Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Structural insights into ligand recognition and regulation of nucleic acid-sensing Toll-like receptors

Toshiyuki Shimizu

Toll-like receptors (TLRs) activate the innate immune system in response to invading pathogens. Although nucleic acids are one of the principal TLR ligands, they are not inherently pathogen-specific and, thus, carry the risk of triggering autoimmunity. There are multiple unique regulatory mechanisms aimed at preventing accidental activation of nucleic acid-sensing TLRs. Recent structural studies revealed that different nucleic acid-sensing TLRs have specific modes of recognizing nucleic acids as ligands regulated by diverse regulation mechanism both at the receptor and ligand levels. This review summarizes structural knowledge on the ligand recognition and regulation mechanism by nucleic acid-sensing TLRs.

Address
Graduate School of Pharmaceutical Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Corresponding author: Shimizu, Toshiyuki (shimizu@mol.f.u-tokyo.ac.jp)

Current Opinion in Structural Biology 2017, **47**:52–59

This review comes from a themed issue on **Protein–nucleic acid interactions**

Edited by **Stephen Cusack** and **Christoph Müller**

http://dx.doi.org/10.1016/j.sbi.2017.05.010

Introduction

The innate immune system senses pathogen-associated or cell damage-associated structurally conserved molecules through various pattern recognition receptors (PPRs) [1]. The activation of membrane-bound PRRs such as Toll-like receptors (TLRs), and cytoplasmic PPRs such as Nod-like receptors (NLRs) and RIG-I like receptors (RLRs) triggers signaling cascades that induce the generation of pro-inflammatory cytokines and interferons. Among PPRs, TLRs, evolutionarily conserved membrane-spanning receptors homologous to the *Drosophila* Toll protein, are mostly expressed in macrophages and dendritic cells, where they recognize various pathogen-specific molecules, including carbohydrates, peptidoglycans, and nucleic acids [2].

Nucleic acids that encode and transmit genetic information in all living organisms, including viral, bacterial, and eukaryotic pathogens, are major substrates detected by the innate immune system. Several types of nucleic acid sensors including members of the TLR, NLR, RLR, and cGAS families have been identified [3–5]. Recent studies of TLR structure in several mammalian species revealed that nucleic acid-sensing TLRs (TLR3, TLR7/8, TLR9, and TLR13) recognize different nucleic acid patterns and transmit ligand-triggered signals through diverse molecular mechanisms (Table 1). In this review, I summarize recent structural data on the ligand recognition and regulation of nucleic acid-sensing TLRs.

General structural features of TLRs

TLRs are type I integral membrane receptors containing N-terminal extracellular leucine-rich repeats (LRRs), a transmembrane domain, and a C-terminal cytoplasmic domain [6,7] (Figure 1a). The LRR domain is responsible for ligand recognition, while the cytoplasmic domain known as the Toll/IL-1 receptor (TIR) region activates downstream signaling cascades by interacting with adaptor proteins such as MyD88, Mal, TRIF, and TRAM [8]. Biochemical and structural studies on TLR extracellular domains [9–13] suggest the following mechanism of TLR activation. Inactive TLRs exist as monomers and ligand binding induces dimerization of the extracellular TLR domain, producing a typical 'm' shaped structure, where the intracellular C-terminal regions of the two TLR protomers are positioned in close proximity [14]. Subsequent dimerization of the intracellular TIR domain is followed by the recruitment of adaptor proteins which execute signal transduction.

Nucleic acid-sensing TLRs and their pattern recognition mechanism

Viral/bacterial nucleic acids are potent stimulators of innate immunity. In humans, nucleic acids are recognized by TLR3 and TLRs 7–9; among them, TLR3 is activated by double-stranded (ds)RNA, while TLR7 and TLR8, which are closely related, recognize single-stranded (ss) RNA [15–17], and TLR9 senses DNA with unmethylated cytosine phosphate-guanosine (CpG) motifs [18–20]. In addition, mouse TLR13 detects bacterial 23S ribosomal (r)RNA [21,22] (Figure 1b). Contrary to other TLRs, TLR3 and TLRs 7–9 are expressed on endosomal membranes [23], which limits the recognition of self-derived ligands [24]. In addition, the members of the TLR7 subfamily comprising TLRs 7–9 contain a characteristic long inserted loop region (known as the

Table 1

Structures of nucleic acid-sensing TLRs

TLR	Species	Ligand/comment	Resolution (Å)	PDB ID	Refs.
TLR3	Human	−/apo form	2.1, 2.4	1ZIW, 2A0Z	[25,26]
	Mouse	−/apo form	2.7	3CIG	[11]
		dsRNA(46 bp)	3.4	3CIY	
TLR7	Monkey	G + polyU	2.5	5GMF	[31•]
		Loxoribine + polyU	2.6	5GMG	
		R848	2.2	5GMH	
TLR8	Human	−/apo form	2.3	3W3G	[32]
		CL097, CL075,	2.0, 2.3,	3W3J, 3W3K	
		R848 (3 forms)	2.1–2.7	3W3L, 3W3M, 3W3N	
		DS-877	1.8	3WN4	[50]
		DS-802, XG-1-236	2.0, 2.1	4QBZ, 4QC0	[51]
		Hybrid-2	2.1	4R6A	[52]
		N-1-3, N-1-4	2.1, 2.1	5AWB, 5AWD	[53]
		Agonistic ssRNA[a]	2.0, 2.4, 2.6	4R07, 4R08, 4R09	[30••]
		Uridine	1.9	4R0A	
		Z-loop uncleaved	2.6	5HDH	[48•]
TLR9	Horse	−/apo form	2.4	3WPB	[34•]
		DNA_1668 (12 mer)	1.6	3WPC	
		iDNA4084	2.8	3WPD	
	Bovine	DNA_1668 (12 mer)	2.4	3WPE	
	Mouse	−/apo form	2.0	3WPF	
		iDNA4084, iDNA_super	2.3, 2.3, 2.3	3WPG, 3WPH, 3WPI	
		−/LRR15-25	2.4	4QDH	[54]
TLR13	Mouse	ssRNA (13 mer)	2.3	4Z0C	[35••]

[a] ORN06 (20 mer), ssRNA40 (20 mer), phophorothiated ORN06 (20 mer).

Z-loop) composed of approximately 30 amino acid residues (Figure 1b,c).

TLR3

TLR3 crystal structure was first resolved for the ligand-free form [25,26] and then for the dsRNA-bound form [11]. Two TLR3 protomers sandwich the dsRNA molecule which makes contact with both N- and C-terminal sites on the lateral side, generating an m-shaped structure (Figure 2a). The dsRNA containing 46 bp is approximately 130 Å long, which corresponds to the length of the overall complex (~140 Å) and is consistent with the fact that dsRNA shorter than ~40 bp cannot bind or activate TLR3 [27].

TLR3 protomers bend toward each other at one end to accommodate the dsRNA molecule. As a result, the interaction between the two TLR3 molecules occurs at their C-termini (Figure 2a), which are approximately 25 Å apart. Electrostatic interactions at the ribose-phosphate backbone play the major role in dsRNA recognition by TLR3, and several His residues associated with the dsRNA phosphate groups are essential for TLR3 activation. As protonation of these His residues occurs at low pH, successful binding of dsRNA to TLR3 requires the acidic environment [25,27].

TLR8 and TLR7

Although TLR7 and TLR8 have been considered to primarily recognize ssRNA, they are also activated by small synthetic ligands such as imidazoquinolines and nucleoside analogs [28,29], raising a question about the molecular basis for the recognition of these structurally and chemically distinct ligands. Crystal structure analyses have unraveled this mystery [30••,31•].

First, the structures of unliganded and small synthetic ligand-bound TLR8 have been elucidated [32]. Small synthetic agonists bind the TLR8 dimer at two symmetric positions, triggering the rearrangement of the dimeric configuration, which brings TLR8 C-termini into closer proximity, thus generating a more compact structure than that of the m-shaped dimer. Then, the crystal structure of TLR8 complexed with 20-mer ssRNA has been reported [30••]. Unexpectedly, this study revealed that TLR8, instead of binding the full-length ssRNA molecule, associates with RNA degradation products at two distinct sites (Figure 2b). The first site, which accommodates the ssRNA degradation product uridine, is identical to a previously reported binding site for small synthetic ligands [32]. The second binding site located at the interior of the TLR8 ring structure, is sandwiched between the concave surface and the Z-loop and holds ssRNA molecules longer than two nucleotides.