–556 (2013).
115. Cheng, Y. Membrane protein structural biology in the era of single particle cryo-EM. *Current Opinion Structural Biology* **52**, 58–63 (2018).
116. Saitoh, S.-I. Chaperones and transport proteins regulate TLR4 trafficking and activation. *Immunobiology* **214**, 594–600 (2009).
117. Husebye, H. *et al.* Endocytic pathways regulate Toll-like receptor 4 signaling and link innate and adaptive immunity. *EMBO Journal* **25**, 683–692 (2006).
118. Triantafilou, M. *et al.* Lipoteichoic acid and Toll-like receptor 2 internalization and targeting to the golgi are lipid raft-dependent. *Journal of Biological Chemistry* **279**, 40882–40889 (2004).
119. Baldissera, M. D. *et al. Pseudomonas aeruginosa* strain PA01 impairs enzymes of the phosphotransfer network in the gills of *Rhamdia quelen*. *Veterinary Microbiology* **201**, 121–125 (2017).
120. Di Girolamo, N. *et al.* Computed tomography–guided bone biopsies for evaluation of proliferative vertebral lesions in two boa constrictors (*Boa constrictor imperator*). *Journal of Zoo and Wildlife Medicine* **45**, 973–978 (2014).
121. Folkesson, A. *et al.* Adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis airway: an evolutionary perspective. *Nature Reviews Microbiology* **10**, 841–851 (2012).
122. Seixas, R. *et al.* Severe fibrinonecrotic enteritis caused by *Pseudomonas aeruginosa* in a captive monitor lizard (*Varanus niloticus*). *Journal of Zoo and Wildlife Medicine* **45**, 410–412 (2014).
123. Walker, S. E., Sander, J. E., Cline, J. L. & Helton, J. S. Characterization of *Pseudomonas aeruginosa* isolates associated with mortality in broiler chicks. *Avian Diseases* **46**, 1045–1050 (2002).
124. Zimmerman, L. M., Vogel, L. A. & Bowden, R. M. Understanding the vertebrate immune system: insights from the reptilian perspective. *Journal of Experimental Biology* **213**, 661–671 (2010).
125. Keestra, A. M., de Zoete, M. R., van Aubel, R. A. M. H. & van Putten, J. P. M. Functional characterization of chicken TLR5 reveals species-specific recognition of flagellin. *Molecular Immunology* **45**, 1298–1307 (2008).
126. Tahoun, A. *et al.* Host species adaptation of TLR5 signalling and flagellin recognition. *Scientific Reports* **7**, 17677 (2017).
127. Oblak, A. & Jerala, R. The molecular mechanism of species-specific recognition of lipopolysaccharides by the MD-2/TLR4 receptor complex. *Molecular Immunology* **63**, 134–142 (2015).
128. Oblak, A. & Jerala, R. Species-specific activation of TLR4 by hypoacylated endotoxins governed by residues 82 and 122 of MD-2. *PLOS ONE* **9**, e107520 (2014).

129. Ohto, U., Fukase, K., Miyake, K. & Shimizu, T. Structural basis of species-specific endotoxin sensing by innate immune receptor TLR4/MD-2. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 7421–7426 (2012).
130. Keestra, A. M. & Putten, J. P. M. van. Unique properties of the chicken TLR4/MD-2 complex: selective lipopolysaccharide activation of the MyD88-dependent pathway. *Journal of Immunology* **181**, 4354–4362 (2008).
131. Ishida, H., Ohto, U., Shibata, T., Miyake, K. & Shimizu, T. Structural basis for species-specific activation of mouse Toll-like receptor 9. *FEBS Letters* **592**, 2636–2646 (2018).
132. Keestra, A. M., Zoete, M. R. de, Bouwman, L. I. & Putten, J. P. M. van. Chicken TLR21 is an innate CpG DNA receptor distinct from mammalian TLR9. *Journal of Immunology* **185**, 460–467 (2010).
133. Clark, J. M. Crocodilian Biology and Evolution. *Journal of Paleontology* **77**, 597–597 (2003).
134. Green, R. E. *et al.* Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. *Science* **346**, 1254449 (2014).
135. Montminy, S. W. *et al.* Virulence factors of *Yersinia pestis* are overcome by a strong lipopolysaccharide response. *Nature Immunology* **7**, 1066–1073 (2006).
136. Hajjar, A. M. *et al.* Humanized TLR4/MD-2 mice reveal LPS recognition differentially impacts susceptibility to *Yersinia pestis* and *Salmonella enterica*. *PLOS Pathogens* **8**, e1002963 (2012).
137. Novoa, B. & Figueras, A. Zebrafish: model for the study of inflammation and the innate immune response to infectious diseases. In: Lambris J., Hajishengallis G. (eds) Current Topics in Innate Immunity II. Advances in Experimental Medicine and Biology, vol 946. Springer, New York, NY (2012).
138. Ivičak-Kocjan, K., Forstnerič, V., Panter, G., Jerala, R. & Benčina, M. Extension and refinement of the recognition motif for Toll-like receptor 5 activation by flagellin. *Journal of Leukocyte Biology* **104**, 767–776 (2018).
139. Koymans, K. J. *et al.* Molecular basis determining species specificity for TLR2 inhibition by staphylococcal superantigen-like protein 3 (SSL3). *Veterinary Research* **49**, 115 (2018).

Summary

All animals critically rely on their immune system for protection against infectious microorganisms and disruption of homeostasis. An animal's immune response typically starts with receptors that detect a danger signal and subsequently initiate the expression and/or release of effectors. The effectors then clear the source of the danger signal to restore homeostasis. One family of receptors that play a key role in the immune system are the Toll-like receptors (TLRs). TLRs are expressed in endolysosomal compartments or on the surface of cells where they scan the extracellular environment for the presence of molecules from microorganisms or damaged tissues. Upon detection of such molecules, TLRs activate transcription factors that drive the expression of immune system effectors. As TLRs can detect both microbial and endogenous molecules, these receptors are strongly involved in many infectious and non-infectious diseases including auto-immune diseases and cancer. This has marked TLRs as attractive targets for pharmacological modulation. Yet despite the interest in TLRs, multiple aspects of these receptors, for example their evolutionary history, mode of activation and interaction with ligand, are not yet fully understood. The incomplete knowledge about multiple aspects of TLR biology impairs comprehending the precise role that TLRs play in diverse diseases and may be one of the reasons why therapeutics aimed at altering the function of TLRs are not yet effective in mitigating diseases. To improve this situation it is important to gain a deeper understanding of the fundamental principles that underlie the biology of these receptors.

TLRs are present in almost all animals and during animal evolution the TLR gene repertoire has expanded and diversified. This has resulted in a large family of receptors of which the size and composition differ between animal species. New species are formed by an evolutionary force that is driven by the necessity to adapt functions to changing environmental challenges. In the case of TLRs this means that receptors of different species may have undergone species-specific adaptations that alter the receptor's function to meet a specific challenge. Comparing TLRs of different species therefore holds discriminatory power to discover features of the receptor that are important for its function. In this thesis we have implemented this concept by studying TLR biology using a species-comparative strategy. With this strategy our work has resulted in novel insights in multiple aspects of fundamental TLR biology.

Understanding the evolutionary history of a protein can greatly aid in interpreting its role in the biology of an organism. In **Chapter 2** we reviewed the evolution of TLRs. These receptors, and their orthologs called Toll receptors, are ancient as the ancestral receptor originated in the eumetazoan ancestor roughly 600 million years ago. TLRs and Toll receptors have been highly conserved ever since in all deuterostomian (e.g. vertebrates) and protostomian (e.g. insects) animals. In almost all animals, TLRs as well as some of the Toll receptors are involved in the sensing of molecules derived from microorganisms. The ancestral form of these receptors therefore likely evolved to play a role in the immune system. While the general structure of the different TLRs and Toll receptors has remained highly similar, species-specific evolutionary requirements have resulted in large differences in the number of receptor genes present in animals. For example, the nematode *C. elegans* only has one Toll receptor while the purple sea urchin has more than 250 TLRs encoded in its genome. The expansion and diversification of TLRs throughout animal evolution is partly driven by the microbes that co-evolve with the animal host. Some bacteria have evolved strategies to evade detection by TLRs and thus avoid alerting the host immune system. Such microbial strategies, as well as animal-specific co-evolution with distinct microbes, forces a selective pressure onto TLRs that drives either adaptive evolution, seen in TLRs that detect structures specific to bacteria, or purifying evolution seen in TLRs that detect nucleic acids which are not specific to microorganisms. Finally, in protostomian animals some Toll receptors are involved in for instance embryogenesis. Evidence is accumulating that TLRs in deutrostomian animals, by detecting endogenous ligands, are also involved in physiological processes other than immune responses.

An ancient TLR family member that has remained highly conserved throughout animal evolution is TLR5. In mammals and birds, TLR5 detects the bacterial protein flagellin and in mammals this receptor is important for detecting pathogenic bacteria and shaping the microbial community in the intestine. Prior to this thesis it was not known whether the direct detection of flagellin by TLR5 was conserved among species other than mammals and birds. The common ancestors of mammals and birds were early reptiles and reptiles thus take a central position in vertebrate evolution. Studying reptiles may hence aid in understanding host-microbe co-evolution but the immune system of reptiles is strongly understudied compared to that of other vertebrates. In **Chapter 3** we provided the first functional characterization of TLR5, and with that of any TLR, from a reptile. We observed that cells of an iguana lizard activated the NF-κB transcription factor in response to bacterial flagellin. In another reptilian species, the green anole lizard, gene

expression of the flagellin receptor TLR5 was demonstrated throughout the body. Functional studies with recombinant anole TLR5 showed that this receptor detects bacterial flagellin and induced activation of NF-κB in reptile and human cells, indicating strong evolutionary conservation of both TLR5 ligand binding and signaling capacity. The anole TLR5 was found to recognize the D1 domain in flagellin, just as human TLR5. Yet, given their long independent evolution we questioned whether anole and human TLR5 had developed a differential sensitivity to bacterial flagellins. While anole TLR5 was similarly sensitive as human TLR5 to flagellin of *Salmonella*, the reptile receptor proved more sensitive to flagellin of *Pseudomonas*, which is an opportunistic pathogen to reptiles and humans. These findings indicate that the recognition of a conserved domain in flagellin by TLR5 has remained important for more than 320 million years of vertebrate evolution. In addition, our findings illustrate that a selective pressure exerted by distinct microbes drove different adaptations among reptiles and mammals thereby leading to species-specific recognition of the danger signal flagellin.

The species-specific recognition of flagellin indicates that there are functionally important differences between animals in the flagellin-binding site of their TLR5. Flagellin directly binds to a site in the extracellular domain of TLR5 and for the rational design of strategies to therapeutically target TLR5 in the future it is imperative to understand the exact molecular interactions of flagellin binding to TLR5. A previously published crystal structure of flagellin in complex with TLR5b of the zebrafish is used as a model to understand the interaction of flagellin with TLR5 from other species, including human TLR5. However, activation of zebrafish TLR5b by flagellin had never been reported. In addition, while the zebrafish is a widely used model animal in biomedical science, it is an exception in terms of TLRs as the zebrafish carries several additional copies of TLR genes including two *tlr5* genes. The role of these two zebrafish TLR5 paralogs in flagellin detection was unknown. In **Chapter 4** we discovered that the zebrafish TLR5 paralogs TLR5a and TLR5b, unlike TLR5 of any other species, do not detect flagellin as conventional homodimers but instead evolved to physically cooperate and form a heterodimeric flagellin receptor. We also identified that the TLR chaperone UNC93B1 of the zebrafish strongly contributed to TLR5 heterodimerization by facilitating transport of both receptor paralogs to intracellular vesicles. To better understand what heterodimerization of zebrafish TLR5 meant for the receptor interaction with flagellin, we performed a detailed functional analysis using chimeric receptors based on the TLR5b crystal structure. This analysis showed that there are subtle but functionally important differences in the structure of TLR5a, TLR5b and human TLR5 which cannot be explained using

the TLR5b crystal structure as a model. Further analysis with the use of constructed chimeric receptors revealed that to be functional, zebrafish TLR5a and TLR5b must complement each other across multiple regions of the dimeric complex suggesting that a transregional conformational change, possibly via rotation, underlies the TLR5 activational mechanism.

Within the structure of each TLR multiple distinct regions can be identified and for most regions a function has been assigned. The TLR extracellular domain binds ligands, the transmembrane region is necessary for embedment in the membrane and involved in receptor dimerization, and the TIR domain facilitates signal induction. In **Chapter 5** we discovered that the C-terminal tail region of human TLR5, a region without prior known function, is necessary for receptor localization and ligand-induced signaling. TLRs are localized to different cellular compartments but the receptor features that direct transport towards these compartments are poorly defined. When we removed the C-terminal tail of human TLR5 the receptor no longer localized at the plasma membrane and did not respond to flagellin stimulation. Surprisingly, the somewhat evolutionarily conserved charged amino acids in the TLR5 tail were not involved in receptor localization and function. To determine whether potential localization motifs were hidden in the tail, we randomly scrambled its sequence which resulted in the blocking of receptor trafficking and function, indicating that the role of the tail was sequence dependent. However, replacement of the human TLR5 tail with the tail from zebrafish TLR5b, which is also a considerably different sequence, still enabled the chimeric receptor to reach the plasma membrane and respond to flagellin. Interestingly, both the human and zebrafish TLR5 tail sequence were predicted to be phosphorylated at threonine residues and scrambling of the human TLR5 tail sequence weakened this prediction. These novel findings reveal a critical contribution of the TLR5 tail region to receptor localization and function and point towards evolutionarily conserved threonine phosphorylation as a potential mechanism.

The conservation of TLR5 in almost all vertebrate animals indicates that TLR5 evolved very stably and that the detection of flagellin by TLR5 has likely remained an important feature throughout vertebrate evolution. Not all TLRs evolved as stably as TLR5. Members of the TLR1 subfamily evolved much more dynamically across vertebrates and show losses or duplications of TLR genes in diverse clades of animals. TLR15 is a member of the TLR1 subfamily that was previously found to be exclusively present in birds and reptiles. Yet, the evolutionary history of this receptor and whether its function as a receptor for microbial proteases remained

conserved between birds and reptiles was still unclear. **Chapter 6** describes our investigation of the functional evolution of avian and reptilian TLR15. After analyzing a large collection of TLR sequences from diverse vertebrate species, we were surprised to find a TLR15 ortholog to be present in a shark species. This finding reversed the understanding of TLR15 evolution as it indicated that TLR15 is actually an ancient TLR and not an invention in the bird and reptile lineage. Although ancient, the function of TLR15 was likely redundant multiple times throughout evolution as we found that the *tlr15* gene had been lost from the genomes of many vertebrates, including turtles. Still, functional analysis with the TLR15 of a lizard and two crocodile species showed that these receptors detected fungal proteolytic activity, just like the chicken TLR15 ortholog, indicating conservation of TLR15 function through more than 280 million years of evolution. We observed a peculiar difference in protein expression efficiency among the functional reptile and chicken TLR15s which revealed another interesting feature about TLR15 evolution; the codon usage among the unstably evolving *tlr15* genes was highly variable, much more than among *tlr* genes that show a more stable evolution such as TLR5. Combined, these findings indicate that TLR15 evolved far more dynamically than most TLRs and that species-specific codon bias is an important determinant in TLR expression and potentially useful as a prediction parameter for the evolutionary fate of TLRs.

All TLRs recognize conserved microbial ligands that are important for microbe viability. The identification of bird and reptilian TLR15 as a receptor for secreted microbial proteases is therefore unusual and raised questions about the range of micro-organisms that can be detected and whether the detection of microbial proteolytic activity by TLR15 is species-specific. In **Chapter 7** we observed that TLR15 of an alligator could respond to secreted fractions of fungal, yeast as well as bacterial pathogens while TLR15 of the anole lizard only responded to the fungal pathogen. Chicken and crocodile TLR15 also showed variable responses to these pathogens suggesting that unlike most other TLRs, ligand detection by TLR15 developed highly variable throughout host-microbe co-evolution. Testing of the secreted fraction of the pathogenic bacterium *Pseudomonas aeruginosa* revealed an unexpected difference in sensitivity between the highly similar alligator and crocodile TLR15, suggesting that only very few, possibly even one, amino acid conveys species-specificity in TLR15 function. Finally, in search for a specific agonist that activated TLR15 we tested several *P. aeruginosa* protease mutants. This identified the multifunctional virulence factor LasB as a likely activator of TLR15. As a receptor to microbial proteases, TLR15 is unique among

the TLRs which generally detect highly conserved microbial structures. The unstably evolving nature of its ligand, which include pathogen-specific proteases, may partly explain the observed high level of species-specific functionality and evolutionary regression of TLR15.

In **Chapter 8** our novel findings are integrated in a general discussion about different aspects of TLR biology. One major topic includes the factors that shape receptor-ligand evolution, with TLR5 and TLR15 as attractive examples due to their opposite evolutionary histories (stable vs. dynamic). Other topics are discussed from a cell biological perspective following the cellular life cycle of a TLR, i.e. receptor expression, maturation, transport and ligand interaction. Lastly, the evolution driven species-specificity of ligand recognition and its diverse implications are discussed and future perspectives on TLR research are presented.

Collectively, the work described throughout the chapters of this thesis demonstrate the strength of evolution based, species-comparative research for better understanding the principles behind the function of a protein. Hopefully, the novel insights in TLR biology gained through this approach and described herein will serve future research on TLR biology.

Samenvatting in het Nederlands

Alle dieren zijn afhankelijk van hun immuunsysteem voor de bescherming tegen ziekmakende micro-organismen en verstoringen in het interne milieu. Een reactie van het immuunsysteem bestaat grofweg uit twee fasen. In fase één worden moleculen die een gevaar aanduiden, ook wel liganden genoemd, (bijvoorbeeld specifieke liganden afkomstig van een bacterie) gedetecteerd door zogenaamde receptoren. De receptoren activeren vervolgens het immuunsysteem. In fase twee leidt deze activering tot de productie van afweer-moleculen die de bacterie, schimmel, virus of parasiet uitschakelen. Ieder dier heeft heel veel verschillende receptoren om een breed scala aan micro-organismen te herkennen. Een van de belangrijkste receptoren van het immuunsysteem zijn de Toll-like receptoren (TLRs). [TLRs zijn eiwitten]. Er zijn veel varianten van TLRs die samen de TLR familie vormen. Sommige TLR varianten zitten in de membraan van cellen terwijl andere in speciale compartimenten van de cel zitten. Vanaf deze posities scannen de TLRs de extracellulaire omgeving op de aanwezigheid van micro-organismen of liganden van beschadigde weefsels. Als een TLR dergelijke liganden detecteert worden transcriptie factoren geactiveerd. Deze zorgen ervoor dat speciale genen tot expressie komen die nodig zijn voor het bestrijden van het gedetecteerde gevaar. Omdat TLRs zowel liganden van micro-organismen als van beschadigde cellen kunnen detecteren spelen deze receptoren een rol in diverse ziekten waaronder infectieziekten, auto-immuunziekten en kanker. Om deze reden zijn TLRs een aantrekkelijk doelwit voor de ontwikkeling van geneesmiddelen die als doel hebben de functie van TLRs aan te passen. Echter, na ruim 20 jaar onderzoek en ondanks de grote interesse in TLRs bestaan nog vele vraagtekens over bijvoorbeeld hun evolutie, manier van activatie en herkenning van ligand. Door dit gebrek aan kennis wordt het exact begrijpen van de rol die TLRs spelen in diverse ziektes belemmert. De incomplete kennis over TLR biologie is waarschijnlijk ook de reden waarom geneesmiddelen gericht op TLRs vaak niet effectief blijken. Het doel van ons onderzoek is om een beter begrip te krijgen van de fundamentele principes die ten grondslag liggen aan de biologie van deze receptoren.

TLRs zijn aanwezig in bijna alle dieren en gedurende de evolutie van dieren is het aantal TLR varianten uitgebreid en diverser geworden. Dit heeft geleid tot een grote familie van receptoren waarvan de omvang en samenstelling per diersoort kan verschillen. Ook de receptoren zelf verschillen per diersoort. Deze verschillen zijn ontstaan als gevolg van diersoort specifieke aanpassingen aan de functie van de receptor. Het vergelijken van het functioneren van TLRs van verschillende diersoorten biedt daarom een onderscheidend vermogen waarmee eigenschappen

van de receptor ontdekt kunnen worden die een rol spelen bij zijn functie. In dit proefschrift staat beschreven hoe wij middels dit concept van diersoort vergelijkend onderzoek nieuwe inzichten hebben verschaft in meerdere fundamentele aspecten van TLR biologie.

Weten hoe een eiwit, zoals een TLR, geëvolueerd is kan sterk bijdragen aan het begrijpen van de rol van het eiwit in de biologie van een organisme. In **Hoofdstuk 2** van dit proefschrift beschrijven we de huidige kennis van de evolutie van TLRs. Deze receptoren zijn ongeveer 600 miljoen jaar geleden ontstaan in de voorouder van bijna alle dieren. Sinds het moment van ontstaan zijn TLRs altijd bewaard gebleven in alle dieren van de deuterostomia (bijvoorbeeld gewervelde dieren) en protostomia (bijvoorbeeld insecten) groepen. Omdat in al deze dieren de TLRs betrokken zijn bij het detecteren van liganden van micro-organismen is het waarschijnlijk dat de vooruderlijke variant van de TLR is ontstaan om een rol te vervullen in het immuunsysteem. Tijdens de evolutie van dieren is het aantal receptoren per diersoort flink gaan verschillen. De nematode *C. elegans* bijvoorbeeld heeft maar één TLR terwijl de paarse zee-egel meer dan 250 TLR genen heeft. De uitbreiding en diversificatie van TLRs gedurende de evolutie van dieren is deels gedreven door de micro-organismen waarmee ieder dier is geëvolueerd. Sommige micro-organismen waaronder bepaalde bacteriën hebben strategieën ontwikkeld om niet herkend te worden door TLRs en zo activatie van het immuunsysteem te voorkomen. Dergelijke strategieën als ook diersoort specifieke co-evolutie met bepaalde micro-organismen oefenen een evolutionaire druk uit op een diersoort om TLRs aan te passen of om de receptor juist te behouden zoals hij was in de voorouder. In sommige dieren, met name insecten, is al langer bekend dat TLRs ook een rol spelen in systemen anders dan het immuunsysteem, bijvoorbeeld in de ontwikkeling van embryo's. Er wordt steeds meer bewijs gevonden dat ook in andere dieren, waaronder zoogdieren, TLRs meer functies vervullen dan aanvankelijk gedacht werd. Hoe deze nieuwe functies van TLRs de evolutie van de receptoren beïnvloed hebben is een interessant onderwerp dan nader onderzocht dient te worden.

Een van de TLRs die sterk behouden is gedurende de evolutie van dieren is TLR5. In zoogdieren en vogels herkent TLR5 het bacteriële eiwit flagelline en dit is belangrijk voor het herkennen van ziekmakende bacteriën en het vormen van de juiste microbiota in de darm. Voordat dit proefschrift tot stand kwam was het niet bekend of TLR5 van dieren anders dan zoogdieren en vogels ook flagelline kan herkennen. De voorouders van vogels en zoogdieren waren prehistorische reptielen

en reptielen hebben daarom een centrale positie in de evolutie van gewervelde dieren. Het bestuderen van reptielen kan zodanig een bijdrage leveren aan het begrijpen van de co-evolutie van dier en micro-organisme maar van alle gewervelde dieren is er het minst bekend over het immuunsysteem van reptielen. In **Hoofdstuk 3** hebben wij voor het eerst de functie van een TLR van een reptiel beschreven. We vonden dat cellen van een leguaan reageren op flagelline en dat TLR5 tot expressie komt in veel verschillende weefsel van de groene anolis (een hagedis). Functionele studies met TLR5 van de groene anolis lieten zien dat de receptor flagelline kan herkennen wanneer de receptor in zowel reptielen and menselijke cellen tot expressie werd gebracht. Ook vonden we dat TLR5 van de groene anolis net als TLR5 van mensen hetzelfde specifieke onderdeel van flagelline herkend. Echter, gezien de lange en onafhankelijke evolutie van reptielen en zoogdieren vroegen we ons af of de gevoeligheid voor het herkennen van flagelline tussen de TLR5 van reptielen en zoogdieren zou verschillen. We vonden dat TLR5 van de groene anolis en van de mens even gevoelig waren voor flagelline van de bacterie *Salmonella*. Maar de TLR5 van de groene anolis bleek gevoeliger te zijn dan humaan TLR5 voor flagelline van een andere ziekmakende bacterie namelijk *Pseudomonas aeruginosa*. Tezamen betekenen deze bevindingen dat de herkenning van flagelline als ligand en de activatie van het immuunsysteem door TLR5 belangrijk zijn gebleven gedurende meer dan 320 miljoen jaar sinds het moment dat reptielen en zoogdieren gescheiden zijn. Daarnaast illustreren onze resultaten dat een evolutionaire druk uitgeoefend door verschillende bacteriën heeft geleid tot verschillende aanpassingen onder reptielen en zoogdieren waardoor de herkenning van flagelline diersoort specifiek te geworden.

De diersoort specifieke herkenning van flagelline toont aan dat er belangrijke functionele verschillen zijn tussen de TLR5 van verschillende diersoorten in de manier waarop flagelline en TLR5 een interactie aangaan. Het flagelline eiwit bindt aan het extracellulaire deel van TLR5 en om geneesmiddelen te ontwikkelen die de functie van TLR5 kunnen beïnvloeden is het van belang om exact te weten hoe flagelline bindt aan TLR5. In het verleden is de eiwit structuur gepubliceerd waarin te zien is hoe flagelline exact bindt aan de TLR5b van de zebravis. Deze structuur wordt gebruikt als model om te voorspellen hoe bijvoorbeeld ook flagelline aan humaan TLR5 zou kunnen binden. Echter, het is nooit aangetoond dat de TLR5b van de zebravis een functionele receptor is. Daarnaast blijkt de zebravis, welke een belangrijk model dier is in biomedische wetenschap, een uitzondering te zijn met betrekking tot zijn TLRs. De zebravis heeft namelijk meer TLRs dan de meeste dieren en zo ook heeft de zebravis twee TLR5 genen, TLR5a en TLR5b, terwijl bijna alle andere dieren maar één TLR5 gen hebben. Wat de rol is van de twee

zebravis TLR5 genen in de herkenning van flagelline was niet bekend. In **Hoofdstuk 4** hebben wij ontdekt dat allebei de TLR5 varianten van de zebravis nodig zijn om flagelline te herkennen als een zogenaamd heterodimer complex. Dit is wezenlijk anders dan flagelline herkenning door een TLR5 homodimer complex zoals dat werkt in zoogdieren, vogels en reptielen. Om beter te begrijpen hoe de twee TLR5 varianten van de zebravis tezamen als heterodimeer complex dezelfde functie kunnen hebben als het homodimeer complex van andere diersoorten, hebben we een reeks experimenten uitgevoerd met kunstmatige combinaties van receptoren (zogenoemde chimeric receptoren) waarvan het ontwerp was gebaseerd op de eerder gepubliceerde structuur van zebravis TLR5b. Dit onderzoek toonde aan dat er subtiele verschillen zijn tussen zebravis TLR5a en TLR5b en humaan TLR5 die de herkenning van flagelline beïnvloeden maar hoe dit exact werkt kon op basis van de gepubliceerde structuur niet worden ontrafeld. Dit geeft aan dat het huidige model van flagelline binding aan zebravis TLR5b niet erg geschikt is om te voorspellen hoe TLR5 van bijvoorbeeld de mens een interactie aangaat met flagelline. Additionele experimenten met de gemaakte receptoren lieten ook zien dat meerdere regionen in de zebravis TLR5a en TLR5b receptoren in een heterodimeere configuratie moeten staan om het receptor complex functioneel te krijgen. Deze bevinding suggereert dat voor de activatie van TLRs een multiregionale verandering in de structuur van het complex, misschien middels rotatie van de receptoren, moet plaatsvinden om uiteindelijk het immuunsysteem te activeren.

TLRs zijn eiwitten die bestaan uit meerdere regionen. Van de meeste regionen binnen een TLR is de functie al bekend. Zo is de extracellulaire regio betrokken bij het binden van het ligand en de intracellulaire regio bij het doorgeven van signalen binnen de cel. In **Hoofdstuk 5** kwamen we erachter dat het allerlaatste stuk van de humane TLR5, wat het staartstuk wordt genoemd en waarvan voorheen geen functie bekend was, belangrijk is voor de receptor om op de juiste locatie in de cel te komen en het immuunsysteem te activeren. Wanneer we het staartstuk van de receptor verwijderde werd de receptor niet meer op de cel membraan geplaatst en kon het niet meer reageren op flagelline. Hetzelfde gebeurde toen we de aminozuren van het staartstuk in een andere, willekeurige, volgorde plaatsten. Vervanging van het humane staartstuk van TLR5 voor het staartstuk van zebravis TLR5b, wat ook een andere volgorde heeft maar evolutionair relevant is, herstelde de juiste lokalisatie van de receptor en de reactie op flagelline. Deze bevindingen geven aan dat de functie van het staartstuk van TLR5 afhankelijk is van een bepaalde volgorde van de aminozuren. We vonden dat het TLR5 staartstuk van

mens en zebravis allebei potentieel gefosforyleerd konden worden terwijl de voorspelling van fosforylatie van het willekeurig veranderde staartstuk een stuk zwakker was. Onze resultaten tonen hiermee aan dat het staartstuk van TLR5 een belangrijke rol speelt in de lokalisatie en functie van de receptor en mogelijk wordt dit bewerkstelligd door fosforylering van het staartstuk. Deze kennis biedt nieuwe kansen om partner moleculen van TLRs te identificeren en om de TLR5 functie te beïnvloeden door het staartstuk van de receptor te manipuleren.

Het behoud van TLR5 in bijna alle vertebrate dieren geeft aan dat TLR5 heel stabiel geëvolueerd is en dat flagelline herkenning een belangrijke eigenschap is die niet gemist kan worden. Niet alle TLRs zijn zo geëvolueerd als TLR5. Leden van de TLR1 subfamilie zijn meer dynamisch geëvolueerd in vertebrate dieren en sommige leden van deze subfamilie zijn verloren gegaan of juist gedupliceerd in bepaalde dieren. TLR15 is zo'n lid van de TLR1 subfamilie. Eerder onderzoek suggereerde dat TLR15 alleen maar aanwezig is in vogels en reptielen. Echter, veel was nog onbekend over de evolutionaire geschiedenis van TLR15. Ook was het niet bekend of de functie van TLR15 als receptor voor eiwit vernietigende enzymen, zogenaamde proteases geproduceerd door micro-organismen, behouden is gebleven tussen vogels en reptielen. **Hoofdstuk 6** beschrijft ons onderzoek naar de evolutie van TLR15 van vogels en reptielen. Na het analyseren van een grote collectie TLR eiwitten van verschillende vertebrate dieren waren we verrast door de bevinding dat TLR15 ook aanwezig is in een haaien soort. Aangezien TLR15 niet gevonden werd in zoogdieren, amfibieën en vissen betekent de aanwezigheid van TLR15 in een haai dat de receptor door heel veel diersoorten is verloren. Deze bevinding toont het tegenovergestelde aan van de eerdere veronderstelling dat TLR15 ontstaan zou zijn in de voorouder van vogels en reptielen. Het verlies van TLR15 in veel diersoorten, waaronder ook schildpadden, kan aangeven dat TLR15 niet heel belangrijk is geweest voor de overleving van die soorten. Desondanks lieten onze functionele studies zien dat de receptoren van de groene anolis hagedis, een krokodil en een alligator geactiveerd kunnen worden door proteases van een schimmel, net als dat het geval is voor de TLR15 van de kip. Deze resultaten geven aan dat de functie van TLR15 behouden is gebleven gedurende 280 miljoen jaar sinds vogels en reptielen gescheiden zijn. We vonden ook een verschil in de efficiëntie waarmee de cel TLR15 eiwitten van verschillende diersoorten maakt. Daar waar TLR15 van de kip en de groene anolis efficiënt geproduceerd werden was dat niet het geval voor TLR15 van de alligator en de krokodil. Dit leidde tot de ontdekking dat het codon gebruik in TLR15 genen erg variabel is tussen verschillende diersoorten, veel meer dan het geval is voor andere TLRs. Tezamen laten deze bevindingen zien dat TLR15 veel dynamischer is geëvolueerd dan

andere TLRs en dat een verschil in codon gebruik mogelijk gebruikt kan worden als parameter om de stabiliteit van de evolutie van een gen te bepalen.

Alle TLRs zijn zo geëvolueerd dat ze liganden herkennen van micro-organismen die heel belangrijk zijn voor de levensvatbaarheid van de micro-organismen. Op deze manier is de kans klein dat micro-organismen hun ligand verliezen en daardoor onzichtbaar worden voor TLRs. Om die reden is het opvallend dat TLR15 van vogels en reptielen proteases van micro-organismen herkennen omdat deze geen essentiële componenten zijn van de micro-organismen. Dit leidde ons tot vragen over het scala van micro-organismen die herkend kunnen worden door TLR15 en ook of er verschillen zijn in de herkenning van proteases tussen de TLR15 van verschillende dieren. In **Hoofdstuk 7** vonden we dat TLR15 van de alligator reageerde op ziekmakende bacteriën, een schimmel en een gist terwijl TLR15 van de groene anolis alleen reageerde op de schimmel. TLR15 van de kip en krokodil lieten ook verschillende reacties zien op deze micro-organismen. Deze resultaten geven aan dat de herkenning van proteases door TLR15 heel erg diersoort specifiek is. Een van de bacteriën die we hadden getest was goed in staat om TLR15 van de alligator te activeren maar niet TLR15 van de krokodil. Dit was erg interessant omdat de TLR15 van een alligator en krokodil bijna identiek zijn. Dit wijst erop dat een paar veranderingen van aminozuren in TLR15 al kan leiden tot een andere gevoeligheid voor het ligand, en dat terwijl TLR15 uit ongeveer 875 aminozuren bestaat. Onze zoektocht naar een specifieke protease die TLR15 kon activeren leidde tot de identificatie van een wel bekende protease genaamd LasB van de ziekmakende bacterie *Pseudomonas aeruginosa*. De instabiele evolutie van het ligand voor TLR15, welke als niet essentiële component ook niet altijd aanwezig is in alle micro-organismen, kan deels verklaren waarom de gevoeligheid van TLR15 voor zijn ligand heel diersoort specifiek is en waarom TLR15 niet altijd belangrijk is geweest in alle dieren.

In **Hoofdstuk 8** zijn onze nieuwe bevindingen opgenomen in een algemene discussie over verschillende aspecten van TLR biologie. Een van de besproken onderwerpen zijn de factoren die receptor-ligand evolutie beïnvloeden met TLR5 en TLR15 als aantrekkelijke voorbeelden vanwege hun tegenovergestelde evolutie (stabiel en dynamisch). Andere TLR aspecten worden beschreven vanuit een cel biologisch perspectief en zijn o.a. receptor expressie, maturatie, transport en ligand interactie. Ook wordt de diersoort specificiteit van ligand herkenning besproken en de diverse implicaties die dit met zich mee brengt. Het hoofdstuk wordt afgesloten met perspectieven voor toekomstig onderzoek naar TLRs.

Tezamen genomen demonstreert het werk dat beschreven staat in de hoofdstukken van dit proefschrift de kracht van een op evolutie gebaseerd, diersoort vergelijkend onderzoek voor het beter begrijpen van de principes achter de functie van een eiwit. Hopelijk zullen de nieuwe inzichten die middels dit onderzoek behaald zijn anderen dienen en inspireren om de kennis van TLR biologie uit te breiden en toe te passen.

Dankwoord / Acknowledgements

En voordat ik het mij volledig realiseerde was het proefschrift dan toch echt af. Ik zou nog wel jaren door kunnen gaan op dit lab met het vullen van een proefschrift over TLR's en dat komt bovenal door alle mensen om mij heen. Zonder hen was dit proefschrift er nooit geweest. Daarom volgt hier een dankwoord aan hen die direct bijgedragen hebben aan de totstandkoming van dit werk alsook aan iedereen die mij vanaf de zijlijn heeft gesteund.

Allereerst de Chiefs: Jaap en Jos. **Jaap**, eigenlijk is mijn hele PhD avontuur begonnen met jouw welwillendheid om mij in 2009 een kans te geven als student-medewerker in jouw lab. Ik was toen nog maar een tweede jaars HLO student en had nooit verwacht dat ik in de zomer van 2010, en de daaropvolgende zomers, terug mocht komen om het Campylobacter werk voort te zetten. Ik wil je in het bijzonder bedanken voor je vertrouwen in mij, zeker toen ik naar het lab van Martin Blaser in New York mocht, dat was echt super gaaf! Ook wil ik je bedanken voor alle goede tijden die we samen met Jos op de vrijdagmiddagen hebben gehad tijdens de werkbesprekingen. Jouw gevoel voor humor en voor goede science maakten dat ik altijd zin had in de werkbesprekingen. Je werklust en betrokkenheid bij iedereen om je heen zijn heel inspirerend. Ik ben blij dat ik van jou heb mogen leren en ik hoop dat onze paden elkaar in de toekomst nog vaak kruisen!

Jos, dankzij jou is dit PhD avontuur een succes geworden. Bedankt dat je dit avontuur met mij aangedurfd hebt. Ik weet nog goed dat ik in het begin bij jou aan tafel zat en zei dat dat moleculaire werk niet bij mij past en dat ik daar eigenlijk niks van wou weten. Gelukkig wist jij wel beter. Jouw passie voor het begrijpen van moleculaire mechanismen is aanstekelijk gebleken, ik zou nu niets anders willen bestuderen! Ik vind je een uitzonderlijk goede wetenschapper en mentor, en heb waanzinnig veel van je geleerd, zowel wetenschappelijke als sociale skills. Alleen heb ik de "mezelf verkopen" skill nooit volledig onder de knie gekregen, maar dat vind ik niet zo heel erg ;). Ik vind het erg fijn dat je altijd super toegankelijk bent en kijk met veel plezier terug op onze vele, altijd interessante gesprekken die gingen over baanbrekende science maar ook over politiek, geschiedenis, toekomst en nog veel meer. Mede door jouw brede interesse heb ik geleerd om altijd een open-mind te houden wanneer je iets gaat onderzoeken. Bedankt dat je van mij een echte wetenschapper hebt gemaakt!

Nancy, ik ben heel blij dat ik samen met jou op een lab heb gezeten en wil je bedanken voor de ontelbare antwoorden op vragen en goede suggesties om mijn experimenten te verbeteren. Ook heb ik genoten van onze vele gesprekken over

van alles en dan met name over onze vogels. Voor jou is 2019 ook een bijzonder jaar en ga je met je welverdiende pensioen. Ik zou niet weten hoe het lab zonder jou nog functioneert, je bent echt onmisbaar. Geniet van al je vrije tijd straks, maak mooie reizen, bak lekkere dingen en geniet van je mooie tuin met al het leven daarin!

Linda, jou wil ik bedanken voor alle goede gesprekken en alle lol die we hebben gehad. Vooral Lieneke was een dankbaar slachtoffer voor de vele grappen waar we uiteindelijk met z'n allen heel veel om hebben gelachen. Je bent een bijzonder persoon met een uitzonderlijk goed gevoel voor mensen en dat is belangrijk voor het lab, meer dan je zelf zou denken.

Marc, ondanks dat we best wel verschillende onderwerpen bestuderen heb je altijd interesse getoond in mijn werk en had je vaak goede suggesties voor mijn experimenten vanuit een andere invalshoek waarmee ik weer vooruit kon, veel dank daarvoor! Ik ga ook onze leuke en diepgaande gesprekken missen over dieren, de toekomst van de wetenschap en zelfs de toekomst van de aarde die mijn visie hebben verbreed. Ik hoop dat veel van de scenario's die we besproken hebben geen werkelijkheid worden en dat we nog lang van de natuur kunnen genieten. Ik wens je nog heel veel mooie ontdekkingen aan Campylobacter toe!

Karin, ook wij hebben veel leuke gesprekken gehad over dieren, met name over de kikkers bij jullie in de vijver en hoe wonderlijk hun metamorfose is. Bedankt voor je begeleiding tijdens mijn master stage en alle keren dat ik bij je aan mocht kloppen voor leuke plasmides of advies. Ik vind je een hele goede wetenschapper en bewonder je kennis, creativiteit, drive en doorzettingsvermogen met de mucines, wat toch eigenlijk wel *%$&! eiwitten zijn. Ik hoop dat onze (veel te laat gestarte) samenwerking met Xinyue aan MUC1 en TLR5 iets moois gaat opleveren. Heel veel succes met de mucines en dat er nog maar veel mooie papers uit mogen komen!

Marcel, ik vind het ontzettend cool dat jij in 2015 terug kwam naar dit lab! Jouw onderzoekslijn heeft het lab een hele nieuwe en super interessante richting opgestuurd en het heeft in ieder geval mijn interesse heel sterk beïnvloed. Ik denk met veel plezier terug aan onze vluchtige brainstorm sessies over microbiota en alle ideeën die daar ontstonden, we hebben de basis gelegd voor tientallen Nature papers ;). Ik vind het jammer dat we nooit echt samen hebben gewerkt, hopelijk gebeurd dat nog eens en hopelijk gaan we ook nog onze quattro colori reporter verwezenlijken. Ik kijk erg tegen je op als wetenschapper en wil je bedanken voor al je advies, overgedragen kennis en de leuke gesprekken over de collectieve intelligentie van mensen... Ik kan niet wachten op dat Nature paper!

Medi, of **Captain!** met jou was ieder gesprek interessant of het nou ging over Campylobacter, kippen, Iran, wereldpolitiek of onze vriend in de USA, ik vond het

altijd leuk en leerzaam, dank! Ik wens je heel veel succes met je nieuwe baan en hoop dat je toch ooit nog een kans vind om alle positieve effecten van opium nader te onderzoeken ;)

Lieneke, zonder jou was ik nooit aan dit PhD project begonnen. Jouw begeleiding heeft mij een hele goede head-start gegeven. Ik heb waanzinnig veel van je geleerd, dank daarvoor! Zoals jij ook in jouw proefschrift hebt vermeld hebben wij allebei een sterke mening wat tot veel interessante en leerzame discussies heeft geleid. Maar we hebben ook heel veel gelachen, ik ken niet veel mensen met zo'n goed gevoel voor humor en met name het over en weer citeren van South Park vond ik fantastisch!

Kasia, I felt quit sad when you left the lab because your presence gave the whole lab a good vibe. I have never before met somebody with such a positive spirit and the ability to see a good thing in every bad thing. I have learned a lot from you, especially because you are a great teacher. Thank you for all your insights and advice about the PhD life and the quest for post-doc jobs, it has helped me a lot! I wish you all the best at ETH and wherever you go next and can't wait to see some of your great Cryo-EM structures published!

Claudia, ook jou aanwezigheid had een sterk positief effect op het lab, met jou erbij was het nooit saai. Ik heb veel van je geleerd, zowel op technisch niveau, waaronder alle western blot tips en tricks als op persoonlijk niveau. Ik bewonder enorm hoe jij je PhD afgerond hebt. Je bent voor mij het toonbeeld van doorzettingsvermogen! Ook wil ik je bedanken voor alle leuke I&I borrels en voor het introduceren van mij bij de borrelcommissie. Zonder jou was ik daar zeker weten nooit beland… ;)

AX, ik ben erg blij dat wij een kamer hebben gedeeld. Dankzij jou heb ik het kloneren onder de knie gekregen en heb ik ook geleerd om software een stuk efficiënter te gebruiken… Ik ben altijd erg onder de indruk geweest van jouw werk, hoe je het uitvoerde en hoe je het presenteerde en ook hoe je je PhD zo soepel doorlopen hebt. Ik herinner mij vooral veel leuke discussies over sport. Hoe je altijd met allerlei feiten en weetjes mij probeerde te overtuigen dat de Tour de France echt wel leuk is… bedankt voor je volharding daarin! Ik wens je heel veel succes in Bordeaux en kijk uit naar je papers!

Guus, dude, jij bent voor mij het bewijs dat er ook PhD studenten zijn die wel een leven hebben. Ik benijd hoe jij naast al je werk ook nog gaat squashen, klimmen, spinnen, fietsen etc. waar haal je de tijd en energie vandaan? Misschien een optimaal microbioom? Ik ga onze gesprekken over gut microbes missen en hoop toch echt nog een keer samen te gaan gamen. Bedankt dat je mijn paranimf wou zijn! Het duurt niet lang meer voordat jij aan de beurt bent. Ik weet zeker dat

je een super interessant proefschrift af gaat leveren en terecht komt op een mooie plek, misschien wel nabij de Rockies!

Xinyue, my favorite colleague! Thank you for being my other paranimph, you and Guus will do great. I enjoyed your company in the lab very much. Our many, often long conversations have cost me a lot of time that I could have spend otherwise but I would not have missed them for anything! You may not see it or like it but you are on your way to become an excellent researcher and I hope you will combine this with your fun and kind character to continue to strengthen science in the future! It will not be long before you will hand in your thesis and I know it will be great. Hopefully we can add a chapter about the interaction between MUC1 and TLR5, one of which is the best protein in the world, and it is not MUC1 ;).

Jiannan, I am very happy that you joined our lab! You are a very nice and helpful guy. I am very impressed by your skills in the lab, you perform experiments in the way that everybody should. You are also very tidy, which is good because I know exactly where all your stuff is in the freezer... You will become a great researcher but also remember to enjoy other things like basketball, maybe you can even win some money with it to pay me for all those expensive questions you asked ;).

Yaro, de gesprekken met jou waren altijd leuk en interessant. Vooral als je weer los ging met je ideeën om problemen op te lossen of systemen te verbeteren. Jouw creativiteit is inspirerend! Ik vind het knap dat je de keuze hebt gemaakt om de pilli van B. frag achter je te laten en je meer te richten op dat waar je gelukkig van wordt, namelijk uitvinden. Ik weet zeker dat Yaro Inc. werkelijkheid wordt en dat wat je dan ook gaat uitvinden het lab leven dragelijker gaat maken, succes!

Xuefeng, what's up brahhh?! I am glad to have you as a roommate, every afternoon you cook some very good smelling food in the microwave which makes me very hungry and eager for some good Chinese food. Thanks for sharing your knowledge about K. Raikonen and Manchester United, those are topics about which I literally know nothing... It is very nice to see you getting better and better in the lab and I think your ideas about Campylobacter host-cell interaction are very interesting and worthwhile investigating. Keep going strong brah and remember to chill!

Ook wil ik iedereen van de MHD waaronder **Henk**, **Edwin**, **Martin**, **Albert**, **Hanne**, **Soledad**, **Tryntsje**, **Maarten**, **Maaike**, **Lianci** en **Melanie** bedanken voor alle leuke labuitjes en interessante WIP's. **Victoria**, jou wil ik bijzonder bedanken voor al je adviezen en onze leuke gesprekken, ik hoop dat we die nog eens voort kunnen zetten, en dan graag opnieuw in Alpbach! **Roel**, jij ook in het bijzonder

bedankt voor je interesse in de TLRs en de goede tijden tijdens de bordspelletjes en het op sleeptouw nemen van ons met boulderen!

Ook veel dank aan iedereen van de KLIF, en in het bijzonder **Rolf**, **Frans**, **Niels** en **Arjen** voor alle humor en goede gesprekken, **Aldert** voor de bioinformatica lessen, **Birgitta**, **Koen** en **Linda** voor alle begeleiding bij het labwerk, **Haitske** voor het leuke veldwerk bij de boeren en **Marta** en **Ricardo** for the good PhD talks!

En ik zal ook zeker niet vergeten om de hardcore van het VMDC: **Carolien**, **Anky**, **Marianne**, **Wim**, **Sylvia** en **Wendy** te bedanken voor alle goede tijden en begeleiding in het begin en het regelmatig kweken van bijzondere stammen voor mijn onderzoek.

En dan mijn mede herpetofielen. **Maarten**, een simpele groet aan jou in de wandelgangen liep altijd uit op een boeiend gesprek van minimaal een half uur over mooie oorden vol reptielen en amfibieën. Ik ben nog altijd jaloers op je voor het hebben van twee geweldige banen waarin je enerzijds nieuwe Campylobacter soorten ontdekt en anderzijds via RAVON de kennis over de Nederlandse herpetofauna vergroot. Ik vind het super cool wat je doet! Heel veel plezier in Australië, ik kom je zeker weer opzoeken om je verhalen te horen. **Marja**, ik was in het begin een beetje bang voor je omdat je nogal direct bent. Maar gaande weg onze samenwerking aan het Anolis TLR5 paper kwam ik erachter dat je eigenlijk waanzinnig aardig bent. Ik waardeer enorm je directheid en eerlijkheid, bij jou zijn dingen gewoon duidelijk! Ook geniet ik altijd van je verhalen over Aruba en wat je meemaakt in je praktijk. En wat fijn dat ik altijd bij je terecht kan als ik weer in lichte paniek ben vanwege een schildpad of hagedis waar iets mee is, heel veel dank daarvoor! **Bernardo**, I am very glad that I got to know you during our PhD. You are a super nice guy and I cherish all our interesting conversations. I deeply respect how you went through your PhD and defended your thesis under those unique circumstances. I am also a huge fan of your work, which I think makes a real contribution to the understanding and conservation of amphibian ecology. I hope you will have the chance to continue this work in the future and I wish you and your family all the best! **Victor**, ik ben heel erg blij dat je toendertijd contact met mij zocht voor een SEM studie van de eischalen van C. signatus. Ik heb veel geleerd van dit project met jou en je hebt er een heel mooi paper van gemaakt. Ik bewonder jou enorm, je bent voor mij het ultieme voorbeeld van een echte, door passie en nieuwsgierigheid gedreven wetenschapper. Jouw studies naar de ecologie van schildpadden vind ik toonaangevend en van groot belang. Bedankt dat ik met je mee mocht naar Zuid-Afrika, het was een waar avontuur dat mij nog lang zal heugen, zowel in negatieve maar vooral in positieve zin!

Ruben, **Manon**, **Nienke** en **Lars**, jullie heel erg bedankt voor alle fun tijdens meetings en het pompoensnijden, sumoworstelen, spelletjesavonden, paaschallenge, de buitenbioscoop (mislukt...), bier en wijn tests en uiteraard de PubQuizes. Ik had nooit verwacht dat in een borrelcommissie zitten zo leuk zou zijn en dat is 100% aan jullie te wijden. Bedankt ook dat jullie mij altijd de Japanse mix gunde, ik weet dat jullie het ook heel lekker vinden en het moet niet makkelijk zijn geweest om ieder keer mij die kilo te zien weg werken!

Ook veel dank aan mij studentes **Thessa**, **Sara** en **Caren**! Ik vond het heel leuk om jullie te begeleiden en heb daar heel veel van geleerd, dank voor jullie geduld met mij en ik wens jullie allen het beste toe!

Richard, **Esther** en **Rob**, jullie hulp met en adviezen over de microscopen hebben mijn werk naar een hoger niveau gebracht. Richard, jouw kennis en enthousiasme maken ieder gesprek interessant en maakte dat ik meteen weer cellen wou gaan uitzaaien op een glaasje. Esther, bedankt voor al je uitgebreide en zorgvuldige uitleg van de apparatuur, door jou heb ik mooie plaatjes leren maken. Rob, heel veel dank dat je iedere vrijdagochtend de microscoop alvast aanzette, dat heeft toch een hoop tijd gescheeld. Ook veel dank voor de leuke gesprekken over parkieten!

Also special thanks to **Martin**, **Guillermo** and **Sabine** for helping me in your lab and making me feel at home in New York!

Many thanks to **Mark**! I am very happy that you were interested in starting a collaboration with us and were so fast with sending the crocodile DNA samples. Thanks to the material you send, chapter 6 became a very interesting story with some great starring species!

Jean-Paul, **Raymond** en **Ronald**, mijn ex-docenten van de HU, ik wil jullie bedanken voor alle tijd die jullie hebben genomen om al mijn talloze vragen te beantwoorden alsook voor het aanmoedigen van mij om door te studeren aan de universiteit. Jean-Paul, ik kom zeker nog een keer bij je langs om van je te leren hoe ik een marine aquarium moet opzetten en onderhouden!

Geert, **Maria** en **Huub**, mijn zeer gewaardeerde collega's van Wageningen University, bedankt voor de goede tijden tijdens mijn master en dank dat ik mijn stages in Utrecht heb mogen doen. Ook bedankt voor de uitnodigingen om als gastspreker bij jullie te mogen presenteren, ik voel me nog altijd vereerd. Ik hoop dat we elkaar in de toekomst nog vaker treffen!

Gelukkig heb ik tijdens mijn PhD de nodige afleiding kunnen vinden in het maken van vette muziek met veel hele coole en getalenteerde muzikant-vrienden. Heel veel dank voor alle goede muzikale tijden aan: **Lennart**, **Dirk** en **Johan** van Fat Connenction, **Jan**, **Leo** en **pa** van de Jan Deurloo band en **Dick**, **bro** en **pa** van John, Dick and Sons. And a very special thanks to the members of the best band in

the world! **Marcel**, you lay down a sick bass line man, your playing is insanely cool! **Roel**, is there anything you cannot play with those amazing piano fingers while singing great Chinese lyrics? **Yaro**, you know some serious sweet blues licks man, keep playing! **Jiannan**, the man with the iron throat, it is great seeing you go all the way with the Black Leopard song, you are a proper rock star! And finally **Alexia**, the lady with the silk throat, I have never played music with somebody who can sing so good, please don't stop singing, ever!

Wouter, zonder jouw hulp met de wiskunde tijdens mijn pre-master was ik waarschijnlijk nooit toegelaten aan Wageningen University, heel veel dank daarvoor! Ook kijk ik met veel plezier terug op onze avonden kletsen over werkelijk alles, gesprekken met jou zijn altijd interessant en ik hoop dat er nog velen komen.

Kevin, Kim, Caressa, Eva, Rob, Vincent, Eddy, Nel, Anja, Erik en Césaf, het is dan eindelijk zover, ik ben niet meer in opleiding, of met andere woorden, ik ben afgestudeerd. En dat was nooit gelukt zonder jullie. Ik ben heel erg blij met jullie als mijn familie en koester alle herinneringen aan ons samenzijn. Bedankt voor jullie interesse in mijn onderzoek, het zal niet altijd makkelijk zijn geweest om de verhalen over receptor evolutie te volgen maar ik ben blij dat ik het heb mogen vertellen aan jullie. Kevin, jou wil ik in het bijzonder bedanken voor je vertrouwen in mij. Het heeft me altijd veel goeds gedaan om van mijn broer te horen dat ik het wel aankon, dank! Jij bent dit jaar ook officieel "afgestudeerd" en je weet niet half hoe trots ik daarop ben!

Pa, bedankt voor alles. Bedankt dat je mij altijd mijn gang hebt laten gaan en dat je al mijn keuzes gesteund hebt. Bedankt dat je mij hebt geleerd om naar alles nieuwsgierig te zijn en om goed te luisteren en te kijken naar dat wat interessant is. Ik ga het frequent samen muziek maken heel erg missen en vind dat we iedere keer dat we elkaar zien, de gitaren erbij moeten pakken.

En als laatste gaat mijn grootste dank uit naar **Diënty**. Zonder jou was me dit nooit gelukt. Zonder jou was ik allang drie keer gestopt. Bedankt voor je onvoorwaardelijke steun en geduld. Het zal niet altijd leuk zijn geweest om getrouwd te zijn met een PhD student en het spijt me voor alle keren dat ik veel te laat thuis kwam en chagrijnig was omdat het schrijven niet lukte. Maar door jouw begrip en vertrouwen in mij heb ik het gehaald, daarom is dit proefschrift eigenlijk een team-effort. Ik hoop dat je mij de rest van je leven wilt blijven steunen en dat je lekkere dingen wilt blijven maken zoals de anolis-brownies, zebravis-bonbons en krokodillen-taarten om onze prestaties te vieren. Ik heb heel veel zin om met jou het volgende avontuur te beginnen!

Curriculum Vitae

Carlos Voogdt was born on the 18th of January 1987 in Zeist, the Netherlands. He graduated from pre-vocational secondary education at VMBO Schoonoord, Zeist in 2003. From 2004 to 2008 he followed vocational education at Aeres MBO, Barneveld and graduated in 2008 as a paraveterinarian. From 2008 to 2012 he followed higher professional education at the HogeSchool Utrecht and graduated cum laude, obtaining a bachelor degree in Biology and Medical Laboratory Science with a specialization in zoology. During his bachelors he conducted two internships, one at the Vrije Universiteit Amsterdam on the subject of cardiovascular disease in a rat animal model and one at the Utrecht University on the subject of maternally transmitted behavior in laying hen chicks. In the consecutive summers of 2009 until 2012 he worked as a research assistant at the Laboratory of Clinical Infectiology of the Utrecht University under the supervision of Prof. dr. Jaap Wagenaar, helping with large scale culturing of *Campylobacter* species isolated from poultry. In the winter of 2011 Carlos worked for two weeks at the laboratory of Prof. dr. Martin Blaser at the New York University, United States, collecting strains of *Campylobacter fetus*. In 2012, Carlos started the master program Animal Sciences at the Wageningen University, the Netherlands from which he graduated in 2014. During his masters he conducted two internships at the Laboratory of Infection Biology of the Utrecht University under the supervision of Prof. dr. Jos van Putten. One internship focused on the development of a quantitative PCR for immune genes of the green anole lizard. The subject of the other internship was the Toll-like receptor 5 of the green anole lizard. After obtaining his masters degree, Carlos continued studying Toll-like receptors of animals in a PhD project at the Laboratory of Infection Biology, Department of Infectious Diseases and Immunology of the Utrecht University. The results obtained from this PhD project are described in this thesis.

Publications

1. V.J.T. Loehr, **C.G.P. Voogdt**, D. Nolan. Ultrastructure of Eggshells from Wild and Captive Speckled Dwarf Tortoises, *Chersobius signatus*. Herpetologica, 2019 March 75;1:63-68. doi: 10.1655/0018-0831-75.1.63.

2. **C.G.P. Voogdt**, M.E. Merchant, J.A. Wagenaar, J.P.M. van Putten. Evolutionary regression and species-specific codon usage of TLR15. Frontiers in Immunology, 2018 Nov 13;9:2626. doi: 10.3389/fimmu.2018.02626.

3. **C.G.P. Voogdt**, J.A. Wagenaar, J.P.M. van Putten. Duplicated TLR5 of zebrafish functions as a heterodimeric receptor. Proceedings of the National Academy of Sciences of the United States of America (PNAS), 2018 Apr 3;115(14):E3221-E3229. doi: 10.1073/pnas.1719245115.

4. I.R. Fink, D. Pietretti, **C.G.P. Voogdt**, A.H. Westphal, H.F.J. Savelkoul, M. Forlenza, G.F. Wiegertjes. Molecular and functional characterization of Toll-like receptor (Tlr)1 and Tlr2 in common carp (*Cyprinus carpio*). Fish and Shellfish Immunology, 2016 Sept 56:70-83. doi: 10.1016/j.fsi.2016.06.049.

5. **C.G.P. Voogdt**, L.I. Bouwman, M.J. Kik, J.A. Wagenaar, J.P.M. van Putten. Reptile Toll-like receptor 5 unveils adaptive evolution of bacterial flagellin recognition. Scientific Reports, 2016 Jan 7;6:19046. doi: 10.1038/srep19046.

6. **C.G.P. Voogdt**, J.P.M. van Putten. The evolution of the Toll-like receptor system. In D. Malagoli (Ed.), The evolution of the immune system – conservation and diversification, Elsevier, 2016 pages 311-330.

7. **C.G.P. Voogdt.** Schildpadden, krokodillen en slangen, heelmeesters van de toekomst? RAVON, 2016 June 61.

List of abbreviations

AP	Adapter protein
CJM	Cytoplasmic juxta membrane
DDC	Duplication, degeneration, complementation
ECD	Extracellular domain
EJM	extracellular juxta membrane
ICD	Intracellular domain
LAMP-1	Lysosome associated membrane protein 1
LPS	Lipopolysaccharide
LRR	Leucine rich repeat
LRRNT	N-terminal LRR
MAMP	Microbe associated molecular pattern
MYA	Million years ago
MyD88	Myeloid differentiation primary response gene 88
NF-kB	Nuclear factor kappa light chain enhancer of activated B cells
PAMP	Pathogen associated molecular pattern
PAR	Protease activated receptor
RLU	Relative light units
TIR	Toll/Interleukin-1 receptor homology domain
TLR	Toll-like receptor
TM	Transmembrane domain
UNC93B1	UNC-93 homolog B1